Artificial Intelligence, the Hype, the Hope and the Lopsided lore

Dimitar Popov
Institute of Philosophy and Sociology – Bulgarian Academy of Sciences
dimiturpopov1990@gmail.com

Abstract: The article explores the parallel evolution of the technical aspects of artificial intelligence with its technological applications and the growing misinformation or misinterpretation of the scientific results, ostensibly reported by the mass media. It delves into the problem of endemic hype (promoting or publicizing extravagantly) cultivated in our society by the soundly constructed terminology when it comes to what the current state of the field could achieve. The author's premise is that the goal and prospect of every scientific inquiry should not be intentionally enshrine with artificial enigma, in the prospect of drawing public interests and funding, but at the same time risking inflammation of societal expectations and anxieties. This paper will analyze why the research field of artificial intelligence is particularly susceptible to the so-called public hype, and how the hype could be a good index of an upcoming scientific Winter, since the lore of the field has already showed this tendency a couple of times in the past.

Keywords: artificial intelligence, hype, myth, media, hopes.

Since the inception of the scientific revolution during 16th and 17th century, the essence of scientific inquiry has been centered around constructing and testing theories that best explain natural phenomena, by adopting empirical or rationalistic ways of testing hypothesis, formulating mathematical models and conceptualizing explanations. The strict principles of academia, that governs the publication and popularizations of the results achieved during any scientific program are largely vectorized for another field professionals, which by large means that the gap between the successful scientific results and their informative presentment to a broader audience is inherently difficult and prone to misunderstanding. Any listener outside the immediate professional scientific program, must be accommodated to the peculiar terminology and tangibles that presuppose what could be considered as a real progress or an astonishing result. The problem is not new: in 1831, when Michel Faraday publicly demonstrated that moving a magnet inside coils creates electric current, the immediate famous reaction from the audience was "What good is that?". This inability to recognize considerable scientific achievement is totally normal, and in classical sense the role of a scientist adopts two roles, the one of the researcher and the one of the educator or science communicator. This communication has always happened either by giving interviews or lectures, from the most prominent names of the field, when entirely new discovery or hypothesis has taken shape. In the case of Artificial Intelligence (A.I.), the general trend of popularizing any significant accomplishment is usually not just demonstrating the novelty behavior of a proposed engineered methodology or strategy, but also hardening the original intent or idea of the inventor as well, boosting his or hers core believes of how the scientific results should be accepted, analyzed and reported. The best expletory narrative that we could adopt to defend that point is to analyze two famous ideas, which have been reiterated multiple times in the lore of A.I., the idea of self-improving intelligent machines and the forthcoming of technical *singularity*,[1] both of the ideas had been postulated in the early pioneering years the field of A.I. but failed short being constructed as full scientific hypothesis. Mainly because, formulizing any meaningful way of potential experimental falsification mechanism has never been incepted.

The first idea was formulated by I.J. Good, who was working alongside Alan Turing at Bletchley park during the Second World War, and had major contribution to the cryptographic breakthrough which helped the Allied forces defeating Nazi Germany. Good was famous chess player and as such was used to the concept of accumulative improvement of one's strategies, amounting to chance maximization and wining outcome. Good was highly influenced by Turing's idea, that the engineering of autonomous cognition[2] into machinery, should be investigated as the coming up with a design of problem-solving systems. Those systems should be able of doing small incremental steps to get a definite result or solution, when presented with a task. From that viewpoint any task that could potentially be solved by natural intelligence, is to be adopted as union of fragments of the original or parent task. Each fragment containing a piece of the general knowledge, of how the broader parent task is to be resolved. Turing and Good assumed that by encapsulating and modeling human cognition as a mechanical process they could be able to "solve" intelligence in mathematical fashion through an algorithm; that is: if each problem is treated as an object, it should consistently follow from the rules of a higher formal system, and then intelligence could be perceived as purely computational.[3] This corresponded with Good's intuition that chess could also be perceived as purely computational, because to win a game of chess it is not enough to simply follow the rules, you need to know which rules you have to select in a first place. Good, however, added one more detail to the concept of intelligent machines,[4] which he also borrowed from chess, the idea that in time, a player will constantly improve his skills and strategies. His idea was relatively simple: if a machine can reach human-level intelligence, it should be capable to improve on itself and suppress human level intelligence, and consequently every next level of intelligence. For Good it was obvious that, as any human could improve their knowledge and skills overall, a machine that is on a par with any human, should also possess feedback mechanism that would enable the same behavior. This self-improvement would lead new smarter mechanisms to be constructed resulting in exponentially accelerating process, which is nowadays hugely proclaimed as inevitable by famous A.I. researchers, the most prominent among whom is the head of

[1] I follow the original definition of "singularity" by Ray Kurzweil.

[2] Alan Turing believed that cognition could be explained by doing: create a machine that is on a par with a human, and you have explained cognition (cfr. Turing 1936).

[3] This position was incepted in Turing's early years, before obtaining a Ph.D. After the publication of the Gödel's incompleteness theorem, Turing made an attempt in his dissertation to reframe his views, but later ultimately had to abandon them.

[4] Autonomous cognition and intelligence are used interchangeably here: they are not synonyms, but it is largely accepted that in order to be intelligent you must exhibit some sort of cognition as well. The opposite is not true. (See Dumper *et al.*)

the engineering department of Google, Ray Kurzweil, who coined the term "singularity" as the inexorable result of this acceleration leading to change so rapid and deep than human life will be irreversibly transformed (Kurzweil 2005). Both ideas, that of intelligent feedback improvement and exponential intelligence explosion ending with "singularity" are traceable to Good's concept of "ultraintelligence" ("Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an 'intelligent explosion', and the intelligence of man will be left far behind" (Good 1966: 33)) and persisted into Nick Bostrom's modern-day bestseller *Superintelligence: Path, Dangers, Strategies*:

"Before the prospect of an intelligent explosion, we humans are like small children playing with a bomb. Such is the mismatch between the power of our plaything and the immaturity of our conduct. Superintelligence is a challenge for which we are not ready now and will not be ready for a long time. We have a little idea when the detonation will occur, though if we hold the device to our ear we can hear a faint ticking sound." (Bostrom 2015, 319)

Hence, two of the most famous myths were created to capture the lopsided intellectual bias of two scientists gifted with incredible intelligence themselves Good and Turing, and the depth of influence that they have preached has been enough to predestine the immergence of a bestselling book, which although its author claims to be scientific in nature, tells nothing but hypothetical futuristic stories. These two ideas caught incredible amount of media coverage as well. In the early foundational years of A.I., and till present times they are still highly oversold from science publication to sci-fi movies and comprise most of the hype surrounding any new A.I. model that is presented to the public.

Now we will analyze why the hypothesis of self-improving machines and "singularity" are not scientifically sustainable and could only be accepted as myths or as sci-fi literature content. First let us begin with the self-improving mechanism, that allows for design of more intelligent machines, that Good and Bostrom both adopt to argue the possibility of superintelligence. Unsurprisingly, the exact mechanism or the base position from which this extreme self-improvement starts, is never specified in neither mathematical nor empirical way. It seems that the phenomena of cascade intelligence to superintelligence is taken as obvious and therefore not submitted to any need to be even rigidly elaborated upon in technical terms. It is actually pretty obvious on another hand, that simple hardware upgrade, which increase the computation power of a compute does not linearly or exponentially increase or by any mean firstly make a computer intelligent and secondly more intelligent in any way. This is a very trivial conclusion, but if is not the hardware, then either the software or both must evolve in a way that reflects biological evolution. If we consider that in the biological world there are already intelligent agents, there is no evidence that these agents have ever designed more intelligent version of themselves. The reason for this situation is, that to be able to understand how to create artificial brains wielding greater general intelligence than ours, we must first understand our base position, our own cognition, our own reasoning mechanisms, before claiming that an artificial automatic improvement by design is in any way possible. Reflecting again on our own enigma and our own general intelligence, why would we claim that another entity that possesses general intelligence would have any deeper insight to its own cognitive abilities then us? What we have here is not exponential self-improvement towards general intelligence, because only an already generally intelligent agent could have any chance in recognizing what even an overall improvement or increase in general intelligence is.

As for the second myth of “singularity”, it is, to say the least, strictly predetermined by the success of the previous one. “Singularity” is only plausible if exponential growth in self-improving generally intelligent machine is possible. But even in the case that such machine behavior is inevitable Ray Kurzweil (2001) supports his hypothesis by adhering to the Law of Accelerating Returns, by which any scientific advances, proliferated by A.I. must feed on themselves, increasing the rate of further advance, and pushing well past what one might sensibly project by linear extrapolation of current progress. Kurzweil regards the latter explanation as a law but provides no demonstration of it; it is based on the bias of his interpretation, that scientific progress must be viewed as a constant upholding trend. Again, as it was in the case of Good-Bostrom self-improving hypothesis, there is nothing, but the author’s own kaleidoscope that could help us not to dismiss both of the hypothesis, not just as simply being wrong, but also as pseudoscientific. What Kurzweil and his propounders tend to ignore is that even if we have intelligent machine that we could objectively judge as possessing general cognition abilities and abundance of natural resources, so these machines could conduct and carry scientific progress, it is of most crucial interest to be reasonably aware that science has never been an activity, that has a progressive upholding trend, not linear and not ever exponential. Philosophers of science like Thomas Kuhn and Imre Lakatos have considerable contribution to the analysis of the history of science itself. What they suggested is that even if scientific theories really experience regular long spans of classical improvement, meaning that they would be constantly improved on, and their core ideas constantly reused to build better methods for prediction, occasionally every theory amounts to unbreachable problems within itself. When this state ensues, the theory and all of what have been built on top of it will inevitably transit into state of crisis, which could only be circumvent by abandoning the old theory and collectively adopting another, still not well established theory. These cycles of classical advancement and crisis are not just something which happens when already established theories cannot surmount to concrete and tractable prediction; they are also result of the social aspect induced by the scientists themselves. It is the degree of commitment of the scientists to their scientific theory, the scientists’ shared theoretical beliefs, values, instruments and techniques, and even philosophy. The strong or weak commitment to adhere to the current well-established theory or to altogether abandon it, is what also play significant role in how fast transition to and out of a crisis period would be completed. All of that is in stark contrast of that how general intelligent machines are hypothesized to carry science to reach “Singularity”. Scientific progress is simply not uniform and cannot be represented as linear or exponential growing trend; the natural progress of science, from historical perspective is a complex endeavor that is already being performed by generally intelligent entities, and is chaotic with transitional periods from smooth regions to rigid nonlinearity and vice versa. However, we cannot fully dismiss the possibility of “singularity”; what we argue here is that “singularity”, just as a sole result of self-improving generally intelligent machines that cause exponential scientific improvement is not feasible. “Singularity” might very well be the natural convergence of the science itself with or without any use of artificial intelligence.

As we have argued up to here, the myths in A.I. could be traced back to the seminal papers of one of its founding fathers, Alan Turing. The next segment will present an analysis of how the mass media have significantly decreased in their publications the complexity of the scientific achievements in the field of A.I., to such level that the most perpetrated method of injecting any knowledge into the public is, favorably, to compare it with a natural human skill of performing very narrow specific task. This superficial reporting of the field’s achievement is not to be blamed entirely to the journalist community, as we showed in the previous sections, scientists could be strongly committed to their theories. Every scientist who exaggerates a hypothesis or result, is partly inclined to do so on the hope that their work will bear fruition, and partially because of the underling bias towards their cordial theory of adherence. Computer scientists working in the field of artificial intelligence, have one great disadvantage in stark contrast with scientists working, for example, in the field of physics or mathematics, they do not have the luxury of possessing concrete benchmarks while investigating how to artificially recreate any part of the human cognition. Let us take for example image or object recognition as one of the most pursued goals in the field. From 2005 to 2010 the most prominent way of testing new vision A.I. models was the annual PASCAL Visual Object Classes competition,[5] which by 2010 was offering to the competing teams the possibility to test their models against nearly fifteen thousand photographs.

Each photograph has been manually labeled with certain classification string as for example, ‚dog‘, ‘cat’, ‘boat’, ‘building’, ect., the competitors would receive couple of thousand of these photographs in advance, in order to train their classification algorithms, this training set of photographs was available online, and entering the competition was possible for everybody who wanted to participate.[6] After the training phase the judges provided a test set. The best competitor’s algorithm, will be the one which will score the highest percentage of correctly matched labels to images. For example, if hundred images are provided for a test, usually the best algorithm will recognize and correctly label around forty to sixty images from the total one hundred: that was the case before 2014.[7] The general goal behind the competition was to boost researchers, to present solution which would exhibit ability to generalize between the images of the training set, instead, however, the PASCAL competition, although hugely influential, was often criticized that was producing an overfitting effect instead. The researchers were focusing too much on achieving a good score on the particular training and test set that the PASCAL judges were distributing, rather than

[5] Vision models, before 2012, were mainly based on cluster analysis, based on technique called Support Vector Machine (SVM). Fine tuning of SVM is what practically the PASACAL competition was about.
[6] Model(s) and algorithm(s) are used interchangeably here.
[7] See Deng *et al*. 2009.

really creating models that could generalize successfully, outside the PASCAL competition itself. After 2010, that issue was addressed. The solution, however, did not really escape the initial entrenchment. After PASCAL, a much bigger data set of images was created, ImageNet. Up to this day ImageNet collection contains almost 3.2 million images segregated in 5247 different categories and subcategories. So, the general direction of how to address the strive to increase the ability to generalize in the real world, was to increase the number of images from the real world. In the end, ImageNet, has led to the improvement of one particular model, with has architecture based on artificial neural networks (ANN), namely ConvNet (O'Shea, Nash 2015). ConvNet, was particularly well suited to be trained on large sets of image data, which ImageNet provided. The result reported in 2015 was impressive: 98% over the final test set, which meant millions of correct predictions. The huge amount of training examples automatically increased the capability of ConvNet to generalize and correctly classify objects in the real world, simply because more data meant more freedom for the model to pick variations between the training data points. In the end, the success of ConvNet was harnessed by huge corporations like Microsoft, Google, and Facebook to improve their automatic photo recognition systems (photo tagging) and search engines. The glamorous success of ConvNets was the reason a very infamous paper to be published by the Chinese company Baidu, titled "Deep Image: Scaling up Image Recognition", in which they claimed that their particular version of ConvNet is better than humans in object recognition. This paper was later retracted. About the same time, Microsoft published in a research blog "…major advance in technology designed to identify the objects in a photograph or video, showcasing a system whose accuracy meets and sometimes exceeds human level performance" (Linn 2015, 1).

Both companies included an important disclaimer that they are talking about accuracy specifically on ImageNet. The mass media were not, however, so cautious, resulting in headlines such as "Computers Now Better Than Humans At Recognizing And Sorting Images" (Hern 2015). This assertion is based on one particularly lopsided claim that has also been injected into the lore of A.I from the media. The claim is that the human error rate on image recognition in ImageNet is 5%, which make it worst in comparison to the 2% error rate in ConvNet. What the media has omitted to analyze is what is the actual format of a ConvNet prediction. One may think that if a tester "shows" a picture of a dog to the algorithm the corresponding output should be a string label "dog", the reality is somehow different: what the algorithms are producing is a statistical measure of the top-five most likely answers. So, the output is actually a list of the form "cat, dog, sofa, house, person", if the real answer happens to be in the list, the current test instance is counted as correct. The position of the objects in the list has a meaning, the object in the beginning of the list is the one which, according to CovnNet, has the greatest statistical chance of being the correct label. If, however, we only consider the first object as solely counted as the final prediction, not the whole list, the performance of ConvNet drops to 82% as reported in 2017 (Zoph *et al.* 2017). The second, peculiar part is how the 5% error rate in human was measured. The experiment, which gave this result, was performed by Andrej Karpathy, who reportedly tried to classify 1500 images from ImageNet and was wrong in 75 of them or 5% (Karrpathy 2014).

There are many other media titles that are designed to shock the readers. What seems to capture at best the readers attention is a subtle claim of predictions accuracy, that seems impossible to be carry on rationally by humans, but taken as a hunch, if our prediction happens to be the correct one. Examples of such bogus titles are Next Web's "This scary A.I. has learned how to pick out criminals by their faces" (Mix 2016) and Newsweek's "A.I. Can Tell If You're Gay: Artificial Intelligence Predicts Sexuality from One Photo with Startling Accuracy" (Ahmed 2017). These two titles are not designed to inform the readers with well-constructed and balanced report, although they are citing scientific papers as their original sources, more careful reading of the science papers themselves reveals that the authors went in great depth of technical terms and mathematical equation to frame the term "accuracy" in connotation that overstate their model performance. The result then is not as quite impressive as it was reported by the media. Models cannot predict our sexual or crime inclinations, what they do is to overfit over an already known test set, in which it has already been well established which image relates to a person who is gay or criminal.

The latter two media articles are yet another showcase of the misuse of subtle and faint information coming from the research field of A.I., which thanks to the Internet is extremely easy to spread. Nowadays is almost effortless to cause immediate hype on a given topic but if scientists are intentionally or unintentionally biased about their research and journalists only care to generate enough clicks on their cities, what chance is there, for the average reader, to subdue this glaring articles to scrutiny?

We argue that this behavior is not entirely intentional showing experimental results that praise scientists hard work is always favorable instead of showing what was oversell as promising research. Unfortunately, the outcome is a pseudo-scientific sci-fi expectation in our society and as the history has already shown to us, if extremely hyped expectations are not met, funds and new projects are cut off extremely abruptly. In 1974, the field of A.I. entered its first Winter, after the famous Lighthill report (see Agar 2020). The main point of the report was that due to the problems of combinatorial explosion and intractability, the research was deemed as utter failure to achieve its grandiose objectives. The same happened in the 1990s with the collapse of the so called 'expert systems' (Brock 2018), which had one excruciating weak point, the inability to learn. That meant that every new situation needed to be updated constantly by encoding new symbolic rules, which drove the price for maintaining to astronomical figures. In 2022, by the time of writing this article, most of the previous problems have been significantly mitigated, but not entirely resolved. Although generalization was massively increased by the incorporation of artificial neural nets and the creation of deeper and deeper architectures with many hidden layers,[8] those models cannot generalize outside their respective training set, which means that any new knowledge has to be incorporated in the existing training set and relearn from a scratch, this unfortunate phenomena causing neural nets to 'catastrophically forget' is on a par of maintaining coast with the 'expert system' in the past. It will not be fair to leave out and not to praise the success of neural networks: they have launched a new

[8] The so-called Deep Learning paradigm in Machine Learning. Machine Learning is subfield of artificial intelligence concentrating of how to create more robust algorithms when it comes to induce new knowledge in computation machines.

era in A.I. research and ultimately products based on that technology are now almost impossible to exclude from our everyday life (to name but a few: search engines, text-autocomplete, voice and speech recognition, machine translation, autonomous vehicle, recommendation systems and many more). It is easy not to be particularly skeptical to what neural networks may or may not deliver, since only in ten years between 2012 and 2022 most of the technical world now has been completely merged with that technology. Although the amount of success could easily render the readers lightheaded, we as a society must hold these new technologies in high scrutiny and be mindful of the failures, which the field of A.I. also suffered in the last ten years. This is the best approach to balance and regulate the endemic hype that, we are seeing accompanying the enigmatic words ‘artificial intelligence’. After the success of ConvnNet, there were plethora of scandalous setbacks which significantly increased the negative attention towards A.I. research. In 2015 Google’s product Google Photos labeled Jacky Alcine and his black friends as “gorillas”, this scandal induced Google to reportedly revise and fix their model, but investigation by *Wired,* showed that Google only managed to block its image recognition algorithm from identifying gorillas altogether (Hern 2018). This incident is in result of yet another mechanism or problem that is even more dangerous that the intractability and the combinatory explosion that killed the research trends back in the 70s. Nowadays, neural network models cannot be investigated in order to answer the question as to why they have produced a certain output. This is famously called “the explainable problem in neural nets”. Researchers cannot examine their post trained models on the base of tightly formulated mathematical equation, for, the connections that have been created during training, remain hidden into the deep obscure layers of those nets. This problem resembles the past ones in the sense that it is inseparable due the sole nature of the core technology. As long as you are using neural nets, you will end up with unexplainable black box predictions.

Summing up, the glaring upholding hype trend, the bias and the overselling from the media and scientists, the impassible internal problem of explainability – all this is more or less what has already been observed, before the past two Winters. These details might very well suggest that we are in front of another Winter and the way to avert is to commit to reporting more balanced interpretations of the results in the field and demand well-structured communication from the science communicators with the broad public. It is very important to check and double check every shocking media publication in relation to A.I, and to compare it carefully to what is being already established as stable from more than one source. To be mindful of the scientific biases, media inclined hype trends, and the general detriments which comes out-of-the-box with any particular technology. Following these simple recommendations, we hope that the general trend of glaring mass media titles about A.I. kept at bay.

In conclusion: The field of A.I. is currently experiencing its unprecedented levels of unfiltered mass media news publications. What they omit to inform the broad public is why and how those results should be submitted to scrutiny. We are witnessing proclamation of almost supernatural abilities being attained by computer models, but upon closer inspection the results are usually framed in a way to serve a subtle bias toward the preferable end goal of the researchers. What is important for the reader is to observe the signs which suggest that “accuracy” might have been reframed, that the opposite viewpoint has been omitted and that criticisms have not been issued from the journalists towards the scientists. We must always be vigilant about the fact that, due to the complexity of artificial neural networks, a large portion of what is assumed as great scientific success never survives its first contact with reality. Armed with that overview of the nature of the science behind A.I., the general public must try to inoculate itself to the media drifts which are causing out of control hype and threatening dawning of a new Winter.

References

Agar, J. 2020. What is Science for? The Lighthill Report on Artificial Intelligence Reinterpreted. *The British Journal for the History of Science*, 53(3), pp. 289-310.

Ahmed, T. 2017. AI Can Tell if You’re Gay: Artificial Intelligence Predicts Sexuality from One Photo with Startling Accuracy. *Newsweek* [online], 8 September. Available at: https://www.newsweek.com/ai-can-tell-if-youre-gay-artificial-intelligence-predicts-sexuality-one-photo-661643 [Accessed 21 September 2022].

Brock, D.C. 2018. Learning from Artificial Intelligence’s Previous Awakenings: The History of Expert Systems. *AI Magazine*, 39(3), 3-15. https://doi.org/10.1609/aimag.v39i3.2809.

Deng, J., Dong, W., Socher, R., L.-J. Li, Kai Li and Li Fei-Fei. 2009. ImageNet: A Large-Scale Hierarchical Image Database. In: *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (CVPR Workshops), Miami, FL, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

Dumper, K. *et al*. Introduction to Psychology: Chapter 7: Cognition & Intelligence Overview. In: Dumper, K. et al. *Introductory Psychology*. Pressbooks [online]. Available at: https://opentext.wsu.edu/psych105/chapter/chapter-7-cognition-intelligence/ [Accessed 26 December 2022].

Good, I.J. 1966. Speculation Concerning the First Ultraintelligent Machine. *Advances in Computers*, Volume 6, pp. 31-88.

Hern, A. 2015. Computers Now Better than Humans at Recognising and Sorting Images. *The Guardian* [online], 13 May. Available at: https://www.theguardian.com/global/2015/may/13/baidu-minwa-supercomputer-better-than-humans-recognising-images [Accessed 21 September 2022].

Hern, A. 2018. Google’s Solution to Accidental Algorithmic Racism: Ban Gorillas. *The Guardian* [online], 12 January. Available at: https://www.theguardian.com/technology/2018/jan/12/google-racism-ban-gorilla-black-people [Accessed 21 September 2022].

Karrpathy, A. 2014. What I Learned from Competing against a ConvNet on ImageNet. *Andrej Karpathy blog* [online], 2 September. Available at: https://www.karpathy.github.io/2014/09/02 [Accessed 21 September 2022].

Kurzweil, R. 2001. Tracking the Acceleration of Intelligence. The Law of Accelerating Returns. *Kurzweil* [online]. Available at: https://www.kurzweilai.net/the-law-of-accelerating-returns [Accessed 21 September 2022].

Kurzweil, R. 2005. *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking.

Linn, A. 2015. Microsoft Researchers Win ImageNet Computer Vision Challenge. *The AI Blog* [online], 10 December. Available at: https://blogs.microsoft.com/ai/microsoft-researchers-win-imagenet-computer-vision-challenge/ [Accessed 26 December 2022].

Mix. 2016. This Scary Artificial Intelligence Has Learned How to Pick Out Criminals by Their Faces. *The Next Web* [online], 25 November. Available at: https://thenextweb.com/news/artificial-intelligence-criminals-face [Accessed 21 September 2022].

O'Shea, K., Nash, R. 2015. An Introduction to Convolutional Neural Networks. *arXiv* [online]. Available at: https://arxiv.org/abs/1511.08458 [Accessed 26 December 2022].

Turing, A. 1936. On Computable Numbers, with an Application to the *Entscheidungsproblem*. *Proceedings of the London Mathematical Society*, 42(2), pp. 230-265.

Zoph, B. *et al*. 2017. Learning Transferable Architectures for Scalable Image Recognition. *arXiv* [online]. Available at: https://arxiv.org/abs/1707.07012 [Accessed 26 December 2022].