

Essay

# The Singularity Isn't Simple! (However We Look at It) A Random Walk between Science Fiction and Science Fact

Vic Grout 

Department of Computing, Wrexham Glyndŵr University, Wrexham LL11 2AW, UK.;  
v.grout@glyndwr.ac.uk; Tel.: +44-1978-293-203

Received: 11 April 2018; Accepted: 18 April 2018; Published: 19 April 2018



**Abstract:** It seems to be accepted that *intelligence*—*artificial* or otherwise—and ‘the singularity’ are inseparable concepts: ‘The singularity’ will apparently arise from AI reaching a, supposedly particular, but actually poorly-defined, level of sophistication; and an empowered combination of hardware and software will take it from there (and take over from us). However, such wisdom and debate are simplistic in a number of ways: firstly, this is a poor definition of the singularity; secondly, it muddles various notions of intelligence; thirdly, competing arguments are rarely based on shared axioms, so are frequently pointless; fourthly, our models for trying to discuss these concepts at all are often inconsistent; and finally, our attempts at describing any ‘post-singularity’ world are almost always limited by anthropomorphism. In all of these respects, professional ‘futurists’ often appear as confused as storytellers who, through freer licence, may conceivably have the clearer view: perhaps then, that becomes a reasonable place to start. There is no attempt in this paper to propose, or evaluate, any research hypothesis; rather simply to challenge conventions. Using examples from science fiction to illustrate various assumptions behind the AI/singularity debate, this essay seeks to encourage discussion on a number of possible futures based on different underlying metaphysical philosophies. Although properly grounded in science, it eventually looks beyond the technology for answers and, ultimately, beyond the Earth itself.

**Keywords:** futurism and futurology; hard science fiction; artificial intelligence; models of consciousness; intelligent machines; machine replication; machine evolution and optimization; technological singularity

## 1. Introduction: Problems with ‘Futurology’

This paper will irritate some from the outset by treating ‘serious’ academic researchers, professional ‘futurists’ and science fiction writers as largely inhabiting the same techno-creative space. Perhaps, for these purposes, our notion of science fiction might be better narrowed to ‘hard’ sci-fi (i.e., that with a supposedly realistic edge) loosely set somewhere in humanity’s (rather than anyone else’s) future; but, that aside, the alignment is intentional, and no apology is offered for it.

Moreover, it is justified to a considerable extent: purveyors of hard future-based sci-fi and professional futurism have much in common. They both have their work (their credibility) judged in the here-and-now by known theory and, retrospectively, by empirical evidence. In fact, there may be a greater gulf between (for example) the academic and commercial futurist than between either of them and the sci-fi writer. Whereas the professional futurists may have their objectives set by technological or economic imperatives, the storyteller is free to use a scientific premise of their choosing as a blank canvas for *any* wider social, ethical, moral, political, legal, environmental or demographic discussion. This 360° view is becoming increasingly essential: asking technologists their view on the future of technology makes sense; asking their opinions regarding its wider impact may not.

### 1.1. Futurology 'Success'

Therefore, for our purposes, we define an alternative role, that of '*futurologist*' to be any of these, whatever their technical or creative background or motivation may be. A futurologist's predictive success (or *accuracy*) may be loosely assessed by their performance across three broad categories: *positives*, *false positives* and *negatives*, defined as follows:

- **Positives:** predictions that have (to a greater or lesser extent) come to pass within any suggested time frame [*the futurologist predicted it and it happened*];
- **False positives:** predictions that have failed to transpire or have only done so in limited form or well beyond a suggested time frame [*the futurologist predicted it but it didn't happen*];
- **Negatives:** events or developments either completely unforeseen within the time frame or implied considerably out of context [*the futurologist didn't see it coming*].

Obviously, these terms are vague and subject to overlap, but they are for discussion only: there will be no attempt to quantify them here as metrics. (However, see [1] for an informal attempt at doing just this!) Related to these is the concept of justification: the presence (or otherwise) of a coherent argument to support predictions. Justification may be described not merely by its presence, partial (perhaps unconvincing) presence or absence but also by its form or nature; purely scientific, for example, or based on wider social, economic, legal, etc. grounds. The ideal would be a set of predictions with high accuracy (perhaps loosely the ratio of positives to false positives and negatives) together with strong justification. The real world, however is never that simple. Although unjustified predictions might be considered merely guesses, some guesses are lucky: another reason why the outcomes of factual and fictional writing may not be so diverse. Finally, assessment of prediction accuracy and justification rarely happen together. Predictions made *now* can be considered for their justification but not their accuracy: that can only be known later. Predictions made in the past can have their positives and negatives scrutinised in detail but their justification has been made, by the passage of time and the comfort of certainty, less germane.

This flexible assessment of accuracy (positives, false positives and negatives) and justification can be applied to any attempt at futurology, whatever its intent: to inform, entertain or both. We start then with one of the best known examples of all.

Star Trek [2,3] made its TV debut in 1966. Although five decades have now passed, we are still over two centuries from its generally assumed setting in time. This makes most aspects of assessment of its futurology troublesome but, as a snapshot exercise/example in the here-and-now, we can try:

- **Positives:** (as of 2018) voice interface computers, tricorders (now in the form of mobile phones), Bluetooth headsets, in-vision/hands-free displays, portable memory units (now disks, USB sticks, etc.), GPS, tractor beams and cloaking devices (in limited form), tablet computers, automatic doors, large-screen/touch displays, universal translators, teleconferencing, transhumanist bodily enhancement (bionic eyes for the blind, etc.), biometric health and identity data, diagnostic hospital beds [4];
- **False Positives:** (to date) replicators (expected to move slowly to positive in future?) warp drives and matter-antimatter power, transporters, holodecks (slowly moving to positive?), the moneyless society, human colonization of other planets [5];
- **Negatives:** (already) The Internet? The Internet of Things? Body Area Networks (BANs)? Personal Area Networks (PANs)? Universal connectivity and coverage? [6] (*but all disputed* [7]);
- **Justification:** pseudo-scientific-technological in part but with social, ethical, political, environmental and demographic elements?

A similar exercise can be (indeed, informally has been) attempted with other well-known sci-fi favourites such as Back to the Future [8] and Star Wars [9] or, with the added complexity of insincerity, Red Dwarf [10].

### 1.2. Futurology ‘Failure’

In Star Trek’s case, the interesting, and contentious, category is the negatives. Did it really fail to predict modern, and still evolving, networking technology? Is there no Internet in Star Trek? There are certainly those who would defend it against such claims [11] but such arguments generally take the form of noting secondary technology that could *imply* a pervasive global (or universal) communications network: there is no first-hand reference point anywhere. The ship’s computer, for example, clearly has access to a vast knowledge base but this always appears to be held locally. Communication is almost always point-to-point (or a combination of such connections) and any notion of distributed processing or resources is missing. Moreover, in many scenes, there is a plot-centred (clearly intentional) sense of isolation experienced by its characters, which is incompatible with today’s understanding and acceptance of a ubiquitous Internet: on that basis alone, it is unlikely that its writers intended to suggest one.

This point regarding (overlooking) ‘negatives’ may be more powerfully made by considering another sci-fi classic. Michael Moorcock’s *Dancers at the End of Time* trilogy [12] was first published in full in 1976 and has been described as “one of the great postwar English fantasies” [13]. It describes an Earth millions of years into the future in which virtually anything is possible. Through personal ‘power rings’, its cast of decadent eccentrics have access to almost limitless capability from technology established in the distant past by long-lost generations of scientists. As the hidden technology consumes vast, remote stellar regions for its energy, the inhabitants of this new world can create life and cheat death. It is an excellent example of the use of a simple sci-fi premise as a starting point for a wider social and moral discussion. Setting the scene takes just a few pages but then it gets interesting: *how do people behave* in a world where *anything* is possible?

Yet, there is no Internet; or even anything much by way of mobile technology. If ‘The Dancers’ want to observe what might be happening on the other side of the planet, they use their power rings to create a plane (or a flying train) and go there at high speed: fun, perhaps, but hardly efficient! A mere two decades before the Internet became a reality for most people in the developed world, Moorcock’s vision of an ultimately advanced technological civilization did not include such a concept.

Of course, many writers from E.M. Forster (in 1928) [14], through Will Jenkins (1946) [15], to Isaac Asimov (1957) [16], even Mark Twain (much earlier in 1898) [17], have described technology with Internet-like characteristics so there can be no concluding that imagining any particular scientific innovation is necessarily impossible. However each of these, whilst pointing the way in this particular respect to varying degrees, are significantly adrift in other areas. In some there is an archaic reliance on existing equipment; in others, a network of sorts but no computers. Douglas Adams’s *Hitchhiker’s Guide to the Galaxy* (1978) [18] is a universal knowledge base without a network. There is no piece of futurology from more than two or three decades ago which portrays the current world particularly accurately in even most, let alone all, respects. Not only does technological advancement have too many individual threads, the interaction between them is hugely complex.

This is not a failing of fiction writers only. To give just one example, in 1977, Ken Olsen, founder and CEO of DEC [19], made an oft-misapplied statement, “*there is no reason for any individual to have a computer in his home*”. A favourite of introductory college computing modules, supposedly highlighting the difficulty in keeping pace with rapid change in computing technology, it appears foolish at a time when personal computers were already under development, including in his own laboratories. The quote is out of context, of course, and applies to Olsen’s scepticism regarding fully-automated assistive home technology systems (climate control, security, cooking food, etc.) [20]. However, as precisely these technologies gain traction, there may be little doubt that, if he stands unfairly accused of being wrong in one respect, time will inevitably prove him so in another. This form of ‘it won’t happen’ prediction (but it then does) can be considered simply as an extension of the *negatives* category for futurology success or, if necessary, as an additional false negatives set; but this would digress this discussion unnecessarily at this point.

Therefore, the headline conclusion of this introduction is, of course, the somewhat obvious ‘futurology is difficult’ [21]. However, there are three essential, albeit intersecting, components of this observation, which we can take forward as the discussion shifts towards the eponymous technological singularity:

- *Advances in various technologies do not happen independently or in isolation:* Prophesying the progress made by artificial intelligence in ten years’ time may be challenging; similarly, the various states-of-the art in robotics, personal communications, the Internet of Things and big data analytics are difficult to predict with much confidence. However, combine *all* these and the problem becomes far worse. Visioning an ‘always-on’, fully-connected, fully-automated world driven by integrated data and machine intelligence superior to ours is next to impossible;
- *A dominant technological development can have a particularly unforeseen influence on many others:* Moorcock’s technological paradise may look fairly silly with no Internet but the ‘big thing that we’re not yet seeing’ is an ever-present hazard to futurologists. As we approach (the possibility of) the singularity, what else (currently concealed) may arise to undermine the best intentioned of forecasts?
- *Technological development has wider influences and repercussions than merely the technology:* Technology does not emerge or exist in a vacuum: its development is often driven by social or economic need and it then has to take its place in the human world. *We* produce it but it then changes *us*. Wider questions of ethics, morality, politics and law may be restricting or accelerating influences but they most certainly cannot be dismissed as irrelevant.

These elements have to remain in permanent focus as we move on to the main business of this paper.

## 2. Problems with Axioms, Definitions and Models

Having irritated some traditional academics at the start of the previous section, we begin this one in a similar vein: by quoting Wikipedia [22].

*“The technological singularity (also, simply, the singularity) [23] is the hypothesis that the invention of artificial superintelligence will abruptly trigger runaway technological growth, resulting in unfathomable changes to human civilization [24]. According to this hypothesis, an upgradable intelligent agent (such as a computer running software-based artificial general intelligence) would enter a “runaway reaction” of self-improvement cycles, with each new and more intelligent generation appearing more and more rapidly, causing an intelligence explosion and resulting in a powerful superintelligence that would, qualitatively, far surpass all human intelligence. Stanislaw Ulam reports a discussion with John von Neumann “centered on the accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue” [25]. Subsequent authors have echoed this viewpoint [22,26]. I. J. Good’s “intelligence explosion” model predicts that a future superintelligence will trigger a singularity [27]. Emeritus professor of computer science at San Diego State University and science fiction author Vernor Vinge said in his 1993 essay *The Coming Technological Singularity* that this would signal the end of the human era, as the new superintelligence would continue to upgrade itself and would advance technologically at an incomprehensible rate [27].*

*At the 2012 Singularity Summit, Stuart Armstrong did a study of artificial general intelligence (AGI) predictions by experts and found a wide range of predicted dates, with a median value of 2040 [28].*

*Many notable personalities, including Stephen Hawking and Elon Musk, consider the uncontrolled rise of artificial intelligence as a matter of alarm and concern for humanity’s future [29,30]. The consequences of the singularity and its potential benefit or harm to the human race have been hotly debated by various intellectual circles.”*

Informal as much of this is, it provides a useful platform for discussion and, although there may be more credible research bases, no single source will be the consensus of such a large body of opinion: in this sense, it can be argued to represent at least the common view of the Technological Singularity (TS). Once again, we see sci-fi suggestions of the TS (Vinge) pre-dating academic discussions (Kurzweil [31]) and consideration of the wider social, economic, political and ethical impact also follows close behind: these various dimensions cannot be considered in isolation; or, at least, should not be.

### 2.1. What Defines the 'Singularity'?

However, before embarking on a discussion of what the TS actually is, if it might happen, how it might take place and what the implications could be, a more fundamental question merits at least passing consideration: is the TS really a 'singularity' at all?

This is not an entirely simple question to answer. Different academic subjects, from various fields in mathematics, through the natural sciences, to emerging technologies have their own concept of a 'singularity' with most dictionary definitions imprecisely straddling several of them. Sci-fi [27] continues to contribute in its own fashion. However, a common theme is the notion of the rules, formulae, behaviour or understanding of any system breaking down or not being applicable at the point of singularity. ('We understand how the system works everywhere except here') A black hole is a singularity in space-time: the conventional laws of physics fail as gravitational forces become infinite (albeit infinitely slowly as we approach it). The function  $y = 1/x$  has a singularity at  $x = 0$ : it has no (conventional) value. On this basis, is the point in our future, loosely understood to be that at which machines start to evolve by themselves, really a singularity? To put it another way, does it suggest the sort of discontinuity most definitions of a singularity imply?

It may not. Whilst there may be little argument with there being a period of great uncertainty following on from the TS, including huge questions regarding what the machines might do or how long (or just how) humans, or the planet, might survive [31], it is not clear that the TS marks an impassable break in our timeline. It is absolutely unknown what happens to material that enters a black hole; it may appear in some other form elsewhere in the universe but that is entirely speculative: there is no other side, of which we know. The  $y = 1/x$  curve can be followed smoothly towards  $x = 0$  from either (positive or negative) direction but there is no passing continuously through to the other. However, at the risk of facetiousness, if we go to bed one night and the TS occurs while we sleep, we will still most likely wake up the following morning. The world may indeed have become a very different place but it will probably continue, for a while at least. It is possible that we sometimes confuse 'uncertainty' with 'discontinuity' so perhaps 'singularity' is not an entirely appropriate term?

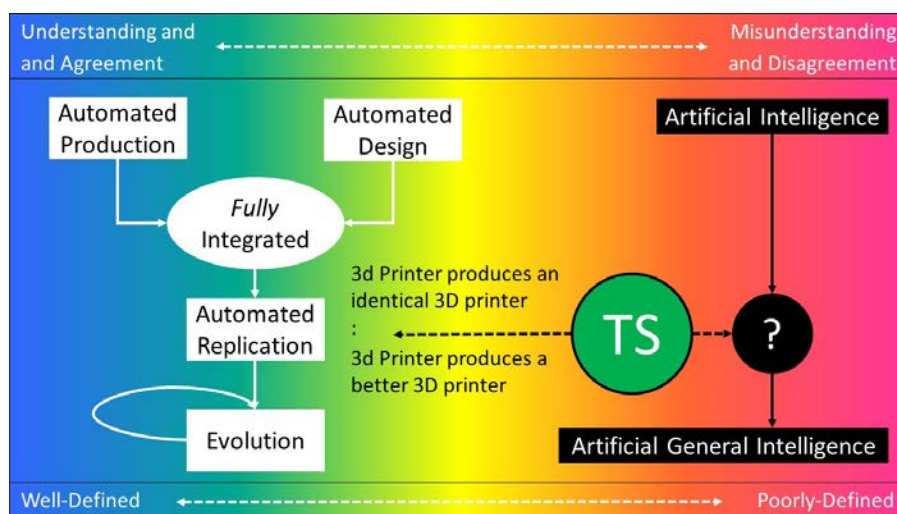
Returning, to less abstract matters, we attempt to consider what the TS might actually be. Immediately, this draws us into difficulty as we begin to note that most definitions, including the Wikipedia consensus, conflate different concepts:

- *Process-based*: (forward-looking) We have a model of what technological developments are required. When each emerges and combines with the others, the TS should happen: 'The TS will occur *when* ...'; [This can also be considered as a '*white box*' definition: we have an idea of how the TS might come about];
- *Results-based*: (backward-looking) We have a certain outcome or expectation of the TS. The TS has arrived when technology passes some *test*: 'The TS will have occurred *when* ...'; [Also a '*black box*' definition: we have a notion of functionality but not (necessarily) operation];
- *Heuristic*: ('rule-of-thumb', related models) Certain developments in technology, some themselves defined better than others, will continue along (or alongside) the road towards the TS and may contribute, directly or indirectly, to it: 'The TS will occur as ...'; [Perhaps a '*grey box*' definition?].

Figure 1 attempts to distil these overlapping principles in relation to the TS. Towards the left of the diagram, more is understood (with clearer, concrete definitions) and more likely to achieve consensus



among futurologists. Towards the right, concepts become more abstract, with looser definitions and less agreement. Time moves downwards.



**Figure 1.** Combined (but simplified) models of the technological singularity. TS, Technological Singularity.

## 2.2. How Might the ‘Singularity’ Happen?

Any definition of the TS by reference to AI, itself not well defined, or worse, to AGI (*Artificial General Intelligence*: loosely the evolutionary end-point of AI to something ‘human-like’ [32]) is, at best, a heuristic one. Not only does this not provide a deterministic characterization of the TS, it is not *axiomatically* dependent upon it. Suppose there was an *agreed* formulaic measurement for AI (or AGI), which there is not, what *figure* would we need to reach for the TS to occur? We have no idea and it is unlikely that the question is a functional one. In addition, *is* it unambiguously clear that this is the *only* way it could happen? Not if we base our definitions on precise *expectations* rather than *assumptions*.

A process-based model (white box), built on what we already have, is at least easier to understand. *Evolution* is the result of iterated replication within a ‘shaping’ or ‘guiding’ environment. In theory, at least, (automatic) replication is the automated combining of *design* and *production*. We already see examples of both automated production and automated design at work. Production lines, including robotic hardware, have been with us for decades [33]: in this respect, accelerated manufacturing began long ago. Similarly, software has become an increasingly common feature of engineering processes over the years [34] with today’s algorithms being very sophisticated indeed [35]. Many problems in electronic component design, placement, configuration and layout [36,37], for example, would be entirely impossible now without computers taking a role in optimizing their successors. In principle, it is merely a matter of connecting all of this seamlessly together!

Fully integrating automated design and automated production into automated replication is far from trivial, of course, but it does give a better theoretical definition of the TS than vague notions of AI or AGI. It is, however, also awash with practical problems:

- 1 [Design] Currently, the use of design software and algorithms is piecemeal in most industries. Although some of the high-intensity computational work is naturally undertaken by automated procedures, there is almost always human intervention and mediation between these sub-processes. It varies considerably from application to application but nowhere yet is design automation close to comprehensive in terms of a complete ‘product’.
- 2 [Production] A somewhat more mundane objection: but even an entirely self-managed production line has to be fed with materials; or components perhaps if production is multi-stage, but raw materials have to start the process. If this part of the TS framework is to be fully automated

- and *sustainable* then either the supply of materials has to be similarly seamless or the production hardware must somehow be able to source its own. Neither option is in sight presently.
- 3 [Replication] Fully integrating automated design with production is decidedly non-trivial. Although there are numerous examples of production lines employing various types of feedback, to correct and improve performance, this is far short of the concept of taking a *final* product design produced by software and immediately implementing it to hardware. Often, machinery has to be reset to implement each new design or human intervention is necessary in some other form. Whilst 1 and 2 require considerable technological advance, this may ask for something more akin to a human ‘leap of faith’.
  - 4 [Evolution] Finally, if 1, 2 and 3 were to be checked off at some point in the future (*not* theoretically impossible), the transition from static replication to a chain of improvements across subsequent iterations requires a further ingredient: *the guiding environment, or purpose, required for evolution across generations*. Where would this come from? Suppose it requires some extra program code or an adjustment to the production hardware to produce the necessary improvement at each generation, how would (even could) such a decision be taken and what would be the motivation (of the hardware and software) for taking it? Couched in terms of an optimization problem [38], what would the objective function be and where would the responsibility lie to improve it?

Although consideration of 4 does indeed appear to lead us back to some notion of intelligence, which is why the heuristic (grey box) definition of the TS associated with AI or AGI is not outright wrong, it is simply unhelpful in any practical sense. Exactly *what* level of *what* measurement of intelligence would be necessary for a self-replicating system to effect improvement across generations? It remains unclear. Having broken a process-based TS down into its constituents, however, it does suggest an alternative definition.

Because a results-based (black box) definition of the TS could be something much easier to come to terms with. For example, we might understand it to be the point at which a 3D printer autonomously creates an identical version of itself. Then, presumably, with some extra lines of code, it creates a marginally better version. This superior offspring has perhaps slightly better hardware features (precision, control, power, features or capabilities) and improved software (possibly viewed in terms of its ‘operating system’). As the child becomes a parent, further hardware and software enhancements follow and the evolutionary process is underway. Even here we could dispute the *exact* point of the TS but it is clearly framed somewhere in the journey from the initial parent, through its extra code, to the first child. ‘Intelligence’ has not been mentioned.

The practical objections above still remain, of course. We can perhaps ignore 1 [Design] by reminding ourselves that this is a black box definition: we can simply claim to recognize when the outcomes have satisfied our requirements. Similarly 2 [Production] might be overlooked but it does suggest that the 3D printer might have to have an independent/autonomous supply network of (not just present but all possible future) materials required: *very* challenging, arguably *non-deterministic*; the alternative, though, might be that the device has to be *mobile*, allowing it to source or forage for its needs! Finally, although we can use the black box defence against being able to describe 3 [Replication] in detail, the difficult question of 4 [Evolution] remains: *Why would the printer want to do this?* Therefore, this might bring us back to intelligence after all; or does it?

Because, if we are forced to deal with intelligence as a driver towards the TS (even if only to attempt to answer the question of whether it is even required), we encounter a similar problem in what may be a recurring pattern: we have very little idea (understanding or agreement) on what ‘*intelligence*’ is ether, which takes us to the next section.

### 3. Further Problems with ‘Thinking’, ‘Intelligence’, ‘Consciousness’, General Metaphysics and ‘The Singularity’

The irregular notion of an intelligent machine is found in literature long before academic non-fiction. Although it may stretch definitions of ‘sci-fi’, *Talos* [39] appears in stories dating back to the third

century BC. A giant bronze automaton, patrolling the shores of Europa (Crete), he appears to have many of the attributes we would expect of an intelligent machine. Narratives vary somewhat but most describe him as of essentially non-biological origin. Many other examples follow over the centuries with Mary Shelley's monster [40] being one of the best known: Dr. Frankenstein's creation, however, is the result of (albeit corrupted) natural science. These different versions of ('natural' or 'unnatural') 'life' or 'intelligence' continue apace into the 21st century, of course [41–44], but, by the time Alan Turing [45] asked, in 1950, whether such things were possible in practice, sci-fi had already given copious judgement in theory.

How important is this difference between 'natural' and 'unnatural' life? Is it the same as the distinction between 'artificial' and 'real' intelligence? Who decides what is natural and unnatural anyway? (The concepts of 'self-awareness' and 'consciousness' are discussed later.) As Turing pointed out tangentially in his response to 'The Theological Objection' [45], this may not be a question we have the right, or ability, to make abstract judgements on. (To paraphrase his entirely atheist argument: 'even if there were a God, who are *we* to tell Him what he can and cannot do?') However, even in the title of this seminal work ('Computing Machinery and Intelligence'), followed by his opening sentence ('*I propose to consider the question, "Can machines think?"*'), he has apparently already decided that 'thinking' and 'intelligence' are the same thing. Although he immediately follows this with an admission that precise definitions are problematic, whether this equivalence should be regarded as axiomatic is debatable.

### 3.1. 'Thinking' Machines

However, something that Turing [45] forecasts with unquestionable accuracy is that, *'I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted'*. It cannot be denied that today we do just this (and have done for some time): at the very least, a computer is 'thinking' while we await its response. It might be argued that the use of the word in this context describes merely the delay, or even that it has some associated irony attached, but we still do it, even if we do not give any thought to what it means. Leaving aside his almost universally misunderstood 'Imitation Game' ('Turing Test': see later) discussion, in this the simplest of predictions, Turing was right.

(His next sentence is then often overlooked: *'I believe further that no useful purpose is served by concealing these beliefs. The popular view that scientists proceed inexorably from well-established fact to well-established fact, never being influenced by any improved conjecture, is quite mistaken. Provided it is made clear which are proved facts and which are conjectures, no harm can result. Conjectures are of great importance since they suggest useful lines of research.'* This very much supports the guiding principles of this paper: so long as we can recognize the difference between science-fact and anything short of this, conjecture, and even fiction, have value in promoting discussion. Perhaps the only real difference between the sci-fi writer and the professional futurist is that the latter expect us to believe them?)

### 3.2. 'Intelligent' Machines

Therefore, if 'thinking' is a word used too loosely to have much value, what of the perhaps stronger 'intelligence'? Once again, the dictionary is of little help, confusing several clearly different concepts. 'Intelligence', depending on the context, implies a capacity for: *learning, logic, problem-solving, reasoning, creativity, emotion, consciousness or self-awareness*. That these are hardly comparable is more than abstract semantics: does 'intelligent' simply suggest 'good at something' (number-crunching, for example) or is it meant to imply 'human-like'? By some of these definitions, a pocket-calculator is intelligent, a chess program more so. However, by others, a fully-automated humanoid robot, faster, stronger and superior to its creators in every way would not be if it lacked consciousness and/or self-awareness [46]. What is meant (in Star Trek, for example) by frequent references to 'intelligent life'? Is this tautological or are there forms of life lacking intelligence? (After all, we often say this of people we lack respect for.) Are the algorithms that forecast the weather or run our economies



intelligent? (They do things with a level of precision that we cannot.) If intelligence can be anything from simple learning to full self-awareness, at which point does AI become AGI, or perhaps ‘artificial’ intelligence become ‘real’ intelligence. There sometimes seem to be as many opinions as there are academics/researchers. And, once again, we return to what increasingly looks to be a poorly framed question: *how much ‘intelligence’ is necessary for the TS?*

Perhaps tellingly, Turing himself [45] made no attempt to define intelligence, either at human or machine level. In fact, with respect to the latter, accepting his conflation of intelligence with thought, he effectively dismisses the question: he precedes his ‘*end of the century*’ prediction above by, ‘*The original question, “Can machines think?” I believe to be too meaningless to deserve discussion.*’ Instead, he proposes his, widely quoted but almost as widely misunderstood, ‘*Imitation Game*’, eventually to become known as the ‘Turing Test’. A human ‘interrogator’ is in separate, remote communication with both a machine and another human and, through posing questions and considering their responses, has to decide which is which. It probably makes little difference whether either is allowed to lie. (Clearly then something akin to ‘intellectually human-like’ is thus being suggested as intelligence here.) In principle, if the interrogator simply cannot decide (or makes the wrong decision as often as the right one) then the machine has ‘won’ the game.

Whatever the thoughts of sci-fi over the decades [47,48] might be, this form of absolute victory was not considered by Turing in 1950. Instead he made a simple (and quite possibly off-the-cuff) prediction that ‘*I believe that in about fifty years’ time it will be possible, to programme computers, . . . , to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning.*’ Today’s reports [49] of software ‘passing the Turing Test’ because at least four people out of ten chose wrongly in a one-off experiment are a poor reflection on this original discussion. We should also consider that the sophistication of the interrogators and their questioning may well increase over time: both pocket calculators and chess machines seemed unthinkable shortly before they successfully appeared but are now mainstream.

However, the key point regarding this formulation of an ‘intelligence test’ is that it is only loosely based on knowledge. No human knows nothing but some know more than others. Similarly, programming a computer with knowledge bases of various sizes is trivial. The issue for the interrogator is to attempt to discriminate on the basis of how that knowledge is manipulated and presented. ‘Q. What did Henry VIII die from?’ ‘A. Syphilis’, requires a simple look-up but ‘Q. Name a British monarch who died from a sexually transmitted disease’, is more complex since neither the king nor the illness is mentioned in the original question/answer. As a slightly different illustration, asking for interpretations of the signs, ‘Safety hats must be worn’, and ‘Dogs must be carried’, is challenging: although syntactically similar, their regulatory expectations are very different.

### 3.3. ‘Adaptable’ Machines

This, in turn, leads us to the concept of adaptability, often mooted in relation to intelligence, by both human-centric and AI researchers. Some even take this further to the point of using it to define ‘genius’ [50]. The argument is that *what* an entity *knows* is less important than how it *uses* it, particularly in dealing with situations where that knowledge has to be applied in a different context to that in which it was learned. This relates well to the preceding paragraph and perhaps has a ring of truth in assessing the intelligence of humans or even other animals; but what of AI? What does adaptability look like in a machine?

In fact, this is very difficult and rests squarely on what we think might happen as machines approach, then evolve through, the TS. If our current models of software running on hardware (as separate entities) survive, then adaptability may simply be the modification of database queries to analyse existing knowledge in new ways or an extension to existing algorithms to perform other tasks. Thus, in a strictly evolutionary sense, it may, for example, be sufficient for one generation to autonomously acquire better software in order to effect hardware improvements in the next. This is essentially a question of extra code being written, which is becoming simple enough, even

automatically [51], if we (again) put aside the question of what the machine's motivation would be to do this.

All these approaches to machine 'improvement', however, remain essentially under human control. If, in contrast, adaptability in AI or machine terms implies fundamental changes to how that machine actually operates (perhaps, but not necessarily, within the same generation), or we expect the required 'motivation for improvement' to be somehow embedded in this evolutionary process, then this is entirely different, and may well have to appear independently. The same conceptual difficulty arises if we look to machines in which hardware and software are merged into an inseparable operational unit, as they appear to be in our human brains [52]. If the future has machines such as this: machines that have somehow passed over a critical evolutionary boundary, beyond anything we have either built or programmed into them, then something very significant indeed will have happened. What might cause this? To an extent, our accepted framework within which machines/computers operate may have been replaced by something conceptually beyond our control (possibly our understanding), one of the deeper concerns regarding the TS. However, where would this come from? Well, there may be numerous possibilities, so this is by no means an irrefutable logical progression we follow now but, having postponed the discussion to this point, it seems appropriate to deal with 'consciousness' or 'self-awareness'.

### 3.4. 'Conscious' Machines

Not all writers and academics, from neuroscientists to philosophers [53], consider these terms entirely synonymous but the distinction is outside the scope of this paper. However, whether consciousness necessarily implies intelligence, particularly how that might play out in practice, is certainly contentious and returning to sci-fi illustrates this point well. If, only temporarily, we follow a (largely unproven) line of reasoning that consciousness is related to 'brain' size, that is if we create sufficiently complex hardware or software then sentience emerges naturally, then, rather than look to stand-alone machines or even arrays of supercomputers, our first sight of such a phenomenon would presumably come from the single largest (and most complex) thing that humanity has ever created: the *Internet*. Whilst, in this context, Terminator's 'Skynet' [54] is well known, the considerable variants on this theme are best illustrated by two lesser-known novels.

Robert J. Sawyer's *WWW* Trilogy [41–43] describes the emergence of '*Webmind*', an AI apparently made from discarded network packets endlessly circling the Internet topology. Leaving aside the question of whether the 'time-to-live' field [55] typically to be found in such packets might be expected to clear such debris, the story suggests an interesting framework for emergent consciousness based purely on software. Sawyer loosely suggests that a large enough number of such packets, interacting in the nature of cellular automata [42] would eventually achieve sentience. For any real evidence we have to the contrary, perhaps they could; there is nothing fundamentally objectionable in Sawyer's pseudo-science.

However, *Webmind* exhibits an equally serious problem we often observe in trying to discuss AI, probably AGI, in fact, or any post-TS autonomous computer or machine: the tendency to be unrealistically anthropocentric, or worse. In *WWW*, not only is *Webmind* intelligent by any reasonable definition, it is also almost limitlessly powerful. Once it has mastered its abilities, it is quickly able to 'do good' as it sees it, partially guided by a (young) human companion. It begins by eliminating electronic spam, then finds a cure for cancer and finally overthrows the Chinese government. It would remain to be seen, of course, quite how 'good' any all-powerful AI would really regard each of these objectives. At best, the first is merely a high-level feature of human/society, the second would presumably improve human lives but the overall balance of nature is unknown (even if the AI would really care much about *that*) and the last is the opinion of a financially-comfortable, liberally-minded western author: not *all humans* would necessarily share this view, let alone an emergent intelligence! Frankly, whatever A(G)I may perceive its *raison d'être* to be, it is unlikely to be this!

In marked contrast, in the novel, *Conscious*, Vic Grout's 'It' [44], far from being software running on the Internet, is simply the Internet *itself*. The global combination of the Internet and its supporting power networks have reached such a level of scale and connection density, they automatically acquire an internal control imperative and begin behaving as a (nascent form of) independent 'brain'. Hardware and software appear to be working inseparably: although its signals can be observed, they emanate from nowhere other than the hardware devices and connections themselves. Humans have supplied all this hardware, of course, and although, initially, *It* works primitively with human software protocols, these are slowly adapted to its own (generally unfathomable) purposes. *It* begins its life as the physical (wired) network infrastructure before subsuming wireless and cellular communications as well. Eventually, *It* acquires total mastery of the 'Internet of Everything' (the 'Internet of Things' projected a small number of years into the future) and causes huge, albeit random, damage.

Crucially, though, Grout quite deliberately offers no insight into the workings of *Its* 'mind'. Although *It* is almost immeasurably powerful (its peripherals are effectively everything on Earth: human transport, communications, climate control, life-support, weaponry, etc.), it remains throughout no more than embryonic in its sophistication, even its intelligence. If it *can* be considered brain-like, it is a brain in the very earliest stages of development. *It* is undoubtedly *conscious* in the sense that it has acquired an independent control imperative, and, as the story unfolds, there is clear evidence of learning: at least, becoming familiar with its own structure and (considerable) capabilities, but in no sense could this be confused with being 'clever'. There is no human attempt to communicate with *It* and not the slightest suggestion that this could be possible. Although *It* eventually wipes out a considerable fraction of life on Earth, it is unlikely that it understands any of this in any conventional sense. *It* is as much simply a 'powered nervous system' as it is a brain. *Conscious* [44] is only a story, of course, but, as a model, if actual consciousness could indeed be achieved without real intelligence (perhaps it requires merely a simple numerical threshold, neural complexity, for example, to be reached) then it would be difficult to make any confident assertions regarding AI or AGI.

### 3.5. Models of 'Consciousness'

Grout's framework for a conscious Internet is essentially a variant of the philosophical principle of panpsychism [56]. (Although 'true' panpsychism suggests some level of universal consciousness in both living and non-living things, Grout's model suggests a need for external 'power' or 'fuel' and an essential 'tipping point', neural complexity here, at which consciousness either happens or becomes observable.) The focus of any panpsychic model is essentially based on hardware. However, as with Sawyer's model of a large software-based AI, for all we really know of any of this, it is as credible as any other.

As with 'intelligence', there are almost as many models of 'consciousness' [53] as there are researchers in the field. However, loosely grouping the main themes together [57], allows us to make a final, hopefully interesting, point. To this end, the following could perhaps be considered the main 'types' of explanations of where consciousness come from:

1. Consciousness is simply the result of neural complexity. Build something with a big enough 'brain' and it will acquire consciousness. There is possibly some sort of critical neural mass and/or degree of complexity/connectivity for this to happen.
2. Similar to 1 but the 'brain' needs energy. It needs power (food, fuel, electricity, etc.) to make it work (Grout's 'model' [44]).
3. Similar to 2 but with some symbiosis. A physical substrate is needed to carry signals of a particular type. The relationship between the substrate and signals (hardware and software) takes a particular critical form (maybe the two are indistinguishable) and we have yet to fully grasp this.
4. Similar to 3 but there is a biological requirement. Consciousness is the preserve of carbon life forms, perhaps. How and/or why we do not yet understand, and it is conceivable that we may not be able to.

5. A particular form of 4. Consciousness is somehow separate from the underlying hardware but still cannot exist without it (Sawyer's 'model' [41–43]?).
6. Extending 5. Consciousness is completely separate from the body and could exist independently. Dependent on more fundamental beliefs, we might call it a 'soul'.
7. Taking 6 to the limit? Consciousness, the soul, comes from God.

Firstly, before dismissing any of these models out-of-hand, we should perhaps reflect that, if pressed, a non-negligible fraction of the world's population would adopt each of them (which is relevant if we consider *attitudes* towards AI). Secondly, each is credible in two respects: (1) they are logically and separately non-contradictory and (2) the human brain, which we assume to deliver consciousness, could (for all we know) be modelled by any of them. In considering each of these loose explanations, we need to rise above pre-conceptions based on belief (or lack of). In fact, it may be more instructive to look at *patterns*. With this in mind, these different broad models of consciousness can be viewed as being in sequence in two senses:

- In a simple sense, they (Models 1–7) could be considered as ranging from the 'ultra-scientific' to the 'ultra-spiritual'.
- Increasingly sophisticated machinery will progressively (in turn) achieve some of these. 1 and 2 may be trivial, 3 a distinct possibility, and 4 is difficult to assert. After that, it would appear to get more challenging. In a sense, we might disprove each model by producing a machine that satisfied the requirements at each stage, yet failed to 'wake up'.

However, the simplicity of the first 'progression' is spoilt by adding 'pure' panpsychism and the second by Turing himself.

Firstly, if we add the pure panpsychism model as a new item zero in the list:

*Everything* in the universe has consciousness to some extent, even the apparently inanimate: humans are simply poor when it comes to recognizing this, then this fits nicely before (as a simpler version of) 1 whilst, simultaneously, the panpsychic notion of a 'universal consciousness' looks very much like an extension of 9. (It closely matches some established religions.) The list becomes cyclic and notions of 'scientific' vs. 'spiritual' are irreparably muddled.

Secondly, whatever an individual's initial position on all of this, Turing's original 1950 paper [45] could perhaps challenge it ...

### 3.6. A Word of Caution ... for Everyone

In suggesting (loosely) that some form of A(G)I might be possible, Turing anticipated 'objections' in a variety of forms. The first, he represented thus:

*"The Theological Objection: Thinking is a function of man's immortal soul. God has given an immortal soul to every man and woman, but not to any other animal or to machines. Hence no animal or machine can think."*

He then dealt with this robustly as follows:

*"I am unable to accept any part of this, but will attempt to reply in theological terms. I should find the argument more convincing if animals were classed with men, for there is a greater difference, to my mind, between the typical animate and the inanimate than there is between man and the other animals. ... It appears to me that the argument quoted above implies a serious restriction of the omnipotence of the Almighty. It is admitted that there are certain things that He cannot do such as making one equal to two, but should we not believe that He has freedom to confer a soul on an elephant if He sees fit? We might expect that He would only exercise this power in conjunction with a mutation which provided the elephant with an appropriately improved brain to minister to the needs of this sort. An argument of exactly similar form may be made for the case of machines. It may seem different*

*because it is more difficult to 'swallow'. But this really only means that we think it would be less likely that He would consider the circumstances suitable for conferring a soul. . . . In attempting to construct such machines we should not be irreverently usurping His power of creating souls, any more than we are in the procreation of children: rather we are, in either case, instruments of His will providing mansions for the souls that He creates."*

Turing is effectively saying (to 'believers'), 'Who are you to tell God how His universe works?'

This is a significant observation in a wider context, however. Turing's caution, although clearly derived from a position of unequivocal atheism, can perhaps be considered as a warning to any of us who have already decided what is possible or impossible in regard to A(G)I, then expect emerging scientific (or other) awareness to support us. That is not the way anything works; neither science nor God are extensions of ourselves: they will not rally to support our intellectual cause because we ask them to. It makes no difference what domain we work in; whether it is (any field of) science or God, or both, that we believe in, it is *that* that will decide what can and cannot be done, not *us*.

In fact, we could at least tentatively propose the notion that, ultimately, this is something we *cannot* understand. There is nothing irrational about this: most scientific disciplines from pure mathematics [58], through computer science [59] to applied physics [60] have known results akin to 'incompleteness theorems'. We are entirely reconciled to the acceptance that, starting from known science, there are things that cannot be reached by human means: propositions that cannot be proved, problems that cannot be solved, measurements that cannot be taken, etc. However, with a median predicted date little more than 20 years away [22], can the reality of the TS really be one of these undecidables? How? Well, perhaps we may simply never get to find out . . .

#### 4. The Wider View: Ethics, Economics, Politics, the Fermi Paradox and Some Conclusions

The 'Fermi Paradox' (FP) [61] can be paraphrased as asking, on the cascaded assumptions that there are a quasi-infinite number of planets in the universe, with a non-negligible, and therefore still extremely large, number of them supporting life, why have none of them succeeded in producing civilizations sufficiently technologically advanced to make contact with us? Many answers are proposed [62] from the perhaps extremes of 'there isn't any: we're alone in the universe' to suggestions of 'there has been contact: they're already here, we just haven't noticed', etc.

It could be argued or implied, however, that the considerable majority of possible solutions to the FP suggest that *the TS has not happened on any other planet in the universe*. Whatever, the technological limitations of its 'natural' inhabitants may have been, whatever may or may not have happened to them (survival or otherwise) following their particular TS, whatever physical constraints and limitations they were unable to overcome, their new race(s) of intelligent machines, evolving at approaching an infinite rate, would surely be visible by now. As they do not appear to be, we are forced into a limited number of possible conclusions. Perhaps the machines have no desire to make contact (not even with the other machines they presumably know must exist elsewhere) but, again, this (machine motivation) is simply unknown territory for us. Perhaps these TSs are genuinely impossible for all the reasons discussed here and elsewhere [63]. Or perhaps their respective natural civilizations never get quite that far.

##### 4.1. Will We Get to See the TS?

This, indeed, is an FP explanation gaining traction as our political times look more uncertain and our technological ability to destroy ourselves increases on all fronts. It has been suggested [64] that it may be a natural (and unavoidable) fate of advanced civilizations that they self-destruct before they are capable of long-distance space travel or communication. Perhaps this inevitable doom also prevents any TS? We should probably not discount this theory lightly. Aside from the obvious weaponry and the environmental damage being done, we have other, on the surface, more mundane, factors at work. Social media's ability, for example, to threaten privacy, spread hatred, strew false information, cause



division, etc., or the spectre of mass unemployment through automation, could all destabilize society to crisis point.

Technology has the ability to deliver a utopian future for humanity (with or without the TS occurring) but it almost certainly will not, essentially because it will always be put to work within a framework in which the primary driving force is profit [65]. A future in which intelligent machines do all the work for us could be good for everyone; but for it to be, fundamental political-economic structures would have to change, and there is no evidence that they are going to. (Whether ‘not working’ for a human is a good thing has nothing to do with the technology that made their work unnecessary but the economic framework everything operates under. If nothing changes, ‘not working’ will be socially unpleasant, as it is now: technology cannot change that, but there could soon be more out of work than in.) Under the current political-economic arrangement, nothing really happens for the general good, whatever public-facing veneer may be applied to the profit-centric system. However, such inequality and division may eventually (possibly even quickly) lead to conflict? The framework itself, however, appears to be very stable so it may be that it preserves itself to the bitter end. All in all, the outlook is a bleak one. Perhaps the practical existence of the TS is genuinely unknowable because it lies just beyond any civilization’s self-destruct point, so it can never be reached?

#### 4.2. Can We Really Understand the TS?

Coming back to more Earthly and scientific matters, we should also, before ending, note something essential regarding evolution itself: it is imperfect [66]. The evolutionary process is, in effect, a sophisticated local-search optimization algorithm (variants are applied to current solutions in the hope of finding better ones in terms of a defined objective). The objective (for biological species, at least) is survival and the variable parameters are these species’ characteristics. The evolutionary algorithm can never guarantee ‘perfection’ since it can only look (for each successive generation) at relatively small changes in its variables (mutations). As Richard Dawkins puts it [67] in attempting to deal with the question of ‘*why don’t animals have wheels?*’

*“Sudden, precipitous change is an option for engineers, but in wild nature the summit of Mount Improbable can be reached only if a gradual ramp upwards from a given starting point can be found. The wheel may be one of those cases where the engineering solution can be seen in plain view, yet be unattainable in evolution because it lies the other side of a deep valley, cutting unbridgeably across the massif of Mount Improbable.”*

Therefore, advanced machines, seeking evolutionary improvements, may not achieve perfection. Does this mean they will remain constrained (and ‘content’) with imperfection or will/can they find a better optimization process? There are certainly optimization algorithms superior to local-search (often somewhat cyclically employing techniques we might loosely refer to as AI) and this/these new races(s) of superior machines might realistically have both the processing power and logic to find them. In the sense we generally apply the term, machines may not *evolve* at all: they may do something *better*. There may or may not be the same objective. Millions of years of evolution could, perhaps, be bypassed by a simple calculation? However, whatever happens, the chances of us being able to understand it are minimal.

#### 4.3. Therefore, Can the TS Really Happen?

Returning finally to the main discussion, if we persist in insisting that the TS can only arise from the evolution of AI to AGI, or worse that machines have to achieve self-awareness for this to happen then we are beset by difficulties: we really have little idea what this really means so a judgement on whether it can happen is next to impossible. There are credible arguments [68] that machines can never achieve the same form of intelligence as humans and that Turing’s models [41] are simplistic in this respect. If both the AGI requirement for the TS and its (AGI’s) impossibility to achieve are simultaneously correct then, trivially, the TS cannot happen. However, if the TS can result simply from

the conjunction of increasingly complex, but fully-understood, technological processes, then perhaps it can: and a model of what this might look like in practical terms is, not only possible but, independent of definitions of intelligence. Finally, other predictions of various crises in humanity's relationship with technology and its wider impact [65] within the framework of existing social, economic and political practice could yet render the debate academic.

However, in many ways, the real take home message is that many of us are not on the same page here. We use terms and discuss concepts freely with no standardization as to what they mean; we make assumptions in our self-contained logic based on axioms we do not share. Whether across different academic disciplines, wider fields of interest, or simply as individuals, we have to face up to some uncomfortable facts in a very immediate sense: that many of us are not discussing the same questions. On that basis it is hardly surprising that we seem to be coming to different conclusions. If, as appears to be the case, a majority of (say) neuroscientists think the TS cannot happen and a majority of computer scientists think it can then, assuming an equivalent distribution of intelligence and abilities over those disciplines, clearly they are not visioning the same event. *We need to talk.*

One final thought though, an admission really, as this paper has been written, probably in common with many others in a similar vein, by a middle-aged man having worked through a few decades of technological development: sometimes with appetite, but at other times with horror ... could this possibly be just a 'generational thing'? Just as Bertrand Russell [69] noted that philosophers themselves are products of, and therefore guided by, their position and period, futurologists are unlikely to be (at best) any better. This may yet all make more sense to generations to come. There is no better way to conclude a paper, considering uncertain futures and looking to extend thought boundaries through sci-fi, than to leave the final word to Douglas Adams [70].

*"I've come up with a set of rules that describe our reactions to technologies:*

1. *Anything that is in the world when you're born is normal and ordinary and is just a natural part of the way the world works.*
2. *Anything that's invented between when you're fifteen and thirty-five is new and exciting and revolutionary and you can probably get a career in it.*
3. *Anything invented after you're thirty-five is against the natural order of things."*

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Can Futurology 'Success' Be 'Measured'? Available online: <https://vicgrout.net/2017/10/11/can-futurology-success-be-measured/> (accessed on 12 October 2017).
2. Goodman, D.A. *Star Trek Federation*; Titan Books: London, UK, 2013; ISBN 978-1781169155.
3. Siegel, E. *Treknology: The Science of Star Trek from Tricorders to Warp Drive*; Voyageur Press: Minneapolis, MN, USA, 2017; ISBN 978-0760352632.
4. Here Are All the Technologies Star Trek Accurately Predicted. Available online: <https://qz.com/766831/star-trek-real-life-technology/> (accessed on 20 September 2017).
5. Things Star Trek Predicted Wrong. Available online: <https://scifi.stackexchange.com/questions/100545/things-star-trek-predicted-wrong> (accessed on 20 September 2017).
6. Mattern, F.; Floerkemeier, C. From the Internet of Computers to the Internet of Things. *Inf. Spektrum* **2010**, *33*, 107–121. [CrossRef]
7. How Come Star Trek: The Next Generation did not Predict the Internet? Is It Because It Will Become Obsolete in the 24th Century? Available online: <https://www.quora.com/How-come-Star-Trek-The-Next-Generation-did-not-predict-the-internet-Is-it-because-it-will-become-obsolete-in-the-24th-century> (accessed on 20 September 2017).
8. From Hoverboards to Self-Tying Shoes: Predictions that Back to the Future II Got Right. Available online: <http://www.telegraph.co.uk/technology/news/11699199/From-hoverboards-to-self-tying-shoes-6-predictions-that-Back-to-the-Future-II-got-right.html> (accessed on 20 September 2017).

9. The Top 5 Technologies That ‘Star Wars’ Failed to Predict, According to Engineers. Available online: <https://www.inc.com/chris-matyszczyk/the-top-5-technologies-that-star-wars-failed-to-predict-according-to-engineers.html> (accessed on 21 September 2017).
10. How Red Dwarf Predicted Wearable Tech. Available online: <https://www.wareable.com/features/rob-grant-interview-how-red-dwarf-predicted-wearable-tech> (accessed on 21 September 2017).
11. Star Trek: The Start of the IoT? Available online: <https://www.ibm.com/blogs/internet-of-things/star-trek/> (accessed on 21 September 2017).
12. Moorcock, M. *Dancers at the End of Time*; Gollancz: London, UK, 2003, ISBN 978-0575074767.
13. When Hari Kunzru Met Michael Moorcock. Available online: <https://www.theguardian.com/books/2011/feb/04/michael-moorcock-hari-kunzru> (accessed on 19 September 2017).
14. Forster, E.M. *The Machine Stops*; Penguin Classics: London, UK, 2011, ISBN 978-0141195988.
15. Leinster, M. *A Logic Named Joe*; Baen Books: Wake Forest, CA, USA, 2005, ISBN 978-0743499101.
16. Asimov, I. *The Naked Sun*; Panther/Granada: London, UK, 1973, ISBN 978-0586010167.
17. Twain, M. *From the ‘London Times’ of 1904*; Century: London, UK, 1898.
18. Adams, D. *The Hitch Hiker’s Guide to the Galaxy: A Trilogy in Five Parts*; Heinemann: London, UK, 1995, ISBN 978-0434003488.
19. Rifkin, G.; Harrar, G. *The Ultimate Entrepreneur: The Story of Ken Olsen and Digital Equipment Corporation*; Prima: Roseville, CA, USA, 1990, ISBN 978-1559580229.
20. Did Digital Founder Ken Olsen Say There Was ‘No Reason for Any Individual to Have a Computer in His Home’? Available online: <http://www.snopes.com/quotes/kenolsen.asp> (accessed on 22 September 2017).
21. The Problem with ‘Futurology’. Available online: <https://vicgrout.net/2013/09/20/the-problem-with-futurology/> (accessed on 26 October 2017).
22. Technological Singularity. Available online: [https://en.wikipedia.org/wiki/Technological\\_singularity](https://en.wikipedia.org/wiki/Technological_singularity) (accessed on 16 January 2018).
23. Cadwalladr, C. Are the robots about to rise? Google’s new director of engineering thinks so . . . . *The Guardian (UK)*, 22 February 2014.
24. Eden, A.H.; Moor, J.H. *Singularity Hypothesis: A Scientific and Philosophical Assessment*; Springer: New York, NY, USA, 2013, ISBN 978-3642325601.
25. Ulam, S. Tribute to John von Neumann. *Bull. Am. Math. Soc.* **1958**, *64*, 5.
26. Chalmers, D. The Singularity: A philosophical analysis. *J. Conscious. Stud.* **2010**, *17*, 7–65.
27. Vinge, V. The Coming Technological Singularity: How to survive the post-human era. In *Vision-21 Symposium: Interdisciplinary Science and Engineering in the Era of Cyberspace*; NASA Lewis Research Center and Ohio Aerospace Institute: Washington, DC, USA, 1993.
28. Armstrong, S. How We’re Predicting AI. In *Singularity Conference*; Springer: San Francisco, CA, USA, 2012.
29. Sparkes, M. Top scientists call for caution over artificial intelligence. *The Times (UK)*, 24 April 2015.
30. Hawking: AI could end human race. *BBC News*, 2 December 2014.
31. Kurzweil, R. *The Singularity is Near: When Humans Transcend Biology*; Duckworth: London, UK, 2006, ISBN 978-0715635612.
32. Everitt, T.; Goertzel, B.; Potapov, A. *Artificial General Intelligence*; Springer: New York, NY, USA, 2017, ISBN 978-3319637020.
33. Storr, A.; McWaters, J.F. *Off-Line Programming of Industrial Robots: IFIP Working Conference Proceedings*; Elsevier Science: Amsterdam, The Netherlands, 1987, ISBN 978-0444701374.
34. Aouad, G. *Computer Aided Design for Architecture, Engineering and Construction*; Routledge: Abingdon, UK, 2011, ISBN 978-0415495073.
35. Liu, B.; Grout, V.; Nikolaeva, A. Efficient Global Optimization of Actuator Based on a Surrogate Model. *IEEE Trans. Ind. Electron.* **2018**, *65*, 5712–5721. [[CrossRef](#)]
36. Liu, B.; Irvine, A.; Akinsolu, M.; Arabi, O.; Grout, V.; Ali, N. GUI Design Exploration Software for Microwave Antennas. *J. Comput. Des. Eng.* **2017**, *4*, 274–281. [[CrossRef](#)]
37. Wu, M.; Karkar, A.; Liu, B.; Yakovlev, A.; Gielen, G.; Grout, V. Network on Chip Optimization Based on Surrogate Model Assisted Evolutionary Algorithms. In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, Beijing, China, 6–8 July 2014; pp. 3266–3271.
38. Sioshansi, R.; Conejo, A.J. *Optimization in Engineering: Models and Algorithms*; Springer: New York, NY, USA, 2017, ISBN 978-3319567679.

39. Rhodius, A. *The Argonautica*; Reprinted; CreateSpace: Seattle, WA, USA, 2014, ISBN 978-1502885616.
40. Shelley, M. *Frankenstein: Or, the Modern Prometheus*; Reprinted; Wordsworth: Ware, UK, 1992, ISBN 978-1853260230.
41. Sawyer, R.J. *WWW: Wake*; The WWW Trilogy Part 1; Gollancz: London, UK, 2010, ISBN 978-0575094086.
42. Sawyer, R.J. *WWW: Watch*; The WWW Trilogy Part 2; Gollancz: London, UK, 2011, ISBN 978-0575095052.
43. Sawyer, R.J. *WWW: Wonder*; The WWW Trilogy Part 3; Gollancz: London, UK, 2012, ISBN 978-0575095090.
44. Grout, V. *Conscious*; Clear Futures Publishing: Wrexham, UK, 2017, ISBN 978-1520590127.
45. Turing, A.M. Computing Machinery and Intelligence. *Mind* **1950**, *49*, 433–460. [CrossRef]
46. How Singular Is the Singularity? Available online: <https://vicgrout.net/2015/02/01/how-singular-is-the-singularity/> (accessed on 16 February 2018).
47. Dick, P.K. *Do Androids Dream of Electric Sheep*; Reprinted; Gollancz: London, UK, 2007, ISBN 978-0575079939.
48. Poeltl, M. *A.I. Insurrection: The General's War*; CreateSpace: Seattle, WA, USA, 2018, ISBN 978-1981490585.
49. Computer AI Passes Turing Test in 'World First'. Available online: <http://www.bbc.co.uk/news/technology-27762088> (accessed on 19 February 2018).
50. Adaptability: The True Mark of Genius. Available online: [https://www.huffingtonpost.com/tomas-laurina-vicius/adaptability-the-true-mar\\_b\\_11543680.html](https://www.huffingtonpost.com/tomas-laurina-vicius/adaptability-the-true-mar_b_11543680.html) (accessed on 16 February 2018).
51. Our Computers Are Learning How to Code Themselves. Available online: <https://futurism.com/4-our-computers-are-learning-how-to-code-themselves/> (accessed on 19 February 2018).
52. Why Your Brain Isn't a Computer. Available online: <https://www.forbes.com/sites/alexknapp/2012/05/04/why-your-brain-isnt-a-computer/#3e238fdc13e1> (accessed on 19 February 2018).
53. Schneider, S.; Velmans, M. *The Blackwell Companion to Consciousness*, 2nd ed.; Wiley-Blackwell: Hoboken, NJ, USA, 2017, ISBN 978-0470674079.
54. August 29th: Skynet Becomes Self-Aware. Available online: <http://www.neatorama.com/neatogeek/2013/08/29/August-29th-Skynet-Becomes-Self-aware/> (accessed on 23 March 2018).
55. Time-to-Live (TTL). Available online: <http://searchnetworking.techtarget.com/definition/time-to-live> (accessed on 23 March 2018).
56. Seager, W. *The Routledge Handbook of Panpsychism*; Routledge Handbooks in Philosophy; Routledge: London, UK, 2018, ISBN 978-1138817135.
57. "The Theological Objection". Available online: <https://vicgrout.net/2016/03/06/the-theological-objection/> (accessed on 27 March 2018).
58. Gödel, K. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. *Mon. Math. Phys.* **1931**, *38*, 173–198. (In German)
59. Turing, A.M. On computable numbers, with an application to the Entscheidungsproblem. *Proc. Lond. Math. Soc.* **1937**, *42*, 230–265. (In German) [CrossRef]
60. Heisenberg, W. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Z. Phys.* **1927**, *43*, 172–198. (In German)
61. Popoff, A. *The Fermi Paradox*; CreateSpace: Seattle, WA, USA, 2015, ISBN 978-1514392768.
62. Webb, S. *If the Universe is Teeming with Aliens ... Where is Everybody? Fifty Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life*; Springer: New York, NY, USA, 2010, ISBN 978-1441930293.
63. Braga, A.; Logan, R.K. The Emperor of Strong AI Has No Clothes: Limits to Artificial Intelligence. *Information* **2017**, *8*, 156. [CrossRef]
64. Do Intelligent Civilizations Across the Galaxies Self Destruct? For Better and Worse, We're the Test Case. Available online: <http://www.manyworlds.space/index.php/2017/02/01/do-intelligent-civilizations-a-cross-the-galaxies-self-destruct-for-better-and-worse-were-the-test-case/> (accessed on 28 March 2018).
65. Technocapitalism. Available online: <https://vicgrout.net/2016/09/15/technocapitalism/> (accessed on 16 February 2018).
66. The Algorithm of Evolution. Available online: <https://vicgrout.net/2014/02/03/the-algorithm-of-evolution/> (accessed on 2 April 2018).
67. Why don't Animals Have Wheels? Available online: <http://sith.ipb.ac.rs/arhiva/BIBLIoteka/270ScienceBooks/Richard%20Dawkins%20Collection/Dawkins%20Articles/Why%20don%3Ft%20animals%20have%20wheels.pdf> (accessed on 2 April 2018).

68. Logan, R.K. Can Computers Become Conscious, an Essential Condition for the Singularity? *Information* **2017**, *8*, 161. [[CrossRef](#)]
69. Russell, B. *History of Western Philosophy*; reprinted; Routledge Classics: London, UK, 2004, ISBN 978-1447226260.
70. Adams, D. *The Salmon of Doubt: Hitchhiking the Galaxy One Last Time*; reprinted; Pan: London, UK, 2012, ISBN 978-0415325059.



© 2018 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).