GUEST EDITORIAL

Novel Beings: Moral Status and Regulation

David Lawrence[1]* and Sarah Morley[2]

[1]Centre for Biomedicine, Self and Society, Usher Institute, University of Edinburgh, Edinburgh EH8 9LN, Scotland
[2]Newcastle Law School, Newcastle University, Newcastle Upon Tyne NE1 7RU, United Kingdom
*Corresponding author. Email: d.lawrence@ed.ac.uk

A world of 'posthumans,' synthetic human embryos, perhaps even thinking artificial intelligences (AI) becomes increasingly inevitable. A vast range of technologies, including genomics, synthetic biology, advanced pharmaceuticals, neurotechnologies, and breakthroughs in computer science and artificial intelligence research promise to create new forms of sentient, even sapient intelligent life in the not-too-distant future; and perhaps change our understanding of life itself.

These types of technology promise significant effects on our way of life, of working, and of interacting with others— perhaps even as significant as in the science fiction worlds they were once relegated to. This special section explores these effects, and what they may mean for us in terms of our understandings of ourselves and the assumptions on which we base so many facets of our lives and society.

We are already capable of the fundamentals of creation, products of artifice which may in future possess moral value or approach forms of sentient or even sapient 'life.' These developments promise to draw into question the nature of what it is to be human and what it is to be a person before the law. These are concepts upon which many significant institutions are founded. The advent of these new intelligent life forms – 'novel beings'—may require us to re-evaluate *Homo sapiens*' position as the sole bearer of human rights; and will pose disruptive global challenges for society and the law regarding their moral status, protections, and freedoms, as well as their obligations, and ours towards them. Should a synthetic animal fall under 'natural' animal protection provisions? Should AI who fulfil the criteria for moral value to greater degrees be afforded greater consideration? Reasoning creatures resulting from either techno-logical route may equally deserve their own protections and freedoms; they may be moral agents, or perhaps they should share in natural personhood—i.e. to be 'legally human.'[1] It is vitally important to establish whether existing law will remain sufficient to manage the potential development of new types of intelligent being from these technologies, and if not, how it ought to be adapted to meet the requirements of the future.

Critics of our stance suggest that these are not problems for today; but are instead so distant that we ought to be concerned with more immediate social issues. Certainly, there is some merit to this point. However, it is a continuous cycle. There is always a current need that deserves to be addressed, and so we leave the future problems until they themselves are current—which may be too late. The advent of the novel being is an event of such magnitude that it warrants breaking this cycle.

The law, particularly in a common law system, has a predilection for reactivity rather than being proactive. Doubtless, many legal experts will criticize so sweeping a statement, and there is increasing legal scholarship supporting the *ex ante* view.[2] Nonetheless, law as an entity necessarily must act in retrospect in an attempt to prevent given situations, actions, or circumstances from recurring. This is a sensible and logical approach: only through experience can we know for certain what weaknesses the system may have, and thus seek to patch them with new rulings or even legislation. It would not do to legislate blindly in fear of the unforeseen; that waylays totalitarianism and dystopia. For this reason, the law often trails behind the world—eventually papering the cracks, but leaving them to grow in the meantime. Some of these cracks are unfortunate, even unsightly, but to push a metaphor, not structurally

threatening. These cracks emerge with greater speed and frequency in certain sectors, notably in law relating to science and technology.

Unfortunately, reactionary behaviour is likely too entrenched and too fundamentally a part of the structure of lawmaking, to be easily remedied. It could be argued that the system has functioned well enough for centuries and has proven itself broadly up to the task of responding robustly before too much suffering results. By any moral regime, it would be considered right and good to ensure the best course of action was being settled upon rather than acting rashly and without consideration. However, it is the case that we have already seen a range of issues and even harms caused by the precursors to sapient technology. With these stakes being so high, it is imperative that we at least consider the implementation of minimum standards before we can allow reactionary law to take its course.

In recent years, for example, there has been a scramble to legislate for the advent of Mitochondrial Replacement Therapies (MRTs), which were made legal under license in 2015[3] to great controversy of public opinion. However, the technology was known and discussed in scientific circles for quite some time beforehand;[4] and it bears saying that it is possible that cases of mitochondrial disease—with the potential for terrible suffering—could have been avoided in the four year interim between the technology being deemed 'ready'[5] and the first clinical license being granted.[6]

We might consider even more foreseeable, broader circumstances such as the explosion in the last decade and a half of the internet and internet-connected aspects of daily life. It is nigh impossible to live a modern lifestyle without constant interaction with the internet in one way or another, and the situation developed so freely that the internet itself is something of a 'wild west' (or is, at least, beneath the surface). There are domestic laws applied in order to restrict access to specific illegal content in various ways, but it is the case that any user with a modicum of knowledge and ability can bypass blocks on sites with ease. One of the few examples in the United Kingdom of specific legislation to control internet content, the Digital Economy Act 2010[7] (which addresses issues such as copyright infringement and media policy), had several statutory elements never pass into law thereby reducing its utility; which was widely criticised both academically and by the very 'digital economy' it was aimed to support.[8] More to the point, it is a prime example of the aforementioned papering over the cracks. Due to the lateness of the Act coming into being, the structure (for want of a better term) of the internet was well-established, and so the demands and results of the Act are easily circumnavigated. Reliable statistics prove difficult to find, but piracy, money laundering, and access to forbidden content remain an insurmountable problem online. This is to say nothing of the sweeping controls proposed by the Conservative government led by Theresa May, which were largely agreed to be impossible to implement.[9]

It is simply too late to impose regulation that would actually effectively moderate the content or reach of the internet; and given the exponentially greater effect on society that new intelligent life is likely to have, it is vital that we avoid the same scenario. In 1962 Marshall McLuhan famously predicted that "Instead of tending towards a vast Alexandrian library the world has become a computer, an electronic brain, exactly as an infantile piece of science fiction,"[10] and described the 'global village' with which we are now intimately familiar. McLuhan was by no means the only thinker with this kind of foresight, even if the issues were not ones riding high in the public discourse. Had decisive action to control and regulate the internet been taken earlier, perhaps the current situation would be different and a comfortable balance could have been struck between freedom of speech and safeguarding users. We find ourselves with precisely this opportunity in regard to novel beings, to bring together life and computer sciences with the humanities in a new way which focuses on redefining, and providing a holistic alternative to, the piecemeal approach that is currently used in the regulation of emerging technologies. The papers included in this Special Section aim to set the scene for this future work.

To begin, we propose uniting a number of fields that are often held as disparate, but which are all fundamental to understanding what it is to be human- and how we might apply this understanding to our creations. The present authors are part of a previously established network incorporating expertise from company law, medical law, bioethics, philosophy, biomedical sciences, neuroscience, computer science, regulatory theory, science and technology studies, and clinical practice.

We will need to consider in more detail the interplay between consciousness, responsibility, and liability, and to attempt to provide a basis for developing workable legal definitions of consciousness that can inform a regulatory and social policy response. This foundation will be essential for assessing levels of accountability required by companies (as the developers), or the novel beings themselves and any 'human' rights that might be bestowed upon them. Morally significant questions are currently not answered by the law; for example, as the law stands, the developers of these technologies are not required to consider whether these beings have a right to life, to liberty, or to self-ownership, nor to the impacts its existence and operations may have on society. It is the role of the humanities, particularly those that most influence societal perspectives, to help determine the correct course of action.

Our preliminary work in this area, supported by the Wellcome Trust,[11] was the first of its kind to explore how to regulate advanced morally significant technology and to consider the suitability of company law in specifically regulating the conduct of corporations in the development, operation and disposal of these technologies. The paper we (Lawrence and Morley) include here is a summation of some of our findings, and those which follow examine fascinating conundrums and elements which arose over the course of our discussions with the authors, between the authors themselves, and with the many participants in our project.

Daniel Tigard, in his contribution, explores the idea of moral responsibility. By examining the tripartite theory promoted by David Shoemaker (amongst other conceptions), he makes a case that its attributability to a so-called 'artificial' moral agent is not necessarily a new problem. There are, already, pluralistic means by which we tend to hold non-agent machines responsible; and if we accept that moral agency does not automatically equate to responsibility, it follows that these mechanisms may have a significant part to play. This sentiment, that new kinds of agent (or novel beings) are not entirely unprecedented problems, echoes throughout the papers that follow. Tigard's paper elicited a reply from Marc Champagne, in which he questions the very possibility of artificial responsibility.

Henry Shevlin's paper applies this notion to moral patiency, which he locates around personhood, preference, and the capacity to suffer. He suggests that this approach may no longer suffice when we cannot be sure of recognising these traits, but proposes that we rely instead on a more intuitive strategy of cognitive equivalence. Namely, Shevlin argues that artificiality is not what matters in determining patiency, but rather evident behaviors and their causal cognitive characteristics.

Alex McKeown, meanwhile, takes this notion further by highlighting that it is not only cognitive behaviour with which we ought be concerned in determining moral agency, but that a broader, embodied conception might be more useful. We cannot necessarily assume that we will be able to understand the abstract 'mind' of an entirely new life form, as we have a tendency to rely on recognisable physicalities for empathetic purposes. Whilst we may have difficulty engaging with something entirely alien, taking into consideration the ways in which that being is physically instantiated, may be our most successful route to understanding its possible values.

Gardar Arnasson presents a case against the idea that the creation of novel beings—or indeed their likely precursors in cognitively uplifted nonhuman primates—is, in itself, morally problematic. Whilst acknowledging that there may very well be contingent harms by virtue of the likely reasons such a being may be created (such as neurodegenerative research), he points out that these are somewhat divorced from that being's moral status. The issues of novel concern, he argues, will only emerge where the primates are granted characteristics that constitute personhood—i.e., "the capacity for self-consciousness with meta-cognition."

Joshua Jowitt unifies some of the themes of these previous papers, and considers more specifically how we might apply existing legal norms to establish a test for justifying any rights for our novel beings. Using the examples in law of recent great ape rights cases, he demonstrates that legal tendencies towards anthropocentrism are not inviolable and it seems there should be little barrier to the granting of such personhood. His Gewirthian perspective, including the contingent justification for universalising legal rights claims, provides a means for assessment of their legitimacy.

Miranda Mowbray's paper takes a practical approach by considering the limitations of the criteria we may develop for recognising moral status in Novel Beings. In doing so she specifically highlights the ability of developers in AI to satisfy our anthropocentric standards for conscious behaviour. Mowbray explores the many ways in which existing software could be considered to surpass these criteria; and, given this, argues that any reasoning behind the granting of rights to an AI must take account of the potential for human designers to deliberately manipulate these standards, in order to take advantage of any rights claims and shirking of responsibility.

Aisling McMahon builds on this fear of a possible 'abuse of the system' and examines a different aspect of the novel beings question, investigating the control that could be exerted over novel beings through patenting, and the ways that licensing could impinge upon the rights and agency of novel beings. McMahon therefore considers the extent to which having consciousness should be an exclusionary factor for patentability and how much control a developer should have over its creation.

Isra Black looks ahead to an entirely different challenge likely to follow the emergence of novel beings. Black explores the necessary criteria that novel beings would need to fulfil in deciding to end their own lives, and if indeed such considerations are primarily anthropocentric. It may be that a novel being has a significantly different value set from our own, and so an analysis of assisted nonexistence for novel beings should also facilitate our understanding of the choices to end our own lives.

Together, the papers that follow represent only the beginning of our collective thinking in this new area. Our aim now is to broaden this to include a greater wealth of perspectives from the humanities more broadly. We envision including yet more disciplines such as anthropology, sociology, literary studies, and creative arts culture. All these disciplines actively engage with considerations around what it is to be human, and we are confident that the work already being conducted in these areas around AI will lead to new horizons.

Although, not being something we will see today, nor indeed tomorrow, it is vital that we turn our efforts now to considering the novel being—even if it is to protect ourselves. In so doing, we are likely to discover some fundamental truths about who—and what—we value.

Notes

1. Lawrence DR, Brazier M. Legally human? The status and challenge of novel consciousness. *Law Medical Law Review* 2018;**26**(2): 309–327.
2. Notably including an Opinion of the European Commission. See: European Economic and Social Committee The Proactive Law Approach: A Further Step Towards Better Regulation at EU Level (2009) *Official Journal of the European Union* 2009/C175/05.
3. UK Parliament. The Human Fertilisation And Embryology (Mitochondrial Donation) Regulations 2015 No. 572 (2015).
4. Newcastle University having applied for a research license in 2004: Randerson J. Scientists seek to create 'three-parent' babies. *New Scientist* 19 Oct 2004; available at: https://www.newscientist.com/article/dn6547-scientists-seek-to-create-three-parent-babies/ (last accessed 28 Nov 2019).
5. Connor S. UK becomes first country In world to approve IVF using genes of three. *The Independent* 27 June 2013; available at: http://www.independent.co.uk/news/science/uk-becomes-first-country-in-world-to-approve-ivf-using-genes-of-three-parents-8677595.html (last accessed 28 Nov 2019).
6. Gallagher J. Three-person baby licence granted. *BBC News* 16 Mar 2017; available at: http://www.bbc.co.uk/news/health-39292381 (last accessed 28 Nov 2019).
7. Digital Economy Act 2010 (c.24) (as amended by Digital Economy Act 2017).
8. Dutton W. Aiming at copyright infringers and hitting the digital economy. *Prometheus* 2010;**28**: 385; Andersen B. Shackling the digital economy means less for everyone: The impact on the music industry. *Prometheus* 2010;**28**: 375.

9. Onwurah C. Theresa May's plans to regulate the internet are late, unworkable and wrong. *Newstatesman.com* 6 Jun 2017; available at: http://www.newstatesman.com/politics/staggers/2017/06/theresa-mays-plans-regulate-internet-are-late-unworkable-and-wrong (last accessed 28 Nov 2019).

10. McLuhan M. *The Gutenberg Galaxy: The Making Of Typographic Man*. Toronto, ON: University of Toronto Press; 1962: 32.

11. (WT 208871/Z/17/Z).

Raoul Hausman (1886-1971), The Spirit of our Age (Mechanical Head). Location: Musee National d'Art Moderne, Centre George Pompidou, Paris, France. Photo Credit: CNAC/MNAM/Dist. RMN-Grand Palais/ Art Resource, NY and Artists Rights Society NY, Reproduced by Permission.