

Framing the Future: The *Foundation* Series, Foundation Models and Framing AI

Clare Williams

University of Kent, United Kingdom

Abstract

Are we at risk of preserving in digital aspic current hegemonic ways of doing, talking and thinking about law and economy? Our dominant frames are those of neoclassical economics and doctrinal law, which, combined, align with neoliberal assumptions and preferences. Invisible through their ubiquity, these partially reveal and conceal problems and solutions, perpetuating inequalities. And yet, unlike social biases, our framing tends to fly largely under the radar, as we see in Isaac Asimov's *Foundation* series.

Taking three themes from Asimov's work that align with current hegemonic frames, namely the centrality of the individual as the subject, a belief in scientism and technological progress to save humanity, and a dissatisfaction with the limitations of language, this paper explores what, how and why we might exercise caution in the development of natural language processing (NLP) artificial intelligence (AI). By drawing parallels with Asimov's work and the development of foundation models that underpin all NLP AI systems, this paper asks to what extent neoliberal framing and the inequalities perpetuated therein might come to be embedded in future technologies.

Keywords: Science fiction; Artificial intelligence (AI); foundation models; Natural Language Processing (NLP); Asimov.

1. Foundations: Asimov's *Foundation* Series and AI's Foundation Models

Unlike Hari Seldon, we cannot predict the future. On planet Earth in the twenty-first century, simply interrogating the present can be a complex undertaking. Hari is a mathematician and our first protagonist in Asimov's *Foundation* series.¹ Having pioneered the field of 'psychohistory', which resembles sociology on steroids in its quantification and prediction of the sum of human behaviour, Hari believes that he can not only foretell future events but can also intervene and mitigate the impending collapse of the Galactic Empire. And yet, we can find echoes of similar approaches throughout the twentieth and twenty-first centuries, in both fact and fiction. Insights from Hari's work can offer us three core themes for exploring the implications for society of future technologies and how we might regulate them—specifically, natural language processing (NLP) artificial intelligence (AI). In a tale of life imitating art, these themes characterise the trajectory of both law and economics in the twentieth and early twenty-first centuries, and now comprise the analytical and normative backdrop against which we seek to regulate NLP AI.

Fiction is not anathema to critical explorations in the social sciences.² Economists frequently start out with the proposition 'Imagine...', usually following this entreaty with a simplified, fictionalised account of real-world interactions.³ Lawyers are similarly prone to asking 'what if?' and to using hypothetical situations to explore the limits of legal doctrine.⁴ In querying

¹ Asimov, *Foundation*.

² Vaihinger's 'as if', or 'fiction' theory is noteworthy here. See *inter alia* Kelsen, "On the Theory of Juridic Fictions"; Samuel, "The Epistemological Challenge."

³ See, for example, Coase, "The Problem of Social Cost." We might note that the concept of the free, self-regulating market is little more than a utopian fiction. See, for example, Polanyi, *The Great Transformation*.

⁴ Western legality is founded on a series of fictions, from the legal person to John Locke's mental gymnastics creating *terra nullius* that justified settler-colonialism, land expropriation and extraction. Even legal interpretation by courts of contracts relies on questions of what the parties might have intended, blurring the line between fact and fiction. Kramer, "Common Sense Principles of Contract Interpretation."



Except where otherwise noted, content in this journal is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/). As an open access journal, articles are free to use with proper attribution. ISSN: 2652-4074 (Online)

future imaginaries, especially in regard to future technologies, we are already in the realm of science fiction.⁵ Unlike law and economics, which are (or at least should be) drawn back to their real-world applications, science fiction offers a realm from which there is no imperative to return to reality. By staying with the unreal, with the speculative, we can explore future demands that might be made of both law and economics through cultural lenses.⁶ We can ask what kind of society we might want, what perils this society might face from emerging technologies, and how these might be averted, modified, or tempered. This might, in turn, extend to the conceptualisation of law, the prevalence of the social or cultural, and the perceived relations between the law, the economy and society. An exploration of law and technology literature already notes that such endeavours can be intrinsically limited by how the inquiry is framed, noting the tendency towards doctrinal, positivistic scholarship—a hegemonic approach that elides the normative.⁷ What tends to be lacking, to date, in both the law and technology and AI literatures is a deep dive into the importance of framing—that is, the vocabularies and grammars that we use to understand and perform legal and economic phenomena, and how these, in turn, perform us.

The problem, though, as we can see from even a cursory reading of Asimov’s *Foundation* series, is that to be able to interrogate and imagine alternatives, we need an awareness that things can change. We need a starting point that acknowledges the temporal, spatial and cultural contingencies of both law and economics; that our ways of regulating behaviour and allocating resources are not immutable facts of nature but can, do and indeed ought to change.

As a thought experiment, consider the differences between our notions of law and economy and those of the Roman Empire some 2,500 years ago. While there are some parallels—many of our legal concepts evolved from origins in Roman law—centuries of evolution, along with some revolutions in legal and economic thought, mean that these contemporary social institutions now look quite different to their Roman counterparts, despite a modest time difference of some 2,500 years.⁸ By contrast, Asimov’s *Galactic Empire* is set tens of thousands of years hence, and billions of miles away. We might, then, expect the legal and economic institutions described and performed by characters in the *Foundation* series to have evolved drastically from their earthbound origins. Yet, they are curiously unchanged from their manifestations in the 1940s United States: characters refer to legal treaties, rights to land ownership and resource extraction, and describe commercial transactions. This copy-and-paste of the social structures so central to the book can be familiar and yet jarring, given the radical scientific and technological innovations against which they are counterposed.⁹ Could this have been a deliberate ploy on the part of the author? Or was the wholesale import of legal and economic conceptual and linguistic tools simply an ‘unthought known’, those deep rationalities we are so frequently unaware of?¹⁰ Were ways of doing, talking and thinking about law and economy simply so obvious, so ubiquitous and so unchangeable to Asimov because no other way of framing these things had ever been put to him?¹¹

One answer might be that Asimov imported the social institutions of the twentieth-century United States into the Galactic Empire to interrogate them against an unfamiliar backdrop. A fictionalised examination allows ‘for a cultural response, consideration, and critique of events before they occur’, allowing Asimov to interrogate possible future technologies and their impacts on United States society.¹² However valuable this approach might have proven, we are potentially at risk of making the same mistakes in the context of AI by failing to interrogate our ‘unthought knowns’.¹³ In interrogating the risks and rewards for humanity of AI, our task is similar to Seldon’s in predicting and mitigating harm. The development of NLP AI, which depends on a small number of foundation models, is transforming virtually all walks of life, from healthcare, to education, to law. Foundation models, as monolithic entities that are expensive to devise and train, are relatively few but underpin almost all NLP AI systems. If you are talking to a chatbot about your television insurance or a health complaint, or using a legal algorithm to draft a contract, you are, most likely, using an adapted and applied foundation model. Google searches: foundation model.¹⁴ Siri or Alexa: foundation model. The list goes on. The point is that we tend to use these tools every day without realising. These models are trained using existing data that were originally created by (a subset of) humans—flawed, biased humans with

⁵ Tranter, “The Speculative Jurisdiction.”

⁶ Travis, “Making Space.”

⁷ Tranter, “The Laws of Technology”; Tranter, “The Law and Technology Enterprise”; Tranter, “The Speculative Jurisdiction”; Travis, “Making Space.”

⁸ Samuel, “Classification of Obligations and the Impact of Constructivist Epistemologies”; Samuel, “Have There Been Scientific Revolutions in Law?”; Samuel, Rethinking Legal Reasoning; Samuel, “Causation and Casuistic Reasoning in Roman and Later Civil Law.”

⁹ Asimov does not spend much time on character or plot development in the books. Nevertheless, the books comprise core science fiction canon.

¹⁰ Bernstein, “There is No Local Here, Love”; Frerichs, “Re-embedding Neoliberal Constitutionalism.”

¹¹ Haraway, *Staying with the Trouble*; Norman, *Posthuman Legal Subjectivity*.

¹² Travis, “Making Space: Law and Science Fiction,” 248.

¹³ Bernstein, “There is No Local Here, Love.”

¹⁴ Uszkoreit, “Transformer”; Nayak, “Understanding Searches.”

preferences and assumptions. And these biases are not only preserved but can also be amplified over future generations of foundation models.¹⁵

Foundation model de-biasing initiatives are becoming increasingly sophisticated in identifying and fixing biases found in foundation models and their downstream adaptations, as later sections discuss in more detail. But the more nuanced, subtler framing of legal and economic phenomena that can encode preferences and perform inequalities is still passing, largely, unnoticed. We are, then, potentially embedding hegemonic notions of law and economy into our future AI systems. And in doing so, we are potentially placing these mainstream conceptual and linguistic tools further beyond the reach of current or future interrogation, and beyond the reach of natural linguistic evolution. But the way we talk matters. Like a torch shining across a topographical landscape, the frames we deploy both reveal and conceal by highlighting some voices and interests while casting others into deep shadow. And yet they tend to be those taken-for-granted parts of speech that fly largely under the conscious radar. While technological progress has innumerable benefits to offer humanity, by preserving hegemonic framing, we risk preserving in digital aspic preferences for those values and interests that are highlighted by that frame. And these tend to be those that align with dominant frames: neoclassical economics and doctrinal approaches to law.

1.1 Three Themes

Perhaps—subplot twist—Asimov foretold the rise of NLP AI systems and the foundation models that underpin them. Perhaps he meant to transplant directly the legal and economic institutions familiar to us today in the global north and the (neo-)liberal social context in which they emerged. Perhaps Asimov intended to warn us about the dangers of encoding social institutions and transposing them from one social context to another without a second thought, bound up as they are in our metaphors and linguistic fictions. However unlikely this may be, the *Foundation* series gives us three sub-themes that align with hegemonic framing of law and economy to explore the potential pitfalls, and opportunities, of AI systems. Through these, we can also explore how we might be preserving the conceptual and linguistic tools of neoliberalism (neoclassical economics and doctrinal or positivist law) in future foundation models.

In turning first to individualism, a core tenet of both neoclassical economic theory and doctrinal legal scholarship, the paper will suggest that Asimov's atomistic and individualistic characters like Hari Seldon and his fellow scientists are simply *homines economici* and *homines juridici*.¹⁶ They are, and act as, the social atoms that comprise the focus of mainstream approaches in the social sciences. Using insights from performativity and the cognitive sciences, we can see how these caricatured models that have developed in both economics and the law have a performative effect on their subjects. Over time, we come to perform as the model, much as the inhabitants of Asimov's First Galactic Empire do, and this assumption underpins many of the conceptual and linguistic tools on which foundation models are trained.

Then, second, focusing on belief in technological progress and scientism to save humanity, the following section sets out how a reliance on quantification in the social sciences (exemplified by the Law and Economics movement and seen in the World Bank's *Doing Business* series) attempted to realise just such a move. While characters in the *Foundation* series engage in 'symbolic logic' to boil down textual treaties to mere symbols, we can see parallels in the quantification, mathematisation and comparison of the law and of society, as well as in the encoding of language in models for the development of NLP.¹⁷ These correspond, respectively, to the empirical, analytical and normative tools of neoclassical economics and have been widely applied to virtually every area of social life.¹⁸ Being preserved, unwittingly, in future NLP foundation models raises questions about the encoding and amplification of bias and resulting inequality.¹⁹

Third, the paper turns to Asimov's (spoken through Hardin) dissatisfaction with the rhetoric and obfuscation of language.²⁰ Human thought processes are largely metaphorical, and our ways of talking about law and economy are no different. But baked

¹⁵ Bommasani, "On the Opportunities."

¹⁶ Latin plurals of the more familiar *homo economicus* and *homo juridicus* are borrowed from MacKenzie, "Constructing a Market," 139.

¹⁷ The European Commission has suggested that '[c]ombining symbolic reasoning with deep neural networks may help us improve explainability of AI outcomes' and defines 'symbolic approaches' as those rules created through human intervention. See European Commission, White Paper, 5.

¹⁸ Perry-Kessaris, "Prepare Your Indicators."

¹⁹ The OECD Recommendation of the Council on Artificial Intelligence advocates for the preservation of 'economic incentives' for innovation within AI, while recommending that AI development is based on key principles such as 'inclusive growth' and 'human-centred values and fairness'. See OECD "Recommendation of the Council on Artificial Intelligence."

²⁰ See Hardin's comments in *Foundation* as well as his description of Holk's 'symbolic logic' analysis that 'succeeded in eliminating meaningless statements, vague gibberish, useless qualification – in short, all the goo and dribble', which resulted in finding that 'he had nothing left. Everything cancelled out'. We can see Asimov's distaste for, and distrust of, verbiage and loquacity and his preference for the

into our metaphors are assumptions, preferences and biases about law, economy and society that we frequently lose sight of, despite the conceptual heavy lifting we ask them to perform. As future foundation models are trained on our metaphors, no amount of contextual understanding can compensate for overlooking the importance of framing and the inequalities that might result. Given that the political right has, for decades, been setting the frames, this seems a curious point in history to develop AI systems that might preserve in digital aspic our hegemonic neoliberal conceptual and linguistic tools and the voices and values they privilege. Perhaps forever.

The final section provocatively asks at what point Asimov's Zeroth Law of robotics might come into play if AI systems were permitted to entrench inequalities through legal and economic frames. Before turning to these three themes though, a word about metaphors is useful. But before that, a short introduction to foundation models is instructive.

2. Foundation Models, Natural Language and Artificial Intelligence

Foundation models require huge investments of time and money. They are the results of complex training involving 'self-supervised learning' on the part of the algorithm.²¹ So, for example, to train one foundation model (with the friendly name of BERT), a 'masked language modelling task' is used, where the algorithm is asked to fill in the blank, using only the surrounding context for guidance.²² An example is '*I like _____ sprouts*'. As humans, we might note that the most obvious word here is 'eating'. But we might also be able to think of dozens of other verbs conveying what we might like to do with a sprout. The point is that we use the surrounding words as context to deduce the most likely word. Tasks like this are more scalable as they rely on unlabelled data and, as they force the model to predict parts of the inputs, they result in models that are potentially richer and more useful.²³

Foundation models are intermediary assets rather than end products; they require adaptation and application to realise real-world impact. Nevertheless, their value comes from their 'emergent qualities', or skills that these foundation models can develop on their own without prompts or commands from a human. This derives from the 'self-supervised' learning that foundation models undertake, which grew out of work on 'word embeddings', which 'associated each word with a context-independent vector' and provided the basis for a wide range of NLP models. Context was then added through the 'autoregressive' approach (e.g., predict the next word given the previous words), and this produced models that represented words in context. The next wave in self-supervised learning built on the Transformer architecture (the 'T' in BERT), which incorporated 'more powerful deep bidirectional encoders of sentences' that again allowed for scaling up to larger models and datasets. Accordingly, the BERT, GPT-2, RoBERTa, T5 and BART models followed.

The 'sociological inflection' at this point, and the realisation that a small number of foundation models could be adapted and applied to an infinite variety of tasks, led to the era of the foundation model.²⁴ Now, '[a]lmost all state-of-the-art NLP models are [...] adapted from one of a few foundation models' such as BERT.²⁵ This has benefits: homogenisation means that the costly and time-consuming training required to get BERT up to a certain standard can be applied to a wide range of downstream scenarios and contexts. But, conversely, 'all AI systems might inherit the same problematic biases of a few foundation models'.²⁶ Like an echo in a cave that repeats countless times as it bounces off the cave walls, foundation models have the potential to echo throughout the applications to which they are adapted, repeating and re-entrenching intrinsic biases to an ever-increasing range of contexts. Two problems arise.

First, the next generation of foundation models will likely be trained on data curated from interactions with the current generation, which already encodes bias. The echo chamber of foundation model homogenisation, then, could easily see not only the preservation but also the augmentation of intrinsic bias over generations of foundation models. Second, foundation models are too complex and too monolithic to be adjustable following their release, despite the acknowledgement that 'any

clear, neat, precision of numerical values that finds echo in the quantified preferences and binaries of neoclassical economics. Asimov, *Foundation*, 64, 67.

²¹ Bommasani, "On the Opportunities," 4. The term 'foundation' also connotes 'the significance of architectural stability, safety, and security: poorly-constructed foundations are a recipe for disaster and well-executed foundations are a reliable bedrock for future applications': Bommasani, "On the Opportunities," 7.

²² BERT, or Bidirectional Encoder Representations from Transformers, is deeply bidirectional, making it able to draw on context to determine task response. Devlin, "Open-Sourcing BERT"; Devlin, "BERT."

²³ Bommasani, "On the Opportunities," 4.

²⁴ Bommasani, "On the Opportunities."

²⁵ Bommasani, "On the Opportunities," 5.

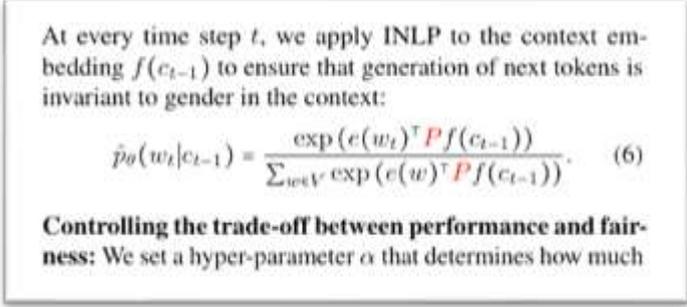
²⁶ Bommasani, "On the Opportunities," 5.

flaws in the [foundation] model [will be] blindly inherited by all adapted models'.²⁷ Therefore, social or representational biases discovered must be remedied *post hoc* and, accordingly, a de-biasing industry has sprung up, concerned both with identifying biases and then devising the best ways of prompting foundation models to engage in self-supervised remedial study.²⁸ Explicit social bias can be identified easily. But more commonly, bias is concealed in layers of metaphor and implication, meaning that while each of the words might be uncontroversial, when put together in a certain context, their combined meaning can produce inequalities. If, for example, the model tends to assume that doctors are male and nurses are female, we can appreciate a subtle bias, despite all the individual words being unproblematic. Nevertheless, our growing reliance on AI is a grand social experiment that is being conducted on a largely unsuspecting public in real time, without Hari's psychohistorical tools to calculate and predict future benefits and harms.

3. Language, Meaning and Metaphor

As a uniquely complex phenomenon, language, rhetoric, or 'sweet talk' is at once a cognitive and cultural system of encoded preferences and assumptions.²⁹ This holds as much for small talk as for social institutions like law and economy. And yet language is more than a system of signs and sounds.³⁰ Thus, what it means 'to mean' or 'to understand' has long troubled philosophers, sociologists, linguists and now computer programmers working in AI.³¹ Foundation models can not only process but also recreate original text that can pass, for the most part, for human-created content.³² But to ask whether a foundation model 'understands' what it has just written inquires into the temporal, spatial, cultural and contextual contingencies of language, as well as the subjective interpretation of symbols and sounds. While foundation models process words in context, they lack an awareness of the language's subtext, of social context and of the full pragmatic impact of those words. Foundation models remain amoral, acontextual and unaware of the meaning that lies within the (very convincing) texts they can generate.³³ Thus, while for humans, language is a key to unlock communication, for foundation models, language is a mathematical process that can unlock engagement with humans.

Hari and some of Asimov's other characters are unimpressed with the vagaries of language, engaging in a process called 'symbolic logic' to boil down the text of legal treaties to their core assertions. They find, to their collective horror, that a treaty in fact stated the opposite of their assumptions.³⁴ We might note parallels here with NLP AI in the way that language modellers quantify linguistic tokens, breaking down language into constituent parts that can then be processed by an algorithm, as seen in Figures 1 and 2. These illustrate how we might quantify language and how we might represent these data visually. As section 5 discusses, the quantification of language and the human interactions it facilitates reflects a twentieth-century belief in the potential of mathematics to encode all social phenomena.



At every time step t , we apply INLP to the context embedding $f(c_{t-1})$ to ensure that generation of next tokens is invariant to gender in the context:

$$\hat{p}_{\theta}(w_t|c_{t-1}) = \frac{\exp(e(w_t)^T P f(c_{t-1}))}{\sum_{w \in V} \exp(e(w)^T P f(c_{t-1}))} \quad (6)$$

Controlling the trade-off between performance and fairness: We set a hyper-parameter α that determines how much

Figure 1. Screenshot from Liang, "Towards Understanding," 5, showing the quantification of language as tokens that can be processed by an algorithm.

²⁷ Bommasani, "On the Opportunities," 6. Liang, "Towards Understanding."

²⁸ This tends to be through a series of language prompts and bias classifiers. See Liang, "Towards Understanding."

²⁹ McCloskey, "The Rhetoric of Law and Economics"; Shariatmadari, "Don't Believe a Word," 64.

³⁰ See further with regard to the law Goodrich, "Rhetoric, Semiotics, Synaesthetics."

³¹ Kramer, "Common Sense Principles of Contract Interpretation."

³² This is not foolproof though, and as Bender et al. note, foundation models can still lapse into 'stochastic parrots', spewing out nonsense. Bender, "On the Dangers of Stochastic Parrots." On AI-generated essays, see Sharples, "New AI Tools."

³³ Sharples, "New AI Tools."

³⁴ Asimov, *Foundation*, 65–67.

model	offensive 42.2% pos. (dev.)			intent 44.8% pos (dev.)			lewd 3.0% pos (dev.)			group 66.6% pos (dev.)			in-group 5.1% pos (dev.)		
	F_1	pr.	rec.	F_1	pr.	rec.	F_1	pr.	rec.	F_1	pr.	rec.	F_1	pr.	rec.
SBF-GPT ₁ -gdy	75.2	88.3	65.5	74.4	89.8	63.6	75.2	78.2	72.5	62.3	74.6	53.4	-	-	-
dev. SBF-GPT ₂ -gdy	77.2	88.3	68.6	76.3	89.5	66.5	77.6	81.2	74.3	66.9	67.9	65.8	24.0	85.7	14.0
SBF-GPT ₂ -smp	80.5	84.3	76.9	75.3	89.9	64.7	78.6	80.6	76.6	66.0	67.6	64.5	-	-	-
test SBF-GPT ₂ -gdy	78.8	89.8	70.2	78.6	90.8	69.2	80.7	84.5	77.3	69.9	70.5	69.4	-	-	-

Table 4: Experimental results (%) of various models on the classification tasks (gdy: argmax, smp: sampling). Some models did not predict the positive class for “in-group language,” their performance is denoted by “-”. We bold the F_1 scores of the best performing model(s) on the development set. For easier interpretation, we also report the percentage of instances in the positive class in the development set.

classification, commonsense generation, and conditional generation (Bosselet et al., 2019; Keskar et al., 2019).

Training We cast our frame prediction task as a hybrid classification and language generation task, where we linearize the variables following the frame hierarchy.⁷ At training time, our model takes as input a sequence of N tokens:

$$\mathbf{x} = \{[\text{STR}], w_1, w_2, \dots, w_n, [\text{SEP}], w_{[\text{lewd}]}, w_{[\text{off}]}, w_{[\text{int}]}, w_{[\text{grp}]}, [\text{SEP}], w_{[\text{g}]}_1, w_{[\text{g}]}_2, \dots, [\text{SEP}], w_{[\text{s}]}_1, w_{[\text{s}]}_2, \dots, [\text{SEP}], w_{[\text{ing}]}, [\text{END}]\} \quad (1)$$

We minimize the cross-entropy of the contextual probability of the correct token in our full linearized frame objective (of length N):

$$\mathcal{L} = -\frac{1}{N} \sum_i \log p_{\text{GPT}}(w_i | w_{0:i-1})$$

During training, no loss is incurred for lower-level variables with no values, i.e., variables that cannot take values due to earlier variable values (e.g., there is no targeted group for posts marked as non-offensive).

In our experiments we use pretrained versions of OpenAI’s GPT and GPT2 (Radford et al., 2018, 2019) for our model variants, named SBF-GPT₁ and SBF-GPT₂, respectively. While their architectures are similar (stack of Transformers), GPT was

Figure 2. Screenshot from Sap, “Social Bias Frames,” 5482, showing how we might quantify language, and how this might be processed by an algorithm. There are parallels with Asimov’s concept of ‘symbolic logic’ in the Foundation series.

However, problems can arise when we lose sight of the language we are quantifying and how our metaphors construct our social worlds. Metaphor is one of the key ways in which we not only add colour and vivacity to our language, making it more memorable and visual, but with which we can also mask or shade intent. Our ‘ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature’, from the ‘most mundane details’ through to the concepts that ‘structure what we perceive, how we get around in the world, and how we relate to other people’.³⁵ Metaphor is not merely an ‘optional, rhetorical flourish’ but our ‘most pervasive means of ordering our experience into conceptual systems’, and our ways of talking about the law and the economy are no different.³⁶ Metaphors, then, are those linguistic tropes to which we default when speaking about social phenomena—for law and economy are just that.³⁷ We might think of them as collections of social behaviour that we lump together for linguistic convenience. Neither exists as a tangible entity but can be observed in how we do, talk and think. For example, if I ask you to point to ‘the law’, you might point to a courtroom, a judge, or a statute. What you are pointing to is, respectively, a building, a person and a piece of paper. Similarly, if I ask you to point to ‘the economy’,

³⁵ Lakoff, *Metaphors We Live By*, 3.

³⁶ Bandes, “Empathy,” 361; Levit, “Legal Storytelling”; Lakoff, *Metaphors We Live By*, 25; Geary, *I is an Other*.

³⁷ Hari deploys metaphor without a second thought, referring at his trial to the destruction of the ‘fabric of society’. Asimov, *Foundation*, 30. On metaphor as a trope or, more accurately, a class inclusion statement, see Ortony, *Metaphor and Thought*.

you might point to a bank, some coins, or even cryptocurrency. But, again, you are pointing to a building, some small pieces of metal and digitally encoded data. And, yet, undeniably, a common meaning is communicated: we understand references to collections of social behaviours that we can identify in terms of their use, abuse and avoidance.³⁸ In line with George Lakoff's observation then, we might note that 'the law' and 'the economy' are 'ontological metaphors'; non-entities that we refer to every day, generally without further consideration.³⁹

But we build on this linguistically. We say that the 'law has made an ass of us all', or that 'the economy has taken a nosedive'. In personifying these primary ontological metaphors, we can identify a second layer of metaphor that sits atop the first, adding colour and vivacity. We can then note a third layer of metaphor that discusses the relationship of two ontological metaphors to each other; for example, that 'the law is embedded in society', or 'the economy must be re-embedded in society'. Here, embeddedness operates as a tertiary level ontological metaphor, visually describing the relationship between two primary, ontological metaphors. The embeddedness metaphor, then, implies that, in the first example, the law and society (yet another ontological metaphor) coexist in space and time. In the second example, we understand that the economy should be brought back to coexist in space and time with society.

And yet, peering through the layers of metaphor, we can appreciate that describing the relationship of two non-entities to each other is disingenuous. If both law and economy are social phenomena, coming into existence only when performed in an interaction (synchronous or asynchronous), it is nonsense to speak of one being 'embedded' within the other. Neither the law or the economy can exist without society. But our layers of metaphor can mask this, and herein lies the problem. When our linguistic and conceptual tools for law and economy are conceptually inconsistent (describing socially constructed phenomena as ontologically distinct entities), two things occur. First, embeddedness talk maintains the metaphorical fiction of the ontologically separate law, economy and society. Second, it further entrenches the hegemonic approaches within each sphere—doctrinal law and neoclassical economics, respectively. Our throwaway metaphor, then, not only masks the socially constructed nature of legal and economic phenomena, but elides more imaginative, sociological responses that we might explore to the crashes, crises and catastrophes that we are currently facing. To be clear, this is the result of the disconnect between our linguistic tools (metaphorically saying that the law is embedded in society) and what we do in practice (socially constructing the law through interactions). The problems arise when we lose sight of our metaphors, the concepts baked into them and the heavy lifting we are asking these concepts to do. Then, having foreclosed sociological responses to crashes, crises and catastrophes through incompatible conceptual and linguistic tools, we are left with no other option than to repeat the concepts that led us into those dilemmas in the first place.

But there is more: metaphors can operate generatively, as self-fulfilling prophecies, shaping not only our worlds but also ourselves. Coupled with insights from linguistics, psychology and sociology, this might give us cause to sit up and take note about the emergence of self-supervised learning in AI or the development of foundation models.⁴⁰ In 2006, researchers investigating attitudes to animal welfare across Europe made a mistake.⁴¹ In some of the telephone interviews, respondents were referred to as citizens. In other interviews, respondents were referred to as consumers. The survey did not set out to test the performativity of respondent labelling, but when the data were analysed and all other variables were controlled, analysts found that those respondents who had been identified as citizens had a preference for government to regulate animal welfare. In contrast, respondents identified as consumers tended to show a preference for more market-based interventions such as product labelling and price. While more research is needed, insights from the cognitive sciences indicate that how we talk about social institutions can affect our self-perception and social positionality.⁴² These insights indicate that switching up the framing can indeed have significant impacts on associated phenomena: how we perceive and perform them.

This is particularly relevant for law and economics as they have their own caricatures that embody how mainstream legal and economic theories understand the world. *Homo juridicus*, the reasonable man of law who contracts to realise his interests, is reproduced in mainstream legal language of property, contract and rights-based discourse. *Homo economicus*, the rational man of economics, contracts efficiently to maximise his utility, and we can note the language of 'productivity', 'efficiency', 'value' and 'exchange' as common terms that encode this caricature; terms that have seeped into wider discourse and shaped broader social narratives.⁴³ For if we only have the language of *homo economicus-juridicus*, we can only talk like *homo economicus-juridicus*, think like *homo economicus-juridicus* and act like *homo economicus-juridicus*. Asimov's cast of characters in the

³⁸ Perry, "Use, Abuse and Avoidance." Accordingly, we can identify the performance of 'law' as much through avoidance of the formal legal system as engagement with it.

³⁹ Lakoff, *Metaphors We Live By*.

⁴⁰ Schön, "Generative Metaphor."

⁴¹ Eurobarometer, Special EB; Law, "Seeing Like a Survey."

⁴² MacKenzie, "Constructing a Market"; Callon, "What Does it Mean?"; MacKenzie, *Do Economists Make Markets?*

⁴³ We can think of Giddens' 'double hermeneutic' here. Giddens, *The Constitution of Society*.

Foundation series exemplify this, trading in property rights and land ownership for extraction, discussing the economic value of a planet and the productivity of its workforce, and relying on legalistic treaties that encode post-Westphalian notions of sovereignty and statehood that have little relevance to the Galactic Empire. These characters cannot escape the mental models of their forebears, despite being in completely different time, space and cultural settings. But, like us, they are for the most part unaware of the ways in which their vocabularies and grammars, bound up in layers of metaphor, mould their social (inter)actions.

It would seem odd, then, to think that training an algorithm to reproduce our current ways of doing, talking and thinking could be anything but as deeply flawed as our current day-to-day conceptual and linguistic tools. And yet, that is precisely what we are doing. The following three sections take three sub-themes from Asimov's *Foundation* series to interrogate potential pitfalls of preserving our current frames in foundation models.

4. Framing Foundation Models (1): The Atomistic Individual

Asimov was not writing about a future society but about a collection of individuals who live in the future, some thousands of millennia hence. This is no coincidence. The individual actor is central to neoclassical economics and doctrinal law, and, as the hegemonic frames of the 1940s United States, these were transposed directly into Asimov's Galactic Empire. The result is a focus on individual accountability and the justification of any attempt to maximise individual utility. *Homo economicus*, in mainstream economics, and his lawyerly cousin *homo juridicus*, embody the bundle of features and desires typically assumed by each discipline in their calculations about human society. Neither are particularly moral souls, nor are they socially networked. The focus sits squarely on the individual actor or agent, although this is gradually being challenged by post-individualist scholarship.⁴⁴

Our current conceptual and linguistic tools, much as those deployed by Hardin and his co-protagonists in the *Foundation* series, assume the rational individual agent, and assume that he—for he is invariably a he—is solely responsible for maximising his own utility. Asimov's storytelling sets the perfect fictionalised render of an atomistic social conglomeration. We are not enlightened about the relationships between characters. Nor are we told much about their emotional or mental states. The characters are, for the most part, singular automata that deliver their set lines to the page and go about constructing the narrative. Neoclassical economics paints a similar picture of society, although more recent insights from behavioural economics and psychology have called out *homo economicus* for the 'rational fool' he was⁴⁵ and still is. For 'he' still exerts a surprising influence over how we do, talk and think about law and economy, and how they, in turn, think about us.

We might ask, then, if our assumptions about the atomistic individual, with their fixed preferences and perfect information, who generally disregards the wishes of others and the wider moral context in which he acts, should be preserved in future foundation models. What might we lose sight of in the process? Might this leave us with conceptual and linguistic tools that de-prioritise notions of community, networks and interactions?

5. Framing Foundation Models (2): Progress and Scientism

We see, through Hari Seldon's eyes, that the First Galactic Empire is dying, collapsing under the weight of a bloated and overbearing bureaucracy on the ruling planet of Trantor. Nevertheless, there is hope that the impending Dark Ages can be shortened.⁴⁶ Seldon gives an account of his psychohistorical analysis that not only calculates the present but also predicts the future. The quantification, mathematisation and comparison of social phenomena reflect the empirical, analytical and normative tools, respectively, of neoclassical economics.⁴⁷ Asimov might have been writing science fiction, and might even have believed in the possibility of quantifying the social realm, but the late twentieth and early twenty-first centuries saw this happen.⁴⁸ A

⁴⁴ See vulnerability theory, work on the Anthropocene, and sociological approaches such as economic sociology of law, relational work and actor-network theory.

⁴⁵ Thaler, "From Homo Economicus to Homo Sapiens"; Thaler, *Nudge*; Thaler, *Misbehaving*. Far from being a panacea, behavioural economics is the subject of significant scholarly challenge. See Frerichs, "Putting Behavioural Economics in its Place"; Hockley, "Do Nudges Work?"

⁴⁶ Through Hari Seldon's psychohistorical analysis of the downfall of the Galactic Empire due to its bureaucracy, we glimpse Asimov's distaste for big state. Technological innovation is counterposed with statehood and bureaucracy, echoing Hayekian economics. Reading back through a Polanyian lens by contrast, we can appreciate the fictional account of Hayek's self-regulating market economy. See Polanyi, *The Great Transformation*.

⁴⁷ See Perry-Kessaris, "Recycle, Reduce and Reflect"; Perry-Kessaris, "Prepare Your Indicators."

⁴⁸ The apocryphal apogee here is probably Gary Becker's economic analysis of children, but Daren Acemoglu's use of proxy indicators (death rates to infer quality of institutions) runs a close second. Nevertheless, the 'LLSV' literature, in setting an agenda for the quantification of

good example is the field of Law and Development, which saw the application of neoclassical economics theories and methods to the law, instrumentally designing legal reforms and interventions to realise economic goals.⁴⁹ This appropriation of the tools of neoclassical economics continued a trend in the twentieth century of ‘physics envy’ throughout the social sciences.⁵⁰ Economics aped the methods of the natural sciences, and law, in turn, aped economics. In seeking out quantifiable methodologies and eschewing the uncertainties and grey areas so central to the messy realities of social life, law sought the same reverence and respect accorded to the natural sciences and, latterly, economics.⁵¹

One result was the World Bank’s *Doing Business* indicators, which quantified aspects of the legal system of each country, assigning them a ranking that indicated how desirable to foreign investors each system was. But, as with all indicators, once you create the ‘law reform Olympics’ and start keeping score, everybody wants to win.⁵² Countries engaged in legal reform for no more reason than jumping in the World Bank rankings, in effect transforming those indicators from mere measurements into ‘technologies of governance’.⁵³ The fact that local communities on the ground might be disadvantaged in the process, or that no one had asked foreign investors what they really wanted, was beside the point. It seemed intuitive that a strong rule of law would enable economic development.

The *Doing Business* debacle of 2019, in which the whole programme came crashing down amid accusations of bullying, corruption and data manipulation, caused the World Bank to terminate the programme in 2021.⁵⁴ Nevertheless, a belief in the power of quantification, and of indicators to inform, enlighten and regulate, underpins the World Bank’s new *Business Enabling Environment*, which replaces the *Doing Business* series and will once again rank countries according to the features of their legal systems.⁵⁵ Indicators are just as rhetorical as words. And yet, having the appearance of objectivity and scientism, the analytical and normative undercurrents tend to be pushed further beyond sight and interrogation. Sometimes, though, having a clear answer can be more important than being right.⁵⁶ Neat, simplified, quantifiable answers can be as preferable for time-poor policymakers on planet Earth as they are for the Board on the planet of Terminus, as we see reflected in the Board’s engagement in symbolic logic to reduce down legal treaties to their core principles.

Asimov’s belief in the potential of scientific and technological progress to save humanity—ostensibly from itself—reflects a hegemonic viewpoint that persists in society some 80 years after the publication of the *Foundation* series.⁵⁷ Technological progress must undoubtedly play a part in our responses to the crashes, crises and catastrophes facing humanity. But social and behavioural shifts are also necessary, and to imagine and accomplish these, we need new ways of doing, talking and thinking about those ubiquitous social institutions that Asimov overlooked in his *Foundation* series. For example, thinking of the environmental catastrophe of climate change, we can appreciate that our conceptual and linguistic tools, in reproducing social patterns of behaviour that recreate *homo economicus-juridicus*, recreate the conditions that led us into the climate crisis in the first place.⁵⁸ *Homo economicus-juridicus* cares little for the fate of the planet—he simply does not have the conceptual or linguistic tools to imagine acting otherwise. And yet, these conceptual and linguistic tools comprise the data on which we are training future AI systems.

legal analysis, had an undeniably profound impact on legal scholarship. See Acemoglu, *Institutions as the Fundamental Cause of Long-Run Growth*; La Porta, “Legal Determinants of External Finance”; La Porta, “The Quality of Government”; La Porta, “The Economic Consequences of Legal Origins”; Becker, “An Economic Analysis of Fertility.”

⁴⁹ Perry-Kessaris, “Prepare Your Indicators.”

⁵⁰ Mirowski, “Do Economists Suffer from Physics Envy?”

⁵¹ Barber, “Absolutization of the Market”; Hirsch, “Clean Models vs. Dirty Hands.”

⁵² World Bank, *Rule of Law as a Goal of Development Policy*; World Bank, *Rule of Law and Development*; Taylor, “The Law Reform Olympics.”

⁵³ Davis, *Indicators as a Technology of Global Governance*.

⁵⁴ Machen, *Investigation of Data Irregularities*.

⁵⁵ See <https://www.worldbank.org/en/programs/business-enabling-environment>.

⁵⁶ Hirsch, “Clean Models vs. Dirty Hands.” See Part 2 of Asimov’s *Foundation* series. Curiously, the notion of a treaty with Trantor invokes notions of post-Westphalian state sovereignty that is clearly anathema to Asimov’s Galactic Empire.

⁵⁷ Asimov appears to be working with a linear concept of time, with progress moving in only one direction: a Fukuyama-esque reading of the human condition. John Gray, by contrast, urges us to think of human history as cyclical, ebbing and flowing. See Gray, *Heresies*; Fukuyama, *The End of History*.

⁵⁸ MacKenzie and Millo suggestively propose imagining their socially situated, morally networked cousin, *homo sociologicus*, who is freed from the assumptions, preferences and constraints of mainstream economics and law. MacKenzie, “Constructing a Market,” 141.

6. Framing Foundation Models (3): Natural Language and Its Limitations

If I ask you not to think of an elephant, you will, inevitably, think of an elephant.⁵⁹ This is entirely natural and unavoidable; it is simply how the human brain works. But there are profound implications here for how we talk and think about law and economy. As Lakoff and Johnson note, '[t]he people who get to impose their metaphors on the culture get to define what we consider to be true'.⁶⁰ This extends to legal and economic phenomena, as the discussion above set out with reference to embeddedness talk. So, then, is the economy embedded in society? Or has society become embedded in the economy? What about the law—can we say that the law is embedded in society?

Since the financial crisis of 2008, embeddedness has become a fashionable metaphor for describing the dysfunctional relationships between law, economy and society; notably non-entities that we can observe only through their social construction. Following the 2008 financial crisis, a consensus emerged that the three spheres had drifted apart, leaving some communities underserved and others elided altogether. Previously, a preference for scientism had seen the three spheres separate out into silos of research and activity, reflected in our conceptual and linguistic tools for each. We talk about 'the economy' *as if* it were a real thing.⁶¹ Similarly, we refer to the 'the law' *as if* it were an entity, using ontological metaphors to lump together useful collections of social behaviours for linguistic convenience. This fiction has the effect of masking the social construction of both law and economy and of maintaining the metaphorical fiction of the ontologically separate law, economy and society. As the above discussion noted, problems arise when we lose sight of the divergence between our linguistic fictions and our social facts, the concepts baked into these fictions and the heavy lifting we are asking them to do. At the same time, our linguistic and conceptual tools encode preferences and assumptions that have become invisible through ubiquity and familiarity but tend to reproduce neoclassical and doctrinal approaches. These, in turn, generate inequalities. For those whose interests align with the hegemonic values of dominant legal and economic frames, the outcome is privilege. In other words, our linguistic tools pick winners. For those whose interests do not align with the dominant rationalities of neoclassical law and economics, simply finding the vocabularies and grammars to express their needs can prove an exclusionary hurdle. Inequalities that result tend to remain largely ignored by AI de-biasing work, which, for the most part, remains focused on gendered and racialised language. By contrast, the inequalities resulting from our linguistic framing remain, like the frames that give rise to them, largely invisible.

For decades, while the political right has been asking us to think of elephants, the political left has been asking us *not* to think of elephants.⁶² While the political right has funded research posts, think tanks and university chairs, aligning rhetoric with their political ideals, the left has failed to respond adequately. In reframing the debate, both in the United States and, to a lesser extent, Europe, the right has devised such phrases as 'tax relief'.⁶³ This phrase implies that tax, being a burden, is something the government can offer relief from. Such shifts in framing tend to occur at the micro- and meso-levels, with one word here and one word there. But the net impact occurs at the macro-levels of social institutions and the meta-levels of our wider shared conceptual tools.⁶⁴

In simply arguing that 'tax relief is a bad idea', the political left repeats, re-entrenches and reinforces frames set by the political right, ironically strengthening the political right's view of the world through repetition. Instead, to challenge a frame, we must *reframe*, and this might involve replacing the phrase 'tax relief' with, for example, 'reduced investment'.⁶⁵ Challenging frames, then, demands a deep understanding of the linguistics and psychology behind the effectiveness of framing.⁶⁶ Reframing is not a quick fix but propagated over decades; the cumulative realisation of myriad micro-level shifts in language that add up to wider changes in how we perceive legal, economic, social and political phenomena. In the meantime, we are left with the conceptual and linguistic tools of *homo economicus-juridicus*, for the most part, like Asimov, unable to imagine anything different. As MacKenzie and Millo suggestively propose, what if we were able to imagine the socially networked, morally

⁵⁹ Lakoff, Don't Think of an Elephant.

⁶⁰ Lakoff, cited in Geary, *I is an Other*, Loc. 1950.

⁶¹ Kelsen, *On the Theory of Juridic Fictions*; Samuel, *Rethinking Legal Reasoning*.

⁶² Lakoff, Don't Think of an Elephant.

⁶³ Example from Lakoff, Don't Think of an Elephant.

⁶⁴ See Frerichs, *Studying Law, Economy, and Society*.

⁶⁵ Lakoff, Don't Think of an Elephant.

⁶⁶ The linguistic turn that had its impact on law in the 1980s and motivated much of the critical legal studies movement 'suffered from the dominance of analytic linguistic philosophy and the continuing drive to science'. While semiotics has offered support for positivist jurisprudence and legal semiotics came to be an exercise in mapping the linguistic presuppositions of legal theory, the role of rhetoric was overlooked. See Goodrich, "Rhetoric, Semiotics, Synasthetics," 153–154.

constructed notion of *homo sociologicus*?⁶⁷ What vocabulary might they deploy? And which interests and values might they be able to balance?

While *homo sociologicus* remains a pipe dream, and while our conceptual and linguistic tools are, to a large extent, those aligning with neoliberal voices and values, this might appear a curious time to be curating data that will train future foundation models. At issue is the wider lack of awareness about the preferences and biases baked into our language that fall below the level of discriminatory or biased discourse. If our frames are those of the political right and if these are being curated into foundation models, they are likely to be reproduced in all downstream applications, be these in healthcare, education, or legal services, perpetuating the assumptions of neoclassical economics and doctrinal law wherever AI is deployed. In the process, the voices, interests and values that align with hegemonic frames will enjoy privilege, while those that do not will continue to be sidelined or silenced. The net result, perpetuated through future AI systems, is the framing of our future actions, and our future selves, as *homines economici-juridici*.

7. Harm and Asimov's Zeroth Law: Language as Technology

Before being let loose on the public, all new doctors must swear the Hippocratic Oath: a pledge of ethics in which they promise, above all else, to do no harm. Similarly, Asimov's Zeroth Law of robotics states that a robot may not harm humanity or, by inaction, allow humanity to come to harm. While we can appreciate that explicitly discriminatory language can cause harm, to what extent can we say the same for inequalities perpetuated by our framing of legal and economic phenomena?

In many contexts, 'machine learning has been shown to contribute to, and potentially amplify, societal inequality', and foundation models 'may extend this trend'.⁶⁸ And at each step, the conceptual and linguistic tools that encode hegemonic assumptions and preferences become further entrenched, more ubiquitous and less available to challenge. But there is startlingly little awareness of the use of Transformers such as BERT and GPT-3 and how they might be shaping societies and ourselves in the process. For example, were you aware of Google's upgrade to BERT in 2019?⁶⁹ Have you consented for your data to be used to train future generations of foundation models?

While de-biasing initiatives tend to focus on racialised and gendered language, and work on implicit bias is increasing, our framing remains largely unnoticed, as little more than tacit, taken-for-granted background noise of social institutions.⁷⁰ In preserving these, we risk perpetuating structural inequalities that can produce tangible harms for sections of the population. At what point, then, might an AI system be compelled to intervene to adjust the damaging effects of the conceptual and linguistic tools baked into its foundation model? And if it determines the harm too great, then what?

A starting assumption here is the acknowledgement that entrenched and systemic socio-economic disadvantage might constitute harm.⁷¹ Of course, neoliberal frames are able to justify (to human minds at least) the inequalities they produce according to their own internal logics. So, a lack of access to justice, for example, might be explained as the result of an individual's inability to maximise their own utility, meaning that they do not have sufficient funds to hire a lawyer.⁷² Might an AI system agree, or take a more critical approach?

As previous sections noted, any challenge to existing frames depends on an acknowledgement that our conceptual and linguistic tools are not immutable rules of nature. Yet even the framing of AI as 'natural' language processing forecloses the idea that we might consider alternatives. Our current linguistic tools are the product of careful curation and design over generations by those seeking to achieve political ends, namely the minimisation of the state, prioritisation of the individual and centrality of the self-regulating market. In short, the language on which BERT was trained is no more 'natural' than space travel or the internet.⁷³

⁶⁷ MacKenzie, "Constructing a Market," 141.

⁶⁸ Bommasani, "On the Opportunities," 19; Blodgett, "Language (Technology) is Power"; Bender, "On the Dangers of Stochastic Parrots"; Sap, "Social Bias Frames."

⁶⁹ Uszkoreit, "Transformer"; Nayak, "Understanding Searches."

⁷⁰ Sap, "Social Bias Frames"; Liang, "Towards Understanding."

⁷¹ Currently debates are framed in terms of socio-economic disadvantage and inequalities. See Layard, "Vehicles for Justice."

⁷² Socio-economic justice is critical, given the overlaps between procedural and substantive inequalities. However, legal duties to address this remain inadequate. See Layard, "Vehicles for Justice." Causation, along with standing, is problematic, and parallels can be seen in climate litigation where legal tools to address clear dangers appear lacking. Nevertheless, the application of AI to legal fora promises to improve access to information and justice by reducing costs. See Dale, "Law and Word Order."

⁷³ Note that the data on which foundation models are trained are curated. Humans determine what is important and what is not, what is relevant and what is not, thus (normatively) determining the language on which BERT and its descendants will be trained.

Yet in describing the language BERT and its peers can process as ‘natural’, we elide the normative and analytical choices that created this language. In losing sight of the rhetoric of all language, we simultaneously risk placing notions like *homo sociologicus* permanently beyond imagination.

Additionally, this elides the contingencies and contestation that occurs in language about which ideas we realise (or aspire to), which voices we hear and which interests we highlight. Preferences for the self-regulating market, and even the opposition of market and state as a binary, are bound up with and reproduced in our wider linguistic tools, becoming encoded in the data traces we leave every time we interact online. Like the indicators noted above in section 5, we might think of language *as* technology, with the potential to regulate interactions and discipline those who dare to deviate.⁷⁴ For example, have you ever altered your spelling or grammar on a computer simply to make those little red squiggles disappear? In homogenising language and in labelling it ‘natural’, the normative binary (is your spelling right or wrong) disciplines deviation, marking out divergence as other. In doing so, pluralistic voices, interests and values are minimised.

We undoubtedly need a greater awareness of how we do, talk and think about legal and economic phenomena—that is, how we frame law and economy. This paper has suggested a research agenda for AI developers and socio-legal researchers to look beyond social and representational biases to the deep, subconscious structures and institutions that we perform through our linguistic tools. Our hegemonic conceptual and linguistic tools have intrinsic limitations, prioritising some interests while minimising others. Challenging this status quo demands a greater appreciation of how we frame—and are framed by—law and economy on a daily basis. In the meantime, we risk sleepwalking into preserving our current conceptual and linguistic tools, with their baked-in assumptions and preferences, in digital aspic—perhaps forever. In his *Foundation* series, Asimov shows us what might happen when we lose sight of the social construction of law and economy. Our risks are greater though. While Asimov penned novels that occasionally jar, we risk the potential entrenchment of inequalities in future AI systems. In short, the way we talk matters.

Acknowledgements

Grateful thanks to the UK’s Economic and Social Research Council (ESRC) SeNSS (South East Network for Social Sciences) for funding. I would also like to thank the three reviewers for their detailed and helpful comments and suggestions on the paper, and the copyeditor for such useful suggestions and edits. All errors are my own.

⁷⁴ See Chu, “Overcoming Built-In Prejudices.”

Bibliography

- Acemoglu, Daron, Simon Johnson, and James Robinson. *Institutions as the Fundamental Cause of Long-Run Growth*. National Bureau of Economic Research Working Paper Series No. 10481, May 2004.
- Asimov, Isaac. *Foundation*. Kindle. HarperVoyager, 2018.
- Bandes, Susan. "Empathy, Narrative, and Victim Impact Statements." *University of Chicago Law Review* 63, no 2 (1996): 361–412.
- Barber, Bernard. "Absolutization of the Market: Some Notes on How We Got from There to Here." In *Markets and Morals*, edited by Gerald Dworkin, Gordon Bermant, and Peter G. Brown, 15–32. Washington DC: Hemisphere, 1977.
- Becker, Gary. "An Economic Analysis of Fertility." In *Demographic and Economic Change in Developed Countries*, 209–240. National Bureau of Economic Research, Columbia University Press, 1960.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models be Too Big?" FAccT '21, March 3–10, 2021, Virtual Event, Canada ACM, March 2021. <https://doi.org/10.1145/3442188.3445922>
- Bernstein, Rebecca, Antonia Layard, Martin Maudsley, and Hilary Ramsden. "There is No Local Here, Love." In *After Urban Regeneration: Communities, Policy and Place*, edited by Dave O'Brien and Peter Matthews, 95–110. Policy Press, University of Bristol, 2015.
- Blodgett, Su Lin, Solon Barocas, Hal Daumé III, and Hanna Wallach. "Language (Technology) is Power: A Critical Survey of 'Bias' in NLP." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5454–5476. Association for Computational Linguistics, July 5, 2020. <https://doi.org/10.18653/v1/2020.acl-main.485>
- Bommasani, Rishi, and Percy Liang. "On the Opportunities and Risks of Foundation Models." Centre for Research on Foundation Models (CRFM), Stanford Institute for Human-Centered Artificial Intelligence, 2021.
- Callon, Michel. "What Does it Mean to Say That Economics is Performative?" In *Do Economists Make Markets?: On the Performativity of Economics*, edited by Donald MacKenzie, Fabian Muniesa, and Lucia Siu, 309–355. Princeton University Press, 2008.
- Chu, Ann Gillian. "Overcoming Built-In Prejudices in Proofreading Apps." *SRHE News Blog* (blog), May 24, 2022. <https://srheblog.com/2022/05/24/overcoming-built-in-prejudices-in-proofreading-apps/>
- Coase, Ronald. "The Problem of Social Cost." *Journal of Law and Economics* 3 (1960): 2–144. <https://doi.org/10.1086/466560>
- Dale, Robert. "Law and Word Order: NLP in Legal Tech." *Natural Language Engineering* 25, no 1 (December 19, 2018): 211–217. <https://doi.org/10.1017/S1351324918000475>
- Davis, Kevin E., Benedict Kingsbury, and Sally Engle Merry. *Indicators as a Technology of Global Governance*. NYU Law and Economics Research Paper No. 10-13, NYU School of Law, Public Law Research Paper No. 10-26, April 2010.
- Devlin, Jacob, and Ming-Wei Chang. "Open-Sourcing BERT: State-of-the-Art Pre-Training for Natural Language Processing." *Google AI Blog* (blog), November 2, 2018. <https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html>
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "BERT: Pre-Training of Deep Bi-Directional Transformers for Language Understanding." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186. Minneapolis, MN: Association for Computational Linguistics, 2019.
- Eurobarometer. *Special EB on Attitudes of European Towards Animal Welfare* (Ref 442), March 2016. <https://europa.eu/eurobarometer/surveys/detail/2096>
- European Commission. *White Paper on Artificial Intelligence – A European Approach to Excellence and Trust*. European Commission Brussels, 19.2.2020 COM(2020) 65 final, February 19, 2020. https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
- Frerichs, Sabine. "Re-embedding Neoliberal Constitutionalism: A Polanyian Case for the Economic Sociology of Law." In *Karl Polanyi, Globalisation, and the Potential of Law in Transnational Markets*, edited by Christian Joerges and Josef Falke, 65–84. Oxford, UK: Hart, 2011.
- Frerichs, Sabine. *Studying Law, Economy, and Society: A Short History of Socio-Legal Thinking*. Helsinki Legal Studies Research Paper No. 19, University of Helsinki, March 15, 2012.
- Frerichs, Sabine. "Putting Behavioural Economics in its Place: The New Realism of Law, Economics, and Psychology and its Alternatives" *Northern Ireland Legal Quarterly* 72, no 4 (2020): 651–681. <https://doi.org/10.53386/nllq.v72i4.920>
- Fukuyama, Francis. *The End of History and the Last Man*. Penguin, 1993.
- Geary, James. *I is an Other: The Secret Life of Metaphor and How It Shapes the Way We See the World*. Harper Collins eBook, 2011.
- Giddens, Anthony. *The Constitution of Society: Outline of the Theory of Structuration*. Oxford, UK: Polity, 1984.

- Goodrich, Peter. "Rhetoric, Semiotics, Synaesthetics." In *Research Handbook on Critical Legal Theory*, edited by Emiliios Christodoulidis, Ruth Dukes, and Marco Goldoni, 151-165. Research Handbooks in Legal Theory Series. Elgar, 2019.
- Gray, John. *Heresies: Against Progress and Other Illusions*. Granta, 2004.
- Harraway, Donna. *Staying with the Trouble: Making Kin in the Chthulucene*. Duke University Press, 2016.
- Hirsch, Paul, Stuart Michaels, and Ray Friedman. "Clean Models vs. Dirty Hands: Why Economics is Different from Sociology." In *Structures of Capital: The Social Organization of the Economy*, edited by Sharon Zukin and Paul DiMaggio, 39-57. Cambridge University Press, 1990.
- Hockley, Tony. "Do Nudges Work? Debate Over the Effectiveness of 'Nudge' Provides a Salutory Lesson on the Influence of Social Science." *LSE Impact Blog*, July 29, 2022. <https://blogs.lse.ac.uk/impactofsocialsciences/2022/07/29/do-nudges-work-debate-over-the-effectiveness-of-nudge-provides-a-salutory-lesson-on-the-influence-of-social-science/>
- Kelsen, Hans, and Christoph Kletzer. "On the Theory of Juridic Fictions. With Special Consideration of Vaihinger's Philosophy of the As-If." In *Legal Fictions in Theory and Practice*, edited by Maksymilian Del Mar and William Twining, 1-20. Switzerland: Springer International Publishing, 2015.
- Kramer, Adam. "Common Sense Principles of Contract Interpretation (and How We've Been Using Them All Along)." *Oxford Journal of Legal Studies* 23, no 2 (2003): 173-196. <https://doi.org/10.1093/ojls/23.2.173>
- La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. "Legal Determinants of External Finance." *Journal of Finance* 52, no 3 (1997): 1131-1150. <https://doi.org/10.1111/j.1540-6261.1997.tb02727.x>
- La Porta, Rafael, Florencio Lopez-de-Silanes, and Andrei Shleifer. "The Economic Consequences of Legal Origins." *Journal of Economic Literature* 46, no 2 (2008): 285-332.
- La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. "The Quality of Government." *Journal of Law, Economics and Organization* 15, no 1 (1999): 222-279. <https://doi.org/10.1093/jleo/15.1.222>
- Lakoff, George. *Don't Think of an Elephant: Know Your Values and Frame the Debate*. Kindle. Chelsea Green Publishing, 2014.
- Lakoff, George, and Mark Johnson. *Metaphors We Live By*. University of Chicago Press, 1980.
- Law, John. "Seeing Like a Survey." *Cultural Sociology* 3, no 2 (2009): 239-256. <https://doi.org/10.1177/1749975509105533>
- Layard, Antonia. "Vehicles for Justice: Buses and Advancement." *Journal of Law and Society* 49, no 2 (2022): 406-429. <https://doi.org/10.1111/jols.12368>
- Levit, Nancy. "Legal Storytelling: The Theory and the Practice - Reflective Writing across the Curriculum." *The Journal of the Legal Writing Institute* 15 (2009): 259.
- Liang, Paul Pu, Louis-Philippe Morency, and Ruslan Salakhutdinov. "Towards Understanding and Mitigating Social Biases in Language Models." In *Proceedings of the 38th International Conference on Machine Learning, PMLR 139*, 6565-6576. 2021.
- Machen, Ronald C., Matthew T. Jones, George P. Varghese, and Emily L. Stark. *Investigation of Data Irregularities in Doing Business 2018 and Doing Business 2020 Investigation Findings and Report to the Board of Executive Directors*. Wilmer Cutler Pickering Hale and Dorr LLP, September 15, 2021.
- MacKenzie, Donald, and Yuval Millo. "Constructing a Market, Performing Theory: The Historical Sociology of a Financial Derivatives Exchange." *American Journal of Sociology* 109, no 1 (July 2003): 107-145. <https://doi.org/10.1086/374404>
- MacKenzie, Donald, Fabian Muniesa, and Lucia Siu. *Do Economists Make Markets? On the Performativity of Economics*. Princeton University Press, 2008.
- McCloskey, Donald. "The Rhetoric of Law and Economics." *Michigan Law Review* 86, no 4 (February 1988): 752-767.
- Mirowski, Philip. "Do Economists Suffer from Physics Envy?" *Finnish Economic Papers* 5, no 1 (Spring 1992): 61-68.
- Nayak, Pandu. "Understanding Searches Better than Ever Before." Google. *The Keyword* (blog), October 25, 2019. <https://blog.google/products/search/search-language-understanding-bert/>
- Norman, Jana. *Posthuman Legal Subjectivity: Reimagining the Human in the Anthropocene*. Routledge, 2022.
- OECD. "Recommendation of the Council on Artificial Intelligence." OECD/LEGAL/0449. 2022.
- Ortony, Andrew. *Metaphor and Thought*. 2nd ed. Cambridge University Press, 2008.
- Perry, Amanda J. "Use, Abuse and Avoidance: Foreign Investors and the Legal System in Bangalore." *Asia Pacific Law Review* 12, no 2 (2004): 161-189. <https://doi.org/10.1080/18758444.2004.11788134>
- Perry-Kessaris, Amanda J. "Prepare Your Indicators: Economics Imperialism on the Shores of Law and Development." *International Journal of Law in Context* 7, no 4 (2011): 401-421. <https://doi.org/10.1017/S174455231100022X>
- Perry-Kessaris, Amanda J. "Recycle, Reduce and Reflect: Information Overload and Knowledge Deficit in the Field of Foreign Investment and the Law." *Journal of Law and Society* 35, no 1 (2008): 65-75. <https://doi.org/10.1111/j.1467-6478.2008.00425.x>
- Polanyi, Karl. *The Great Transformation: The Political and Economic Origins of Our Time*. 2nd ed. Princeton, NJ: Beacon Press, 1944.
- Samuel, Geoffrey. "Causation and Casuistic Reasoning in Roman and Later Civil Law." n.d.
- Samuel, Geoffrey. "Classification of Obligations and the Impact of Constructivist Epistemologies." *Legal Studies* 17 (1997): 448-482. <https://doi.org/10.1111/j.1748-121X.1997.tb00416.x>

- Samuel, Geoffrey. "Have There Been Scientific Revolutions in Law?" *Journal of Comparative Law* 11, no 2 (2016): 186–213.
- Samuel, Geoffrey. *Rethinking Legal Reasoning*. Edward Elgar, 2018.
- Samuel, Geoffrey. "The Epistemological Challenge: Does Law Exist?" In *Rethinking Comparative Law*, Edward Elgar, 2021
- Sap, Maarten, Saadia Gabriel, Lianhui Qin, Dan Jurafsky, Noah A Smith, and Yejin Choi. "Social Bias Frames: Reasoning about Social and Power Implications of Language." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5477–5490. Association for Computational Linguistics, 2020.
- Schön, Donald A. "Generative Metaphor: A Perspective on Problem-Setting in Social Policy." In *Metaphor and Thought*, edited by Andrew Ortony, 2nd ed., 137–163. Cambridge University Press, 2012.
- Shariatmadari, David. *Don't Believe a Word; The Surprising Truth About Language*. Kindle. Weidenfeld & Nicolson, 2019.
- Sharples, Mike. "New AI Tools that can Write Students Essays Require Educators to Rethink Teaching and Assessment." *LSE Impact Blog* (blog), May 17, 2022. <https://blogs.lse.ac.uk/impactofsocialsciences/2022/05/17/new-ai-tools-that-can-write-student-essays-require-educators-to-rethink-teaching-and-assessment/>
- Taylor, Veronica L. "The Law Reform Olympics: Measuring the Effects of Law Reform in Transition Economies." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, January 15, 2005.
- Thaler, Richard H. "From Homo Economicus to Homo Sapiens." *Journal of Economic Perspectives* 14, no 1 (Winter 2000): 133–141.
- Thaler, Richard H. *Misbehaving: The Making of Behavioural Economics*. Penguin, 2016.
- Thaler, Richard H., and Cass R. Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Penguin, 2009.
- Tranter, Kieran. "The Law and Technology Enterprise: Uncovering the Template to Legal Scholarship on Technology." *Law, Innovation and Technology* 3, no 1 (2011): 31–83. <https://doi.org/10.5235/175799611796399830>
- Tranter, Kieran. "The Laws of Technology and the Technology of Law." *Griffith Law Review* 20, no 4 (2011): 753–762. <https://doi.org/10.1080/10383441.2011.10854719>
- Tranter, Kieran. "The Speculative Jurisdiction: The Science Fictionality of Law and Technology." *Griffith Law Review* 20, no 4 (2011): 817–850. <https://doi.org/10.1080/10383441.2011.10854722>
- Travis, Mitchell. "Making Space: Law and Science Fiction." *Law & Literature* 23, no 2 (2011): 241–261. <https://doi.org/10.1525/lal.2011.23.2.241>
- Uszkoreit, Jacob. "Transformer: A Novel Neural Network Architecture for Language Understanding." *Google AI Blog* (blog), August 31, 2017. <https://ai.googleblog.com/2017/08/transformer-novel-neural-network.html>
- World Bank. *Rule of Law and Development*. The World Bank, n.d.
- World Bank. *Rule of Law as a Goal of Development Policy*. The World Bank, n.d.