Original Research Article

What are neural networks *not* good at? On artificial creativity

Anton Oleinik

Abstract

This article discusses three dimensions of creativity: metaphorical thinking; social interaction; and going beyond extrapolation in predictions. An overview of applications of neural networks in these three areas is offered. It is argued that the current reliance on the apparatus of statistical regression limits the scope of possibilities for neural networks in general, and in moving towards artificial creativity in particular. Artificial creativity may require revising some foundational principles on which neural networks are currently built.

Keywords

Metaphor, social action, prediction, regression analysis, neural networks, creativity

Introduction

Significant progress in the development and practical implementation of neural networks leaves little doubt about their potential. 'A neural network is a massively parallel distributed processor made up of simple processing units that has a natural propensity for storing experiential knowledge and making it available for use' (Haykin, 2009: 1; see also Goodfellow et al., 2016: 438; Thaler, 2016a: 138). This basic 'connectivist' principle allows for dividing complex problems into multiple simple tasks that can be performed with the help of standard mathematical models,[1] with regression models being among the most commonly used.

Distributed processing paves the way for achieving important breakthroughs. Modelling of the XOR logical operation (the output is true only when inputs differ), in contrast to the AND or OR operations, has no linear solution. With the help of a neural network, the XOR operation can be broken down into several components that may have a linear solution (Haykin, 2009: 141; Jurafsky and Martin, 2018: chapter 8). Pattern recognition represents a particularly promising application of neural networks.

The breakthroughs come at a cost, however. Exact specifications of mathematical models that underpin neural networks are difficult to extract and make explicit. Synaptic connections joining processing units (neurons) are depositories of information about networks (Haykin, 2009: 2, 171; Nielsen, 2015; Thaler, 2016a: 138). Connection weight values can be compared with coefficients in regression models. In contrast to regression and other models, however, connection weight values are difficult to identify and, hence, to enter in a formal 'model', especially in the case of multi-layered networks with 'hidden' neurons. Similar to a 'black box', inputs (training data) and outputs (revealed patterns) of a neural network are known, whereas many internal parameters remain hidden. The relative lack of transparency complicates the formulation of 'if/then' rules with the help of which one spots and describes patterns in data (Thaler, 1998: 22). It also slows down progress in understanding limits of the reliance on neural networks.

This paper offers a view of a social scientist, albeit with a background in mixed research methods, on internal limitations of neural networks. It contains a critical reading of the recent developments in neural

Memorial University of Newfoundland, Canada; The Central Economics and Mathematics Institute of the Russian Academy of Sciences, Moscow, Russia

Corresponding author:
Anton Oleinik, Memorial University of Newfoundland, St. John's, NL, A1C 5S7, Canada.
Email: aoleynik@mun.ca

networks that may be of interest to scholars of computational culture.[2] The article does not have an ambition to offer a comprehensive 'archaeology' of neural networks (cf. Mackenzie, 2017), being focused rather on prospects for achieving artificial creativity with their help. Nor does this article offer a comprehensive overview of limitations of artificial creativity – a broader topic in artificial intelligence. The focus is placed on the inherent limits of neural networks as a driver of artificial creativity. The research question addressed in what follows is how sociological thinking helps better understand the limits of artificial creativity based on neural networks. It is argued that neural networks get us very close to artificial creativity without quite being able to perform several creative tasks. In the words of a sociologist, a robot powered by neural networks may be a good Actor, i.e. someone who closely follows the script, but not a Subject, i.e. someone who meaningfully changes and re-writes the imposed rules (Touraine, 1992: 238–239).

More specifically, this paper aims to explore some limits in the further evolution of neural networks that derive from the foundational principles on which they are based. Limits in the development of a technology can be either internally or externally set. Andrei Sakharov, a Soviet physicist, who worked on military applications of nuclear technologies in the 1950s, subsequently became an advocate for placing limits – by political and civic means – on their further improvement. The development of neural networks has also been influenced by military needs: namely, by a high priority given to target-recognizing ('smart') bombs (Thaler, 1998: 21). Attempts to ban the work on killer robots as the other military application of artificial intelligence in general and neural networks in particular represents an example of externally imposed limits. In August 2017, a group of leading robotics and artificial intelligence experts wrote an open letter to the United Nations calling for a ban on lethal autonomous weapons (Gibbs, 2017).

Particularities of technical solutions implemented at early stages in the development of a technology lie at the origin of the other – internal, set of limits. They determine the scope of the possible and cause 'lock-in' effects. Early adopters face significant costs if they wish to switch to a competing technology later on: they are 'locked in', literally (Arthur, 1989). The reliance on statistical regression, linear and logistic, limits the scope of possibilities for neural networks. Neural networks are very good at identifying patterns, but only if training data has a structured character.[3] Internal limitations of neural networks take particularly manifest forms when they deal with pattern changes in addition to pattern recognition. Artificial creativity powered by neural networks has a problematic character as a result.

Boden (1999: 351) defines creativity as 'the generation of ideas that are both novel and valuable'. She further differentiates (2009: 24–25) between three types of creativity: combinatorial (production of new combinations of known elements); exploratory (probing the boundaries of what is possible without radically changing existing procedures and approaches); and transformational (transforming these boundaries by altering procedures and approaches). Boden made contradictory statements as to what kind of artificial creativity is most difficult to achieve: combinatorial (2009: 25) or transformational (1999: 365). Whichever it may be, the key outstanding issue is the determination of what counts as a 'valuable' novelty.

Thaler (1998, 2016a) discusses more specifically artificial creativity driven by neural networks, seeing its source in perturbations and noise. According to him, creativity is the opposite of the 'training' of a network – something that shakes its operation by altering connection weight values. In other words, creative shocks undermine the 'normal' operation of neural networks.

Take the example of self-driving cars. Driverless cars powered by neural networks can be seen on the streets, but their performance is not yet very impressive. On the one hand, rules of the road are more difficult to define exhaustively: the traffic code represents their subset only. A number of situations remain unspecified by the law, such as an obstructed line with visibility to the horizon, with no oncoming vehicles but the solid centre line. Any reasonable driver would cross into the oncoming lane to pass the obstacle. A driverless car may well stop and wait there before the obstacle is cleared (Riley, 2017). On the other hand, as long as humans continue to drive cars – some of them are unpredictable drivers – neural networks will have a hard time to adjust to all manoeuvres of human drivers. Training data tend to be 'messy' as a result.

The paper has two sections, along with the introduction and conclusion. Several areas in which neural networks underperform compared with humans are identified in the first section. The second section highlights consequences of this relative underperformance on artificial creativity powered by neural networks.

Areas in which neural networks underperform

When assessing the performance of neural networks, it is conventionally measured against the results achieved by humans as a 'gold standard' (Amini et al., 2011: 1574; DiMaggio, 2015: 1; Huang et al., 2012: 1601; Jurafsky and Martin, 2018). Following this line of thinking, artificial creativity could be compared with human creativity. Instead of asking the question of whether a neural network could ever be 'really' creative,

a more relevant inquiry is whether neural networks under- or over-perform in creativity compared with humans. Attempts to develop artificial creativity cover a wide range of areas: literature, music, visual arts and problem solving (Computational creativity, n.d.). This article focuses on limits of artificial creativity based on neural networks in the identification and interpretation of symbols, proactive coordination and prediction.

Identification and interpretation of symbols

Symbols are conventionally understood as meaningful objects, i.e. objects to which a specific meaning is attached. Driving a car requires the ability to identify and read relevant symbols: road marking, road symbol signs and traffic lights. Objects with a symbolic value in one situation do not necessarily have it in another. Perlovsky and Ilin (2012: 807) use the term situation-symbol in this regard: 'situation "office" is characterized by the presence of a chair, a desk, a computer, a book, a book shelf. Situation "playground" is characterized by the presence of a slide, a sandbox, etc. The principal difficulty is that many irrelevant objects are present in every situation'. In order to navigate various situations, one needs to correctly establish their type (road, office, or playground) and then to identify relevant objects that have meaning in this context. Is a car model relevant for the 'road' situation, i.e. does it have an impact on successful (collision-free) driving?

Situation-symbols and object-symbols are socially embedded. There is no 'universal' system of symbols pertaining to driving. There are symbols specific to left- and right-hand traffic. In some countries, licence plates acquire a symbolical value: they serve to signal the privileged status of the car's user (Oleinik, 2016: 20). The same goes for the 'office' situation. This situation cannot be taken for granted everywhere and at any point in time. It results from the process of modernization and rationalization accompanied by 'the separation of the bureaucratic office as a "vocation" from private life, the bureau from the private household' (Weber, 1968: 379). Before this separation occurred, the office and the household shared a similar set of relevant objects. Even mathematical symbols, apparently the most context-free, have multiple interpretations; the meaning attached to them varies across particular communities (Kripke, 1982: 106–110).

It follows that identification and correct interpretation of symbols requires social competence with sensitivity to the context as one of its dimensions. It is here that semiotics provides a useful point of reference. Semiotics studies symbols regardless of their form, material, visual or textual, in a systematic manner. Semioticians consider symbols to be 'the most stable elements of the cultural continuum' (Lotman, 1990: 104). In a sense, symbols could be compared with 'plot-genes' of a culture. The evolution of these plot-genes depends on the specific context understood here as a set of other symbols.

Machine learning of symbols with the help of neural networks offers several solutions to the problem of identification and interpretation of symbols in general, and textual symbols in particular. Topic analysis and regression analysis based on Bayesian (conditional) probability are two of them. They both derive from the idea of co-occurrence of words in a text as a key for identification and interpretation of symbols that it contains.

Specific assumptions, on which these methods are based, and procedures, which they use, differ, however. In the case of topic analysis, principles of Latent Dirichlet Allocation are used for comparing a random allocation of a given 'bag' of words across various topics in a set of texts with the frequencies actually observed. Significant departures from the conditions of a random allocation are indicative of the existence of a topic understood as a constellation of words in which a symbol is embedded (Bail, 2014: 472; Evangelopoulos et al., 2012: 72). From this point of view, topic analysis allows for contextualizing a symbol. A topic embodies a particular combination of words and cannot be separated from those words even at the most basic, methodological level. This particular combination of words constitutes an interpretive frame for attaching a meaning to a person, event, organization, practice, condition, or situation (DiMaggio et al., 2013: 593).

Compared with a human reader of a text, topic analysis has a number of shortfalls. Its critics point to the arbitrariness in the choice of the number of topics to be discovered in a set of texts with the help of machine learning. This parameter has to be specified at the very beginning, similar to cluster analysis. Results are sensitive to the a priori assumptions about the number of possible topics (Bail, 2014: 472). 'Many people utilize topic models in an inductive manner that resembles reading tea leaves' (Bail, 2015: 2; see also Wiedemann, 2013: para 48). The absence of clear criteria for validating outcomes of topic analysis does not help either. 'The standard for selecting a solution is not so much accuracy as utility: Does the model simplify the data in a way that is interpretable' (DiMaggio et al., 2013: 602). In other words, much depends on human input at the final stage as well.

Regression analysis exploiting the concept of Bayesian probability also aims to identify ideas (symbols) in a text in function of words that it contains. The presence of particular words is deemed to be indicative of 'the range of things that speakers are capable of doing in (and by) the use of words and sentences'

(Skinner, 2002: 3). In terms of semiotics, a word is a signifier, an idea (symbol) is a signified (Derrida, 1967). For instance, if a text contains the words 'pleasure', 'sex' and their derivatives,[4] then it likely conveys the idea of sexuality, Foucault's *Histoire de Sexualité* (1976, 1984a, 1984b) being a prime example. The presence of a specific word in a text thus constitutes a condition that it 'symbolizes', or embodies a particular idea as well.

Bayesian probability helps operationalize the notion of conditionality. N-grams (bigrams, trigrams, etc.) represent its simplest manifestation. The assumption is that the probability of the last word in an N-gram depends on the word(s) that occur(s) immediately before (Jurafsky and Martin, 2018: chapter 3).[5] Along with meaningful N-grams – they can be found with the help of *Google Ngram Viewer*[6] – there are a great many meaningless N-grams, such as 'please turn', 'turn your', or 'your homework'.

The conditional probability of y (idea) given the occurrence of x (word) can be computed with the formula $P(y = y | x = x) = \frac{p(y=y, x=x)}{p(x=x)}$, where $P(y = y, x = x)$ refers to the percentage of texts containing the word x conveying the idea y, $P(x = x)$ – to the percentage of the word x among all words in a set of texts (Goodfellow et al., 2016: 57; Murphy, 2012: 177). The more words-signifiers are taken into account, the more accurately classification of ideas-symbols can be achieved even on the basis of the so-called 'naïve' Bayes assumption, or the assumption of the conditional independence of words accounted for in regression analysis (Evans et al., 2007: 1015; Grimmer and Stewart, 2013: 277; Jurafsky and Martin, 2018: chapter 6; Mackenzie, 2017: 112).[7] Applications of regression analysis designed along those lines are popular in the social sciences, ranging from psychology (Schwartz et al., 2013) to applied economics (Yakovleva, 2017: 6).

The brief outline of two approaches to identifying and interpreting symbols with the help of neural networks suggests that they have some internal limits. On the one hand, these limits derive from attempts to extrapolate patterns found in one set of sources, the training data, to a much larger set. This limit takes a particularly manifest form in the case of selecting the number of topics, but also characterizes any regression. We will return to limits of extrapolation in a subsection that follows. On the other hand, both approaches can allow learning to read manifest meanings, as opposed to reading between the lines. Nothing in this line from Austen's *Pride and Prejudice* (1813), 'to be fond of dancing was a certain step towards falling in love', explicitly indicates a potential connection with the idea of sexuality. Yet dancing may well be perceived as a symbol of sexuality, at least in some contexts. In other words, the proper identification and interpretation of symbols calls for going beyond studying constellations of words and for thinking metaphorically. 'The primary function of metaphor is to provide a partial understanding of one kind of experience in terms of another kind experience' (Lakoff and Johnson, 1980: 154), i.e., sexuality in terms of dancing. This metaphor suggests that movements of the physical body can become a modality of feeling, thinking, and by extension, a form of creativity (Sutil, 2017).[8] The metaphor of dancing helps also capture an interactive nature of sexuality. So the saying, 'It takes two to tango.'

Modelling social action

Modelling interaction, as compared with individual decision-making, represents the other challenge to the application of neural networks. Weber's definition of social action highlights the need for mutual adjustments and adaptation as a second dimension of social competence, in addition to the capacity to identify and interpret symbols discussed in the previous subsection. 'Action is "social" insofar as its subjective meaning takes account of the behavior of others and is thereby oriented in its course' (Weber, 1968: 4).[9]

Computer scientists and specialists in cognitive sciences acknowledge the need for introducing a 'social' dimension into artificial creativity. Creativity involves the production of ideas that are both novel and valuable. Neural networks can be trained to detect and appreciate novel patterns (Schmidhuber, 2012: 324; Thaler, 2016b: 23). The determination of whether novel patterns are also valuable turns out to be more challenging. 'Value is not found by science [or arts] but negotiated by social groups' (Boden, 1999: 351). It follows that the acceptance of novel ideas as valuable requires social, as opposed to individual, action.

At the same time, aspects related to interaction complicate the task of pattern recognition. In the simplest case of object perception, one needs to identify patterns that characterize a stand-alone, stable and simple object. At the next stage, one has to show 'situation awareness' by perceiving multiple objects, some of which are relevant to a situation-symbol, whereas the others are not (Perlovsky and Ilin, 2012: 805). For instance, family pictures placed on an office table or wall do not fit well the 'office' situation. Images, in general, normally contain a number of elements, both relevant and irrelevant to the idea that they symbolize. The recognition of patterns in interactions – social (behavioural) perception – is still more complicated. Such patterns do not remain stable: they evolve as the interaction unfolds. A partial solution would be to reduce other parties in the interaction to the status of objects, artificially renouncing the need for taking into

account their interests and for adjusting to their moves. They are then considered as 'frozen', in the same way as the values within almost all artificial creativity computer programs are 'frozen', i.e. held constant and extrinsic (Boden, 1999: 353). The notion of instrumentally rational action reflects this solution. Instrumentally rational action is 'determined by expectations as to the behavior of objects in the environment and of other human beings; these expectations are used as "conditions" or "means" for the attainment of the actor's own rationally pursued and calculated ends' (Weber, 1968: 24).

Situations studied by game theory probably represent an approximation for the second dimension of social competence. They help to potentially avoid the reductionism of treating interests and acts of other people as simple parameters in the individual's utility-maximization function. A game theory player cannot reach her optimum without paying close attention to what the other player is doing. A combination of game theory and machine learning looks particularly promising. Neural networks pave the way for modelling the opponent's moves, current and future – in the case of repeated games (Azaria et al., 2014; Chen et al., 2015; Gaudesi et al., 2014: 26). These models do not remain stable. They evolve after information about each subsequent round in the play is 'fed' into them.

Feedback loops in the model represent a dynamic parameter. This parameter does not relate to limitations of individual rationality studied by psychologists and experimental economists (Denzau and North, 1994; Kahneman and Tversky, 1982). Instead, it derives from constraints imposed on individual decision-making in the context of interaction. It must be noted that the involvement in interaction can have both negative (constraining) and positive (enabling) effects on the individual (Giddens, 1984). By joining efforts with other people, one attains goals that were otherwise outside of one's reach.

Malsch (2001: 165) differentiates two types of coordination: reactive and proactive or anticipatory. In the former case, one reacts to an obstacle or opportunity created by the other individuals. Weber's notion of instrumentally rational action applies to this situation. In the latter case, one anticipates (with the help of either learning or inferences) possible interference or opportunities. Anticipatory coordination requires more complex models of the other people. The other individuals involved in coordinated action cannot be reduced to 'conditions' or 'means', as in the previous case. For this reason, models that underpin anticipatory coordination are more complex and difficult to build (Castelfranchi, 1998: 166).

The example of killer robots mentioned previously serves as an illustration. Progress in their development is more significant than in other areas of artificial intelligence. It has not escaped the attention of scholars in science and technology studies. Some prototypes, such as the SGR-A1 sentry gun not only exist but are actually used in practice, in the Korean Demilitarized Zone. Along with some external factors, the rapid development of killer robots may be indicative of the lack of some internal technological constraints. Killer robots embody force as a technique of power. According to Weber, power in its various configurations (force, coercion, authority, etc.) represents a particular case of social action. Relationships mediated by force are still interactions, but in their least sophisticated, essentially reactive form. The individual subject to force simply reacts to its application. She is treated as if 'she were no more than a physical object' (Wrong, 1980: 24), read a 'condition' or 'means'. 'To kill somebody is for sure a social action (although not very sociable!) but it neither is nor requires communication' (Castelfranchi, 1998: 164). The lack of in-built constraints calls for endowing robots with a capacity for explanation and justification as a prerequisite for being allowed to make the decision of whether to kill or not (Maher, 2016).

Another violent technique of power – coercion, requires a bare minimum of communication between the parties involved and, hence, more sophisticated models. The coerced individual has a choice, however unattractive (e.g., your money or your life) it may be. Coercion refers to 'social relations in which the threatener engages in communication with the other at the symbolic level' (Wrong, 1980: 25). To be effective, coercion necessitates at least some proactivity: after all, the threatener does not necessary want to kill the coerced, being satisfied with her purse instead. In other words, a battle is the least sophisticated configuration of social action that could be modelled, even using the existing technologies of artificial intelligence. Anticipatory, proactive coordination requires more than that, which explains the relatively slower progress in the development of, for instance, driverless cars. To be effective, they shall be able to anticipate at times 'unpredictable', 'illogical' or even 'illegal' manoeuvres of human drivers.

The simultaneous accounting of human participants and material objects involved in interactions adds an additional layer of complexity to the task of their modelling. Actor-network theory and related approaches highlight particularities of networks comprised of both humans and material objects. Callon's study of innovations (2002) shows that patterns of interactions between actors involved depend on whether or not novel knowledge is embodied in the minds and bodies of researchers and technicians or in instruments and machines.

Prediction

Neural networks are good at discovering existing patterns in data and extrapolating them. Their performance in prediction of pattern changes in the future is less impressive. Keeping in mind that 'prediction seeks to anticipate, whereas inference seeks to interpret' (Mackenzie, 2017: 107), one might say that neural networks are a powerful tool for making inferences, but not predictions. Accordingly, neural networks are more capable of performing classificatory functions directly related to inference than of making predictions. In contrast to pattern recognition performed by neural networks (in hand written digits, in multi-object images, etc.), their predictive power is less frequently problematized and subject to critical scrutiny.

This internal limitation of machine learning powered by neural networks derives from their reliance on regression analysis. 'Linear regression is the "work horse" of statistics and (supervised) machine learning' (Murphy, 2012: 217; see also Mackenzie, 2017: 40, 169; Witten et al., 2017: chapter 4). The fact that neural networks are applied to Big Data as opposed to small and manageable datasets does not change much. Data points are still to be fit into some kind of line, even if 'the line is not fitted [any more] to x–y coordinates but in a multidimensional space reflecting the large number of features that the algorithm is trying to combine' (McQuillan, 2016: 2). McQuillan emphasizes difficulties with visualizing and, hence, interpreting a best-fit line in a multidimensional space. Here, the argument will be focused on the other problem with regression analysis, namely, its inherent proclivity to equate predictions with extrapolation of existing patterns.

In regression analysis, predictions are made under the assumption that existing patterns will not change in the future. They can be held constant, in other words. While not speaking specifically about regression analysis, Keynes highlights this assumption in the much-quoted passage from his *General Theory of Employment, Interest and Money* (1936). 'In practice we have tacitly agreed, as a rule, to fall back on what is, in truth, a *convention*. The essence of this convention – though it does not, of course, work out quite so simply – lies in assuming that the existing state of affairs will continue indefinitely, except in so far as we have specific reasons to expect a change... We are assuming, in effect, that the existing market valuation, however arrived at, is uniquely *correct* in relation to our existing knowledge of the facts which will influence the yield of the investment, and that it will only change in proportion to changes in this knowledge' (Keynes, 2015: 152). This observation was made during the Great Depression, when the assumption that the existing state of affairs would continue indefinitely did not hold.

The 2008 global financial crisis serves as a more recent reminder that 'the extrapolation of past patterns or relationships cannot provide accurate predictions' (Makridakis et al., 2009: 794). Sophisticated forecasting techniques powered by neural networks appeared no more efficient than 'naïve' forecasting strategies (Perera et al., 2018: 271).

Neural networks in their current architecture can hardly overcome this inherent limitation of regression analysis. A direction for searching for a possible solution was suggested also by Keynes. It involves connecting our assessment of probabilities of future events to our state of knowledge, as the last sentence in the quoted passage clearly suggests. A view on probability as being conditioned by our knowledge is closer to the Bayesian interpretation than to the frequentist one. In the first case, 'probability is used to quantify our uncertainty about something; hence it is fundamentally related to information rather than repeated trials', whereas in the second 'probabilities represent long run frequencies of events' (Murphy, 2012: 27). The differences between the Bayesian and frequentist interpretations lie less in the underlying mathematics (Goodfellow et al., 2016: 53) than in the role played by knowledge and information. Knowledge and information do not affect 'frequentist' probabilities that are assumed to have an 'objective' nature. Bayesian probabilities in this sense are more 'subjective' and subject to social influences (Keynes believed that our expectations about the future could and should be manipulated). However it may be, the application of neural networks in this particular area is still at a very early stage.

Limits of artificial creativity powered by neural networks

All three sets of internal limits of neural networks take manifest forms when they are used to achieve artificial creativity. In the context of the present discussion, creativity involves (i) one's capacity to produce symbols and to transfer them from one context to the other; (ii) one's social intelligence, since innovations are often embedded in social connections and relationships, and; (iii) one's capacity to go beyond simple extrapolation of current trends in predictions. The progress towards artificial creativity can be used to assess how hard or soft these three sets of internal limits of neural networks are.

When discussing artificial creativity, one's capacity to transfer symbols from one context to another becomes highly relevant. Thinking metaphorically constitutes a dimension of creativity. It involves establishing links between symbols in various situations (domains) on the basis of their similarity

(Lotman, 1990: 39). In the humanities and social sciences, the importance of metaphors is acknowledged. For example, they are used as indicators for changes in culture (Pasanek and Sculley, 2008: 345). In the natural sciences and engineering, metaphorical thinking plays an equally important role. This fact is rarely openly admitted though. Ideas and solutions in one context offer valuable insights when dealing with problems in another – this principle applies to creativity regardless of the area of human activity.

The history of neural networks is a case in point. The idea of neural networks was originally inspired by a metaphor of the brain.[10] Taken separately, neurons perform very simple, trivial tasks. When they are connected and form networks, their power substantially increases. Tasks that cannot be tackled by a single neuron are easily performed when being distributed among many of them. The parallel between the brain and neural networks is far from being direct: the brain architecture is rather a source of insights than an exact model for neural networks (Goodfellow et al., 2016: 165).

The other parallel, this time between Freud's theory (as strange as it may sound in this context) and links constituting a neural network, inspired the idea of backpropagation. Werbos (2012: 91), its author,[11] explains: 'chronologically, I translated Freud into a way to calculate derivatives across a network of simple neurons (which Harvard simply did not believe), and then proved the more general chain rule for ordered derivatives to prove it and make it more powerful (and to graduate)'. The principle of backpropagation underpins feedback loops connecting elements (such as specific regressions) of a neural network.

The list of metaphors that have influenced the evolution of neural networks does not stop here. Studies in neuroscience suggest that creativity requires the co-activation and communication between regions of the brain that ordinarily are not strongly connected (Cariani, 2012: 406; Heilman et al., 2003: 369; Kowatari et al., 2009). It means that neural networks created in various, not strongly connected, situations (e.g., driving and reading) can be a potentially powerful tool for discovering and exploiting relationships based on similarity (Heilman et al., 2003: 375; Malsch, 2001: 159).

The metaphor of the brain has its limits, nevertheless. It prompts further developments in neural networks in one direction and overshadows promising inquiries if they deviate from the neuropsychological path, for instance, science and technology studies (Munster, 2011). Scholars in science and technology studies consider the capacity to trace linkages between heterogeneous and previously unconnected elements as a distinctive human social activity, 'which is to say

these linkages are associations' (Gehl and Bell, 2012). It means that the apparatus of neuroscience does not suffice to establish connections between a priori heterogeneous elements and to make sense of these associations. The capacity to do so has a social dimension, as Gehl and Bell's case study of the rather unsuccessful implementation of the otherwise very promising operational system, *Microsoft Vista*, suggests.

When speaking of relationships based on similarity, it is necessary to distinguish valid similarity relations from apparent similarity. Mining of Big Data with the help of neural networks creates unprecedented opportunities for discovering unexpected patterns (see Sala-I-Martin, 1997 for an early example in economics). How to make sure that similarity between some of these patterns goes beyond appearances? McQuillan (2016: 6) observes in this regard that 'what most people would see as coincidence a paranoid person may believe was intentional, while the whole of machine learning is based on finding meaning in patterns of coincidence. Paranoia and machine learning weave around each other like the serpents around Hermes' staff' (2016: 6).

Sociological thinking offers one approach for discriminating against simple appearances; building classifications refers to the other. The Thomas theorem is among few 'theorems' known to sociologists. It states that 'if men define situations as real, they are real in their consequences' (Merton, 1995: 380). If only one individual sees a relationship, then it may be indicative of paranoid behaviour. However, if a relationship makes sense to many people, then chances that a valid – and valuable – similarity relation exists tend to be higher.

In the words of Lakoff and Johnson (1980: 157), 'the acceptance of the metaphor, which forces us to focus *only* on those aspects of our experience that it highlights, leads us to view the entailments of the metaphor as being *true*'. Vee (2012) shows how metaphors create a reality using three classificatory metaphors for computer code in US jurisprudence as an example. Continuing this line of reasoning, one can assume that neural networks assist in generating a 'long list' of potential metaphors, whereas concerted human efforts will still be needed for 'short-listing' valid and valuable metaphors. The short-listing has a negotiated and, hence, social character, which preludes the discussion of a second dimension of creativity below.

The search for relationships based on similarity is also facilitated by building classifications, provided that classes are built on relationships among class members. 'The members of classes may be highly similar to one another, but their similarities result from their membership in the same class (i.e., conforming

to class properties), and not the other way around (i.e., similarity alone cannot define class inclusion)' (Berman, 2013: 141). In this case, classes include members with similar functional properties even if at first sight they look dissimilar. The *word2vec* models developed by *Google* illustrate this point. Embedded words are analysed and classified in function of semantic roles that they perform. This approach allows answering questions of the type: Madrid is to Spain as France is to ___? (Paris). Madrid and Paris have similar functional properties with respect to Spain and France, respectively (Evans and Aceves, 2016: 33). Identification of valid metaphors ('Constantinople is the second Rome, Moscow is the third Rome'[12]) in this way requires minimal human input.

To conclude the discussion of metaphorical thinking as a dimension of creativity, it is necessary to differentiate between understanding of metaphors and their generation. Some progress towards training neural networks to understand metaphors has been made, for instance, in linguistics. At the same time, 'much less research has been devoted to generating metaphor' (Gargett and Barnden, 2015: 104). Artificial creativity, first of all, is about generating metaphor since this process requires new ways of thinking and new techniques, or at least fresh ways of using already established techniques.

A second dimension of creativity refers to interaction. Instead of attributing creativity exclusively to a solitary individual, Sawyer (2006) highlights the importance of group and societal forms of creativity. The sociology of science shows that creativity and innovation tend to be embedded in social interactions and networks. 'Intellectual creativity is concentrated in chains of personal contacts, passing emotional energy and cultural capital from generation to generation' (Collins, 1998: 379). Three examples of chains of personal contacts as a source of innovation serve to illustrate this point.

A network of corresponding scholars that emerged in the late XVI-early XVII centuries became known under the name of the Republic of Letters. In contrast to the current situation, research was not exclusively concentrated in the universities that only started to come into existence at that time. Scholars and free thinkers were scattered across vast areas of Europe, the Middle East and North Africa. Nevertheless, they stayed connected and exchanged ideas and the results of their inquiries by way of letters and publications in scholarly journals. 'It is these societies, [these] academies, and these journals which give a concrete form to the relationship between[...] expertise and reading in the free and universal form of the circulation of written discourse' (Foucault, 2011: 8). Without free circulation of written discourse, most innovations and discoveries of the XVII–XVIII centuries would have been simply impossible.

Two other examples refer to smaller scale social networks as a source of creativity. Mikhail Bakhtin made an important contribution to semiotics. His case is interesting since it highlights difficulties with differentiating an individual contribution and a group contribution. Major breakthroughs were achieved in group discussions with his participation: he took part in the operation of three circles that existed in the 1920s in Nevel, Pskov and Leningrad. 'Ideas as [Bakhtin] understood their nature emerge and disappear in the process of dialogue' (Etkind, 1993: 392). For this reason, the authorship of some works and ideas deriving from the Nevel, Pskov and Leningrad circles remain disputed up to now.

Granovetter (1994) considers the case of technological innovations, showing the social embeddedness of their production and diffusion. He discusses Edison and networks that allowed him to promote his innovations in the area of industrial electricity and power generation. These three examples show that creativity involves interactions, being, from this point of view, a particular case of social action. As indicated before, neural networks still underperform in social action compared with humans, especially as far as anticipatory coordination is concerned. The architecture required for artificial intelligence can be compared with a 'federation' of neural networks, or a network of networks.

Artificial intelligence also requires going beyond extrapolation as a basis for forecasts. Under the current arrangements, neural networks are composed of neurons. Each neuron implements a simple model, often a regression model (Mackenzie, 2017: 192). A neuron computes a weighted sum of the sample inputs applied (for instance, the frequencies of particular words in a text) and then perform an activation function to this sum to calculate its output. The sigmoid transfer function, $f(x) = \frac{1}{1+e^{-x}}$, is a popular activation function (Russell et al., 2010; Sharma and Dey, 2010: 69). Adding more neurons, layers or even neural networks to this scheme does not lessen its dependence on the mechanics of regression analysis.

Regression analysis equates prediction with extrapolation whereas creativity is simply impossible without the emergence of new patterns. The principle of induction underpins regression analysis. Induction represents a necessary, but insufficient precondition for creativity. 'While induction by itself is insufficient for creativity, generalizing from data could be a tool of creative thinking' (Schank and Foster, 1995: 133). The conventional structure of scholarly publications reflects the role played by induction in creativity. A literature review section is a necessary component of any scientific

contribution, but only review essays summarizing the state of knowledge contain nothing else.

By paying attention to various forms of creativity, one better understands the relationship between prediction and creativity. Two pairs of concepts, adaptive versus generative (Bown, 2012: 364) and combinatorial versus emergent, or exploratory-transformational (Boden, 1999: 352; Cariani, 2012: 394), help identify forms of creativity that do not require prediction. In contrast to adaptive creativity, generative creativity does not satisfy a specific function and, thus, cannot be derived from characteristics of a system within which it takes place. In a similar vein, emergent creativity cannot be accounted for in terms of existing structures, functions and behaviours, whereas combinatorial creativity can. Prediction does not suffice for achieving generative or emergent creativity, i.e., its particularly valuable forms.

Adding more data points to training data – a usual method for increasing the accuracy of predictions – does not really help. Neural networks perform reasonably well only as long as the assumption that parameters estimated with training data also hold in validation and test data, i.e. in the entire population of relevant events (Burscher et al., 2015: 128; DiMaggio, 2015: 1; Goodfellow et al., 2016: 224; Hopkins and King, 2010: 236; Pavelec et al., 2008: 414; Scharkow, 2013: 763).

Furthermore, the Great Depression of the 1930s and the 2008 global financial crisis remind us that no extrapolation suffices for predicting a 'black swan', an event that either occurs very rarely or has not yet occurred, but may occur in the future. Neither the frequentist approach to probability nor the Bayesian one helps solve the 'black swan' problem. In the former, rare events are treated as 'outliers' (Bail, 2015: 2). In the latter, the conditional probability is only defined when $P(x = x) > 0$. 'We cannot compute the conditional probability conditioned on an event that never happens' (Goodfellow et al., 2016: 57). Both the Great Depression and the 2008 global financial crisis were unique events with no direct precedents. Sophisticated forecasting models (they were particularly numerous around the time of the most recent crisis) failed to predict the crisis coming. Some humans managed to do so nevertheless using creative and innovative thinking, as shown in *The Big Short* movie (2015), for instance.

As a matter of fact, attempts to increase the accuracy of predictions without changing the underlying assumption that prediction equates to extrapolation could well be counter-productive. They may lead to 'overfitting', i.e. to a regression model that fit training data too closely. 'Imagine a very wiggly line connecting a set of more-or-less linear dots – while the line fits the observed dots exactly it will clearly be a poor predictor of future observations' (McQuillan, 2016: 4; see also Mackenzie, 2017: 121). A metaphor seems appropriate. Neural networks could be trained on Picasso's artwork and produce their exact replicas. Some artists are involved in the business of art forgery or imitation too: after a Picasso many Picassos arise (DeFelipe, 2011: 6). However, pieces of imitation are never valued as much as the originals. Imitating and replicating Picasso's artwork is one thing, creating something original based on his achievements is quite the other.

The social sciences help identify conditions under which prediction can indeed be equated with extrapolation and, thus, the internal limitation of regression analysis does not constitute an obstacle in the future progress of neural networks. A well (even over-) ordered society is needed. Data on social action in a well-ordered society has a highly structured character. Everyone sticks to the rules and regulations, which makes patterns in interactions relatively easily discernable. Speaking metaphorically, social data is free of 'clutter' and 'noise'.

Such a society also represents an environment in which the assumption that the existing state of affairs will continue indefinitely holds. McQuillan (2016: 5) describes this eventuality in the following words. 'The risk is that the statistical regression at the heart of machine learning will become an engine of forms that are socially regressive... What forms of governance and government resonate with algorithmic seeing, and what social distortions may result? Big data seeks to bring about a new form of ordered legibility of society according to its own logic of patterns'.

The idea that the statistical regression can generate socially regressive forms resonates with Foucault's thoughts on *État policé* and *société policée*. In French, *être policé* means to be civilized, to reach the conditions of modernity. Foucault (2007: 321) writes that 'police... is administrative modernity par excellence' and that 'police makes statistics necessary, but police also makes statistics possible' (2007: 315). In other words, neural networks, on the one hand, require socially regressive forms in order to be effective and, on the other hand, contribute to their production and prevalence.

The image of a well-ordered, *policé* society may be appropriate in some cases but not in many others. For instance, there is probably nothing wrong if road traffic becomes well ordered. An ideal form of driving, driving by the rules, makes manoeuvres perfectly predictable and easy to model, especially with the help of neural networks. The number of road accidents, as well as the time spent (wasted) on commuting, would be significantly reduced as a result. This ideal situation could probably be achieved by removing humans from the roads as their least predictable element. Humans may

well find other areas that require their continuous attention. The same logic cannot be applied to society as a whole. The ideal of a well-ordered society as a means to increase predictability in human interactions is questionable to say the least. In addition to clear parallels with an Orwellian society, such an arrangement will simply leave no place for creativity. The reason is simple: creativity undermines predictability interpreted as mere extrapolation.

Conclusions

Using the metaphor of human neurons in the design of neural networks is both enabling and constraining. On the one hand, by connecting neurons and building neural networks, one can tackle complex tasks that no neuron taken separately can perform. The following analogy[13] highlights this beneficial effect of connectivity. Arendt (1969: 44) observed that 'power corresponds to the human ability not just to act but to act in concert'. Paraphrasing her definition, one can say that computational power of neural networks is due not just to the computational capacity of neurons but to their capacity to be connected and to perform tasks in concert.

On the other hand, the reliance on statistical regression in each node of a neural network imposes limits on what can be achieved with their help. Namely, neural networks underperform compared with human beings in identifying and interpreting symbols, in acting socially, which requires mutual coordination and adjustments, and in predictions. Underperformance of neural networks in these three areas that were specifically considered in this article leaves the question of whether artificial creativity can catch up to human creativity wide open. Creativity is hardly possible without one's capacity to think metaphorically, to coordinate proactively and to make predictions that go beyond simple extrapolation. Compared with routine tasks, everything novel and unusual calls for creativity.

If brain neurons worked in the same way as neural networks, then how could humans be creative? Studies in neuroscience suggest that in order to enable creativity, neurons have to be connected not only with neighbouring neurons but also with neurons in the other regions of the brain. At the present stage, most progress has been achieved in building domain (or area) specific neural networks. Thus, in order to overcome some internal limitations of neural networks, one would need to learn how to build mega-neural networks connecting more specialized neural networks. Otherwise, artificial imagination is hardly possible: 'imagination... involves seeing one kind of thing in terms of another kind of thing – what we have called metaphorical thought' (Lakoff and Johnson, 1980: 193). Can the establishment of associations be achieved keeping the reliance on statistical regression as the 'work horse' of neural networks? This issue remains unaddressed in a satisfactory manner. The metaphor of the brain also forecloses a promising program of studying the linkages between heterogeneous and previously unconnected elements as the outcomes of a social activity.

It follows that artificial creativity powered by neural networks continues to underperform, at least at this stage. They could greatly assist in preparing a literature review, i.e. in learning and summarizing the existing body of knowledge. But they are a long way from learning how to produce novel and innovative results, let alone major breakthroughs.

Notes

1. The discussion of specific designs of neural networks, multilayer perceptrons versus radial-basis function networks, feedforward versus recurrent, etc. remains outside of the scope of this article.
2. Socionics is the other reference point. 'Socionics seeks to address the question of how to exploit models from the social world for the development of intelligent computer technologies' (Malsch, 2001: 155).
3. Words in clutter are notoriously difficult for the machine to recognize, which explains the popularity of captcha tests as a barrier against internet bots.
4. The list of a word's derivatives is populated by using the operation of lemmatization (Jurafsky and Martin, 2018: chapter 1).
5. The assumption that the probability of a word depends only on the previous word is called a Markov assumption. It is further exploited in Markov chains of variable length.
6. https://books.google.com/ngrams.
7. For instance, the words 'sex' and 'pleasure' are assumed to be independent despite the fact that they often co-occur in texts on sexuality.
8. This is one of the key premises of movement computing, an emergent field of science and technology studies.

9. It comes as no surprise that socionics came into existence in response to the need to model interactions. 'One of the pioneers of distributed artificial intelligence (DAI), Carl Hewitt, in his ground-breaking work on "patterns of passing messages" drew on the social psychology of G.H. Mead to propose... defining the meaning of a message by the reaction it occasioned in the addressee' (Malsch, 2001: 159). Mead's predecessor, the sociologist Cooley, is known for his conception of the looking glass self: the individual is a product of her interactions with other people, of her reactions to their perception of her personality.

10. Advocates of socionics point out to the other potential metaphor at the origin of distributed artificial intelligence: that of a society composed of well-connected and collaborating individuals (Malsch, 2001: 158).

11. Backpropagation, however, may well have been discovered independently by several scholars, as it often happens in modern science (Merton, 1973). Mackenzie (2017: 174) cites at least two other formulations of the same idea made by other scholars in the 1980s.

12. The addition that 'and there will be no fourth' requires going beyond extrapolation, which preludes a third dimension of creativity below.

13. This analogy has roots in socionics.

ORCID iD

Anton Oleinik ⬤ http://orcid.org/0000-0002-5229-1052

References

Amini R, Sabourin C and De Koninck J (2011) Word associations contribute to machine learning in automatic scoring of degree of emotional tones in dream reports. *Consciousness and Cognition* 20(4): 1570–1576.

Arendt H (1969) *On Violence*. New York, NY: Harcourt, Brace & World.

Arthur BW (1989) Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal* 99(394): 116–131.

Azaria A, Rabinovich Z, Goldman CV, et al. (2014) Strategic information disclosure to people with multiple alternatives. *ACM Transactions on Intelligent Systems and Technology* 5(4): article 64.

Bail CA (2014) The cultural environment: Measuring culture with big data. *Theory and Society* 43(3–4): 465–482.

Bail CA (2015) Lost in a random forest: Using Big Data to study rare events. *Big Data & Society* 2(2): 1–3.

Berman JJ (2013) *Principles of Big Data: Preparing, Sharing, and Analyzing Complex Information*. Waltham, MA: Morgan Kaufmann.

Boden MA (1999) Computer models of creativity. In: Sternberg RJ (ed.) *Handbook of Creativity*. Cambridge: Cambridge University Press, pp. 351–372.

Boden MA (2009) Computer models of creativity. *AI Magazine* 30(3): 23–34.

Bown O (2012) Generative and adaptive creativity: A unified approach to creativity in nature, humans and machines. In: McCormack J and d'Inverno M (eds) *Computers and Creativity*. Berlin/Heidelberg: Springer-Verlag, pp. 361–381.

Burscher B, Vliegenthart R and De Vreese CH (2015) Using supervised machine learning to code policy issues: Can classifiers generalize across contexts? *The Annals of the American Academy of Political and Social Science* 659(1): 122–131.

Callon M (2002) From science as an economic activity to socioeconomic of scientific research: The dynamics of emerged and consolidated techno-economic networks. In: Mirowski P and Sent E-M (eds) *Science Bought and Sold: Essays in the Economics of Science*. Chicago, IL/London: The University of Chicago Press, pp. 277–317.

Cariani P (2012) Creating new informational primitives in minds and machines. In: McCormack J and d'Inverno M (eds) *Computers and Creativity*. Berlin/Heidelberg: Springer-Verlag, pp. 383–417.

Castelfranchi C (1998) Modelling social action for AI agents. *Artificial Intelligence* 103(1–2): 157–182.

Chen W, Chen Y and Levine DK (2015) A unifying learning framework for building artificial game-playing agents. *Annals of Mathematics and Artificial Intelligence* 73(3–4): 335–358.

Collins R (1998) *The Sociology of Philosophies: A Global Theory of Intellectual Change*. Cambridge, MA: The Belknap Press of Harvard University Press.

Computational creativity (n.d.) In: *Wikipedia*. Available at: https://en.wikipedia.org/wiki/Computational_creativity (accessed 11 February 2019).

DeFelipe J (2011) The evolution of the brain, the human nature of cortical circuits, and intellectual creativity. *Frontiers in Neuroanatomy* 5: article 29.

Denzau AT and North DC (1994) Shared mental models: Ideologies and institutions. *Kyklos* 47(1): 3–31.

Derrida J (1967) *De la grammatologie*. Paris: Les Éditions de Minuit.

DiMaggio P (2015) Adapting computational text analysis to social science (and vice versa). *Big Data & Society* 2(2): 1–5.

DiMaggio P, Nag M and Blei D (2013) Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics* 41(6): 570–606.

Etkind A (1993) *Eros nevozmozhnogo: istoriya psikhoanaliza v Rossii* [*Eros of the Impossible: A History of psychoanalysis in Russia*]. St. Petersburg: Medusa.

Evangelopoulos N, Zhang X and Prybutok VR (2012) Latent semantic analysis: Five methodological recommendations. *European Journal of Information Systems* 21(1): 70–86.

Evans JA and Aceves P (2016) Machine translation: Mining text for social theory. *Annual Review of Sociology* 42: 21–50.

Evans M, McIntosh W, Lin J, et al. (2007) Recounting the courts? Applying automated content analysis to enhance empirical legal research. *Journal of Empirical Legal Studies* 4(4): 1007–1039.

Foucault M (1976) *Histoire de la sexualité. Volume 1: La volonté de savoir*. Paris: Les Éditions Gallimard.

Foucault M (1984a) *Histoire de la sexualité. Volume 2: L'usage des plaisirs*. Paris: Les Éditions Gallimard.

Foucault M (1984b) *Histoire de la sexualité. Volume 3: Le souci de soi*. Paris: Les Éditions Gallimard.

Foucault M (2007) *Security, Territory, Population. Lectures at the Collège de France, 1977–1978*. Basingstoke/New York, NY: Palgrave Macmillan.

Foucault M (2011) *The Government of Self and Others. Lectures at the Collège de France, 1982–1983*. New York, NY: Picador/Palgrave Macmillan.

Gargett A and Barnden J (2015) Gen-Meta: Generating metaphors by combining AI and corpus-based modeling. *Web Intelligence* 13(2): 103–114.

Gaudesi M, Piccolo E, Squillero G, et al. (2014) TURAN: Evolving non-deterministic players for the iterated prisoner's dilemma. In: *2014 IEEE Congress on evolutionary computation*, 6–11 July 2014, Beijing, China, pp. 21–27.

Gehl RW and Bell SA (2012) Heterogeneous software engineering: Garmisch 1968, Microsoft Vista, and a methodology for software studies. *Computational Culture: A Journal of Software Studies*. Epub ahead of print September 2012.

Gibbs S (2017) Elon Musk leads 116 experts calling for outright ban of killer robots. *The Guardian*, 20 August. Available at: https://www.theguardian.com/technology/2017/aug/20/elon-musk-killer-robots-experts-outright-ban-lethal-autonomous-weapons-war (accessed 14 May 2018).

Giddens A (1984) *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Polity Press.

Goodfellow I, Bengio Y and Courville A (2016) *Deep Learning*. Cambridge, MA: MIT Press.

Granovetter M (1994) Les institutions économiques comme structures sociales: un cadre d'analyse. In: Orléan A (ed.) *Analyse économique des conventions*. Paris: PUF, pp. 79–94.

Grimmer J and Stewart BM (2013) Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* 21(3): 267–297.

Haykin S (2009) *Neural Networks and Learning Machines*. 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Heilman KM, Nadeau SE and Beversdorf DO (2003) Creative innovation: Possible brain mechanisms. *Neurocase* 9(5): 369–379.

Hopkins DJ and King G (2010) A method of automated nonparametric content analysis for social science. *American Journal of Political Science* 54(1): 229–247.

Huang L, Milne D, Frank E, et al. (2012) Learning a concept-based document similarity measure. *Journal of the American Society for Information Science and Technology* 63(8): 1593–1608.

Jurafsky D and Martin JH (2018) *Speech and Language Processing*, 3rd ed. Upper Saddle River, NJ: Pearson-Prentice Hall.

Kahneman D and Tversky A (1982) Judgment of and by representativeness. In: Kahneman D, Slovic P and Tversky A (eds) *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press, pp. 84–100.

Keynes JM (2015) *The Essential Keynes*. London: Penguin.

Kowatari Y, Lee SH, Yamamura H, et al. (2009) Neural networks involved in artistic creativity. *Human Brain Mapping* 30(5): 1678–1690.

Kripke SA (1982) *Wittgenstein on Rules and Private Language*. Oxford: Basil Blackwell.

Lakoff G and Johnson M (1980) *Metaphors We Live By*. Chicago, IL/London: The University of Chicago Press.

Lotman Y (1990) *Universe of the Mind: A Semiotic Theory of Culture*. Bloomington: Indiana University Press.

MacKenzie A (2017) *Machine Learners: Archeology of a Data Practice*. Cambridge, MA/London: MIT Press.

McQuillan D (2016) Algorithmic paranoia and the convivial alternative. *Big Data & Society* 3(2): 1–12.

Maher J (2016) Artificial rhetorical agents and the computing of phronesis. *Computational Culture: A Journal of Software Studies*, January.

Makridakis S, Hogarth RM and Gaba A (2009) Forecasting and uncertainty in the economic and business world. *International Journal of Forecasting* 25(4): 794–812.

Malsch T (2001) Naming the unnamable: Socionics or the sociological turn of/to distributed artificial intelligence. *Autonomous Agents and Multi-Agent Systems* 4(3): 155–186.

Merton RK (1973) Singletons and multiples in science. In: Merton RK (ed.) *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL/London: the University of Chicago Press, pp. 343–370.

Merton RK (1995) The Thomas Theorem and the Matthew effect. *Social Forces* 74(2): 379–424.

Munster A (2011) Nerves of data: the neurological turn in/against networked media. *Computational Culture: A Journal of Software Studies*, Epub ahead of print December 2011.

Murphy KP (2012) *Machine Learning: A Probabilistic Perspective*. Cambridge, MA: the MIT Press.

Nielsen MA (2015) *Neural Networks and Deep Learning*. Determination Press. Available at http://neuralnetworksanddeeplearning.com/ (accessed 15 March 2019).

Oleinik A (2016) Corruption on the road: A case study of Russian traffic police. *IATSS Research* 40(1): 19–25.

Pasanek B and Sculley D (2008) Mining millions of metaphors. *Literary and Linguistic Computing* 23(3): 345–360.

Pavelec D, Justino E, Batista LV, et al. (2008) Author identification using writer dependent and writer independent strategies. In: *Proceedings of the 2008 ACM symposium on applied computing (SAC'08)*, Fortaleza, Brazil, 16–20 March 2008, pp.414–418.

Perera T, Higgins D and Wong WW (2018) The evaluation of the Australian office market forecast accuracy. *Journal of Property Investment & Finance* 36(3): 259–272.

Perlovsky L and Ilin R (2012) Brain. Conscious and unconscious mechanisms of cognition, emotions, and language. *Brain Sciences* 2(4): 790–834.

Riley SJ (2017) Rise of the robots: Automated trucks will transform an industry and put millions out of work. *The Walrus*, 14 November. Available at: https://thewalrus.ca/rise-of-the-robots/ (accessed 8 May 2018).

Russell SJ, Norvig P and Davis E (2010) *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall.

Sala-I-Martin XX (1997) I just ran two million regressions. *The American Economic Review* 87(2): 178–183.

Sawyer RK (2006) *Explaining Creativity: The Science of Human Innovation*. Oxford: Oxford University Press.

Schank RC and Foster DA (1995) The engineering of creativity: A review of Boden's The Creative Mind. *Artificial Intelligence* 79(1): 129–143.

Scharkow M (2013) Thematic content analysis using supervised machine learning: An empirical evaluation using German online news. *Quality & Quantity* 47(2): 761–773.

Schmidhuber J (2012) A formal theory of creativity to model the creation of art. In: McCormack J and d'Inverno M (eds) *Computers and Creativity*. Berlin/Heidelberg: Springer-Verlag, pp. 323–337.

Schwartz HA, Eichstaedt JC, Kern ML, et al. (2013) Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLoS One* 8(9): e73791.

Sharma A and Dey S (2010) A document-level sentiment analysis approach using artificial neural network and sentiment lexicons. *Applied Computing Review* 12(4): 67–75.

Skinner Q (2002) *Visions of Politics*, Vol. 1. Cambridge: Cambridge University Press.

Sutil NS (2017) Section editorial: Human movement as critical creativity: Basic questions for movement computing. *Computational Culture: A Journal of Software Studies*. Epub ahead of print November 2017.

Thaler SL (1998) The emerging intelligence and its critical look at us. *Journal of Near-Death Studies* 17(1): 21–29.

Thaler SL (2016a) Pattern turnover within synaptically perturbed neural systems. *Medical Hypothesis* 94(September): 138–147.

Thaler SL (2016b) Pattern turnover within synaptically perturbed neural systems. *Procedia Computer Science* 88: 21–26.

Touraine A (1992) *Critique de la modernité*. Paris: Les Éditions Fayard.

Vee A (2012) Text, speech, machine: Metaphors for computer code in the law. *Computational Culture: A Journal of Software Studies*. Epub ahead of print September 2012.

Weber M (1968) *Economy and Society: An Outline of Interpretative Sociology*. New York, NY: Bedminster Press.

Werbos PJ (2012) Neural networks and the experience and cultivation of mind. *Neural Networks* 32: 86–95.

Wiedemann G (2013) Opening up to Big Data: Computer-assisted analysis of textual data in social sciences. *Forum: Qualitative Social Research/Forum: Qualitative Sozialforschung* 14(2): article 23.

Witten IH, Eibe F, Hall MA, et al. (2017) *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. Cambridge, MA: Morgan Kaufmann Publishers.

Wrong DH (1980) *Power: Its Forms, Bases and Uses*. New York, NY: Harper Colophon Books.

Yakovleva K (2017) *Otsenka ekonomicheskoi aktivnosti na osnove tekstovogo analiza* [*Assessing economic activity with the help of text analysis*]. *Series of reports on economic studies*. Moscow: Central Bank of the Russian Federation.