Informational Sovereignty

A New Framework for AI Regulation

*Nicky Gillibrand & Chris Draper**

Abstract

Discussions of digital sovereignty predominate artificial intelligence (AI) discourses. However, digital sovereignty has been unable to effectively respond to longstanding concerns regarding the use of AI. These challenges include systemic bias, transparency/accountability and intellectual property infringement/theft. The authors posit an alternative framework – informational sovereignty – encouraging a recalibration of how technological sovereignty is viewed. Through this model emphasis is placed on respecting jurisdictional boundaries and jurisdictionally appropriate information sources to result in representative outcomes for communities rather than the traditional focus on where the data is held and system reliability that has thus far been the subject of much high-profile litigation. The article therefore sets out a quadripartite model of informational sovereignty encompassing concerns regarding population, territory, recognition and regulation of borders, before analysing the place of informational sovereignty in future iterations of AI regulation, including its practical applicability in the European Union Artificial Intelligence Act (EU AI Act).

Keywords: validation of the legal underpinnings of systems, large language models, sovereignty, rule of law, jurisdiction bias, ai risk, ai transparency, ai accountability, pragmatics of adoption, informational sovereignty, digital sovereignty.

1 Introduction

How do we best regulate artificial intelligence (AI)? This is the central question that has predominated AI discourses when considering its societal value. Very few sectors will be left unaffected by the proliferation of accessible and simple tools that can synthesize what would appear to be weeks of research into a response that takes less than 30 seconds to generate. Although ostensibly appearing as a positive addition to knowledge when used correctly, significant challenges arise when reviewing the source of the datasets in terms of adherence to legal sovereignty, rule of law and quality of outcome.

Existing frameworks used for the basis of regulating AI centre around the concept of digital sovereignty and data protection as the focal point. Here, it is

* Nicky Gillibrand, University College Dublin. Chris Draper, Indiana University.

doi: 10.5553/IJODR/235250022023010001007

posited that a more holistic approach is one of *informational sovereignty* that directly addresses the challenges of AI dataset development, including bias, theft and transparency, in the process removing the ability for AI to be used as a liability shield. In doing so, we shift the focus from data to the information itself being the priority. To better represent the challenges posed by LLM tools, a novel quadripartite theory of informational sovereignty is offered, encompassing concerns regarding population, territory, recognition and regulation of borders.

Although informational sovereignty is novel insofar as a conceptualization, it is not without its practical grounding with the European Union Artificial Intelligence Act (EU AI Act) alluding to the importance of being mindful of extra-jurisdictional contributions to LLM training datasets that exist outside of the generally accepted norm of legal sovereignty and, as a result, skewing the application of matter to be outside the acceptable boundaries of the impacted community.

This article will therefore examine the current state of AI including recent litigation that displays its impact on sectors from law to commerce. The legitimacy of different data sources will be reviewed, particularly in light of shifting the onus from system reliability inherent in digital sovereignty to the regulation of the sources of information, for instance, lawyers working within that jurisdiction as members of professional regulatory bodies. Finally, it will be discussed how informational sovereignty can serve as a framework for future iterations of AI regulation to act as a benchmark of striking a balance between economic concerns regarding innovation and constitutional concerns such as the rule of law and fundamental rights to best serve our communities.

2 The Current State of AI

Due in no small part to the rising accessibility and the proliferation of use of AI, considerable literature on the topic continues to emerge at a rapid pace. AI itself is becoming increasingly newsworthy, particularly in the wake of ChatGPT's rise to prominence and its related controversies such as its ban in Italy,[1] among other notable headlines such as its ability to pass the Uniform Bar Examination in the US.[2] While much of the existing literature on the role of AI in the law to this point is optimistic that it may eventually have a positive impact on access to justice, enabling those who cannot afford a legal professional to use accessible technology that can technically attain the level of a trained professional,[3] with some going as far as to state that AI is a prerequisite for social justice.[4] A significant volume of

[1] H. Ruschemeier, "Squaring the Circle", https://verfassungsblog.de/squaring-the-circle/ (last accessed 8 May 2023).

[2] D. Cassens Weiss, "Latest Version of ChatGPT Aces Bar Exam with Score Nearing 90th Percentile", *ABA Journal*, https://www.abajournal.com/web/article/latest-version-of-chatgpt-aces-the-bar-exam-with-score-in-90th-percentile (last accessed 9 May 2023).

[3] J. Villasenor, "How AI Will Revolutionize the Practice of Law", https://www.brookings.edu/blog/techtank/2023/03/20/how-ai-will-revolutionize-the-practice-of-law/ (last accessed 8 May 2023).

[4] A. Buccella, "'AI for All' Is a Matter of Social Justice", *AI and Ethics* (2022).

work also puts forward the idea that we should remain cautious of the sudden rise of AI usage, with it holding the potential to exacerbate structural inequities inherent in society.[5] This is due to the likelihood of the newest LegalTech remaining cost-prohibitive to underserved members of the public, while high-street lawyers representing less wealthy members of society will also be squeezed by LegalTech;[6] therefore, a significant gulf will remain between profit and not-for-profit AI systems.[7]

Failure to regulate the use of AI in the legal profession remains another significant problem, with jurisdictions focusing primarily on the regulation of AI in the case of autonomous vehicles and for the use of national defence.[8] The value of government regulation cannot be understated as even the CEO of OpenAI urged Members of Congress to legislate AI regulation displaying that even those creating AI products understand that if left unchecked, AI can create a large-scale societal danger.[9] Meanwhile, the EU has made concerted efforts to create AI regulation through the AI Act 2021, which will explored further in due course. The AI Act arose out of the Digital Europe Programme 2021 that sought to strengthen digital sovereignty through investing in AI that adheres to ethical standards and trustworthiness, with its legislation planned to set the global standard.[10] As represented by the US-EU comparator, AI regulations vary significantly between jurisdictions, despite the very real risks it represents worldwide.

While the bulk of the literature focuses on how a failure to properly regulate AI can impact the public at an individual level, there is considerably less on the wider impact to the state's jurisdiction and constitutional architecture. Of these, it is said to be pivotally important for the societies to have control over the source code of the AI datasets before it is ceded to private tech corporations who may ultimately regulate AI and subsequently impact the rule of law.[11] The rule of law is said to be challenged in three ways by AI: the aforementioned blurring of the private-public regulatory sphere on fundamental rights; the subsequent failure to demarcate legal certainty within this framework; and the lack of transparency and accountability

[5] H. Kanu, "Artificial Intelligence Poised to Hinder, Not Help Access to Justice", https://www.reuters.com/legal/transactional/artificial-intelligence-poised-hinder-not-help-access-justice-2023-04-25/ (last accessed 8 May 2023).

[6] A. Telang, "The Promise and Peril of AI Legal Services to Equalize Justice", https://jolt.law.harvard.edu/digest/the-promise-and-peril-of-ai-legal-services-to-equalize-justice (last accessed 8 May 2023).

[7] A. Reichman and G. Sartor, "Algorithms and Regulation", in *Constitutional Challenges in the Algorithmic Society*", eds. H-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor and G. De Gregorio (Cambridge University Press, 2022) p. 157.

[8] Law Library: Library of Congress, "Regulation of Artificial Intelligence in Selected Jurisdictions", https://tile.loc.gov/storage-services/service/ll/llglrd/2019668143/2019668143.pdf (last accessed 8 May 2023) pp. 1-2.

[9] *The Guardian*, "The EU Is Leading the Way on AI laws. The US Is Still Playing Catch-up", https://www.theguardian.com/technology/2023/jun/13/artificial-intelligence-us-regulation (last accessed 14 July 2023).

[10] European Parliament, "Shaping the Digital Transformation: EU Strategy Explained", https://www.europarl.europa.eu/news/en/headlines/society/20210414STO02010/shaping-the-digital-transformation-eu-strategy-explained (last accessed 12 July 2023).

[11] S. Rosengrun, "Why AI Is a Threat to the Rule of Law", *Digital Society* (2022), 1(10) p. 10.

doi: 10.5553/IJODR/235250022023010001007

of the mechanisms of decision-making.[12] By challenging the rule of law, one challenges potentially centuries of constitutional tradition that forms the basis of civilized society. As such, the implications may be widespread, with theorists stating that there requires a substantive reconfiguration of the relationship between law, technology and legal culture in order to incorporate algorithmic rationality.[13] If, therefore, LLMs gain a significant role in the legal profession and fail to be representative of legal culture, synonymous to some with the rule of law,[14] this can result in declining public sentiment towards the legal system more generally, which is insurmountably detrimental to the wider functioning of the state.

These discourses are also significantly related to our concerns regarding the impact of LLMs and their datasets on jurisdictional sovereignty which remain largely unaddressed. It is, therefore, of utmost importance to exercise caution when considering the role of LLM tools in the law and consider any substantive advancement in its capacity through the lens of sovereignty discourses. Viewing issues of AI standards, controls and regulation – through the lens of sovereignty, both of the traditional and digital variety – entails a re-examination of the human aspects of these tools which make them simultaneously valuable and unprecedentedly dangerous and is, therefore, the most reasonable approach for ensuring the necessary representation that delivers appropriate outcomes for jurisdictions.

2.1 Perspectives from Case Law

Although much of the academic commentary on AI stems from a place of hope, the practical application has displayed the significant risks associated with greater use of AI. From law to finance, the use of AI in its current form has resulted in lawsuits that display its inappropriateness in its current form to be used as a reliable tool.

In the case of Mata *v*. Avianca Inc.,[15] a brief filed with the court by the plaintiff's lawyer contained multiple citations that were invented by ChatGPT by combining fragments of real training data. When the lawyer in question, who now regards ChatGPT as "unreliable",[16] engaged the program for this research, he asked it to

[12] O. Pollicino and G. De Gregorio, "Constitutional Law in the Algorithmic Society", in *Constitutional Challenges in the Algorithmic Society*, eds H-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor and G. De Gregorio (Cambridge University Press, 2022) p. 7.

[13] M. Catanzariti, "Algorithmic Law: Law Production by Data or Data Production by Law?", in *Constitutional Challenges in the Algorithmic Society*, eds H-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor and G. De Gregorio (Cambridge University Press, 2022) p. 89.

[14] R. Michaels, "Legal Culture", https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=3012&context=faculty_scholarship (p. 1).

[15] "Mata v. Avianca, Inc., No. 1:2022cv01461 – Document 54 (S.D.N.Y. 2023)", https://law.justia.com/cases/federal/district-courts/new-york/nysdce/1:2022cv01461/575368/54/.

[16] *New York Times*, "A Man Sued Avianca Airline. His Lawyer Used ChatGPT", www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html (last accessed 12 July 2023).

doi: 10.5553/IJODR/235250022023010001007

verify the cases as legitimate displaying peripheral concerns about its ability to lie,[17] further removing its legitimacy as a competent and reliable tool.

Another example of ChatGPT's propensity to fabricate information, sometimes to an extremely damaging extent, is represented by the instance of Australian law professor Jonathan Turley's name wrongly appearing on a generated list of legal scholars that had sexually harassed somebody.[18] Once again, similar to Mata case, ChatGPT committed another error by citing a non-existent *Washington Post* article from 2018 with significant detail as its source. Where ChatGPT presents these falsehoods as statements of fact, significant harm can arise to somebody's professional and personal life; this places everybody without discrimination as a potential subject of its damaging false claims.

A prominent theme that re-arises in AI discourses is that of intellectual property infringement. Oftentimes, an author or artist is not consulted when their work is trained into an AI's dataset. One of the most notable instances of this is Sarah Silverman's and other authors' claim that their books were summarized by using illegal shadow libraries,[19] as suggested in a paper by Meta AI.[20] The authors are currently in the process of suing for copyright infringement. As such, AI has displayed itself, though its training data, to act outside the boundaries of intellectual property rights.

In the public sector, within political and financial realm, the Dutch government employed the use of AI to take stock of childcare benefit applications. Although it is not the only case of AI being used in the realm of taxes with the IRS contracting machine learning firm Brillient to automate its documentation processes,[21] it represented a considerable scandal as applications from ethnic minority families were significantly more likely to be flagged as fraudulent and subsequently denied benefits.[22] The Dutch tax scandal, or Kinderopvangtoeslagaffaire, was indicative of the underlying institutional racism within the Dutch tax authority which forced over 20,000 into economic distress as a result of its racial bias. Bias, in addition to theft and transparency, represents the key tenets of what users of AI should remain

[17] Bloomberg, "ChatGPT Can Lie but It's Only Imitating Humans", https://www.bloomberg.com/opinion/articles/2023-03-19/chatgpt-can-lie-but-it-s-only-imitating-humans?leadSource=uverify%20wall (last accessed 14 July 2023).

[18] *The Washington Post*, "ChatGPT Invented a Sexual Harassment Scandal and Named a Real Law Prof as the Accused", https://www.washingtonpost.com/technology/2023/04/05/chatgpt-lies/ (last accessed 12 July 2023).

[19] *The Guardian*, "Sarah Silverman Sues OpenAI and Meta Claiming AI Training Infringed Copying", www.theguardian.com/technology/2023/jul/10/sarah-silverman-sues-openai-meta-copyright-infringement (last accessed 12 July 2023).

[20] H. Touvron et al., "LLaMA: Open and Efficient Foundation Language Models", Meta AI https://arxiv.org/pdf/2302.13971.pdf (last accessed 12 July 2023).

[21] NextGov, "IRS Awards $70 Million Contract for Digital Modernization", https://www.nextgov.com/artificial-intelligence/2022/04/irs-awards-70-million-contract-digital-modernization/363938/ (last accessed 12 July 2023).

[22] Bloomberg Tax, "We Can All Learn a Thing or Two from the Dutch AI Tax Scandal", https://news.bloombergtax.com/tax-insights-and-commentary/we-can-all-learn-a-thing-or-two-from-the-dutch-ai-tax-scandal (last accessed 12 July 2023).

doi: 10.5553/IJODR/235250022023010001007

wary of, particularly until a new system is enacted which sufficiently avoids these present circumstances.

3 Deconstructing AI: The Core Issues

Bias, as displayed in the Dutch tax scandal, can create significant structural inequality. These biases arise as a result of the content of the training data which perpetuates the bias found in the decision-making of those behind the datasets and unrepresentative data sampling. However, without appropriate examination of the data and its implicit and explicit biases, it can be challenging to determine the cause of the bias. Regardless of this, the AI biases left unchecked for longer will cause a further perpetuation of this at great cost to the groups who are the victim of this bias.[23]

As such, one of the most pertinent issues surrounding AI datasets is that poorly constructed AI datasets may provide incorrect information and give rise to considerable bias in decision-making,[24] infringing the rights of individuals and groups with certain characteristics.[25] If used in sentencing, such bias can ultimately result in a deprivation of one's liberty based on these characteristics.[26] As such, warnings have arisen that AI datasets must not only be bigger but also be of better quality, which is generally described as the dataset being unbiased and less expensive and, most importantly, remaining legally compliant,[27] in turn assisting the cultivation of more predictable outcomes.[28] Therefore, the quality of datasets is paramount to AI fulfilling any sort of function and cultivating public trust on AI as an alternative to traditional services.[29]

A human-centric solution to AI bias is posed as ensuring the teams behind dataset development are diverse[30] and, therefore, a more representative microcosm of society, while ensuring that historical inequalities are no longer perpetuated.[31]

[23] P. Hall and D. Ellis, *A Systematic Review of Socio-technical Gender Bias in AI algorithms*. (Emerald Publishing Limited, 2023) p. 1.

[24] C. Gans-Combe, "Automated Justice: Issues, Benefits and Risks in the Use of Artificial Intelligence and Its Algorithms in Access to Justice and Law Enforcement", in *Ethics, Integrity and Policymaking: The Value of the Case Study*, eds D. O'Mathuna and R. Iphofen (Springer, 2022) p. 175.

[25] R. Rodrigues, "Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities", *Journal of Responsible Technology* (2020), 4, 100005

[26] United Nations Office on Drugs and Crime, "Artificial Intelligence: A New Trojan Horse for Undue Influence on Judiciaries", https://www.unodc.org/dohadeclaration/en/news/2019/06/artificial-intelligence_-a-new-trojan-horse-for-undue-influence-on-judiciaries.html (last accessed 9 May 2023).

[27] J. Soh Tsin Howe, "Building Legal Datasets", https://datacentricai.org/neurips21/papers/74_CameraReady_building-legal-datasets-CamReady.pdf (pp. 1-2).

[28] S. Wolfram, "What Is ChatGPT Doing … and Why Does it Work?" (Wolfram Media, 2023).

[29] M. Kusak, "Quality of Data Sets That Feed AI and Big Data Applications Enforcement", *ERA Forum* (2022), 23 p. 209.

[30] P. Hall and D. Ellis, *A Systematic Review of Socio-technical Gender Bias in AI Algorithms* (Emerald Publishing Limited, 2023) p. 12.

[31] World Economic Forum, "White Paper: How to Prevent Discriminatory Outcomes in Machine Learning" (Global Future Council on Human Rights 2016-2018) p. 18.

As such, although at first glance AI could be seen as having the potential to be a great equalizer, at present it is acting as a consolidator of inequalities.

As the antithesis to inequality is fairness, different approaches have emerged to place the necessary importance ensuring just outcomes. These include pre-processing the data to apply biases in counterfactual scenarios where the sensitive attribute may result in an unfair pathway. This can be designed in such a manner that corrects the variables that arise as descendants of attributes that result in unfair outcomes without constraining the parameters of the language model.[32] Alternatively, as a post-processing measure, predictions can be adapted after the fact to satisfy a predefined standard of fairness.[33] Many other mechanisms exist to attempt to combat bias; however, one approach, addressing the explainability of the AI systems, bridges the gap between human-centric and technological responses and the matter of AI bias. If a system can correctly identify the particulars of a decision and the data that led to a result, one may be able to ascertain the source of that bias.[34]

While attempting to find the source of the bias appears to be a reasonable approach, matters of accountability and transparency are plagued by uncertainty in relation to AI. Where in the conventional company structure there is an employee who can usually be pinpointed with responsibility for the particulars of a task, who should accept the blame where an AI results in unrepresentative outcomes? This is far more difficult to ascertain. AI runs the risk of becoming a liability shield for the shortcomings of those involved. As such, it is of paramount importance to have clear demarcations within the datasets with humans in the loop as representative data sources and the ultimate authority, particularly when matters pertaining to one's liberty are involved.[35]

Aside from being able to understand the importance of context, having a human tied to the actions of the system[36] is a way in which to deal with transparency particularly in a system which cannot be held to account the way a human can be. The public have come to expect absolute precision and certainty from systems while being able to more easily forgive genuine human error from a place of empathy and solidarity as we hold the ability to consider the consequences of our

[32] S. Chiappa and T. P. S. Gillam, "Path-specific Counterfactual Fairness" ArXiv: 1802.08139, Cornell University, p. 8.

[33] McKinsey & Company, "Tackling Bias in Artificial Intelligence (and in Humans)", https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans (last accessed 12 July 2023).

[34] Ibid.

[35] F. Galli, "Law Enforcement and Data-Driven Predictions at the National and EU Level", in *Constitutional Challenges in the Algorithmic Society*, eds H-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor and G. De Gregorio (Cambridge University Press, 2022) p. 127 – this chapter deals with profiling biases.

[36] A. H. Raymond, C. Draper and D. Coates, "Artificial Intelligence and Governance Policy: A Practical Guide to Identifying, Understanding and Mitigating Legal Risks Associated with AI Integration", in *Leading Legal Disruption: AI Vision for the Future of Artificial Intelligence*, eds G. D'Agostino, A. Gaon and C. Piovesan (Carsewell, 2019) p. 356.

doi: 10.5553/IJODR/235250022023010001007

actions.[37] Since there is no impactful way of holding machines to account, transparency in regards to data and those involved presents itself, at face value, as a reasonable counterbalance.

However, transparency too can prove problematic; excessive transparency leads to obfuscation of functional explainability creating a divide between those who understand AI and those who do not, and, from a political standpoint, such a situation can cause disengagement and disillusion with the problem in question.[38] Furthermore, abundant transparency could reveal important information that people and businesses do not wish to share with others causing data privacy concerns.[39] As such, although transparency in AI systems presents itself as a normative good, there are significant obstacles to ensure this works for both the consumer – as the body providing the sensitive information – and the creator – as the holder of the intellectual property of the AI model. Therefore, a formulation of a new model is required in order to strike a balance between the impacted parties. We posit the appropriate model as our conception of informational sovereignty.

4 Protecting the Rule of Law by Enforcing Jurisdiction

The role of the rule of law within legal systems cannot be understated. The rule of law cemented its place as a foundational principle of constitutional law centuries prior, continuing to predominate until the present day, with its remit extending to contemporary developments such as AI. The rule of law acts as a safeguard against arbitrary power and a maintainer of public order.[40] Also within this, it acts as the bedrock for the formation of laws as the principal consideration on lawfulness on public legal action. In order to protect the rule of the law, a practical restriction exists in terms of each state having responsibility to maintain the quality of the rule of law. Responsibility for this substantially befalls the legal system and, to a degree, the system of government. Both of these are impacted by public values to some extent; the law must adhere to the concerns of public policy and legal culture, as the careers of many of those in the governmental sphere rests firmly upon public opinion.

The rule of law is said to be challenged in three ways by AI: the blurring of the private-public regulatory sphere on fundamental rights; the subsequent failure to demarcate legal certainty within this framework; and the lack of transparency and

[37] A. Reichman and G. Sartor, “Algorithms and Regulation”, in *Constitutional Challenges in the Algorithmic Society*, eds H-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor and G. De Gregorio (Cambridge University Press, 2022) p. 161.

[38] S. Diplock, B. Gosschalk, B. Marshall and K. Kaur-Ballagan, “Non-voters, Political Disconnection and Parliamentary Democracy”, *Parliamentary Affairs* (2002), 55(4) p. 719.

[39] A. Reichman and G. Sartor, “Algorithms and Regulation”, in *Constitutional Challenges in the Algorithmic Society*, eds H-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor and G. De Gregorio (Cambridge University Press, 2022) p. 170.

[40] J. Raz, “The Rule of Law and Its Virtue”, in *The Authority of Law: Essays on Law and Morality* (Oxford University Press, 1979) p. 210.

accountability of the mechanisms of decision-making.[41] All of the above add a layer of obfuscation to a system that is already subject to unintelligibility at the level of a layperson. The result of this would be a more significant gap between the public and those in the legal profession, thus causing disengagement and a subsequent decline in legal culture.

Within the discussion of jurisdictions, a heavier usage of AI LLMs in their current form would result in an incremental decrease in representative legal outcomes. This is through a failure to remain within the confines of established precedents which are intended to promote consistency and predictability in legal outcomes. Several layers of uncertainty arise when differentiating between precedent and persuasive precedent from similar jurisdictions. In addition to this, precedents, and the principles they contain, are not permanent. Instead, they are driven by intertwined community input, often in the form of lawyers, the importance of whom will be stated in due course, and court decisions. The absence of clear direction, and the subsequent damaging inability of individuals and organizations to rely on predictable legal outcomes, would culminate in a decline in legal culture being the primary source of law as it has previously been in common law systems. To uproot a primary source of law particularly through the backdoor, perhaps the one source that the public are undeniably aware of, is incredibly problematic from a democratic perspective. The legal system does not exist in a vacuum; thus, it is incontrovertible that any attempts to incorporate a greater role for AI should not contravene democracy and jurisdictional boundaries.

5 Reconsidering the Legitimacy of Data Sources

Although the aforementioned concerns with the nature of AI are well-documented, what is neglected in the literature are the links to jurisdiction inherent in many of these. Bias arises out of unrepresentative datasets that are subject to a prior level of unrepresentativity where they are initially formed in another state using the data from people, discussions and events that were formed in that jurisdiction. Therefore, it is not only subject to sociocultural lack of representation but also subject to the lack of representation of their jurisdiction in question. These two levels of bias add a level of complexity that requires resolution through a rethinking of the way in which we view digital sovereignty. Representative outcomes should trump the location of the data as long as that data is jurisdictionally appropriate for the community in question.

In terms of other concerns arising from AI usage such as transparency and auditability, this can also be achieved more appropriately through this logic. Where data sources are jurisdictionally representative, it is considerably easier to locate them and hold them accountable. This is due to being more discerning about the datasets used rather than compiling large volumes of data on the basis of quantity rather than quality. While size is important in creating a functional dataset, what

[41] O. Pollicino and G. De Gregorio, “Constitutional Law in the Algorithmic Society”, in *Constitutional Challenges in the Algorithmic Society*, eds H-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor and G. De Gregorio (Cambridge University Press, 2022) p. 7.

doi: 10.5553/IJODR/235250022023010001007

is more important is high-quality data based on accessibility, objectivity and relevancy among other principles.[42] However, missing from these often cited foundations of compiling datasets is appropriate judicial representation; even if progression in AI usability is slowed in the short term by this consideration, it will allow for usable, accurate and appropriate datasets in the future whereby AI can cement itself as a tool to assist in determining outcomes rather than being the final arbiter.

Where do we get these data sources? This would be the modern role for a lawyer – an important data source. As lawyers are accountable to their state bar or regulatory authority as a core administrative principle, they have the ability to protect proper legal procedure and ethics at risk of being impacted by AI overreach.[43] As such, a replicable system for regulation already exists that can be applied to other informational development – by holding those who provide the information to account through regulatory authorities. By acting within the confines of regulatory bodies and legality, this model would allay significant concerns regarding theft of intellectual property rights as lawyers would be at the forefront as a primary data source.

The practicalities of using LLMs in law require training on vast amounts of textual data representing community interests through the arguments made by the lawyers representing the community. These models use machine learning techniques to identify patterns in the data and develop a set of rules or patterns that can be used to make predictions or generate new text. These predictions and generated text represent the arguments and decisions that would be made or arrived at by the community, so long as the dataset was generated by the community.

As such, if the outputs of the LLM are to be appropriate for a jurisdiction, they must be so on three grounds. The LLM training data must reflect the community bounded by that jurisdiction, meaning the model inputs should only be generated by individuals who have met the standards required of representing the community within that jurisdiction. Second, the datasets must be substantial enough to result in generalizable and predictable outcomes based upon that community's law without reference to law from other jurisdictions that would not ordinarily be cited in traditional legal precedents. And last, operational logic reflecting procedure specific to a jurisdiction must be directly encoded for instances when the law clearly requires a known cause to produce a specific effect.

6 The Insufficiency of Digital Sovereignty

Therefore, placing focus on digital sovereignty as a mechanism of ensuring system reliability through a focus on the location of where the data and hardware are held is therefore somewhat misguided. At present through the aforementioned cases

[42] M. Kusak, "Quality of Data Sets That Feed AI and Big Data Applications Enforcement" *ERA Forum* (2022), 23 p. 212.

[43] A. Reichman and G. Sartor, "Algorithms and Regulations", in *Constitutional Challenges in the Algorithmic Society*, eds H-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor and G. De Gregorio (Cambridge University Press, 2022) p. 174.

the system has displayed itself as far from reliable and the subject of various legal actions. There is an alternative to this approach, which we term informational sovereignty: if we instead refocus on the value of representative informational accuracy in a given jurisdiction, the result will be predictable, accountable, regulated and transparent outcomes, thus circumventing the main concerns of those involved in discourses on AI/ML.

Protecting communities from the potential harm of AI systems often takes the framing of an outside force acting upon the affected population. In the legal technology vertical, this force can often be seen as anything from profit-driven corporations to malevolent State actors.[44] This focus on protection from outside forces drives protection efforts towards the concepts of digital sovereignty, at whose heart is the concept of data sovereignty. While reasonable, AI-driven justice technologies tools push us to realize that these strategies are fundamentally ineffectual.

Digital sovereignty refers to the idea that nations and individuals should have control over their own digital technologies, data and infrastructure. The concept of digital sovereignty is based on the idea that the digital world has become a vital part of modern life and that control over digital technologies and data is essential for maintaining national security, economic competitiveness and personal privacy. In attempts to exert this control, the focus of digital sovereignty can be framed within the remit of traditional geopolitical sovereignty which has been subject to centuries of prior discourse.[45] Here, Krasner's quadripartite conception of sovereignty can be reworked as a basis to incorporate the challenges presented by an increasing use of AI in the legal profession:[46]

– *Population* is conceptualized as control over data. Digital sovereignty emphasizes the importance of individual and national control over personal data and information. This includes data privacy, data protection and the ability to decide how and when data is collected, used and shared.
– *Territory* is conceptualized as control over digital infrastructure. Digital sovereignty also involves control over the infrastructure and systems that support digital technologies. This includes control over networks, servers and other digital hardware and software.
– *Recognition* is conceptualized as control over digital governance. Digital sovereignty emphasizes the importance of national sovereignty in digital governance and regulation. This includes the ability of nations to set their own rules and regulations for digital technologies and data and their ability to enforce those rules and regulations.
– *Regulation of borders* is conceptualized as protection against cyber threats. Digital sovereignty also involves protecting against cyber threats such as

[44] S. Rosengrun, "Why AI Is a Threat to the Rule of Law", *Digital Society* (2022), 1(10) p. 9.

[45] T. Hobbes, *Leviathan* (Harvard Classics, 1651), Chapter 13 Para 10; W. A. Dunning, "Jean Bodin on Sovereignty", *Political Science Quarterly* (1896), 11(1) p. 92.

[46] S. Krasner, *Sovereignty: Organised Hypocrisy* (Princeton University Press, 1999). Within this work, Krasner sets out four variants of sovereignty: domestic (exercise of authority within a territory), interdependence (control over cross-border flow), international legal (recognition of territory by other territories) and Westphalian (non-intervention by others in the affairs of a territory).

doi: 10.5553/IJODR/235250022023010001007

cyber-attacks, cyber espionage and cyber terrorism. This includes developing robust cybersecurity measures and protocols and collaborating with other nations to combat cyber threats.

While traditional sovereignty concepts consider the population to be human individuals, digital sovereignty considers data itself to be the population that must be protected through rigorous control.[47] When defining this data population, the concept of data sovereignty typically features two unique aspects whose reasonableness AI-driven tools directly challenge:

- *Data protection laws*. Many countries have implemented data protection laws that regulate the collection, storage and use of personal data. These laws give individuals control over their personal data and require organizations to obtain consent before collecting and processing personal data.
- *Data localization*. Data localization is the practice of requiring that data be stored in a specific geographic location. This allows countries to maintain control over their citizens' data and protect it from foreign governments and companies.

The focus on these two aspects of data sovereignty are typically implemented by governments through restricting what data generated by one person's existence can be copyrighted by another without the generator's consent and restricting the jurisdiction wherein the silicon upon which the generated dataset must be physically located.

AI tools challenge the reasonableness of modern data sovereignty constructs because, although they must access the data contained on the silicon that is intended to be protected by the concepts of digital and data sovereignty, the information perceived from an AI tool is a by-product of the appropriate relationships interpreted between the training data. For the United States citizens, this can be illustrated by the difference between an integer 123456789, a person defined by social security number 123-45-6789, and a company defined by employer identification number 12-3456789.

The data generated by an individual is an artefact of their existence and cannot recreate a projection of their existence without the context of the individual. The information associated with this contextually derived assembly of the data is what makes any AI or LLM usable. This is why concepts of data sovereignty when considering the regulation of AI for LegalTech uses require a reconfigured, more appropriate 'information sovereignty' concept.

In the same way that the laws of a jurisdiction are only accepted if they reflect the community contained within the jurisdiction, and the laws of a jurisdiction are made by the legal professionals operating within that jurisdiction, an LLM is only appropriate for use within a jurisdiction if the data is assembled in a manner that incorporates the context of the legal professionals from within that jurisdiction. The location of the silicon upon which the data that assemble that data into information, or the location of the stochastic datasets that dynamically deploy that

[47] L. Amoore, "Cloud Geography: Computing, Data, Sovereignty", *Progress in Human Geography* (2018), 42(1) p. 16.

data within an AI tool, does pose a risk in the form of model access or reliability. But the appropriateness of an AI tool is based solely on its ability to represent the information gathered through observation of the population it will serve. This requires that tool suitability is defined by the source of information that was observed through the training of the model.

The fact that any LLM is little more than a technological mimic of the observations it is fed has become more rapidly understood than possibly any comparable revelation for any other transformative technology.[48] This means that, in the same way precedent in a jurisdiction would not be accepted if it was attempted to be made by a legal practitioner who is not authorized to practice in that jurisdiction, an AI LLM that is used by a jurisdiction must be restricted to assemblies of data that are deemed appropriate because they are trained upon observations of practitioners from that jurisdiction. This rethinking of how AI tools should be jurisdictionally restricted leads to a proposal of 'information sovereignty' that could be represented as:

- *Population*. Model training must be limited to observations or interactions with individuals from that jurisdiction.
- *Territory*. The jurisdiction is not geographically constrained but, instead, inclusive of practitioners and systems operating within its represented community.
- *Recognition*. System outputs must be sufficiently auditable to verify that it is consistently reflecting an appropriate representation of community-accepted practitioners.
- *Regulation of borders*. System outputs must be sufficiently immutable to prevent modification when transferred across systems.

In following this structure, AI could be used in such a way that it does not harm the democratic foundations of a community or lead to unfounded or unrepresentative outcomes. Since LLMs are not at the stage where they can appropriately respond to concerns expressed by the legal community, sufficiently considering these four tenets would go a significantly long way in addressing these concerns and fortifying trust in AI. Until this is the case, it would be improper to consider LLMs as a sufficient device to contribute meaningfully towards important sectors such as legal, business and financial on more than just a superficial level. For instance, those who cannot afford traditional legal services still deserve representative legal outcomes and rights to due process. Where a case may hinge on a fine technicality, AI is unlikely to yet have the appropriate level of nuance to effectively respond. While this remains the case, this variety of technology has not yet sufficiently evolved into a trusted tool.

[48] Boost.AI "What Are Large Language Models and How Do They Work?", https://www.boost.ai/blog/llms-large-language-models (last accessed 16 May 2023).

doi: 10.5553/IJODR/235250022023010001007

7 European Union Artificial Intelligence Act (EU AI Act)

The EU AI Act is the first of its kind, marking a significant and pivotal step towards appropriate AI regulation. In an attempt to be a global leader of AI regulation,[49] the Act sets out the ways in which it intends AI to complement humans through a focus on fundamental rights and the needs of society. This is identified by European Commission President Ursula von der Leyen as a "critical area" and a "key political priority".[50] The EU approach is predominantly one based on the concept of digital sovereignty with importance placed on the EU's ability to act autonomously in the digital world.[51]

The EU, although a proponent of digital sovereignty in name, is cognizant of its flaws. As such, the EU AI Act makes important allusions to the notion of informational sovereignty; although not yet conceptualized fully, it shares the priority of jurisdictional appropriateness for the creation of datasets, highlighting this as a core problem to be addressed. This is displayed most prominently in sections 10(4), 12(2) and 61(2).

Section 10 ensures that datasets must pay due regard for the "specific geographical, behavioral or functional setting";[52] the implication of this is to act as a safeguard to ensure representative outcomes. The geographic element in particular highlights the significance of due consideration of jurisdiction in order to achieve representative and appropriate datasets for communities without undue external influence. This elimination of *bias* through a requirement for using jurisdictionally appropriate data sources is a key tenet of informational sovereignty.

Section 12[53] insists upon logging capabilities that offer a level of *auditability* sufficient for evaluating system performance. Specifically requiring these logs "shall ensure a level of traceability of the AI system's functioning … [and] enable the monitoring of the operation of the high-risk AI system with respect to the occurrence of situations that may result in the AI system presenting a risk" acknowledges that traditional means of predicting model suitability is not possible with continually updating probabilistic algorithms. This section builds upon the protection from bias in Section 10 by requiring an observation-based framework for monitoring model cause and effect, like one would when educating and evaluating human performance, since the current code review approaches relied upon by digital sovereignty strategies are not sufficient.

[49] European Parliament, "Artificial Intelligence: The EU Needs to Act as a Global-Standard Setter", https://www.europarl.europa.eu/news/en/press-room/20220318IPR25801/artificial-intelligence-the-eu-needs-to-act-as-a-global-standard-setter (last accessed 12 July 2023).

[50] European Parliament, "Digital Sovereignty for Europe", https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf (last accessed 12 July 2023).

[51] European Parliament, "Digital sovereignty for Europe", https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf (last accessed 12 July 2023).

[52] European Parliament, "Artificial intelligence: The EU Needs to Act as a Global-Standard Setter", https://www.europarl.europa.eu/news/en/press-room/20220318IPR25801/artificial-intelligence-the-eu-needs-to-act-as-a-global-standard-setter (last accessed 12 July 2023); Section 10 of the European Union Artificial Intelligence Act 2021.

[53] Section 12, European Union Artificial Intelligence Act 2021.

doi: 10.5553/IJODR/235250022023010001007

Section 61 states the necessity of data collection through a post-market monitoring system to evaluate continued compliance with the Act;[54] the value of this section is to provide *transparency* and *ownership* for the humans in the process. In this way, the section protects from the governance ills currently producing negative outcomes from *theft* to *negligence* by ensuring AI can no longer be used as a liability shield since as compliance with the Act falls upon those responsible for the monitoring system. This ensures that the implementation of what we describe as informational sovereignty is subject to continued protection through clear processes and recourse rather than the uncertainty presented by the digital sovereignty approach.

The implication of these sections agrees with our premise that the information is the source of the value and requires the protection rather than the data. As such, our conceptualization of information sovereignty has the ability to act as a means to reframe future iterations of the AI Act, both in the EU and beyond in order to protect innovation which represents a key criticism of the Act in an open letter signed by 150 European companies,[55] while expanding protections that directly address the contemporary threats to the rule of law, our communities and businesses.

8 Concluding Remarks

While ostensibly the use of AI tools presents significant opportunities, at present it is plagued with risks, inaccuracies and inconsistencies that have the potential to be damaging in the long term if left unaddressed. For instance, the improper use of AI tools as a replacement for conventional legal services has far-reaching implications, impacting the individual, industry and the traditional conception of the state. It is posited this will transpire primarily through jurisdictional overreach of AI tools that pose the substantial risk of blurring the delimitations of community law through datasets that fail to differentiate along jurisdictional boundaries.

Through examining the most widely accepted sovereignty framework in AI discourses, namely digital sovereignty, and its subsequent shortcomings in addressing the key criticisms of AI such as bias, transparency and theft, a new conception is required. The proposed starting point for a solution is set forth as a new conception of informational sovereignty to act as a bulwark for the protection of democracy and the individual. This is based upon the importance of limiting the model's training to observing individuals from the *population* in question, including the practitioners and systems operating with that *territory*, providing accountability through the *recognition* of reflecting the outputs of practitioners within that community while in doing so providing sufficiently immutable outputs to prevent modification outside *regulated borders*. The process of shifting the focus from digital to informational sovereignty has already begun through the AI Act; in providing a conceptual framework for the training of LLMs to follow, these adapted criteria

[54] Section 61, European Union Artificial Intelligence Act 2021.

[55] *Financial Times*, "European Companies Sound Alarm over Draft AI Law", https://www.ft.com/content/9b72a5f4-a6d8-41aa-95b8-c75f0bc92465 (last accessed 12 July 2023).

doi: 10.5553/IJODR/235250022023010001007

would be significant reassurance for society more broadly to consider the use of appropriate AI tools. In addition, a reframing assists future iterations of AI legalization both in the EU and beyond to have the information to strike an appropriate balance between innovation and addressing threats to the rule of law and fundamental rights. In the long term, these developments would accelerate the use of AI systems by providing appropriate and necessary time for high-quality, jurisdictionally appropriate datasets to be formed. In the meantime, it is unreasonable to expect that AI is ignored; therefore, mitigation of the risks is paramount given the invention of false evidence, or hallucinations, by LLM tools such as ChatGPT, the lack of predictability and accuracy in outcomes, and bias that threaten due process and structural equality.

doi: 10.5553/IJODR/235250022023010001007