A fourth law of robotics?

Citation for published version:
Schafer, B, Komuves, D, Niebla zatarain, J & Diver, L 2015, 'A fourth law of robotics? Copyright and the law and ethics of machine co-production', *Artificial Intelligence and Law*, vol. 23, no. 3, pp. 217-240. https://doi.org/10.1007/s10506-015-9169-7

Digital Object Identifier (DOI):
10.1007/s10506-015-9169-7

Published In:
Artificial Intelligence and Law

A fourth law of robotics? Copyright and the law and ethics of machine co-production

A fourth law of robotics? Copyright and the law and ethics of machine co-production.

Burkhard Schafer,[1] School of Law, SCRIPT Centre, The University of Edinburgh
David Komuves, School of Law, SCRIPT Centre, The University of Edinburgh
Jesus Manuel Niebla Zatarain, School of Law, SCRIPT Centre, The University of Edinburgh
Laurence Diver School of Law, SCRIPT Centre, The University of Edinburgh

Abstract:

Jon Bing was not only a pioneer in the field of artificial intelligence and law and the legal regulation of technology. He was also an accomplished author of fiction, with an oeuvre spanning from short stories and novels to theatre plays and even an opera. As reality catches up with the imagination of science fiction writers who have anticipated a world shared by humans and non-human intelligences of their creation, some of the copyright issues he has discussed in his academic capacity take on new resonance. How will we regulate copyright when robots are producers and consumers of art? This paper tries to give a sketch of the problem and hints at possible answers that are to a degree inspired by Bing's academic and creative writing.

Introduction

Jon Bing was a pioneer not only in the fields of artificial intelligence and law and the legal regulation of technology, but also as an accomplished author of fiction, with an oeuvre spanning from short stories and novels to theatre plays and even opera. In these works of "fabelprosa" (speculative fiction) he often anticipated legal and technological problems decades ahead of their time. In the short story *A meeting in Georgestown* (Bing 1981), for example, he explored the potential of software backdoors and Trojans as an instrument of politics, enabling states to spy not just on adversaries but also on their friends. The Snowden revelations would show three decades later just how prescient this account had been. Bing's academic interest in the legal regulation of technology and his desire as an artist to image possible futures were closely intertwined, and often overlapped. As an academic, he discussed robots and the legal issues of autonomous or semi-autonomous technologies (see for

[1] B.Schafer@ed.ac.uk work on this paper was supported by the RCUK funded CREATE

example Bing 2008; Bing and Sartor 2003). As a novelist, he wrote about robots and their interaction with humans, for example in the 1978 TV play "Stowaway" (*Blindpassasjer*) which he authored together with his long-time collaborator Tor Åge Bringsværd. In an interview in 2007, he says the story illustrated how, in order to function correctly, it is not enough for a robot simply to copy humans; it must adapt, transform and adjust to that which it perceives through its sensors (Bing 2007, p. 24). At this point, a number of potentially interesting legal issues arise. Academic legal discourse has so far focused on the consequences of delictual/tortious liability: if robots learn about the world in a way that is similar to how children do so, by acquiring raw information, then adapting, transforming and adjusting to it, we might then ask if – just as with children – robots eventually transcend their masters' legal liability for their actions (see for example Matthias 2004; Dante and Tamburrini 2006). By contrast, the discussion here focuses on the implications for copyright law, so far a largely overlooked issue (cf. Bing 2004). The input that a robot requires for learning will in some cases be protected by copyright – the rights to the information acquired by its sensors are owned by someone else. The process of adapting, transforming and adjusting the original input might in turn create works capable of copyright protection, either as unauthorised "derivative works" or, if the degree of transformation is sufficient, as wholly new works with a new (and yet to be determined) copyright owner (on derivative work and copyright see, for example, Reese 2004). Not all of these new forms of producing potentially copyrightable work require a robot in the strict sense of the word – an embodied AI with sensors and affectuators integrated by software. Intelligent software programs, some of them capable of running on standard computers, are becoming equally capable of creativity. Rather, robot technology merely adds new dimensions to the challenges to the legal system already caused by creative AIs, and a new sense of urgency. Robots, as we will see, can be deployed in contexts where complex legal rules interact – for instance when a care robot follows a visually impaired person through an art exhibition, reads out to her the description under the paintings, and as an incidental by-result makes a copy of the text. Here, new business models for robots can cause tension between copyright law and anti-discrimination law that policy makers may have to address. Similarly, robot technology may in the future enable radically new forms of copying: an industrial robot could copy the movements of experienced workers and eventually replace them – is “a way to move” something that needs in the future a form of copyright protection? If a local orchestra can only dream of being led by a famous conductor, would a robot that not only looks like him, but also copied his characteristic movements, be a replacement?[2] While some of these issues arise only for embodied AI, underneath most of the copyright challenges posed by robots is a theme they share with unembodied software: what it means for a computer program to be creative. We will therefore discuss in this paper embodied AI (robots strictu sensu) and unembodied AI together, using “robot” as a term for them both. We follow in this also Bing, who highlighted in his discussion of the literary history of the term “robot” the importance of Lovelace’s analytical machine and the early AI (Bing 2008 p.200). However, where legal issues are specific to embodied AI, we will point this out.

When we consider copyright as a question of robot law and robot liability, we enter a new field where the academic and artistic dualism in Bing’s work is readily apparent. As a novelist Bing was of course subject to copyright, and equally, the breadth of his output resulted in him acquiring a wide variety of copyrights himself. As an editor of numerous anthologies he would have been acutely aware of the legal minefields facing anyone involved with re-publication of older works. Finally, as an academic, he wrote extensively and authoritatively on issues of copyright in the digital environment (see, for example, Bing 1997; Bing 2003; Bing 2004; Bing 2009; Bing 2013).

If robots become our co-creators, imbued with genuinely creative capacity, we will be faced with a series of legal, ethical and technological challenges which will affect every step in the creative value chain. Novel legal and technological

---

[2] extrapolating on existing ideas, http://www.telegraph.co.uk/news/1954704/Robot-conducts-orchestra-for-first-time.html

solutions might be needed to shape this new form of co-production. For the purpose of this paper, we follow roughly a model of artistic creation that distinguishes between “Upstream” and “Downstream” problems. “Upstream” covers all those legal issues that arise in order to make a robot capable of producing a new work, such as right to access and learn from older, copyright protected works. It deals mainly, but not exclusively with the *input* that an AI requires. “Downstream” covers all those legal issues that arise from the *output* of robot creativity, especially the legal status of the works that are created by a machine. To ensure that new business model using robot technology are law compliant, both aspects need consideration. “Upstream” this will mean ensuring that robots use only material they have the right to use for their learning inputs. It may be that the scope of material permitted for such use will need to be either widened or narrowed through legislative intervention, in order to find an appropriate balance between the legitimate interests of creators, developers and the users of robots. Downstream, decisions need to be made regarding which (if any) intellectual property rights should be attached to works created with little or no human input, and indeed who the rightholder should be. As we shall see, some of the more promising business models for machine creativity arise in environments where humans will have limited time or resources to evaluate the legal implications of the robot’s actions. Ideally, therefore, robots should be designed to be compliant with copyright law “by design”, perhaps as a fourth law of robotics in addition to Asimov’s three well-known originals. This problem of copyright enforcement will have to be addressed regardless of what the substantive copyright demands. We will focus on the problem of copyright enforcement though code, as it is the most obvious link between the legal and the computational, the fiction in the form of Asimov’s laws and reality in the form of Digital Rights Management (hereinafter DRM). This does not mean that we think our current substantive copyright law is necessarily the final word on the legal protection of computer generated works, far from it. We will indicate some of the changes that may become necessary in our discussion below. In our view, the accommodating UK approach which explicitly provides for copyright in computer generated work is an important step in the right direction for the "downstream problem". At present, it is unique in the international arena, and reconciling it with the strong emphasis of civil law jurisdictions on moral rights may be a challenge. In any case it needs matching provisions "upstream", possibly as a new exception for machine learning, to allow the robotics industry to realise its full potential. We also consider it likely that new forms of licensing content will emerge, driven by industry, which do not so much change the legal regime but use it in new and creative ways. However the new copyright for robots will look like though, enforcement will be a major problem. The number of potential copyright infringers after all will rise exponentially, much faster than the increase of "biological" copyright users could ever accomplish. How to teach robots copyright from copy wrong will therefore become a major challenge, one that in our view can ultimately only be addressed by a form of "computational copyright law".

Bing too spotted the potential of "computational copyright law" to Digital Rights Management. He explored the potential usefulness of the approach in the context of e-governance and e-democracy. In such a setting, he envisaged the use of a DRM-like technology to enable governmental organization to improve their communication with citizens, releasing information through semi-autonomous or fully autonomous ambient intelligence (Abie et. al. 2004). For autonomous systems to handle digital objects in a copyright-compliant manner, they require to carry information about their copyright status in a machine-readable form – something Bing briefly discussed under the heading "copymarks" (Bing 2004).

With these preliminary remarks, we have prepared the ground for our paper. In thinking about the convergence between art, academia, science fiction and technology law that made Bing's work so unique, we explore how we might rethink DRM as "Asimov's fourth law" in a world of robot co-production. In the first two sections we consider robots and copyright "upstream" and "downstream", as well as robots as consumers and creators of art. In the third and final section, we discuss how more recent developments in law and AI could make Bing’s vision of “copymarks” a reality. In our discussion we draw on the philosophical, legal and aesthetic literature generated by depictions of robots in science fiction as much as in the real-world art and creative industries. The result is a somewhat impressionistic account covering emerging business models in the creative industries and the socio-legal and ethical challenges that those raise, as well as the role of technology in addressing them. Our understanding of “art” is intentionally broad, ranging from unique works of great artistic merits to routine press snippets or business reports. The reason is that as we will see, the discussion on copyright and AI has been hampered in the past by a focus on more speculative robotic “high art” by academic writers, while the more pressing and realistic developments took place in more mundane, economically valuable but often overlooked routine production of works.

1. Unmaking existence

In an opinion piece for WebProNews, the novelist and technology writer Jason Lee Miller stated:

> “Chess is one thing, but if we get to the point computers can best humans in the arts—those splendid, millennia-old expressions of the heart and soul of human existence—then why bother existing?”

Many of us feel an instinctual unease when confronted with the possibility of robot-generated art. In “Look to Windward”, the Scottish author Iain M Banks introduces a dialogue between the famous Chelgrian composer Ziller and the powerful AI (a “Mind’) that is in charge of the habitat where Ziller finds himself a refugee:

> 'If you tried, if any Mind tried, could you impersonate my style?' the Chelgrian asked. 'Could you write a piece – a symphony, say – that would appear, to the critical appraiser, to be by me, and which, when I heard it, I'd imagine being proud to have written?'

> The avatar frowned as it walked. It clasped its hands behind its back. It took a few more steps. 'Yes, I imagine that would be possible.'
>
> 'Would it be easy?'
>
> 'No. No more easy than any complicated task.'
>
> 'But you could do it much more quickly than I could?'
>
> 'I'd have to suppose so.'
>
> 'Hmm.' Ziller paused. The avatar turned to face him. Behind Ziller, the rocks gently beneath their feet. 'So what,' the Chelgrian asked, 'is the point of me or anybody else writing a symphony, or anything else?' […] What would be the point for those listening to it?'
>
> 'They'd know it was one of their own species, not a Mind, who created it.'
>
> 'Ignoring that, too; suppose they weren't told it was by an AI, or didn't care.'
>
> 'If they hadn't been told then the comparison isn't complete; information is being concealed. If they don't care, then they're unlike any group of humans I've ever encountered.'

As systematic studies have shown, Banks' intuition is widely shared. Not only are we still capable of identifying computer generated compositions for what they are, the emotional reaction against them from both artists and their audiences is strong (Moffat and Kelly 2006). Yet the question of robot creativity is by no means new. Indeed, the idea of using a combination of random and mechanical processes to generate new works of art precedes the modern computer by centuries. In the 18th and 19th century musical compositions developed with the aid of dices were popular, the earliest example of which is probably Johann Philipp Kirnberger's 1757 "*Der allezeit fertige Menuetten- und Polonaisencomponist*". The best known example was published in 1792 by Nikolaus Simrock (and at some point attributed to Mozart). It was a dice game capable of producing more than forty five trillion different waltzes without, so the publisher proudly proclaimed, “understanding anything about music or composition” (Nierhaus 2009, p. 36). In 1830, Gustav Gerlach published the “art of composing Scottish dances without being musical” (die “*Kunst, Schottische Taenze zu componiren, ohne musicalisch zu sein*” – Haupenthal 1994) thus demonstrating, if proof was needed, that even more demanding forms of art are capable of (near-) random generation. We discuss below a putative attempt to revive the idea of random generated art using modern computer tools.

When computers started to leave the confines of university research laboratories, artists very soon began to explore their potential for creativity. Nicolas Schöffer's CYSP 1 (Cybernetic Spatiodynamic Sculpture) showed in 1956 how the then-resurgent “kinetic art” could put modern machines at the heart of that endeavor (Schöffer 1963, p. 50; see also Kac 1997). Schöffer's interactive work comprised several sensors and analogue electronic components which interacted with observers to produce different kinds of movements. In 1964, Nam June Paik and Shuya Abe's Robot K-456 used robot-generated art to thematise issues of remote control and of freedom, while Edward Ihnatowicz's *Senster* was perhaps the first example in which the issue of robotic behavioral autonomy was brought to the fore. In that work, the robot was assigned one of several possible personalities, which then responded to changing situations autonomously (Reichhard 1978, p. 56). This type of co-production involving both a human and a (partly) unpredictable machine can create potentially interesting copyright issues, but these will be largely of an academic nature: we understand intuitively that Ihnatowicz was using the machine merely as a tool to implement a pre-existing artistic vision; his authorship is ultimately unquestioned.

Examples of computer generated poetry and prose emerged soon afterwards. In 1985, RACTER became the first book (a poetry anthology) written entirely by a computer (RACTER 1985; for a historical overview and examples, see Hartman 1996 and Funkhouser 2007). In poetry, probabilistic and evolutionary generation of word chains can show remarkable results (see for example Jiang and Zhou 2008; for a more recent discussion see Manurung, Ritchie, and Thompson 2012), although there are also more traditional rule-based systems which try to model the technique of writing poetry more faithfully (see for example Gervás 2001; Oliveira 2012).

Of all the arts, music was arguably the field that was changed most quickly, and profoundly, by computer creativity (for an overview, see Rowe 1993; 2001). Reviving the tradition described above of composition by dice using probabilistic algorithms, computer-generated music quickly became the primary domain for computer generated works. In a typical application of the time – and of a form that we will revisit below – Iannis Xenakis's "stochastic music" generated a multitude of possible works using the "Monte Carlo" method, after which the principles of composition were applied to select the best creations (Xenakis 2001). Although there were some early discussions of the implications of such works for copyright law, at the time it was considered that the artistic input of the operator of the programme was significant enough that authorship was not in question (Kostelanetz 1971, p. 243):

> “Since the computer, as an information-processing machine, can act only as an intermediary between the ideas of a composer and their realization, there is no art in the technology itself. For this reason, any work of art employing a computer must inevitably be machine art, rather than an artistic machine, to draw a crucial distinction.” (Kostelanetz 1971, p. 245)

On this view, computers remain from a copyright perspective merely the artist's tool – they are not independent creators. Kostelanetz concludes with an interesting quote from Charles Csuri (Kostelanetz, p. 229):

> "I think that in time the artist will sit down and think about a picture and then a computer will translate his brain impulses into that picture.”

This may seem like a radical proposition, and indeed it does go far beyond what is currently possible. This is despite progress in brain imaging which makes it possible to “read the mind” (see for example Solso 2000), both when appreciating art (see for example Di Dio et. al. 2007; Nadal 2008) and when creating it (see for example Schlegel et. al. 2015). Indeed, it is already possible for a computer linked to an fMRI reader to identify what type of object a person is seeing (Shinkareva et. al. 2008; Kamitani and Tong 2005), what she is visualising in her imagination (Pogue 2012) or even what she is dreaming about (Horikawa et. al. 2013). Possible applications of the science do include the possibility of directing a robot by thought alone, for example in assisting paraplegic patients. Csuri's vision is almost at the point of being realised – computers can now give us a crude description of what a person is (probably) thinking about (Wolpaw et. al. 2002) and allow them to manipulate a robotic arm merely by thinking about the movement (Lee et. al. 2009).

Even though Csuri’s vision was radical and bold, in one crucial respect it was constrained by a traditional conception of artistic endeavor. At its heart is the subjective vision of the human artist – the computer is a mere tool used to translate the internal creative concept or idea into an external embodiment. This leaves the central conceptual vocabulary of copyright law unaffected: we still have the dichotomy between idea and expression, and the physical threshold between the two.

Bing provided us with a slightly different vision of a merger between mind and machine when he proposed applications in the working environment, where the machine would impart information directly to the human brain. In his vision of the workplace of the future, images and information are directly beamed into the human eye (Bing 1998) or, as he suggested in a talk given at IRIS, the human brain. Here too, the issue of robot creativity does not emerge. Yet Bing’s model of computer generated “art perception” raises potentially interesting copyright issues: in this scenario the computer could give a viewer the impression that they are viewing, for example, a famous painting. Yet nowhere in the computer would we find a copy of that painting, not even in a digital form – rather, we would find a complex set of instructions on how to manipulate a set of laser beams, or a set of electrodes, so that our mind creates the *impression* of the painting. Did "copying" of that painting take place, even if the only "copy" is directly assembled in the human brain, and the only traces on the computer are codifed instructions on how to manipulate an external device?

We have seen so far how the idea of computer generated art has a long and venerable tradition. Until recently computers have been treated merely as tools in the creative process: they are the handmaiden of the artist's creativity, but are without their own creative faculties. While this approach has remained dominant in the arts, since the late 1980s academic consideration has been given to the question of "true" computer creativity, particularly within science-based AI research. This discussion had two separate yet interrelated aspects: (i) the semantic question of what it would mean for a machine to be called "creative", and (ii) the technical question of whether machines could be creative in the sense thus defined. The quest for creative robots fulfilled two purposes: it helped in the formulation and refinement of new conceptual questions about the nature of art and (human) creativity, and it led to more convincing examples of machine creativity which challenged the notion of the machine as merely a passive tool intended to help the human artist realise her creative vision. On the first question, the work of Margaret Boden started – and for a long time dominated – a discussion which raised deep conceptual issues around the connection between randomness and creativity (see for example Boden 1994; Boden 1998; Boden 2009; for a more recent overview of the debate see McCormack and d'Inverno 2012). It also enabled the development of rigorous methods for comparing the creative ability of different computer systems as well as classification schemes for them, from the most mundane (such as automatic spell checking) to the most sophisticated (such as a robot doing jazz improvisations – see for example Wiggins 2009; Jordanous 2012). Intuitively we feel that there is a point somewhere on that spectrum where the input of the robot is significant enough that it can be deemed a co-producer of the work, if not the main creator. As we have seen above, the idea of genuinely creative AI is often met with mixed responses, because of the important role of creativity in defining what it is to be human. Indeed, computer scientist Mark Riedl recently suggested that the iconic Turing test be amended to match, or even replaced with, the "Lovelace test", which requires a robot to produce art that cannot be distinguished from human-created art (Riedl 2014). This could give rise to copyright issues, and indeed there have been some highly sophisticated studies of the copyright implications of computer generated art (see for example Bridy 2012; Lee 2012; for an early discussion see Butler 1981).

These legal discussions of copyright implications have, however, been highly abstract and so have remained at the academic fringe. In the absence of court decisions or legislative intervention, they were based mostly on "what if" scenarios (for a recent exception to this, see McCutcheon 2012). We can see some of the reason for this in the above discussion: artistic practice made sure that actual computer input remained limited and was explicitly subordinated to the artist's vision or project. By contrast, academic discussions and projects in the field of computer science focused on the peaks of human creativity, as the robo-Rembrandt or Alighieri constituted a difficult and therefore interesting challenge. While considerable progress was made in machine creativity (see for example Gervas 2000; Legaspi 2007), these approaches typically lacked a specific business model or commercial implementation and thus remained of interest mainly to researchers and small groups of connoisseurs. They did not raise copyright challenges because it was in nobody's interest to do so and no economic interests appeared to be at stake.

In what follows, we argue that this situation is about to change. We are moving into an era where man-machine co-production of creative works will become commercially viable and commonplace. The issue is not "high art", but rather more mundane forms of creativity, for example writing short news stories or short jingles for computer games. Crucially for copyright, some of these will be not just produced but also disseminated automatically, without human intervention or oversight. This removes the last important connection to human artistic creation from the equation – the judgment that something *is* art or that a specific work *is* a successful embodiment of the artist's vision. As early as the 1970s, Kostelanetz had emphasized that while random generators give machines the ability to produce a large number of potential works, only the human judgment that identifies some of these as acceptable can complete the process of artistic creation. Those copyright systems that can cope with Duchamp's *Found-Objects* can also cope with this form of man-machine co-production (on the question of copyright for Found Objects, see for example Ward 1992). They all struggle, however, if even this minimal element of "creation by designation" is lost.

In the next section, we will briefly return to the world of literary fiction to discuss some recent developments which indicate that, as we share our world more and more with autonomous machines, these scenarios are turning from fiction into reality, in the process raising some profound ethical and legal-regulatory issues.

1.2. Downstream: borrowing books in the Library of Babel

In September of 2014, a number of technology blogs reported a potential "copyright apocalypse".[3] A Russian company, Qentis, claimed to have found a way to use computer technology to create not just *some* new and original works – itself a considerable challenge – but *every* possible text, of a given

---

3 See for example https://torrentfreak.com/copyright-apocalypse-trolls-attack-the-net-from-the-future-140928/, last accessed 21 January 2015.

length, that it is possible to write (and in a range of languages).[4] By using statistical and evolutionary algorithms which combine the smallest building blocks of language – individual letters – they claim to have

> "generated and deployed 97.42% of all possible useful texts of ten to 400 words in length (the remaining 2.58% has already been deployed in the last 2000 years) [...] Qentis aims to create 99.2% percent of all target-length Internet text, making it by far the largest copyright holder in the world."

The business model behind this idea is simple: becoming the world's largest copyright troll. Or in their own words:

> As Qentis approaches 100 percent of content generation, all content owners will eventually have to pay royalties to our clients or face massive lawsuits."[5]

Superficially, what Qentis claims to be doing is similar to the dice-generated music that was so popular as a parlour game in the 18th and 19th century. Where it differs is in leveraging the massive computing power available in the 21st century. While a human throwing dice for her entire life, or even the life of this planet, would not be enough time even to dent the number of possible minuets that can be created in this way, thus leaving almost all possible minuets available for others to discover and compose, the claim here is that massively parallelised computing can make not just a dent in the total number of possible works, but can in fact generate all of them. A "brute force" approach that generates all possible combinations of words in English will, eventually, produce every meaningful text there could be.

The good news for artists everywhere is that Qentis is itself "just" a work of art, created by the Vienna-based performance artist Michael Marcovici.[6] As

---

4 http://www.qentis.com, last accessed 21 January 2015.

5 http://www.qentis.com/work/work-13/, last accessed 21 January 2015.

with most of his work, it explores concepts of creativity and copyright through the lenses of technology. Despite its playfulness, the project asks some important questions about the ethics of copyright, computer generated works and business models in the creative economy. It challenges our conceptions of creativity and the economic value that we attach to creative work, and exposes concerns about the logical limits of our legal vocabulary when attempting to conceptualize and resolve the legal tensions that arise from disruptive technologies.

For several reasons, both technological and legal, the Qentis business model would fail (for a more detailed discussion see Komuves et al 2015). For example, copyright does not prevent independent creation of the same work by two people; only direct copying is prohibited. So Qentis would have to show that the alleged infringer had found the work in question on their database and had copied it from there. Now comes a twist: using the very same mathematics that allows Qentis to prove, rigorously, that the piece in question *must* be somewhere in their database, the defendant can also prove that he could not have possible found it there – successfully finding it in such a database, or indeed *any* meaningful portion of text, would require many times the life span of our universe. As Jon Bing, in a very different context, said:

> "To ask why we need libraries at all, when there is so much information available elsewhere, is about as sensible as asking if roadmaps are necessary now that there are so very many roads."

The Qentis database is not a library and has no librarians. To navigate it is to drive on unknown roads, with no roadmap. It contains not just all meaningful texts of up to 400 words; it contains every possible combination of words up to 400, the overwhelming majority of which will be meaningless. "Hds dwh ckfk"

---

[6] http://www.artmarcovici.com/BIOGRAPHY, last accessed 21 January 2015.

is in the library, as is “ssss ssssss sssss”.[7] Qentis is, of course, a technologically-enhanced version of Borges “Library of Babel”, which in turn is based on the famous “infinite Monkey theorem” devised by the French mathematician Émile Borel (Borel 1913): a thousand monkeys, hitting keys at random on a typewriter keyboard for an infinite amount of time, will at some point “almost surely” produce the complete works of William Shakespeare.

This idea led Borges to imagine a “Total Library” (Borges 2007, p. 219):

> “Everything would be in its blind volumes. Everything: the detailed history of the future, Aeschylus' The Egyptians, the exact number of times that the waters of the Ganges have reflected the flight of a falcon, the secret and true nature of Rome, my dreams and half-dreams at dawn on August 14, 1934, the proof of Pierre Fermat’s theorem, […], […] Everything: but for every sensible line or accurate fact there would be millions of meaningless cacophonies, verbal farragoes, and babblings. Everything: but all the generations of mankind could pass before the dizzying shelves — shelves that obliterate the day and on which chaos lies — ever reward them with a tolerable page.”

Borges returns to this idea in his short story “The Library of Babel”. Here too the majority of the books are just meaningless strings of letters. They are not ordered in any logical fashion and are assigned to rooms seemingly at random. Thus, even if the library *necessarily* contains all useful information in existence, including predictions of the future, the sheer amount of

---

7 A rare attempt, by a group of performance artists, to implement the infinite monkey theorem in reality used six Sulawesi macaque monkeys The monkeys produced five pages consisting mainly of the letter “s” before they started destroying the typewriters with a stone and using it as toilet https://web.archive.org/web/20130120215600/http://www.vivaria.net/experiments/notes/publication/NOTES_EN.pdf. In a paper dedicated to the memory of Jon Bing, one should point out that elephants would have been the superior choice, as they outperform monkeys when it comes to create art: http://www.elephantartgallery.com

unstructured information means that it is also and *necessarily* useless for readers, which leaves the librarians in a state of despair.

Bing, as a pioneer of intelligent legal knowledge retrieval who had an acute awareness of the difficulties created by ever-increasing amounts of information (Bing 1984; Bing 1991), would have appreciated the irony of this situation. As long as robots use nothing but brute force to create artistic works, they are no danger to the human artist. Not necessarily because they fail to produce works, but rather because they produce so many that human judgement and ingenuity remains indispensable: judgement is a crucial elements of artistic creation, as mentioned above.

A more complicated question is whether Qentis would have any copyright in the works that its software generates in the first place. In the paradigmatic case of copyright-protected work, a human author takes an abstract idea or thought (unprotected by copyright) and transforms it into a concrete, expressed form. On the one hand the transition from an idea to an embodiment is closely linked with the concept of creativity, and on the other it requires the type of effort and investment that copyright law also aims to reward. These key components are not present in the case of Qentis. It lacks a human author, and the way in which it generates work means that the concept of an idea is absent. Instead it deals only in expressions. But expressions of what, since there is no pre-existing idea? This kind of objection against the very concept of artificial intelligence was raised in particularly poignant form by the philosopher John Searle in his famous “Chinese Room” thought experiment. The experiment suggests that machines can only ever achieve the likeness of knowledge. It posits that an artificial intelligence has knowledge in only the same way as someone simulating the knowledge of Chinese by replacing one set of (what are for her) unintelligible scribbles with another according to a predefined set of rules does, and that this is the best it can ever be expected to achieve. Computers lack intentionality and thus they do not truly communicate; they merely “ape” communication (Searle 1980).

Long before Searle, and directly referencing the infinite monkey theorem, the historian and philosopher Robin Collingwood drew inferences from the theory of literature. For him, anything generated by a random process, even if it looks exactly like a real text, should not be considered an artistic work: there is an ontological difference between "the works of Shakespeare" – which for Collingwood are an abstract, imaginary object closely tied to the idea of emotional expression – and a particular physical combination of letters. He attacks, in scathing terms, those who think random processes can generate works of art (Collingwood 1958, p. 126, footnote 1):

> "But the interest of the suggestion lies in the revelation of the mental state of the person who can identify the 'work' of Shakespeare with the series of letters printed on the pages of a book bearing that phrase as its title: and thinks, if he can be said to think at all, that an archaeologist of 10,000 years hence, recovering a compete text of Shakespeare from the sands of Egypt but unable to read a single word of English would possess Shakespeare's dramatic and poetic works."

Note that Collingwood does not reject the idea that monkeys could, physically, generate a text that is letter by letter identical to Shakespeare's; his claim is instead that this would nonetheless fail to qualify as "Shakespeare's works" because of the lack of intentionality involved in its creation. The music pieces created by the throwing of dice fail to be proper compositions for the same reasons. So far, most legal systems adopt this line of argument. There cannot be a copyright work without a (human) author, and no protected expression without a concomitant idea which it embodies. Consequently, copyright protection for computer-generated works has generally been rejected.

The US Copyright Office guidance states that:[8] "[a]s discussed in Section 306, the Copyright Act protects "original works of authorship, 17 U.S.C. § 102(a)". To qualify as a work of "authorship" a work must be created by a human being. It states further that "[t]he Office will not register works produced by nature, animals, or plants" and gives as examples "a photograph taken by a monkey" or "a mural painted by an elephant". Such discrimination against elephants would undoubtedly have aggrieved Bing. The guidance continues by saying the Office "will not register works produced by a machine or a mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author".

The situation is different in the UK. It has one of the few legal systems in which computer-generated art is explicitly protected (for a comparative discussion, see McCutcheon 2013). Section 9(3) of the Copyright, Designs and Patents Act 1988 (c. 48) specifies that "[i]n the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken." Thus, despite the other obstacles discussed above, Qentis would be the copyright holder under UK law.

Qentis' most realistic and believable characteristic is its restriction to texts of 400 or fewer words – indeed, it is this length of text that is most likely to become generable by computers in the near future. Short, technical articles and notes for online publication, data-driven journalism and summarisation services are most likely to avail themselves of this technology. A typical application could for example harvest customer reviews about holidays in a particular city and rewrite the information into a Wikipedia entry about that city, or take business data and statistics and turn these into a report for

---

[8] http://copyright.gov/comp3/chap300/ch300-copyrightable-authorship.pdf, last accessed 21 January 2015.

shareholders. Services such as *Narrative Science*[9] or *Automated Insight*[10] focus on this segment of the market (for a scientific discussion see for example Lee et al 2012). A different, but even more impressive, approach was taken by Philip M. Parker, professor of management science at *Insead*. For his on demand publishing company, he "authored" over 200,000 books – all written by algorithms which harvested publicly-available information and reshaped it into texts that would typically only be of interest to a very small number of specialists, but who would face significant costs if they had to do the necessary research themselves.[11]

Technologies like these threaten established business models in the creative economy and will devalue certain forms of human creativity.[12] They also disrupt the legal-regulatory machinery. As we have seen, core concepts of copyright law, such as the "idea-expression dichotomy", fail adequately to express the issues that are at stake.

Unlike Qentis, these systems will not simply generate random texts, but will instead learn from and incorporate text written by others. Hugh Petrie argues that the evolution of written ideas can learn from biological evolution in accounting for the role of historical context. He argues that if the aim is to produce "high quality" extended works such as Shakespeare's plays, we must

---

9 http://www.slate.com/articles/technology/future_tense/2012/03/narrative_science_robot_journalists_customized_news_and_the_danger_to_civil_discourse_.single.html, last accessed 21 January 2015.

10 http://towcenter.org/blog/automated-stories-using-algorithms-to-craft-news-content/, last accessed 21 January 2015.

11 http://www.nytimes.com/2008/04/14/business/media/14link.html

12 http://www.cbsnews.com/news/this-post-was-written-by-a-human/, last accessed 21 January 2015.

equip the monkey/computer with not just a typewriter, but what today we would call an expert system that incorporates "whole Elizabethan sentences and thoughts". Furthermore, "It would have to include Elizabethan beliefs about human action patterns and the causes, Elizabethan morality and science, and linguistic patterns for expressing these" (Petrie 1981, p. 132). This approach comes much closer to what many working text generation systems attempt. Rather than using brute force and random generation, they learn, just like human creators do, from what has been done before. In doing so, however, they raise their own copyright issues: the machine is now not only a creator, but also a consumer, of art. As society begins to take seriously the idea of a world shared with intelligent and autonomous machines, the issues that this poses will require consideration. We now give a brief sketch of what the business case for such uses of robotics might look like, and the legal issues that are raised, taking again our lead from science fiction.

1.2 Upstream: Johnny Five needs input

One of the most endearing depictions of a robot in popular fiction is surely Johnny Five, star of the *Short Circuit* films. Although programmed with the ability to learn, he initially lacks the right type of knowledge base to interact properly with his environment and human beings. This changes when the romantic lead of the film, Stephanie, gives him access to her books, including an encyclopaedia. Speed-reading through the pages, Johnny Five not only acquires their knowledge, but also a personality that is shaped by the type of literature he has been given. But what would the copyright situation have been, had this occurred in real life? Did his "reading" of the books not also involve making a digital copy, stored in his memory? In the remainder of the film it becomes obvious that he has not copied (stored) the information, but also "understood" it in the sense that he is capable of applying the information to new situations. At the same time, he continues to give verbatim quotes from the encyclopaedia, indicating that he has kept a copy in his memory storage which he did not delete once the learning process was completed – very much a human behavioural trait.

We may wonder why Jonny Five uses such an antiquated mode of information acquisition – surely, rather than reading the book visually, downloading the information would have been much more efficient? While this may be true in the real world, it matters in the context of science fiction. The argument has been made, quite convincingly, that depicting (anthropomorphic) robots engaged in reading "old fashioned" paper books is an important visual cue to their "human-likeness". This is a recurrent trope found in Andrew, the android in Isaac Asimov's *Bicentennial Man*, and David, the robot in Kubrick's *AI: Artifical Intelligence* which acquires its humanity when its "mother" reads Pinocchio with it (see Hong and Rouget 2014). We showed above how many instinctively recoil at the idea of a creative robot, seeing it as a threat to the "uniqueness" of man. The robot as consumer of art is the flip side of this coin; social acceptance of robots is increased by showing them engaged in one of the most human activities possible. This again demonstrates how central art and its appreciation is for our self-understanding.

Current robots are nowhere near to matching Johnny 5's capabilities. Nonetheless, recent years have seen a considerable push towards commercially viable robotic systems. We discussed briefly above some paradigmatic examples that have recently been brought to market, such as Narrative Science. They can identify relevant inputs, for example market reports, data released by companies, and other less tangible forms of market intelligence such as blog posts or tweets, and turn them into short texts. In theory, at least, this process could be fully automated: the algorithm identifies the correct input(s), transforms them into a text and then releases it to subscribers. The copyright questions are intricate – is the software making potentially illegal copies? And is the output it creates a "work" for the purpose of copyright law? If so, is it a derivative work (which gives the original creator of the system's input a property right) or, because of the degree of permutation and alteration carried out by the machine, is it a work in its own right? If the latter, who is the owner – the company the developed it, or whoever operated the algorithm under a license from them? It might even be the first human reader of the text, if we subscribe to the artistic judgement theory discussed above.

There are, however, other, less obvious copyright issues which a world shared with intelligent machines poses. In particular, care robots which are intended to supporting the aging population and to alleviate strain on health services (see, for example, Sharkey and Sharkey 2012), share some design features with Johnny Five, including their anthropomorphic design. They also share with him

- sensors which take images of their environment,
- a capacity to learn, including in some cases learn from text,
- internal data storage capacity which may or may not have to keep records of the data inputted when learning, and
- a degree of autonomy.

With these capacities, they also create intellectual property issues. Much like Johnny Five they may need to learn from a pre-defined data set, supervised by their creators, or unsupervised according to data they gather directly from their environment. A particular ambitious model called cognitive developmental robotics aims to mimic childhood learning (see, for example, Asada et al 2001; Cangelosi et al 2010). Can – and should – we prevent such robots from learning from copyrighted material for which their developers do not have a license, or should we expand the “education exemption”, which most copyright regimes incorporate, to machine learning? Policy arguments of the type proposed by Chong (2007) for developing countries could be one way to frame this discussion.

Consider a care robot which follows its visually impaired owner to an art gallery. In order to help her, it takes images of the walls to identify the exhibits (and also any obstacles in her path, such as sculptures), and directs her to the paintings she is most interested in. This will typically involve taking images of the paintings for the purpose of pattern recognition, but some of the paintings will be protected by copyright. What if, in addition to storing a copy, the robots also sends the images to other, similar robots in the gallery, in order to increase their collective efficiency through data sharing?

Or consider a robot which accesses the Internet both to "listen" to the news so that it can provide its elderly owner with a summary, and to control the heating based on the weather forecast. Should it matter if the robot is owned by her, acting as her extended eyes and ears only, or if it is owned by a health service providing (possibly paid-for) medical care, so that it returns eventually (along with the data gathered) to a pool of machines under the control of that service?

What these applications have in common is that although copyrighted material is copied and stored it is done so "incidentally", that is not as an end in itself or for dissemination, but for a qualitatively different purpose, for example to facilitate the robot's efficient navigation or more effective service provision. Some of these applications would be perfectly permissible if they were carried out by the robot's human owner, for example looking at a sculpture and, as a result, keeping an image in one's memory. Other processes, such as "reading a book to a child",[13] including "reading the text and skipping the cruel parts", involve more traditional forms of copying and, crucially, also the creation of new works. Even a "dumb" software like the one in the Kindl2 e-reader

---

[13] One of the functions of Pepper, the childminder-robot http://spectrum.ieee.org/automaton/robotics/home-robots/pepper-aldebaran-softbank-personal-robot The Kindle2 has caused similar copyright questions: http://www.fastcompany.com/1160843/authors-guild-says-kindle-2s-text-speech-violates-copyright

already caused concern with rightholders who want to treat this type of transformation as equivalent to a voice performance for an audiobook. This analogy is already stretched – no lasting "work" is created, and the performance is not public. AI adds to this dimension by the ability to adjust and change the text in response to the changing needs of the listener. The connection to the original work, the printed text, becomes more indirect, the result more similar to a derivative work, while at the same time raising the problem of moral rights of the original author who may object against a bowdlerised rendition of her work. Here too we can ask if these acts should not be governed by similar legal norms when undertaken by a humans and robot. If one can ask a human nanny to read a book to a child without legal implication, why should matters be different if that nanny is a robot? That on the one hand a copy of the book is stored (imperfectly) in neural tissue, and on the other (perfectly) in silicon, seems normatively to be irrelevant. The differences between the two are becoming even more blurred as robots begin to use biological material for their processing (for example in DNA computers) and research is carried out into enhancing human memory by implanting chips in the brain. In some applications being considered, the human owner will rely on such a device to mitigate the effects of, for example, dementia. This could raise issues in the interaction between disability discrimination and copyright law. The latter often permits what would otherwise be prohibited copying and transformation of works, or the circumvention of DRM technology to prevent these, if the aim is to make texts accessible for the visually impaired. A robot that stores a text to read it out to its blind owner would be covered by this provision – but the very specific, ad hoc and unsystematic nature of this exception means that the same robot performing the same task for its mobility-impaired owner, who can no longer hold as opposed to see a book, might fall foul of the law (see for example Roos 2005). The basis for a more rational and uniform treatment of assistive robotics and copyright law might come from an unusual source. For a very different purpose, a philosophical theory of consciousness, Chalmers and Clark propose the concept of the "extended mind", using Leonard from the film *Memento* as an illustration. Leonard uses external manifestations – photos – of what he has previously seen in order to enable him, as a sufferer of anterograde amnesia, to think and reason in a way similar to someone whose memory is not affected. “The mind” is not just inside our skull; it is “socially extended” into our environment. On this basis they propose the “principle of parity”:

> “If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is (so we claim) part of the cognitive process.”

We suggest that by giving this principle a normative reading (not one intended by its authors), it becomes easier to construct a principled exception to copyright law that enables assistive technologies by permitting robots to access all the information their owners would be able to access lawfully but for their disability. It might then be suggested that the information that is inevitably copied in this process should therefore be subject to similar limitations as is data stored in (healthy) human memory, and so build in a kind of “forgetfulness by design” as a legally-mandated feature of the robot’s software. In some situations a “private copy” exception, such as the one recently introduced into UK law, will result in a similar outcome. The owner of a book could permit her robot to make a copy of the book to read it out to her – this would be an example of (now-permitted) format shifting. But this construction will not work in all of the scenarios discussed above. For example, if a disabled person is led into an art gallery by her care robot, she she does not normally “acquire a right” to take photos of the exhibits. If, as in the example above, her robot makes a copy in order to allow it to direct her to the exhibits which interest her, this would not constitute the making of a private copy of a copyrighted work that she already owns, and therefore the exception would not apply.

Finally, copyright law also permits the making of a copy if this is necessary to perform some technological tasks, the typical example being the temporary copy of a website stored in a browser's cache. If a robot scours the internet for information about the weather to assist its owner in planning her day, then the texts from which it extracts the information could be said to be stored on the robot as merely transient or incidental copies which exist only for the purpose of transmitting a work across a network between third parties, or alternatively exist only to facilitate a lawful use of the work.

4. A fourth law of robotics – where AI and law meet again

The discussion so far gives only an impressionistic account of the sorts of copyright issues that society will have to address as we share more and more of our lives with intelligent and autonomous machines. Some of the questions are ethical or political in nature – do we *want*, given the centrality of creativity to our self-understanding as humans, to acknowledge robots as creators and thus enable new business models which promise faster and cheaper access to certain types of works, such as news articles? Economically and culturally, could we as a society afford such a development, given that budding artists have often relied in the early stages of their careers on these types of "low level" creative output, from writing short news articles to drawing postcards for tourists? Whatever the answers to these questions will be, enforcing them will be as much of an issue as it is with current copyright law. In an ideal world, robots will only access data which they are entitled to access, make only legally permissible use of it, and protect the economic interests of their owners when releasing work created by them to third parties. How can we ensure, as a design feature, that they will be able to fulfil this role as law-abiding consumers and producers of creative works?

While the capacity of current programming technology to implement practical versions of Asimov's three laws has, with good reason, been called into question (see for example Weld and Etzioni 1994; Pynadath and Tambe 2002), the idea of computational copyright law has a strong pedigree. Lawrence Lessig famously used Digital Rights Management as the paradigmatic example for regulation by code (Lessig 1999) as a new form of regulating online behaviour. The use of DRM to influence *human* behaviour has at best yielded mixed results, being on the one hand incapable of meeting the legitimate expectations of buyers of digital goods with regard to their enjoyment of their purchase (see e.g. Mulligan et al 2003; Favale 2011), and on the other incapable of preventing any truly dedicated human from outsmarting and circumventing software-based access restrictions (see for example Cox and J. Linnartz 1998; Stamp 2003). However, the situation is different at least on this second count if the primary target of DRM is not a human, but a machine. Just as “agent exclusion clauses” have turned out to be an efficient way for website owners to prevent autonomous agents from accessing and indexing their site, DRM restrictions are likely to be highly efficient if they need only to prevent a robot accessing copyrighted material.

In traditional DRM solutions, the creator embeds restrictions into the digital objects she creates. But once we start to think of robots as creators, another meaning of DRM emerges. Here, it could refer to the set of rules that are built into the automated creator, enabling the machine to identify input that it can legitimately access, perform only those transformative operations that the license permits, and eventually release a new work under the appropriate license, all without direct supervision by a human and without exposing its owner to the risk of litigation. In this scenario, DRM changes from a nuisance, a “restrictions management” tool that results in goods that are “defective by design”, into an essential tool for creators which liberates the owner or operator of robot technology and facilitates new business models built on robot co-production.

Bing had anticipated such an “enabling” use of DRM and discussed it in the context of a government agency optimising its communication with its citizens (Abie et al 2004). He also made suggestions on how we can signal the copyright status of a digital object to the world, using “copymarks” based on HTML tags (Bing 2004) to facilitate automated identification of the rights status of objects (Bing 1996). This is important and particularly relevant in the context of robotics, where the owner of the robot will regularly *want* law-compliant behaviour from her machine, but cannot personally control all of its interactions with the Internet. However, in order to work in a fully automated environment where robots are consumers and creators of works, this requires not just that digital objects carry machine-readable information about their copyright status, but also representations of the legal rules that tell them what can and cannot be done lawfully with that object. For this, DRM needs to become smart, by incorporating concepts and ideas from AI and law. In what follows, we will briefly discuss existing research that could be leveraged for such an endeavour, indicate what we perceive the main shortcomings to be, and make an assessment of the viability of such a project in the light of experience gained with AI supported privacy-by-design approaches.

One of the first prominent contributions at the intersection of copyright and AI was the formal representation of an upper level ontology for copyright that constituted one part of a larger investigation into formal models of intellectual property law done by Contissa and Laukyte in 2008. Generally, the emergence of formal ontologies as a building block of the semantic web has given a major boost to the research into formal representations of copyright law, and our own study follows this trajectory. Creative Commons, for example, has made licenses available in RDF ("Resource Description Framework") format, which allows web publishers to embed license information in machine-readable format in web pages, documents and MP3 files. RDF, first published in 1999 and substantially revised in 2014, is a specification for metadata models issued by the World Wide Web Consortium (W3C). It can be used to describe conceptual information about resources which are implemented on the web, and in this way can assist not just information retrieval but also knowledge management and reasoning applications. At its core philosophy is the idea of making statements about web resources in the form of subject-predicate-object expressions, a form of the copymarks anticipated by Bing. This can be used for example to identify the author of a resource or its legal status ("This eBook file has the property of being licensed through a CC-BY license").

An important milestone in developing this approach further and reducing the gap between semantic web and knowledge management applications was a paper by Gordon (2011) which reported results from the MARKOS project. Although only a short paper, it has in the relatively short period since its publication garnered considerable interest and prompted several follow-up studies. It uses the Carneades Argumentation framework to reason about the way in which different copyright licenses interact in complex collaborative projects. Conflicts between Creative Commons licenses are a well-known problem due the "viral" nature of the Share Alike provision, and the handling of the resulting complexities is a natural target for legal-AI analysis. If, as a program developer, I use software from different sources and combine them to form a new product, it can be difficult to determine under what license I am in turn allowed to release the resulting program, and also whether I am in fact permitted to use all the pieces of software I used to build my application. While MARKOS focuses on copyright, the model seems easily transferable to one of the scenarios described above: a robot that collects information from the Internet to rewrite it into a business report or a short news item could easily use a similar system to identify those texts that have an appropriately permissive license, then release its creation under an appropriate license which observes any limitations that might be "inherited" from the source material, all the while optimising its owner's commercial arrests. As long as the license remains sufficiently simple and structured, with sufficiently precise terms, even current systems have shown an ability to handle them correctly. As more and more automated systems become potential users of licensed material, market pressure will push further to simplify and standardise license texts – (legal) reality adjusting in part to the limitations of computer technology.

Unfortunately, our discussion also shows that, while this is an important stepping stone towards AI-facilitated regulatory compliance, it alone will not be enough. The copyright status of a digital object is not only determined by the license that it carries, but also by any rule, or exception from a rule, that might apply as a matter of statute. This requires an appropriate formal representation not just of the license, but also of the relevant statutory regulations and possibly also the case law that interprets them. While it is possible in principle to represent them within the CARNEADES framework as a generic tool to model legal reasoning, the added complexity would mean that the application is unlikely to scale in practice in order that correct conclusions can be arrived at within a reasonable timeframe. All the "hard problems" of legal AI would become involved. Copyright law, to the dismay of many laypeople, uses vague terms routinely. Even a cursory look at the ubiquitous online help fora, for instance the copyright discussions on websites such as Ravelry, are full of "it all depends on the context" advice by the legally trained contributors, advice that is difficult for humans to operationalize and impossible for current software to make sense of. Indeed, the aim of MARKOS is to create "only" a decision support system that assists a human decision maker, rather than replacing her. In a typical application, a human programmer, without significant time pressure and working in her office, will want to check if all the software she has incorporated in her product is freely available, and what type of license her own product should have as a result. This is very different from a situation, such as the one discussed above, where a robot has to make autonomous, unsupervised decisions in real-time about whether it can create and analyse a copy of a photo of the painting on the gallery wall in order to assist its owner.

We can only hint here at future research directions that could begin to address this issue. One draws its inspiration from research into "privacy by design". Data protection law is of a complexity similar to copyright law. In the Internet of Things ("IoT"), it will be crucial to permit only information transfer that is compliant with data protection law. Many IoT devices will be small, with little computing power "on board", and will be aimed at performing tasks under extreme time pressure, for example an insulin pump "talking" to a fridge to determine the necessary dosage based on food consumption. Carrying out a fully autonomous yet comprehensive analysis of the relevant data protection law every time this is needed would be impossible even if we could formally represent all of the applicable regulatory norms. Research for the Smarter Privacy project addresses this problem by separating out the non-time sensitive layer (typically the design stage) of a large ITC project from the time-sensitive events that occur once the system is in situ and operational. A fully-fledged legal reasoner assists the designer of such a system during the less time-sensitive stage, when considering the general infrastructure of the product. Here, abstract decisions are taken about whether the fridge and pump should be connected at all, what types of records are needed when they interact, and how the data flow between them should take place. On the basis of this design, much simpler templates are auto-generated that are computationally cheap, represent only small segments of the legal picture and work "well enough" for most routine transactions (see for example Raabe et al 2010; Oberle et al 2012). On the application layer, the abstract, general rules are then concretised by applying the simplified template to the sensory input of the device. Similarly, we can envisage an approach to copyright and robotics, where the robot developer is assisted by a fully-fledged formal representation of applicable copyright law when she develops a business model and IT infrastructure for a care robot. Part of this process is to anticipate, using human knowledge, as much as possible the likely types of interactions the robot will experience once deployed. The robot then carries with it a much simplified version of the law, a default template, (whose accuracy can nonetheless be traced back to the complete model of the law at the designer's workstation) which it can apply with sufficient accuracy to whatever its sensors detect.

In our view, even this approach could be too ambitious. Raabe and Oberle remain deeply indebted to the formalistic tradition of continental legal reasoning, where reasonably clear rules give rise to clear legal consequences. The crucial transition from the general, abstract planning level to the application-centred templates therefore is only driven by formal, legal considerations. Research currently underway in Edinburgh suggests a slightly different approach that can incorporate an explicit element of reasoning about risk and risk management. In everyday life, we as humans are at best only moderately aware of the technicalities of copyright law that surround us. When taking photos of a castle in Germany, I am vaguely aware that there are laws which give copyright protection to works of architecture. However, despite working professionally in the field, I could not give a full account of the relevant law. Rather, I make a risk assessment: taking this photo may or may not be legal, strictly speaking, but as long as I do not publish it commercially and enjoy great commercial success, it is unlikely that I will be the subject of a law suit. Similarly, the care robot that guides its owner round an art gallery does not need even a scaled down version of the applicable copyright law. Rather, the developer should have carried out a “legal risk assessment” that incorporates not just the legal issues, but also what is known about real litigation risks. This can further simplify the type of rules that the robot, in the application stage, needs to process. From simplified legal rules, they become “dirty” rules of thumb that are also informed by our general knowledge of risks. At the design stage then, the developer would be aided in the process by experience-based meta-rule along the lines of: “if all information is deleted after the visit is over, the gallery owner will not risk bad publicity by suing a disabled visitor”. Complex legal distinctions, such as the difference between a painting and a sculpture, are thus rendered irrelevant. Robots need not be law-compliant all the time; they need, like us, to be compliant only when and to the extent that it matters. One of Asimov’s laws states that a robot must not harm a human through its inaction. It would be the ultimate irony if the inaction were caused by the robot’s need to perform a full legal analysis before deciding whether to act. Unsurprisingly, it was Bing who foresaw the importance of risk modelling for commercially viable and practical legal AI (Mahler and Bing 2006).

With this, the circle closes. Jon Bing dreamt as an artist about a world we share with robots, and as a legal visionary gave us an idea of how we should build the regulatory framework that structures our interactions with them. As his artistic vision gets closer to reality, his ideas on regulation and the interface between law and AI remain as inspiring and topical as they have ever been.

Literature

Abie, H., Foyn, B., Bing, J., Blobel, B., Pharow, P., Delgado, J., ... & Tzovaras, D. (2004). The need for a digital rights management framework for the next generation of e-government services. Electronic Government, an International Journal, 1(1), 8-28.
Asada, M et al. (2001). Cognitive developmental robotics as a new paradigm for the design of humanoid robots. Robotics and Autonomous Systems, 37(2), 185-193.
Bing, J. (1984). Handbook of legal information retrieval. Elsevier Science Inc..
Bing, J. (1991). Legal Text Retrieval Systems the Unsatisfactory State of the Art. JL & Inf. Sci., 2, 1.
Bing, J. (1996). The contribution of technology to the identification of rights, especially in sound and audio-visual works: An overview. International Journal of Law and Information Technology, 4(3), 234-254.
Bing, J., (1997). Availability and copyright." IFLA journal 23.4
Bing, J., (1998). A futuristic exploration of virtual reality. In Jeleff, S. A Virtual New World? Strassburg: Council of Europe
Bing, J., (2003). Deposit, copyright and the notion of a 'document'." Alexandria 15.3 (pp 171-174)
Bing, J. and Sartor, G., (2003). eds. The Law of Electronic Agents. Oslo: Unipubskriftserier
Bing, J. (2004). Copymarks: A suggestion for simple management of copyrighted material. International Review of Law, Computers & Technology, 18(3), 347-374.
Bing, J. (2007). Interview with Erling Rognli, Pegasus 3 (pp.24-35)
Bing, J. (2008). The Riddle of the Robots, J. Int'l Com. L. & Tech., 3, (pp 197 – 206)
Bing, J., (2009). Copyright Protection of Computer Programs, in Estelle Derclaye (ed.), Research Handbook on the Future of EU Copyright, Cheltenham, UK: Edward Elgar, 2009, pp. 401–426
Bing, J., (2013). Celebrating Gnaeus Flavius and Open Access to Law. Journal of Open Access to Law 1.1
Boden, M. A. (1994). Précis of the creative mind: Myths and mechanisms. Behavioral and Brain Sciences, 17(03), 519-531.
Boden, M. A. (1998). Creativity and artificial intelligence. Artificial Intelligence, 103(1), 347-356.
Boden, M. A. (2009). Computer models of creativity. AI Magazine, 30(3), 23.
Borges, J.L., (2007), The Total Library, Penguin: London
Bridy, A. (2012). Coding Creativity: Copyright and the Artificially Intelligent Author. Stanford Technology Law Review, 5, 1-28.

Butler, T. L. (1981). Can a Computer be an Author-Copyright Aspects of Artificial Intelligence. Comm/Ent LS, 4, 707.
Cangelosi, A.et al (2010). Integration of action and language knowledge: A roadmap for developmental robotics. Autonomous Mental Development, IEEE Transactions on, 2(3), 167-195.
Chon, M. (2007). Intellectual property from below: Copyright and capability for education. UC Davis Law Review, 40, 803.
Cox, I. and Linnartz, J. (1998) Some general methods for tampering with water-marks. Selected Areas in Communications, IEEE Journal on, 16(4) (pp587– 593)
Dante, M., and Tamburrini, G. Learning robots and human responsibility. International Review of Information Ethics 6 (2006): 46-51
Di Dio, C et al .(2007). The golden beauty: brain response to classical and renaissance sculptures. PloS one, 2(11), e1201.
Favale, M. (2011). Approximation and DRM: can digital locks respect copyright exceptions? International Journal of Law and Information Technology, 19(4), 306-323
Funkhouser, C., (2007), Prehistoric Digital Poetry: An Archaeology of Forms 1959-1995, Tuscaloosa: University of Alabama Press,
Gervás, P. (2000, April). Wasp: Evaluation of different strategies for the automatic generation of Spanish verse. In Proceedings of the AISB-00 Symposium on Creative & Cultural Aspects of AI (pp. 93-100).
Gervás, P. (2001). An expert system for the composition of formal Spanish poetry. Knowledge-Based Systems, 14(3), 181-188.
Gordon, T. F. (2011, June). Analyzing open source license compatibility issues with Carneades. In Proceedings of the 13th International Conference on Artificial Intelligence and Law ACM (pp. 51-55)..
Haupenthal, G., (1994) Geschichte der Würfelmusik in Beispielen. 2 Vol. Saarbrücken, Univ., Diss.,
Hartman, C.O (1996) Virtual Muse: Experiments in Computer Poetry. Wesleyan University Press
Hong, J and Rouget, T., (2014), Lesende und glaeubige Roboter im Film. In Libert et al (eds). Künstliche Menschen. Würzburg: Koenigshausen & Neumann
Horikawa, T., et al (2013). Neural decoding of visual imagery during sleep. Science, 340(6132), 639-642.
Jiang, L., & Zhou, M. (2008). Generating Chinese couplets using a statistical MT approach. In Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1 (pp. 377-384). Association for Computational Linguistics.
Jordanous, A. (2012). A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. Cognitive Computation, 4(3), 246-279.Kac, E. (1997). Foundation and development of robotic art. Art Journal 56.3)pp 60-67)
Kamitani, Y., Tong, F. (2005), Decoding the visual and subjective contents of the human brain. Nat. Neurosci., 8 (pp. 679–685)
Kostelanetz, R. (1971). MACHINE ART. Centennial Review 15 (pp 229-249)
Komuves, D et al (2015). Monkeying Around with Copyright - Animals, AIS and Authorship in Law' (2015) Jusletter-IT Vol 26 RZ1 - 27
Lee, J. H et al (2009). Brain–machine interface via real-time fMRI: preliminary study on thought-controlled robotic arm. Neuroscience letters, 450(1), 1-6.

Lee, E. "Digital originality." Vanderbilt Journal of Entertainment & Technology Law 14.4 (2012): 919.
Legaspi, R., et al (2007,). Music compositional intelligence with an affective flavor. In Proceedings of the 12th international conference on Intelligent user interfaces ACM. (pp. 216-224).
Lessig, L, (1999) Code and other laws of cyberspace. NY: Basic books
McCormack, J., & d'Inverno, M. (2012). Computers and Creativity. Springer: Berlin
McCutcheon, J. (2012). Vanishing Author in Computer-Generated Works: A Critical Analysis of Recent Australian Case Law, The. Melb. UL Rev., 36, 915.
Mahler, T., & Bing, J. (2006). Contractual risk management in an ICT context–searching for a possible interface between legal methods and risk analysis. Scandinavian Studies in Law, 49, 339-357.
Manurung, R., Ritchie, G., & Thompson, H. (2012). Using genetic algorithms to create meaningful poetic text. Journal of Experimental & Theoretical Artificial Intelligence, 24(1), 43-64
Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. Ethics and information technology, 6(3), (pp175-183)
Moffat, D. C., & Kelly, M. (2006). An investigation into people's bias against computational creativity in music composition. Assessment 13, (pp11 – 17)
Mulligan, D. K., Han, J., & Burstein, A. J. (2003, October). How DRM-based content delivery systems disrupt expectations of personal use. In Proceedings of the 3rd ACM workshop on Digital rights management (pp. 77-89)
Nadal, M., et al (2008). Towards a framework for the study of the neural correlates of aesthetic preference. Spatial vision, 21(3), 379-396.
Oliveira, H. G. (2012, August). PoeTryMe: a versatile platform for poetry generation. In Proceedings of the ECAI 2012 Workshop on Computational Creativity, Concept Invention, and General Intelligence
Oberle, D., Drefs, F., Wacker, R., Baumann, C., & Raabe, O. (2012). Engineering compliant software: advising developers by automating legal reasoning. SCRIPTed, 9(2), 280-313.
Pogue, D, (2012). The Remote Control in Your Mind. Scientific American 307.6 pp 32-32.
Pynadath, D. V., & Tambe, M. (2002). Revisiting Asimov's first law: A response to the call to arms. In Intelligent Agents VIII Springer Berlin Heidelberg. (pp. 307-320)
Raabe, O., Wacker, R., Funk, C., Oberle, D., & Baumann, C. (2010). Lawful Service Engineering–Formalisierung des Rechts im Internet der Dienste. Globale Sicherheit und proaktiver Staat–Die Rolle der Rechtsinformatik, 266, 643-650.
RACTER, (1985) The Policeman's Beard is Half-Constructed: computer prose and poetry by Racter. Warner Books
Reese, R. A. (2008) . Transformativeness and the Derivative Work Right." Columbia Journal of Law & the Arts 31.4 (pp 2012-54)
Reichardt, J. (1978). Robots: fact, fiction+ prediction, Thames & Hudson Ltd., London
Riedl, M. O. (2014). The Lovelace 2.0 Test of Artificial Creativity and Intelligence. arXiv preprint arXiv:1410.6142.
Roos, J. W. (2005). Copyright Protection as Access Barrier for People who Read Differently: the case for an international approach. IFLA journal, 31(1), 52-67.

Rowe, R. (1993). Interactive Music Systems: Machine listening and composing. Cambridge, MA: MIT Press.

Rowe, E. (2001). Machine Musicianship. Cambridge, MA: MIT Press

Searle, J., (1980), Minds, Brains and Programs, Behavioral and Brain Sciences 3 (3) pp 417–457

Sharkey, A., & Sharkey, N. (2012). Granny and the robots: ethical issues in robot care for the elderly. Ethics and Information Technology, 14(1), 27-40.

Shinkareva, S. Vet al . (2008). Using fMRI brain activation to identify cognitive states associated with perception of tools and dwellings. PLoS One, 3(1), e1394.

Solso, R. L. (2000). The cognitive neuroscience of art: a preliminary fMRI observation. Journal of Consciousness Studies, 7(8-9), (pp)75-86.

Stamp, M. (2003). Digital Rights Management: The Technology Behind the Hype. J. Electron. Commerce Res., 4(3), 102-112.

Schlegel, A et al. (2015) The artist emerges: Visual art learning alters neural structure and function. NeuroImage 105 (pp440-451)

Ward, J. A. (1992). Copyrighting Context: Law for Plumbing's Sake. Colum.-VLA JL & Arts, 17, 159.

Weld, D. and Etzioni, O., (1994). The first law of robotics (a call to arms)." AAAI. Vol. 94. 1994 pp. 1042-1047

Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. Knowledge-Based Systems, 19(7), 449-458.

Wolpaw, J. Ret al (2002). Brain–computer interfaces for communication and control. Clinical neurophysiology, 113(6), 767-791.

Xenakis, I., 2001. Formalized Music: Thought and Mathematics in Composition (Harmonologia Series No.6). Hillsdale, NY: Pendragon Press