Law and the Emerging Political Economy of Algorithmic Audits

Petros Terzis*
p.terzis@uva.nl
University of Amsterdam
Institute for Information Law
Amsterdam, the Netherlands

Michael Veale†
m.veale@ucl.ac.uk
University College London
Faculty of Laws
London, United Kingdom

Noëlle Gaumann
noelle.gaumann.20@ucl.ac.uk
University College London
Faculty of Laws
London, United Kingdom

ABSTRACT

For almost a decade now, scholarship in and beyond the ACM FAccT community has been focusing on novel and innovative ways and methodologies to audit the functioning of algorithmic systems. Over the years, this research idea and technical project has matured enough to become a regulatory mandate. Today, the Digital Services Act (DSA) and the Online Safety Act (OSA) have established the framework within which technology corporations and (traditional) auditors will develop the 'practice' of algorithmic auditing thereby presaging how this 'ecosystem' will develop. In this paper, we systematically review the auditing provisions in the DSA and the OSA in light of observations from the emerging industry of algorithmic auditing. Who is likely to occupy this space? What are some political and ethical tensions that are likely to arise? How are the mandates of 'independent auditing' or 'the evaluation of the societal context of an algorithmic function' likely to play out in practice? By shaping the picture of the emerging political economy of algorithmic auditing, we draw attention to strategies and cultures of traditional auditors that risk eroding important regulatory pillars of the DSA and the OSA. Importantly, we warn that ambitious research ideas and technical projects of/for algorithmic auditing may end up crashed by the standardising grip of traditional auditors and/or diluted within a complex web of (sub-)contractual arrangements, diverse portfolios, and tight timelines.

CCS CONCEPTS

• **Social and professional topics → Computing / technology policy**; • **Information systems → Social networking sites**.

KEYWORDS

auditing, algorithmic audits, Digital Services Act, Online Safety Act, political economy

ACM Reference Format:
Petros Terzis, Michael Veale, and Noëlle Gaumann. 2024. Law and the Emerging Political Economy of Algorithmic Audits. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FACCT '24), June 03–06, 2024, Rio de Janeiro, Brazil.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3630106.3658970

*Also with University College London, Faculty of Laws.
†Also with University of Amsterdam, Institute for Information Law.

FACCT '24, June 03–06, 2024, Rio de Janeiro, Brazil

ACM ISBN 979-8-4007-0450-5/24/06
https://doi.org/10.1145/3630106.3658970

1 INTRODUCTION

The idea of auditing online platforms and algorithms as a form of governance is not new nor is its critique [30, 39, 68, 108, 111]. From the drafting of high-level ethical principles and codes of conduct to the technical development of 'fairness toolkits' and standards, scholarship in FAccT and beyond has been focusing on ways to translate normative mandates into actionable corporate practices. Today however, regulations have incorporated forms of these efforts, and we see early signs of the institutional and organisational dynamics that are likely to shape the future trajectories of algorithmic governance.

Looking systematically at the various legal developments in the field of algorithmic auditing in Global North jurisdictions suggests that auditing and certification assessment as a practice of entrusting to third-parties the evaluation of certain properties of AI and IT systems — reliability, security, fairness, transparency and privacy and more — is here to stay. In the US, there have been repeated attempts to introduce federal regulation recommending independent audits for meaningful human rights impact assessments. Proposals have been made at the federal level [4, 6], in local legislation in New York City [7, 63], and the Federal Trade Commission (FTC) has expanded on some of the fundamental principles that such audits shall follow [38, pp. 50–58] (see generally [113]). Canada and Australia are also set to introduce obligations for internal and external audits as accountability mechanisms for various algorithmic systems [51, 58, 92]. In the European Union, this trend is particularly strong. The AI Act introduces a framework for conformity assessment which, for some applications, requires a 'notified body' to audit documentation of compliance [43, 127]. In the same spirit, under art. 15 of the Digital Markets Act (DMA), entities designated as 'gatekeepers' will have to undergo an independent audit for any techniques of consumers' profiling they deploy, whereas under art. 23(3) of the same, the Commission will be able to appoint external auditors as part of its power to conduct inspections. Finally, indicative of the popularity of audit-like mechanisms for technology regulation is the fact that the European Parliament has suggested an AI Act-like conformity assessment procedure for the first (of many) sector-specific regulation for European 'data spaces', namely the proposed European Health Data Space (EHDS) [98, 122].

This complicated regulatory canvas risks blurring the lines of what is expected to be auditable and audited, at which stage of the development, by whom, and to what end [111]. In this frenzy of legal developments, concepts such as risk-assessment and conformity assessment, or compliance and auditing, are often thrown around seemingly interchangeably [52]. Absent clarity, responsibility for

auditing disperses amongst various actors, internal and external forms of control are becoming difficult to differentiate and evaluate, benchmarks and standards vary, and accountability is diluted across a complicated value chain of/for compliance [36]. However, this does not by itself mean that law(s) will not work. On the contrary, it suggests that with legal regimes now in place, the decision-making authority for resolving tensions and inconsistencies is departing from the realm of politics and/or research and migrates to that of law by recognising certain actors as legitimate bearers of the responsibility for/of audit.

To achieve further clarity, this paper systematically reviews the provisions of the UK's Online Safety Act (OSA) 2023, the EU's Digital Services Act (DSA) 2022 and the latter's associated Delegated Regulation on the performance of audits (DRPA) 2023, which taken together arguably present the most comprehensive proposed framework for understanding the dynamics of external auditing of large online platforms and their algorithmic systems. We start from the letter of the law, but do not end there. Instead, through an exploration of consultation submissions of corporate actors, their advertised portfolios and projects, and the industry dynamics in the field of algorithmic auditing, the paper situates the legal analysis within the emerging political economy of algorithmic auditing: Who is likely to occupy this space? What are some political and ethical tensions that are likely to arise? How are the legal mandates of independent auditing likely to play out in practice? Section 2 is descriptive in that it sets out, in a systematic way, the main provisions that will guide the shaping of the algorithmic auditing ecosystem in the EU and the UK moving forward. Section 3 draws on this analysis and observes tensions, challenges, and dilemmas that are likely to arise in practice. And, eventually, Section 4 provides a roadmap of likely scenarios for the future of algorithmic auditing and the uncomfortable reality that researchers in this space might ultimately have to confront and grapple with.

2 COMPARATIVE ANALYSIS OF THE ONLINE SAFETY ACT AND THE DIGITAL SERVICES ACT

2.1 Overview

Throughout their short history, many social media platforms developed their own systems for controlling, monitoring and moderating the flow of content on their platforms, often as a result of controversy, and with a strong role for algorithmic systems [55]. Following a 'grace' period of self-regulation that lasted more than a decade, the rule-making and rule-monitoring dynamics underpinning the functioning of social media platforms look now to change. In this context, the UK's OSA and the EU's DSA have been touted as solutions. They are ambitious instruments that are expected to enable governmental oversight over crucial issues such as the spread of hateful, misinforming, or illegal content, operating by removing aspects of decision-making authority away from private entities and by establishing formal guidelines and processes. Both pieces of legislation can arguably be understood as a by-product of our changed understanding of online platforms and in particular social media platforms as transnational public forums for all matters social, political and otherwise [68, 69]. Both depart from the classic policy hope that rendering firms liable for illegal content upon notice of its existence provides enough incentive for a 'healthier' internet [47, 80],[1] towards direct responsibilities, typically framed as procedural and due diligence duties and obligations [31].

2.1.1 Legislative Background: Online Safety Act. The UK's Online Safety Act 2023 has its roots over 6 years prior in the creation of the 2017 Internet Safety Green Paper [60], itself following pledges around social media in the 2017 Conservative Party Manifesto, and poorly thought through (and thus never-commenced[2]) legislation on website age verification in the Digital Economy Act 2017. Its turbulence has been connected to both the intrinsically contested nature of speech governance and the fact that there have been seven consecutive responsible Secretaries of State for the relevant portfolio (then Digital, Culture, Media and Sport), spanning four Prime Ministers, in the period between the Green Paper and Royal Assent of the Act. The form of the act was initially inspired by the work of Professor Lorna Woods at the University of Essex and Will Perrin at the Carnegie UK Trust, who proposed a 'duty of care' for social media platforms [130], although their proposed draft (with Maeve Walsh) was succinct compared to the final Act, comprising of only 20 sections compared to the passed statute's 241 sections (and 65 pages of 17 further detailed Schedules) [131].

Political opposition to the provisions regulating content that was 'legal but harmful' to adults in the Online Safety Bill as introduced led to their eventual removal. There are multiple reasons for this, but one interpretation is that it clashed with an emerging more divisive, populist form of politics, as such provisions (and their subjective notions of offence or harm) interfere with the perceived political potency of mobilising opinion against marginalised groups, including transgender individuals and migrant communities. The final Act focuses primarily on illegal content in relation to adults, and both illegal content and legal content that may be harmful to children on services which are likely to be accessed by children. However, it is very difficult to create a service that is not 'likely to be accessed by children' without putting in place significant age detection systems — either verification using identity, or 'assurance', a set of highly speculative technologies designed to algorithmically infer individuals' ages from biometrics or service usage data.

The form of the OSA is complex and there are a huge array of provisions and powers — space prohibits us providing a full overview here, least of all a critical one. For our purposes however, the core structure is that online platforms have to assess risks that their services, including their algorithmic systems, facilitate, and after undertaking this assessment, they have certain mitigation duties. The range of risks that need to be assessed and mitigated are significantly wider when platforms have or might have children as users. Differently sized platforms have different obligations, although at the time of writing no threshold or test has been published. This is one of the many parts of the OSA that either are awaiting an executive action (the making of regulations by the Secretary of

[1] Note that the United States is a notable outlier in this, unusual globally in that platforms are immunised against liability for most types of illegal content other than IP infringement regardless of whether they have been given clear and specific notice of its existence. This is the famous 'Section 230' (of the Communications Act, commonly misstated by almost everyone as Section 230 of the Communications Decency Act, where it is actually found in Section 9).

[2] Legislation can be passed, but have a provision that states that certain parts are not active until an order from the Secretary of State — this is called commencement.

State), or can be amended with limited parliamentary oversight by the executive. Faced with the complexity of its tasks, the regulator Ofcom — the existing telecoms and media regulator — is in the process of publishing literally thousands of pages of discussion and guidance for consultation.

2.1.2 Legislative Background: Digital Services Act. The Digital Services Act (DSA) is a Regulation — a type of European law that does not require Member States to manually rewrite ('transpose') it into their national law — passed in 2022, around two years after it was proposed by the European Commission.[3] It responds in part to varying national laws concerning expression on platforms, including Germany's Network Enforcement Act 2017 (Netzwerkdurchsetzungsgesetz, or NetzDG); Austria's Communication Platforms Act; and France's later struck-down 'Avia' Law (loi 2020-766). Indeed, in principle, the EU cannot create legislation such as the Digital Services Act — harmonisation legislation created under the Treaty on the Functioning of European Union (TFEU) art. 114 — without evidence of actual or potential fragmentation of law in Member States which could damage the internal market. The DSA, like much of EU law, has a stated aim in making compliance easier across borders by simplifying and homogenising the rulebook, while also introducing substantively new provisions into many jurisdictions which to date have had limited-to-no domestic platform regulation.

The DSA follows over a decade of European attempts at self- and co-regulation of online services [29, 31, 84], which culminated in the 2016 EU *Code of Conduct on Countering Illegal Hate Speech Online*, the 2017 and 2018 *Communication* and *Recommendation on Illegal Content Online* respectively, and the 2018 *Code of Practice on Disinformation* [83]. The DSA in some ways is more timid than the OSA, replicating nearly verbatim the existing intermediary liability provisions from the e-Commerce Directive 2000, including the prohibition on laws, regulators or court orders obliging general monitoring of all content on a platform.[4] Note that the DSA does not endeavour to create new forms of liability for content — these flow from existing national and Union law — but instead focuses on the information flow and management, including the processes around how this content is promoted, contested and removed. It has some procedural provisions that apply to all online platforms, and then a further range of provisions that apply to only Very Large Online Platforms (VLOPs)/Search Engines (VLOSEs). Many, although not all, of the provisions that we consider to be closely connected to audit are only applicable to these larger entities.

The scope and regulatory remit of the two regulations vary. In particular, where the OSA aspires to make the Internet a 'safer' place by primarily encompassing provisions for safety-by-design and the removal of illegal content online, with a strong emphasis on children's safety, the DSA introduces rules and mechanisms whose constitutive goal is to imbue accountability along with 'greater democratic control and oversight over (…) platforms' [37]. In both regulations the burden of responsibility correlates to the size and scale of the platform. In this direction, both OSA's Category 1 providers as well as the DSA's VLOPs and VLOSEs are/will be determined by secondary legislation and (will) have increased duties and responsibilities. The auditing obligations will follow a similar pattern, increasing in 'severity' and scope along with the scale and size of the platforms involved. However, both regulations have a long tail of platforms in-scope, with the impact assessment of the OSA indicating 25,000 firms will be subject to many of its obligations [61].

The following section sets out the different audit provisions attached to these regimes by looking at internal audits, new audit-like regimes, and external audits. The term 'external' is to be understood quite broadly here as we depart from an established typology of audits (first-party, second-party, and third-party) purely to maintain consistency with the examined legal provisions, and to use them as a starting point to understand how law itself may shape these practices [39, 86].

2.2 Internal audits

Both the OSA and the DSA oblige certain platforms to conduct internal risk assessments on their services. These can be understood in part as forms of internal algorithmic auditing, although the obligations may be wider than that.

Starting from the former, under the OSA there are four different processes of risk assessments that may or may not be applicable according to the nature and size of an online platform. In particular, all platforms must conduct a risk assessment for illegal content (s. 9). Where services are likely to be accessed by children — a test, as mentioned, made very hard for significant platforms to not meet — the respective platform must also conduct a so-called 'children's risk assessment' (s. 12). Finally, the designated 'Category 1' providers will need to perform a 'user empowerment' risk assessment to evaluate the inclusion ('to the extent that it is proportionate to do so') of design features for increasing the control of users on the content delivered through the platform (s. 14). Platforms are instructed specifically to take into account '(in particular) algorithms used by the service' in all these risk assessments (ss. 9(5)(b), 12(6)(b), 14(5(c))).

DSA's internal audits seem to be more expansive in scope and remit, although apply to fewer platforms. VLOPs and VLOSEs are obliged to 'diligently identify, analyse and assess any systemic risk' [emphasis added] stemming from the functioning of their systems and services (art. 34). In particular, recitals 80–83 set out the four main categories of systemic risks that VLOPs and VLOSEs shall take into account, namely: a) the dissemination of illegal content; b) actual or foreseeable impact on the exercise of fundamental rights; c) actual or foreseeable impact on democratic processes; and d) actual or foreseeable negative effect on other social and political rights (including public health or gender-based violence). Again, such risks clearly incorporate many automated and algorithmic systems which will need assessing, and therefore already blur the concept of a risk assessment with algorithmic audit.

Even though these audits are internal, there are many hooks for public participation, seen as both important but under-tooled in the existing algorithmic audit landscape [95]. In many ways, the public participation hook is seems clearer than other regimes, such as the involvement 'where appropriate' for related UK and EU data protection impact assessments under the GDPR [34]. The DSA

[3]While it is confusingly called an 'Act', this is legally meaningless, and is just a name that the EU is branding recently large Regulations with in order to make them sound considerably sexier and more PR-friendly.

[4]In contrast in the UK, this provision was never transposed from EU law into UK law and thus its post-Brexit status is unclear.

invites platforms to engage with groups affected by systemic risks (recital 90), and in a similar context, Ofcom recommends consulting a variety of stakeholders including users and experts when conducting risk assessments, which may also include commissioned research [93].

In terms of method, risks assessments for the purposes of the DSA are expected to look into — amongst others — the data collection and use practices, content moderation strategies, the online interface design, and risks of coordinated manipulation of their services (recitals 84–90). All of these areas increasingly include algorithmic systems in their functioning [56, 75], particularly where minority languages are in use and firms lack staff specialising in those communities [73, 96, 97, 100]. While these regulations require algorithmic assessments, they do not specify methodologies or approaches for understanding challenging, iterative systems, leaving a wide degree of flexibility to organisations [52]. Some organisations within scope of the DSA and OSA may have to deal with potentially illegal AI systems or models being uploaded to platforms, as a form of content, heightening the need for expertise in understanding them as the nature of user-uploaded content changes [57].

2.3 Novel audit-like actors

The DSA introduces novel audit-like mechanisms for large online platforms and entrust particular categories of entities and professionals with audit-like competencies and powers. In particular, two such categories under the DSA are the 'vetted researchers' and the 'trusted flaggers' — both forms of 'outsider oversight'. The former will be independent researchers affiliated with a research organisation and trusted to conduct specific research projects they have applied for in detect, identify, and understand the systemic risk posed by VLOPs'/VLOSEs' systems and services (art. 40 DSA). In particular, according to art. 40 of the DSA, the Digital Services Coordinator or the Commission will be responsible for initiating a data access request that will allow them and/or vetted researchers to monitor and assess compliance of the VLOP or VLOSE in question with the DSA [17]. The OSA does not directly introduce a parallel category to vetted researchers — an issue that was criticised in the Bill's passage [8] — however it obliges the regulator Ofcom to produce a report on researcher data access, and guidance for regulated services, within 18 months from the period when the Secretary of State commences that provision (s. 162). In the DSA, vetted researchers can request data, and platforms either have to provide it, or provide effective alternatives for study. The way that such access might extend to infrastructure — important given the dynamic and difficult to locate 'algorithms' within such platforms [117] — is, however, less clear [42, 44, 125].

'Trusted flaggers' might be conceived of as a form of auditor — organisations of particular expertise and competencies whose notices of illegal content will be dealt as matter of priority and processed by the online platform without undue delay. Their role is envisaged of one of both monitoring and of action, and again, it is the Digital Services Coordinator which can award this status, following the process set out in DSA art. 22. Such flaggers however will struggle to have impact at content moderation scale unless they opt for automation themselves, and in the DSA, do not a direct ability to scrutinise or influence algorithmic systems [16].

In general, the audit-like powers of vetted researchers and trusted flaggers will be task-specific. In particular, the former's powers and scope of action will depend on the Digital Services Coordinator's reasoned request and the data requested therein, whereas trusted flaggers' audit-like powers will depend on the particular expertise of the accredited organisations. In practice, however, it is entirely unclear how challenges for the technical implementation of these powers will be overcome [42, 76].

2.4 External auditors

Finally, the DSA and the OSA introduce important provisions and mandates for the external auditing of large online platforms that, as discussed further in Sections 3 and 4, set the foundations for the emergence of the industry of algorithmic auditing. In this direction, different from art. 40's externally led procedure for monitoring via researcher data access, art. 37 of the DSA sets out a form of annual, mandatory external auditing, to be performed by an independent auditing organisation which will be trusted with assessing with 'a reasonable level of assurance' compliance with all DSA obligations (see discussion below in Section 4.3). The relevant provisions were recently supplemented by the DRPA which includes additional rules and operational details with regard to the performance of these audits.

In a similar but more irregular manner, the OSA introduces the 'skilled person's report', commissioned research either by OFCOM or the provider under scrutiny at OFCOM's request and approval. The scope of such a report will be to help OFCOM in investigating a provider's (possible) failure to comply with OSA mandates and/or to understand the various risks related to the functioning of the provider's systems and services. Alongside this ability to order and commission a 'skilled person's report', OSA empowers OFCOM to conduct its own inspections and audits as part, for example, of an investigation by appointing authorised persons and following the issuance of a respective notice to the provider (OSA s. 107 & sch. 12). A flashpoint in the pipeline for skilled person's reports is any potential attempt to require client-side scanning in messaging or other services in relation to child-abuse imagery, a particularly controversial part of the OSA [79], but one which requires a skilled person's report before any such order can be made (OSA s. 122 (1) and OFCOM Consultation para. 28.22) [11]AppleJoinsOpposition2023OpenLetterSecurity. Insofar as such a report seeks to cove algorithmic detection technologies themselves, it may prove highly political, as previous UK Government funded reports from a research consortium scrutinising entrants to a 'Safety Tech Challenge Fund' on child abuse detection in encrypted systems UK stimulated 'safety tech' for encrypted environments found such tools unable to be effectively audited [99].

3 POTENTIAL TRAJECTORIES

Before delving into the foundations of the emerging political economy of algorithmic auditing, we wish to set out — at a general level of abstraction — three potential trajectories and scenarios for the future of algorithmic auditing. This intellectual exercise is admittedly arbitrary and perhaps oversimplified but in doing so, we wish to illustrate possible futures for transnational algorithmic governance in order to better explore the dynamics of the undergoing material,

institutional, and socio-economic transformations [24, 70]. This is a loose futures exercise [59] which we deploy here in a narrative fashion to showcase and interrogate potential flavours of algorithmic auditing, as socio-technical praxis embedded within a system of legal rules, cultural norms, and material practices, and situated within the broader ecosystem of technology production.

3.1 Convergence with traditional auditing

In this scenario, cultures and methodologies of traditional auditors dominate the market for, and practice of algorithmic auditing. Tech companies and AI start-ups in the field of model validation and assessment are absorbed by the Big 4 — the world's major audit firms, Deloitte, EY, KPMG, and PwC — thereby supplementing their auditing portfolios and augmenting the capabilities of their products and platforms. Nearly all large businesses in the US and UK use these firms: in 2017, 497 of the S&P 500 in the United States, with PwC alone claiming to provide services to 87% of the Fortune 500 [62, 2]. These firms have long sought to use legislation to craft lucrative roles for themselves — both James Deloitte and Edwin Waterhouse (the 'W' in PwC) coauthored Britain's Regulation of Railways Act 1868, which mandated a specified form of double-entry accounting that their firms specialised in — unsurprisingly ending up as major railway auditors as a result [62, 37]. However, while these firms have been behind countless foundational statutes obliging audit to occur, they (and their industries bodies) have for over a century fiercely resisted any effort to legislate *how* they should audit, with their judgment (and they way they choose to exercise it) seen as paramount in their eyes [62, 57-8].

Thus, in this trajectory, the substantive elements of the auditing process, the benchmarks, and methodologies are determined by the negotiations of the Big 4 with their 'clients' and little to no room is left for civil society organisations and researchers to meaningfully engage with, and be informed about the functioning of these systems [74, 102]. Eventually, like traditional audits, algorithmic auditing becomes just another service in the portfolios of multinational corporations conducting statutory audits — just part of the furniture of modern capitalism [62, 55].

3.2 Hybrid practices and hybrid, but separate, cultures

In this scenario, a vibrant start-up ecosystem for the assessment of algorithmic systems emerges. Leveraging its organisational agility and human capital, it overtakes the Big 4 in the offering of services for algorithmic auditing and offers a canvas of different methodologies for audited organisations to choose from. However, the autonomous development of these companies does not by itself suffice to safeguard their organisational independence. Instead, the erratic dynamics of AI-driven venture capital and the constant need for 'healthy' balance sheets have a substantial impact on the direction of these companies. Financial incentives and strategies forge corporate relationships and routines of corporate practice that are both order-enabling, in that they create and stabilise legally constituted forms of private ordering with real-world effects as to how the 'algorithmic society' is measured and evaluated, and norm-shaping, in that auditors and their sub-contractors produce, through their interaction, the substantive indicators based on which algorithmic systems will be eventually tested and measured.

3.3 New forms, cultures, and actors of/for auditing

In this scenario, research-driven, peer-reviewed, and evidence-based auditing methodologies rise as spin-offs from open-source communities, academic institutions, and non-profit organisations to become the new standards for the holistic evaluation of algorithmic systems [41]. Their organisational structure and operation are financially independent from the contractual relationship with the organisations they audit and the fees associated with the undertaken audit are re-directed to other non-profit organisation in the field. The academic credentials they carry along with the scholarly reputation of the people involved renders a collaboration with them strategically imperative for any company wishing to demonstrate its active commitment in building 'responsible AI'. Academic research and journalism investigations are financially supported to become the main drivers of the field whilst civil society organisations, policy fora, and research communities provide the intellectual fuel, human capital, and technical expertise necessary for stabilising relevant routines of socio-technical practice.

4 ELEMENTS OF THE EMERGING POLITICAL ECONOMY OF ALGORITHMIC AUDITING

Where do we seem likely to end up amidst these scenarios? Within a universe of potential trajectories for social action and coordination, law is oftentimes imagined and understood as a system of rules for guiding human behaviour based on certain normative and procedural criteria such as human rights and the rule of law. Although engaging in a legal-theoretical debate remains well out of the scope of this paper, we believe it is important to highlight that our legal enquiry departs from the understanding of law as the form-giving institution that stabilises expectations and exchanges by weaving together the economy with the political system in its state form through formal, often — and importantly for our case — profession–based, organisational arrangements that produce collectively binding decisions within their respective functional areas [72, 19-22]. Oftentimes, and increasingly in the field of technology, in doing so it offloads normative work and routines of such practice to governance regimes that feed on the coordinated interactions of their subjects and imperatives of economic efficiency to (re)produce order that seems and operates beyond law [64, 128]. Throughout this process, law-making as a practice of form-giving and form-shaping renders certain scenarios thinkable and others unthinkable; certain futures imaginable and others unimaginable [120].

In the case of algorithmic auditing, law delegates much of its accountability monitoring to external auditors. Given the ambiguous effectiveness of internal audits and with vetted researchers and trusted flaggers assuming a delegated and task-specific auditing role, auditing work of the kind that civil society and the public at large need in order to understand and scrutinise algorithmic systems and services, is left at the discretion and powers of external auditors. Despite their central role in modern algorithmic governance, however, we argue that these actors will operate in a legal landscape whose institutional framework and safeguards

are not robust enough to confront the powerful economic incentives and inter-organisational dependencies that can develop on the ground. Instead, a closer look at relevant provisions and industry developments may suffice to lower our expectations on the transformative potential of these new laws. In this direction, the provisions around the independence of auditing organisations, the selected standards and methodologies, the broader socio-economic questions that auditing is expected to encompass, as well as the operational details of the actual auditing, altogether create an institutional landscape which remains vulnerable to the stabilising effect of corporate strategies and routines of organisational practice. We will now examine these provisions in order.

4.1 Independence is complex

One of the most crucial safeguards for the effective performance of an audit is the independence of those trusted to conduct it [39]. The legal regimes we study here create confusion and uncertainty around this important audit characteristic.

4.1.1 Unclear OSA Independence. Issues of independence are not deeply considered in the OSA. Much rests on OFCOM's nomination of the 'skilled person' (or approval of the provider's recommendation for that matter) as to safeguard an independent investigation (OSA s 104 (4)-(6)). Furthermore, as both the issuance of a skilled person's report and the 'powers of entry, inspection and audit' will be typically reserved for 'more serious cases' (art. 28.52 OFCOM Consultation), questions of independence of the auditors entangle with the regulatory discretion of which cases warrant such interventions.

4.1.2 DSA Audit Independence, But Blurred Compliance Roles. In contrast, the DSA establishes specific criteria of independence that VLOPs/VLOSEs are required to take into account when appointing an auditor (DSA art. 37 (3)). However, the audit process gets more complicated when considering the compliance function in art. 41, where an officer under that provision is responsible for both identifying and mitigating risks, but also organising and supervising any audits. The independence of the compliance officer (who is expected to be a senior manager employed by the provider or a contracted third party) is expected to be safeguarded by organisational and operational measures, whereas the independence of the auditing organisation by examining potential conflicts of interest (i.e. by the provision of non-audit services related to the audited matters, such as software services, consultancy, training services, or content moderation services [recital 8, DPRA]) and ensuring the absence of fees 'which are contingent on the result of the audit'(art. 37 (3) DSA). Yet we can see indications of which actors desire to claim their role in this space from the final sentence of recital 8 of the DRPA — not part of the initial draft but which found its way to the final text — which reads: '[Provisions on conflicts of interest] should not exclude auditing organisation who have performed statutory financial audits [for the audited organisation]'. This is a clear invitation to the so-called 'Big 4' which have already declared their active interest and manifested their intentions to enter the industry of algorithmic audits [9, 49, 50, 77]. It is not difficult to imagine how the requirements of/for independence will play out in case a VLOP/VLOSE chooses to audit its financial and DSA obligations with the same auditor — will PwC or EY think carefully before publishing a 'negative' audit report for Meta's or Google's recommender systems?

4.1.3 Independence Across the Supply Chain. Auditors under the DSA can contract out part of the auditing process when it is necessary to seek expertise to evaluate, for example, 'the design and functioning of algorithmic systems', 'the risks to fundamental rights', or 'the spread of illegal content' (recital 3 DRPA). The audited organisation remains responsible for ensuring the independence and expertise criteria are met (art. 4 DPRA), but it is unclear on what basis they will have information to do this, given information on the subcontractor and the reasons for their selection will be mediated through the initial auditor, and any subcontractors seem likely feel incentivised to remain within the contractual radar of this auditing ecosystem to ensure steady flow of future projects and partnerships, which may pressure them to give more positive comments than not. The extent to which an independent subcontractor could give a negative opinion which will be faithfully integrated, rather than buried, seems questionable.

These issues echo similar concerns found in the audit of multinational groups, with regulators regularly raising issues about the performance and competence of so-called 'component auditors' which a lead auditor may contract to perform a certain part of an audit, such as a particular national component of a corporate grouping [133]. Literature on issues in component auditors already highlights the difficulties group auditors face in knowing and supervising components, even in cases where they are auditing something in a way that should be familiar to the group auditor (i.e., only the jurisdiction and tax systems differ) [121]. It further highlights that group auditors already struggle with the sociocultural, regulatory, and institutional differences between them and their component auditors, even where these components are a separate entity using their own corporate umbrella and brandname (e.g. *Deloitte Deutschland* or similar) [46]. This seems to be even more difficult when the group auditor is subcontracting out a qualitatively different form of audit — for example, of an algorithmic system, which they may not have expertise to interrogate or validate.

Furthermore, algorithmic systems in areas like content moderation are increasingly servitised into networked supply chains and value chains with many actors and 'many hands' [35, 36]. Assessment of the auditors' and subcontractors' independence given these components — such as content moderation technology providers such as Thorn, Google's Perspective API, Sightengine, Two Hat, and many more — is not clearly guaranteed in the legislation and becomes extremely difficult to assess. In a world of AI supply chains, the reputation of the audited entities' components is perhaps even more important to consider as a corrupting factor than from the audited entity itself [74]. The potential to subcontract draws a thread between the three scenarios illustrated in Section 3, allowing large auditors to retain functional control while using specific expertise of smaller organisations in the process, with unclear consequences for rigour or independence.

4.2 Whose benchmarks? Whose methods?

The baseline criteria and benchmarks based on which an algorithmic audit is (expected to be) performed are an integral part of an auditing process, and determinative of the nature and scope of the methodology [22, p. 24] [30, p. 4] [110]. It can alter the way we think about the very thing we are expected to audit and measure [32, 106, 107]. It will not be news to most readers that translating high-level and generalised principles of fairness, bias, or transparency into an actionable and, by extension, auditable obligation can prove exceptionally challenging [26, 30, 51, 119]. Similarly, negotiating benchmarks for quality assessment internally can itself prove a laborious and demanding process as siloed teams can pose significant communication problems and inefficiencies [66, 82, 89, 116]. Unsurprisingly, there has been a plethora of approaches and toolkits for assessing algorithmic systems, with varied yardsticks for measurement and evaluation depending on the nature of the algorithmic system in question [21, 40, 73, 86, 91, 124, 129].

Benchmarks can vary widely in terms of method, approach, and subject matter [109, 114]. More holistic evaluations will likely draw on qualitative work, including interviews, database and document inspection. Instead, if it is simply the technical specifications of an AI model that are audited, then accepted quantitative methods that, for instance, produce a score for 'robustness' may prove sufficient [33]. Some audits rely on live interaction with the infrastructure, such as sock-puppet audits, i.e. where a 'fake' user is created and researchers observe the user's interaction with the platform [23], or 'bottom up', user-driven auditing [118], supported by researchers [22]. Others are more observational, taking slices of data or code, either from the organisation or via users, for in vitro analysis [18, 22]. These methods incur different costs (and reputational risks), but there are few, if any, studies directly comparing their efficacy and coverage in the context of social media. Effective methods may require creativity and novel methodological generation, rather than turning to an 'accepted' benchmark or approach, particularly in the context of rapidly changing business models and platform practices. As a result, the choice of method is a commercial decision with unclear impacts on audit results.

Benchmarks and methods further reproduce the particular visions and worldview of the auditing entity. For example, METR (previously ARC Evals), a non-profit spin-off of the Alignment Research Center, evaluate AI systems based on their capabilities for 'autonomous replication', meaning whether 'an AI could survive on a cloud server, obtain money and compute resources, and use those resources to make more copies of itself' [2]. In a similar spirit, Apollo Research (or AI Evals), a project sponsored by the Rethink Priorities Initiative, aims at building evaluation models to detect 'deception and potentially other misaligned behavior' [1, 65]. In contrast, and indicatively, Fiddler Auditor tests systems for model robustness based on principles such as transparency, interpretability, fairness, privacy and reliability [13]. This becomes important, as the breadth of systemic risks that the DSA requires to be considered also allows considerable room for framing, bringing the auditor's priorities and worldview and constructing the DSA requirements around them, rather than the other way around. Pushing back against this may be hard for an under resourced Digital Services Coordinator, similarly overwhelmed by the potential breadth of risks that, in slickly written text, seem genuine, that could be considered as part of such an audit.

Under the DRPA, responsibility for the initial formulation of the benchmarks against which compliance is/will be sought, rests with the audited organisation (DRPA art. 5 (1) (a)) and in practice, it is more likely than not that this process will be carried out by the compliance function that presented above (DSA art. 41 and recital 99). Importantly, the evaluation of these benchmarks by the auditing provider is unlikely to have a negative impact on the final outcome of the auditing. Instead, as recital 16 and art. 8 of the DRPA explicitly acknowledge, the audit conclusion should be 'positive with comments' when the auditing organisation considers it necessary to provide further comments on the selected benchmarks in order to 'usefully inform' the future 'benchmarking' of the audited organisation, based on the auditor's knowledge, research, and expertise. This 'agile' back-and-forth between the auditing organisation and the (compliance function of the) audited organisation is likely to have a considerable impact on the way auditing benchmarks are formulated and/or standardised in the future. As the inter-organisational benchmarks consolidate overtime in a form of 'benchmark-as-we-go', the relationship between the 'audited' and the 'auditing' may generate undesirable dependencies [112]. In this interplay, the primary responsibility for the formulation of benchmarks is left to the auditing organisation and the only avenue available for external input in the process is the latter's discretion of using 'information from external sources' in its 'positive with comments' audit opinion (recital 16 DRPA). Eventually, benchmark disparities amongst different auditors may incentivise platforms to choose their assessors and auditors based on their benchmarks (easy or difficult, simple or complicated) and/or intensify institutional and organisational dynamics towards benchmark standardisation, a process that will be in itself extremely difficult to navigate and deliver [90].

4.3 Societal implications and the required expertise

Several studies have discussed the necessity yet complexity of incorporating societal considerations in the assessment of algorithmic systems [20, 21, 88]. The processes for doing so and the effect that organisational structures and practises can have on them are equally well researched and documented empirically [81]. The 'systemic risk' analysis VLOPS/VLOSEs are subject to in the DSA requires such analysis, an obligation fortified by the DRPA, obliging external audits to take into account — amongst other factors — the nature of the audited service and 'the societal and economic context in which the audited service is operated' (DRPA art. 9(4)(a)). When considering this context, the auditing organisation is required to express an opinion with a 'reasonable level of assurance' (DRPA art. 3 and recital 16). Reasonable assurance has no firm definition, but sits in contrast and is weaker than 'absolute' assurance, indicating the auditor is not a guarantor of correctness, and even audits conducted in accordance with given standards may fail to detect material concerns [48]. In financial audits, this stems from sampling, complex estimates that can be changed by fast-moving events, the potential for sophisticated fraud, and the need to make audits economically viable [48].

Unsurprisingly, large multinationals' consultation contributions concerning the DPRA indicates that this level of assurance, which survived to the final version of the DPRA, is one they are unhappy with. Google argued that the novelty of the field and lack of well-stablished rules leads to a material risk of 'the reasonable level of assurance standard' being interpreted and applied inconsistently by different auditing actors [54]. In a similar context, Booking.com argues the reasonable assurance standard in other sectors 'is generally reserved for subject matters that are highly quantitative (and even binary) in nature' [27].

There are clear winners in all of this however — the Big 4, again. Part of the reason why a 'reasonable level of assurance' is hard to explain is because over the last century, 'large accounting firms used their links with regulators and standard-setters to co-produce a coded, excluding but otherwise benign professional language that is rich with acronyms, jargon and euphemisms' [62, 59]. Gow and Kells argue terms like 'reasonable assurance' deliberately 'border on nonsense' [62, 59]. While auditing has become more formalised, with rules describing their form and structure, the discretion has simply been condensed into these arcane and opaque terms, with standards laying out the banal — that audits should have a title page; that they should list the client's instructions, and so on. Consider (as Gow and Kells do) the following description of a 'limited assurance engagement' in a major Australian audit standard:

> *An assurance engagement in which the assurance practitioner reduces engagement risk to a level that is acceptable in the circumstances of the engagement, but where that risk is greater than for a reasonable assurance engagement, as the basis for expressing a conclusion in a form that conveys whether, based on the procedures performed and evidence obtained, a matter(s) has come to the assurance practitioner's attention to cause the assurance practitioner to believe the subject matter information or subject matter is materially misstated. The nature, timing and extent of procedures performed in a limited assurance engagement is limited compared with that necessary in a reasonable assurance engagement but is planned to obtain a level of assurance that is, in the assurance practitioner's professional judgement, meaningful. To be meaningful, the level of assurance obtained by the assurance practitioner is likely to enhance the intended users' confidence about the subject matter information or subject matter to a degree that is clearly more than inconsequential.* [62, 59-60]

If you think that makes little sense — we agree.

Algorithmic auditing takes the now well-known problem of having standards set in an excruciatingly arcane and opaque language to another level, creating a significant new tension in the process. Auditing organisations are expected to assess contextual, societal features, which requires creative and ambitious methods. Yet they should root their analysis in 'proven expertise in the area of risk management, technical competence, and capabilities' (art. 37(3)(b) DSA). Traditional audit methodologies are standardised, repeatable, even if the ways in which they are standardised and repeated are closed industry knowledge. The types of audits that seem to be required by the DSA texts, if they are to be rigorous, are necessarily not.

This does indeed make the standard of the 'reasonable level of assurance' difficult — not because of its ambition of rigour, but because this definition is rooted in the idea of an institutional field, shared (yet often proprietary) understandings, methods and norms between professionals and their organisations in this space [19, 45]. The necessity of contextual, creative analysis leaves key questions to the professional discretion — and socio-political vision — of the auditor [53, 87]. This may not inherently be a bad thing, as long as assumptions are placed on the table. Yet the DSA, and the internal compliance roles it envisages, seem to have few incentives to push auditors away from rote, unambitious, context-free standards towards the creativity and creative rigour needed to undertake sociotechnical analysis with an open-ended list of potential systemic risks. The history of the audit industry indicates that auditors will fiercely protect their processes from external influence. In this context, auditing benchmark and methods seem likely to become just another standard(ised) service in the portfolio of traditional auditing actors from the Global North with their often monolithic cultures and one-dimensional methodologies of/for auditing.

4.3.1 A concerning statutory pressure valve. More worryingly, we can see legally that in the face of an audit system that desires standardisation and repeatability in order to provide 'assurance', auditors are likely to lean on an exit clause in the DSA which enables them to escape complex, value-laden, socio-technical analysis. art. 37(5) of the DSA enables the auditing organisation to avoid auditing specific elements or expressing an opinion on certain aspects of an investigation as long as it includes 'an explanation of the circumstances and the reasons why those elements could not be audited'. Rather than try to reinvent the concept of reasonable assurance such that open-ended challenges can face external scrutiny, as a superficial reading of the DSA might seem to push for, the existence of art. 37(5) provides a way to claim that such areas are simply too hard to analyse, giving a false level of assurance by letting the auditing organisation effectively scope out the qualitative, and potentially most societally crucial, aspects of their mandate.

4.4 Early signs of a new (or old) industry

Today, as the DSA enters into force, industry dynamics have already started taking shape. PwC has launched its Responsible AI Toolkit aiming at offering an 'end-to-end enterprise governance framework' to enable oversight and traceability of a company's AI development lifecycle. EY pitches its 'Trusted AI' tool as a platform capable of providing insights and helping AI design teams in 'quantifying risks' [49]. KPMG's 'Responsible AI' is advertised as a form of AI governance that exposes risks and vulnerabilities without 'compromising on innovation' [9]. Deloitte's 'Trustworthy AI Framework' promises to infuse 'an ethical mindset within [an] organization' by — amongst others — engaging 'external ethics experts and academic institutions to conduct well-rounded client conversations' [77]. Meanwhile, industry partnerships are mushrooming. KPMG and Microsoft have signed a multibillion-dollar agreement for the expansion of their relationship that 'will reshape professional services [...] including [...] [the] use of Artificial Intelligence solutions for clients, industries and society more broadly'.

Deloitte has teamed up with ChatterboxLabs, an AI startup, to enhance model insights offered through its 'Trusted AI toolkit' [78]. Arc Evals has partnered with Anthropic and OpenAI to evaluate their AI systems [2].

In parallel, AI auditing start-ups are attracting finance in the form of donations and venture capital; early signs of a burgeoning ecosystem [52, 71]. These are usually companies specialising in testing the reliability and resilience of an AI model, performing root cause analysis in Large Language Models (LLMs) and troubleshooting for model drifts, and generally monitoring and observing the model's performance. Examples start-ups that have received seed funding include Truera, Arize, Arthur, Hollistic AI, Babl AI, and Aporia whereas Apollo Research and METR, as we have already seen, have spun off from lucrative research groups and initiatives as non-profits. It is highly likely that most of the aforementioned start-ups as well as the traditional actors in the field will strive to claim role not only in the DSA Audit space but potentially in simultaneously in AI conformity assessments under the AI Act (even though third party audits are very, very rarely required by the proposed Regulation) [12, 14, 15]. These early developments seem to indicate a process of professionalisation not dissimilar to the industry transformations that have taken place in the fields of corporate sustainability, finance and accounting [28, 67, 85]. However, given the sheer size of the incumbent auditing players, and their history of aristocratic-style marriages and mergers to consolidate into the concentrated Big 4 we see today [62], we do have to ask how long such small players will even remain independent entities.

4.5 A sharp departure from idealised research-led audits

The terminology 'audit' has a non-conventional use in the academic research community studying algorithmic systems. It has often been used to indicate a study or investigation carried out by researchers, non-profits or by journalists [25, 115]. It has been understood as 'similar in spirit to the well-established practice of bug bounties', even in works that clearly contextualise audits in the context of their practices in other industries [111]. The observations from the emerging political economy of algorithmic auditing above pose challenges for this understanding and this paradigm. This requires particular care to ensure that the term 'audit', with the legitimacy awarded by scholarly communities such as FAccT, is not co-opted to mean something entirely different in practice [132]. At the same time, it requires consideration of how, if at all, this flavour of audit might become sustainable — more than an occasional research project using academics and journalists that soon, due to the incentives of their own roles, move to other challenges.

We have seen that in the world of audit, money and expertise moves erratically, and organisational dependencies, routines of corporate practices, and financial incentives can prevail. Such factors risk hindering the translation of the best analysis methods from scholarship to practice. The most challenging, yet most critical, socio-technical approaches may struggle to penetrate structures tailored to the intentions or desires of the audited or auditing entities [66, 89]. Financialisation and professionalisation may transform an inherently normative enquiry into an iteratively mundane practice of/for calculable, quantified deliverables. The regimes implied by the text of the DSA and OSA further push this direction. There might even be an element of complacency to this — so far, discussions of idealised, methodologically rich and creative algorithmic audits have rarely been linked to potential legislative or regulatory requirements. It is true that such creative audits have been promoted, including as 'soft' industry practices, but with less thought towards how they might practically be mandated.

Furthermore, a focus on public sector algorithmic systems in areas such as policing or welfare might have created confusion in this space. Where audits apply to discrete projects in public bodies, these might be one-off reports where the efficacy and appropriateness can be assessed by a court in case of a judicial review or similar public law challenge. For example, the audit documents accompanying live facial recognition tools in response to the *Bridges* case in England and Wales [3, 5, 103–105] or accompanying the England and Wales Ofqual COVID-19 exam results algorithm [94], did not follow a set standard, instead resembling more ad-hoc quality assurance and modelling criteria familiar from decades of government analytics and modelling practice [101, 123, 126]. But companies, compliance departments, audit industries, and insurers, do not operate in the same way as a public body can when using discretion when discharging their functions If audits are to be carried out regularly, across an entire industry, they institutionally crystallise in ways that make them distinct from the uses of societal or ethical analysis in high-stakes public sector machine learning contexts. This is not to say that public bodies do not find roles professionalised or institutionalised, but that rationality or reasonableness tests lend themselves can, in some ways, lend themselves to more flexible approaches, compared to logics of large multi-national business structures.

5 FOUNDATIONS FOR BETTER AUDIT TODAY

Despite the above, we want to end by highlighting several directions already in the DSA and the OSA which might, if given sharp attention by Digital Services Coordinators, Ofcom or the Commission, help mitigate some of the issues and challenges we have outlined above. Wholesale regulatory change is unlikely at this early stage, so our recommendations focus on what can be done immediately in these early, formative stages of a new regime.

5.0.1 Transparency. The first practicality refers to the transparency and identities of the people involved in the auditing process. A sentence that made it intact to the final text of the DRPA — despite consultation arguments from companies including Google — is art. 7(2). This Article provides that the agreement between the auditor and the audited organisation along with 'any other agreements or engagements letters [between them]' that are relevant to the performance of the audit shall be annexed to the final audit report. In parallel, according to both art. 7 and recital 15 of the DRPA, the auditing organisation is required to specify the details of the staff responsible for carrying out the audit. Such a direction is also hinted in the 'skilled person' profile envisaged by the OSA. Increasing the transparency around more than just the firm or the lead auditor, but the team, and entire supply chain, might go some way to help accountability issues, but only if regulators are willing to provide pressure and scrutiny on the disciplinary mix and methodological capabilities of the teams involved. This could be further supported

by guidelines or informal pressure on the direction of reports from year-to-year. There is also the possibility of undertaking analysis on the tendencies or impartiality of these audit reports, to consider if certain auditors systematically underplay certain risks, and then to provide that as aggregate feedback. Such a focus highlights the role of 'regulatory intermediaries' such as these auditors, both in general and in the algorithmic context [10, 85]. Rules on transparency at the level of auditing organisations are already set out for financial auditors in EU law (Directive 2014/56/EU), yet these would not apply to organisations either solely carrying out, or in respect of, algorithmic audits. As mentioned above, subcontractors are a particular liability in this regard. Legislators may wish to consider such changes depending on the way the industry grows.

5.0.2 Confidentiality. Secondly, contrary to financial auditing that feeds on public statements made by the audited organisations, algorithmic auditing will inevitably deal with information kept in private. In particular, according to art. 5(2) of the DPRA, the audited organisation will provide a wealth of information to the auditor including but not limited to information on 'decision-making structures, [...] relevant IT systems, data sources, [...] as well as explanations of relevant algorithmic systems and their interactions', as well as 'all data necessary for the performance of the audit', including documents, testing environments as well as 'personnel and premises of that provider, and any relevant sub-contractors'. There are explicit references to protections of confidentiality and trade secrets aimed at protecting commercially sensitive information of the audited organisation (art. 37(2) DSA and OFCOM Consultation 28.44-28.48), with Ofcom promising to remain 'mindful of the importance of protecting [such information]' (OFCOM Consultation 28.48)].

Under art. 37(2) DSA, requirements of confidentiality 'shall not adversely affect the performance of the audits'. Yet given how such information might reveal know-how both on underlying services and on the specific risk mitigation factors which other clients may benefit from their own auditors having knowledge of, it is easy to imagine organisations seeking to withhold commercially relevant information or seek extremely strong guarantees, where possible. Unlike in financial audits, where the base documents that may be required are likely to be more rote and predictable than not, the context–specific nature of sociotechnical algorithmic audits indicate that rigorous auditors need to both work out what documents might exist, and push for their release. Given the independence challenges discussed above, it is difficult to imagine this occurring with much ferocity. This is only compounded by art. 37(2) of the DSA, allowing auditors to publish a report while redacting what in their view is 'reasonably be considered to be confidential'.

Regulators have the space to issue guidance on this area that auditors can — and perhaps must — lean upon to do their work. As they gain more knowledge of the types of risks and modalities of analysis, they should publish types of information they expect auditors to have unfettered access to, and require auditors to list documents or other resources they sought access to but were refused. Such small moves might tip balances inside audited–auditing organisation relations to create more externally accountable patterns of engagement, particularly as confidentiality, and especially trade secrets, can be asserted with very limited recourse to externally check whether these claims are even grounded in law.

5.0.3 Timeframe. Finally, in line with the DSA's annual lifecycle for risk assessment, audits are expected to cover a period of one year (DSA art. 34). Given the highly programmable nature of information and computational production, having an annual audit review instead of a 'point-in-time' investigation clearly makes sense as it allows auditors to have a more holistic picture of a system's development lifecycle instead of merely capturing a snapshot of its functioning. However, precisely because of the amount of information that auditors will be expected to process and review, the agile way of internally negotiating and performing changes in an algorithmic system, as well as the tight timelines within which VLOPs/VLOSEs operate, auditors may confront an insurmountable amount of evidence to go through and not enough time. In this regard, neither the DSA nor the DPRA specify an indicative timeframe for the completion of the audit which is left to the discretion of the parties to determine (DRPA art. 7 (1)(d)). We believe that they should, or at least indicate a framework through which such a timeframe can be accountably determined.

6 CONCLUSION

Algorithmic auditing (and the closely associated conformity assessment standardisation 'market' for AI systems) is here to stay — but its form matters and it not yet set in stone. After almost a decade of academic and policy dialogue, we can now witness early signs of the real-world transformation. That is already a big step, but we must not get complacent. A systematic look on the relevant legislative provisions and the undergoing industry developments allows us to foresee (perhaps 'with a reasonable level of assurance'?) the actors involved in the burgeoning field of algorithmic auditing as well as the routines and cultures of professional practice that are likely to shape the field's future. In an ecosystem driven largely by a blend of traditional auditors, standardisation dynamics, share prices and venture capital, law is inevitably translated into (inter-)organisational projects. In this direction, audit independence, benchmark and methodological selection, the evaluation of the socio-economic context of the algorithm, and the disclosure of data to the auditor, are all becoming increasingly stabilised objects of/for negotiations amongst interdependent organisations. What used to be a space for research and contestation, thus transforms into a set of deliverables for professionals. Witnessing this reality unfolding might give researchers in the field a good reason to reflect on, and review the way they study, talk, and (co-) think about 'fairness, accountability, and transparency' in relation to the political economy of technology, and whose vision of 'auditing' they have in mind when they use that term. There is hope — and a legislative hook — for better practices of audit, but they cannot be taken for granted. In this paper, we have sought to lay out some practical critiques and directions for the near-term development and potential repositioning of this field, and hope that audit proponents can take these insights to make strategically powerful interventions in their own jurisdictions and spheres of scholarly and applied influence.

ACKNOWLEDGMENTS

This work is supported by the UKRI under Grant No.: EP/V00784X/1 (https://tas.ac.uk) Trustworthy Autonomous Systems Hub.

REFERENCES

[1] [n. d.]. Apollo Research. https://www.apolloresearch.ai/
[2] [n. d.]. ARC Evals - METR. https://metr.org/
[3] 2020. R (Bridges) v. Chief Constable of South Wales Police [2020] EWCA Civ 1058. https://caselaw.nationalarchives.gov.uk/ewca/civ/2020/1058
[4] 2022. Algorithmic Accountability Act 2022.
[5] 2022. *Standard Operating Procedure (SOP) for the Overt Deployment of Live Facial Recognition Technology (LFR).* Technical Report. https://www.met.police.uk/SysSiteAssets/media/downloads/force-content/met/advice/lfr/policy-documents/lfr-sop.pdf
[6] 2023. Algorithmic Accountability Act of 2023. *U.S.C.* 5, H.R. 5628 (21 Sep 2023), 553. https://www.govinfo.gov/app/details/BILLS-118hr5628ih
[7] 2023. Local Law Int. No. 1894-A (Automated Employment Decision Tools Law ("AEDT")), New York City.
[8] 2023. Online Safety Bill, Hansard (HL Vol 831, Col 385). https://hansard.parliament.uk/lords/2023-06-22/debates/C956350F-D70E-4409-8A4D-4ED05488C6DE/OnlineSafetyBill
[9] 2023. Responsible AI. https://kpmg.com/nl/en/home/services/advisory/trusted-enterprise/responsible-ai.html
[10] Kenneth W. Abbott, David Levi-Faur, and Duncan Snidal. 2017. Theorizing Regulatory Intermediaries: The RIT Model. *The ANNALS of the American Academy of Political and Social Science* 670, 1 (01 Mar 2017), 14–35. https://doi.org/10/f953b4
[11] Hal Abelson, Ross Anderson, Steven M. Bellovin, Josh Benaloh, Matt Blaze, Jon Callas, Whitfield Diffie, Susan Landau, Peter G. Neumann, Ronald L. Rivest, Jeffrey I. Schiller, Bruce Schneier, Vanessa Teague, and Carmela Troncoso. 2021. Bugs in Our Pockets: The Risks of Client-Side Scanning. arXiv:2110.07450v1 https://arxiv.org/abs/2110.07450v1
[12] BABL AI. n.d.. AI Audits. https://babl.ai/ai-audits/
[13] Fiddler AI. n.d.. AI Observability, ML Model Monitoring, and Explainable AI. https://www.fiddler.ai/
[14] Holistic AI. [n. d.]. Digital Services Act Audit - Solution. https://www.holisticai.com/digital-services-act-audit
[15] Holistic AI. n.d.. EU AI Act Readiness - Solutions. https://www.holisticai.com/eu-ai-act-readiness
[16] Naomi Appelman and Paddy Leerssen. 2022. On "Trusted" Flaggers. *Yale J.L. & Tech.* 24, 1 (2022), 452–475. https://yjolt.org/trusted-flaggers
[17] Jef Ausloos and Siddharth de Souza (Eds.). forthcoming. *Researcher Access to Digital Infrastructures.* Cambridge University Press, Cambridge.
[18] Jef Ausloos and Michael Veale. 2020. Researching with Data Rights. *Technology and Regulation* 1 (2020), 136–157. https://doi.org/10.26116/techreg.2020.010
[19] C. Richard Baker, Jean B´edard, and Christian Prat dit Hauret. 2014. The Regulation of Statutory Auditing: An Institutional Theory Approach. *Managerial Auditing Journal* 29, 5 (01 Jan 2014), 371–394. https://doi.org/10.1108/MAJ-09-2013-0931
[20] Agathe Balayn and Seda Gürses. 2021. Beyond Debiasing: Regulating AI and Its Inequalities. https://perma.cc/4UAV-3UFB
[21] Agathe Balayn, Mireia Yurrita, Jie Yang, and Ujwal Gadiraju. 2023. "Fairness Toolkits, A Checkbox Culture?" On the Factors That Fragment Developer Practices in Handling Algorithmic Harms. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. Association for Computing Machinery, New York, NY, USA, 482–495. https://doi.org/10.1145/3600211.3604674
[22] Jack Bandy. 2021. Problematic Machine Behavior: A Systematic Literature Review of Algorithm Audits. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1 (22 Apr 2021), 74:1–74:34. https://doi.org/10.1145/3449148
[23] Nathan Bartley, Andres Abeliuk, Emilio Ferrara, and Kristina Lerman. 2021. Auditing Algorithmic Bias on Twitter. 65–73.
[24] Silke Beck, Sheila Jasanoff, Andy Stirling, and Christine Polzin. 2021. The Governance of Sociotechnical Transformations to Sustainability. *Current Opinion in Environmental Sustainability* 49 (Apr 2021), 143–152. https://doi.org/10.1016/j.cosust.2021.04.010
[25] Abeba Birhane, Vinay Uday Prabhu, and John Whaley. 2022. Auditing Saliency Cropping Algorithms. 4051–4059. https://openaccess.thecvf.com/content/WACV2022/html/Birhane_Auditing_Saliency_Cropping_Algorithms_WACV_2022_paper.html
[26] Su Lin Blodgett, Solon Barocas, Hal Daum´e III, and Hanna Wallach. 2020. Language (Technology) Is Power: A Critical Survey of "Bias" in NLP. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 5454–5476. https://doi.org/10.18653/v1/2020.acl-main.485
[27] Booking.com. 2023. *Submissions by Booking.Com in Response to the European Commission's Request for Feedback on the Draft DRPA.* Technical Report. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13626-Digital-Services-Act-conducting-independent-audits/F3424053_en
[28] Sebastian Botzem. 2012. *The Politics of Accounting Regulation: Organizing Transnational Standard Setting in Financial Reporting.* Edward Elgar Publishing.
[29] Ian Brown and Christopher T. Marsden. 2013. *Regulating Code: Good Governance and Better Regulation in the Information Age.* MIT Press, Cambridge, MA.
[30] Shea Brown, Jovana Davidovic, and Ali Hasan. 2021. The Algorithm Audit: Scoring the Algorithms That Score Us. *Big Data & Society* 8, 1 (2021), 2053951720983865.
[31] Miriam C. Buiten. 2022. The Digital Services Act: From Intermediary Liability to Platform Regulation. *JIPITEC* 12, 5 (2022). https://www.jipitec.eu/issues/jipitec-12-5-2021/5491
[32] Sarmad Chandio, Daniyal Pirwani Dar, and Rishab Nithyanand. 2023. How Auditing Methodologies Can Impact Our Understanding of YouTube's Recommendation Systems. *arXiv:2303.03445* (2023). https://doi.org/10.48550/arXiv.2303.03445
[33] Rumman Chowdhury. 2023. What's an Audit? https://www.get-parity.com/raiblog/whats-an-audit
[34] Athena Christofi, Jonas Breuer, Ellen Wauters, Peggy Valcke, and Jo Pierson. 2022. Data Protection, Control and Participation beyond Consent - Seeking the Views of Data Subjects in Data Protection Impact Assessments. In *Research Handbook on EU Data Protection Law*. Edward Elgar Publishing, 503–529. https://doi.org/10.4337/9781800371682.00029
[35] Jennifer Cobbe and Jatinder Singh. 2021. Artificial Intelligence as a Service: Legal Responsibilities, Liabilities, and Policy Challenges. *Computer Law & Security Review* 42 (2021), 105573. https://doi.org/10/gmq8jm
[36] Jennifer Cobbe, Michael Veale, and Jatinder Singh. 2023. Understanding Accountability in Algorithmic Supply Chains. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, New York, NY, USA, 1186–1197. https://doi.org/gsb98p
[37] European Commission. [n. d.]. The EU's Digital Services Act. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en
[38] Federal Trade Commission. 2022. Combatting Online Harms Through Innovation. https://perma.cc/2S5Y-WQKX
[39] Sasha Costanza-Chock, Inioluwa Deborah Raji, and Joy Buolamwini. 2022. Who Audits the Auditors? Recommendations from a Field Scan of the Algorithmic Auditing Ecosystem. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, 1571–1583. https://doi.org/10.1145/3531146.3533213
[40] Alexander D'Amour, Hansa Srinivasan, James Atwood, Pallavi Baljekar, D. Sculley, and Yoni Halpern. 2020. Fairness Is Not Static: Deeper Understanding of Long Term Fairness via Simulation Studies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, New York, NY, USA, 525–534. https://doi.org/10.1145/3351095.3372878
[41] Mark Dangelo. 2023. Auditing AI: The emerging battlefield of transparency and assessment. https://www.thomsonreuters.com/en-us/posts/technology/auditing-ai-transparency/
[42] Philipp Darius, Daniela Stockmann, Joanna Bryson, Luciana Cingolani, Rachel Griffin, Gerhard Hammerschmid, Maximilian Kupi, Haytham Mones, Simon Munzert, and Rónán Riordan. 2023. Implementing Data Access of the Digital Services Act: Collaboration of European Digital Service Coordinators and Researchers in Building Strong Oversight over Social Media Platforms. https://perma.cc/P2MM-QWUP
[43] Katarina Demetzou. 2022. Introduction to the Conformity Assessment under the Draft EU AI Act, and How It Compares to DPIAs. https://fpf.org/blog/introduction-to-the-conformity-assessment-under-the-draft-eu-ai-act-and-how-it-compares-to-dpias/
[44] Daria Dergacheva, Christian Katzenbach, Sebastian Felix Schwemer, and João Pedro Quintais. 2023. Improving Data Access for Researchers in the Digital Services Act. *SSRN Scholarly Paper* (01 Jun 2023). https://doi.org/10.2139/ssrn.4465846
[45] Paul J. DiMaggio and Walter W. Powell. 1983. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review* 48, 2 (1983), 147. https://doi.org/10.2307/2095101
[46] Denise Hanes Downey and Kimberly D. Westermann. 2021. Challenging Global Group Audits: The Perspective of US Group Audit Leads. *Contemporary Accounting Research* 38, 2 (2021), 1395–1433. https://doi.org/10.1111/1911-3846.12648
[47] Lilian Edwards. 2019. 'With Great Power Comes Great Responsibility?': The Rise of Platform Liability. In *Law, Policy, and the Internet*, Lilian Edwards (Ed.). Hart Publishing, 253–289.
[48] Randal J. Elder, Mark S. Beasley, Chris E. Hogan, and Alvin A. Arens. 2020. *Auditing and Assurance Services: International Perspectives* (17 ed.). Pearson.
[49] Ernst and Young. [n. d.]. EY Trusted AI Platform. https://www.ey.com/en_gl/consulting/trusted-ai-platform
[50] Ilana Golbin. 2021. Algorithmic Impact Assessments: What Are They and Why Do You Need Them? https://perma.cc/6CFY-MYNZ

[51] Ellen P. Goodman and Julia Tr´ehu. 2022. AI Audit-Washing and Accountability. https://www.gmfus.org/news/ai-audit-washing-and-accountability

[52] Ellen P. Goodman and Julia Tr´ehu. 2023. Algorithmic Auditing: Chasing AI Accountability. *Santa Clara High Technology Law Journal* 39, 3 (2023), 289.

[53] Charles Goodwin. 1994. Professional Vision. *American Anthropologist* 96, 3 (1994), 606–633. https://www.jstor.org/stable/682303

[54] Google. 2023. Submissions by Google in Response to the European Commission's Request for Feedback on the Draft DRPA. https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13626-Digital-Services-Act-conducting-independent-audits/F3424072_en

[55] Robert Gorwa. 2024. *The Politics of Platform Regulation: How Governments Shape Online Content Moderation.* Oxford University Press, Oxford, UK.

[56] Robert Gorwa, Reuben Binns, and Christian Katzenbach. 2020. Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance. *Big Data & Society* 7, 1 (2020), 2053951719897945. https://doi.org/10/ggsfrk

[57] Robert Gorwa and Michael Veale. 2024. Moderating Model Marketplaces: Platform Governance Puzzles for AI Intermediaries. *Law, Innovation and Technology* 16, 2 (2024). https://doi.org/10.31235/osf.io/6dfk3

[58] Australian Government. 2023. Royal Commission into the Robodebt Scheme. https://www.pmc.gov.au/sites/default/files/resource/download/gov-response-royal-commission-robodebt-scheme.pdf

[59] HM Government. 2017. The Futures Toolkit: Tools for Futures Thinking and Foresight across UK Government. https://assets.publishing.service.gov.uk/media/5a821fdee5274a2e8ab579ef/futures-toolkit-edition-1.pdf

[60] HM Government. 2017. Internet Safety Strategy – Green paper. https://www.gov.uk/government/consultations/internet-safety-strategy-green-paper

[61] HM Government. 2022. Impact Assessment: The Online Safety Bill. https://publications.parliament.uk/pa/bills/cbill/58-02/0285/210285en.pdf

[62] Ian D. Gow and Stuart Kells. 2018. *The Big Four: The Curious Past and Perilous Future of the Global Accounting Monopoly.* Berrett-Koehler Publishers, a BK Business book, Oakland.

[63] Lara Groves, Jacob Metcalf, Alayna Kennedy, Briana Vecchione, and Andrew Strait. 2024. Auditing Work: Exploring the New York City Algorithmic Bias Audit Regime. In *Proceedings of the 2024 Conference on Fairness, Accountability, and Transparency in Algorithmic Systems (FAccT'24).* ACM, Rio de Janeiro, Brazil. https://doi.org/10.48550/arXiv.2402.08101

[64] Eva Hartmann and Poul F. Kjaer (Eds.). 2015. *The Evolution of Intermediary Institutions in Europe.* Palgrave Macmillan UK, London. https://doi.org/10.1057/9781137484529

[65] Marius Hobbhahn. 2021. Announcing Apollo Research. https://forum.effectivealtruism.org/posts/ysC6crBKhDBGZfob3/announcing-apollo-research

[66] Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daum´e, Miro Dudik, and Hanna Wallach. 2019. Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3290605.3300830

[67] Christopher Humphrey, Anne Loft, and Margaret Woods. 2009. The Global Audit Profession and the International Financial Architecture: Understanding Regulatory Relationships at a Time of Financial Crisis. *Accounting, Organizations and Society* 34, 6-7 (2009), 810–825. https://doi.org/10.1016/j.aos.2009.06.003

[68] Basileal Imana, Aleksandra Korolova, and John Heidemann. 2023. Having Your Privacy Cake and Eating It Too: Platform-supported Auditing of Social Media Algorithms for Public Interest. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–33.

[69] Petros Iosifidis. 2011. The Public Sphere, Social Networks and Public Service Media. *Information, Communication & Society* 14, 5 (2011), 619–637.

[70] Sheila Jasanoff and Sang-Hyun Kim. 2015. *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power.* University of Chicago Press, Chicago, IL. https://doi.org/10.7208/chicago/9780226276663.001.0001

[71] Kate Kaye. 2022. A new wave of AI auditing startups wants to prove responsibility can be profitable. *Protocol* (03 Jan 2022). https://web.archive.org/web/20230416215440/https://www.protocol.com/enterprise/ai-audit-2022

[72] Poul F. Kjaer (Ed.). 2020. *The Law of Political Economy: Transformation in the Function of Law.* Cambridge University Press, Cambridge. https://doi.org/10.1017/9781108675635

[73] Allison Koenecke, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Toups, John R. Rickford, Dan Jurafsky, and Sharad Goel. 2020. Racial Disparities in Automated Speech Recognition. *Proceedings of the National Academy of Sciences* 117, 14 (07 Apr 2020), 7684–7689. https://doi.org/10.1073/pnas.1915768117

[74] Johann Laux, Sandra Wachter, and Brent Mittelstadt. 2021. Taming the Few: Platform Regulation, Independent Audits, and the Risks of Capture Created by the DMA and DSA. *Computer Law & Security Review* 43 (2021), 105613.

[75] Paddy Leerssen. 2023. *Seeing What Others Are Seeing: Studies in the Regulation of Transparency for Social Media Recommender Systems.* Ph. D. Dissertation. University of Amsterdam.

[76] Paddy Leerssen, Amélie P. Heldt, and Matthias C. Kettemann. 2023. Scraping By? Europe's Law and Policy on Social Media Research Access. In *Challenges and Perspectives of Hate Speech Research*, Christian Strippel, Sünje Paasch-Colberg, Martin Emmer, and Joachim Trebbe (Eds.). Vol. 12. Berlin, 405–425. https://doi.org/10.48541/dcr.v12.24

[77] Deloitte Consulting LLP. 2020. Deloitte Introduces Trustworthy AI Framework to Guide Organizations in Ethical Application of Technology in the Age of With. https://perma.cc/RZ3Q-UD47

[78] Deloitte Consulting LLP. 2021. Deloitte AI Institute Teams With Chatterbox Labs to Ensure Ethical Application of AI. https://perma.cc/DC8D-QAS5

[79] Natasha Lomas. 2023. Security Researchers Latest to Blast UK's Online Safety Bill as Encryption Risk. https://techcrunch.com/2023/07/05/uk-online-safety-bill-risks-e2ee/

[80] Daithí Mac Síthigh. 2020. The Road to Responsibilities: New Attitudes Towards Internet Intermediaries. *Information & Communications Technology Law* 29, 1 (2020), 1–21. https://doi.org/10/ggkhjn

[81] Michael A. Madaio, Luke Stark, Jennifer Wortman Vaughan, and Hanna Wallach. 2020. Co-Designing Checklists to Understand Organizational Challenges and Opportunities around Fairness in AI. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376445

[82] Katherine R. Maffey, Kyle Dotterrer, Jennifer Niemann, Iain Cruickshank, Grace A. Lewis, and Christian Kästner. 2023. MLTEing Models: Negotiating, Evaluating, and Documenting Model and System Qualities. In *2023 IEEE/ACM 45th International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER).* 31–36. https://doi.org/10.1109/ICSE-NIER58687.2023.00012

[83] Chris Marsden, Trisha Meyer, and Ian Brown. 2020. Platform Values and Democratic Elections: How Can the Law Regulate Digital Disinformation? *Computer Law & Security Review* 36 (01 Apr 2020), 105373. https://doi.org/10.1016/j.clsr.2019.105373

[84] Christopher T. Marsden. 2011. *Internet Co-Regulation: European Law, Regulatory Governance and Legitimacy in Cyberspace.* Cambridge University Press.

[85] Kira J. M. Matus and Michael Veale. 2022. Certification Systems for Machine Learning: Lessons from Sustainability. *Regulation & Governance* 16, 1 (2022), 177–196. https://doi.org/10.1111/rego.12417

[86] Anna-Katharina Meßmer and Martin Degeling. 2023. Auditing Recommender Systems–Putting the DSA into Practice with a Risk-Scenario-Based Approach. *Stiftung Neue Verantwortung* (2023). https://doi.org/10.48550/arXiv.2302.04556

[87] Jacob Metcalf, Emanuel Moss, and Danah Boyd. 2019. Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics. *Social Research: An International Quarterly* 86, 2 (2019), 449–476. https://doi.org/10.1353/sor.2019.0022

[88] Shira Mitchell, Eric Potash, Solon Barocas, Alexander D'Amour, and Kristian Lum. 2021. Algorithmic Fairness: Choices, Assumptions, and Definitions. *Annu. Rev. Stat. Appl.* 8, 1 (07 Mar 2021), 141–163. https://doi.org/10.1146/annurev-statistics-042720-125902

[89] Nadia Nahar, Shurui Zhou, Grace Lewis, and Christian Kästner. 2022. Collaboration Challenges in Building ML-enabled Systems: Communication, Documentation, Engineering, and Process. In *Proceedings of the 44th International Conference on Software Engineering.* Association for Computing Machinery, New York, NY, USA, 413–425. https://doi.org/10.1145/3510003.3510209

[90] Luca Nannini, Agathe Balayn, and Adam Leon Smith. 2023. Explainability in AI Policies: A Critical Review of Communications, Reports, Regulations, and Standards in the EU, US, and UK. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency.* Association for Computing Machinery, New York, NY, USA, 1198–1212. https://doi.org/10.1145/3593013.3594074

[91] Chris Norval, Richard Cloete, and Jatinder Singh. 2023. Navigating the Audit Landscape: A Framework for Developing Transparent and Auditable XR. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency.* Association for Computing Machinery, New York, NY, USA, 1418–1431. https://doi.org/10.1145/3593013.3594090

[92] Government of Canada. 2022. Summary: Regulatory Powers. https://www.canada.ca/en/canadian-heritage/campaigns/harmful-online-content/summary-session-four.html

[93] Ofcom. 2023. Protecting People from Illegal Harms Online - Summary of Each Chapter. https://www.ofcom.org.uk/__data/assets/pdf_file/0030/270948/chapter-summaries-illegal-harms-consultation.pdf

[94] Ofqual. 2020. *Awarding GCSE, AS, A Level, Advanced Extension Awards and Extended Project Qualifications in Summer 2020: Interim Report.* Technical Report. https://perma.cc/8QY4-VW4G

[95] Victor Ojewale, Ryan Steed, Briana Vecchione, Abeba Birhane, and Inioluwa Deborah Raji. 2024. Towards AI Accountability Infrastructure: Gaps and Opportunities in AI Audit Tooling. https://doi.org/10.48550/arXiv.2402.17861 arXiv:2402.17861

[96] Chinasa T. Okolo, Nicola Dell, and Aditya Vashistha. 2022. Making AI Explainable in the Global South: A Systematic Review. In *Proceedings of the 5th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies.* Association for Computing Machinery, New York, NY, USA, 439–452. https:

//doi.org/10.1145/3530190.3534802
[97] Rock Yuren Pang, Jack Cenatempo, Franklyn Graham, Bridgette Kuehn, Maddy Whisenant, Portia Botchway, Katie Stone Perez, and Allison Koenecke. 2023. Auditing Cross-Cultural Consistency of Human-Annotated Labels for Recommendation Systems. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, New York, NY, USA, 1531–1552. https://doi.org/10.1145/3593013.3594098
[98] European Parliament. 2023. European Health Data Space Regulation (Amendments Adopted by the European Parliament). https://www.europarl.europa.eu/doceo/document/TA-9-2023-0462_EN.html
[99] Claudia Peersman, Jos´e Tomas Llanos, Corinne May-Chahal, Ryan McConville, Partha Das Chowdhury, and Emiliano De Cristofaro. 2023. Towards a Framework for Evaluating CSAM Prevention and Detection Tools in the Context of End-to-end Encryption Environments: A Case Study. https://www.rephrain.ac.uk/safety-tech-challenge-fund/
[100] Authoriteit Persoonsgegevens. 2024. AI & Algorithmic Risks Report Netherlands - Winter 2023 2024. https://perma.cc/M2JU-GCVU
[101] Arthur C. Petersen, P. H. M. Janssen, JP van der Sluijs, J. S. Risbet, J. R. Ravetz, J. Arjan Wardekker, and H. Martison Hughes. 2013. *Guidance for Uncertainty Assessment and Communication.* PBL Netherlands Environmental Assessment Bureau, The Hague, NL. https://perma.cc/N9FJ-RGXU
[102] Joseph Phiri and Pinar Guven-Uslu. 2019. Social Networks, Corruption and Institutions of Accounting, Auditing and Accountability. *Accounting, Auditing & Accountability Journal* 32, 2 (2019), 508–530.
[103] Metropolitan Police. 2022. *Understanding Accuracy and Bias.* Technical Report. https://www.met.police.uk/SysSiteAssets/media/downloads/force-content/met/advice/lfr/other-lfr-documents/lfr-accuracy-and-demographic-differential.pdf
[104] South Wales Police. 2021. Equality Impact Assessment: A Practical Tool to Identify Discrimination. https://perma.cc/E7AU-VFNJ
[105] South Wales Police. 2023. *Policy Document for the Overt Deployment of Live Facial Recognition (LFR) Technology.* Technical Report. https://perma.cc/H5NF-D9QT
[106] Martha Poon. 2007. Scorecards as Devices for Consumer Credit: The Case of Fair, Isaac & Company Incorporated. *The Sociological Review* 55, 2_suppl (2007), 284–306. https://doi.org/10/bbzb34
[107] Martha Poon. 2009. From New Deal Institutions to Capital Markets: Commercial Consumer Risk Scores and the Making of Subprime Mortgage Finance. *Accounting, Organizations and Society* 34, 5 (2009), 654–674. https://doi.org/10/cm8g3x
[108] Inioluwa Deborah Raji. 2022. From Algorithmic Audits to Actual Accountability: Overcoming Practical Roadblocks on the Path to Meaningful Audit Interventions for AI Governance. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. Association for Computing Machinery, New York, NY, USA, 5. https://doi.org/10.1145/3514094.3539566
[109] Inioluwa Deborah Raji. 2022. It's Time to Develop the Tools We Need to Hold Algorithms Accountable. https://foundation.mozilla.org/en/blog/its-time-to-develop-the-tools-we-need-to-hold-algorithms-accountable/
[110] Inioluwa Deborah Raji, Timnit Gebru, Margaret Mitchell, Joy Buolamwini, Joonseok Lee, and Emily Denton. 2020. Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing. https://doi.org/10.48550/arXiv.2001.00964 arXiv:2001.00964 preprint.
[111] Inioluwa Deborah Raji, Andrew Smart, Rebecca N. White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. 2020. Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, New York, NY, USA, 33–44. https://doi.org/10.1145/3351095.3372873
[112] Bogdana Rakova, Jingying Yang, Henriette Cramer, and Rumman Chowdhury. 2021. Where Responsible AI Meets Reality: Practitioner Perspectives on Enablers for Shifting Organizational Practices. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1 (22 Apr 2021), 7:1–7:23. https://doi.org/10.1145/3449081
[113] Varun Ramdas. 2022. Identifying an Actionable Algorithmic Transparency Framework: A Comparative Analysis of Initiatives to Enhance Accountability of Social Media Platforms. *Nat'l LU Delhi Stud. LJ* 4 (2022), 74.
[114] Howard Rosenbaum and Pnina Fichman. 2019. Algorithmic Accountability and Digital Justice: A Critical Assessment of Technical and Sociotechnical Approaches. *Proceedings of the Association for Information Science and Technology* 56, 1 (2019), 237–244.
[115] Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. In *Data and Discrimination: Converting Critical Concerns into Productive: A Preconference at the 64th Annual Meeting of the International Communication Association. Seattle*, Vol. 22. 4349–4357.
[116] D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-Francois Crespo, and Dan Dennison. 2015. Hidden Technical Debt in Machine Learning Systems. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*. MIT Press, Cambridge, MA, USA, 2503–2511. https://proceedings.neurips.cc/paper_files/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf
[117] Nick Seaver. 2017. Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems. *Big Data & Society* 4, 2 (2017), 205395171773810. https://doi.org/10/gd8fdx
[118] Hong Shen, Alicia DeVos, Motahhare Eslami, and Kenneth Holstein. 2021. Everyday Algorithm Auditing: Understanding the Power of Everyday Users in Surfacing Harmful Algorithmic Behaviors. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–29.
[119] Jessie J. Smith, Anas Buhayh, Anushka Kathait, Pradeep Ragothaman, Nicholas Mattei, Robin Burke, and Amy Voida. 2023. The Many Faces of Fairness: Exploring the Institutional Logics of Multistakeholder Microlending Recommendation. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery, New York, NY, USA, 1652–1663. https://doi.org/10.1145/3593013.3594106
[120] Kevin B. Sobel-Read. 2020. Reimagining the Unimaginable: Law and the Ongoing Transformation of Global Value Chains into Integrated Legal Entities. *European Review of Contract Law* 16, 1 (07 Apr 2020), 160–185. https://doi.org/10.1515/ercl-2020-0009
[121] Dan Sunderland and Gregory M. Trompeter. 2017. Multinational Group Audits: Problems Faced in Practice and Opportunities for Research. *Auditing: A Journal of Practice & Theory* 36, 3 (2017), 159–183. https://doi.org/10.2308/ajpt-51667
[122] Petros Terzis and (Enrique) OE Santamaria Echeverria. 2023. Interoperability and Governance in the European Health Data Space Regulation. *Medical Law International* 23, 4 (2023), 368–376. https://doi.org/10.1177/09685332231165692
[123] HM Treasury. 2015. *The Aqua Book: Guidance on Producing Quality Analysis for Government.* HM Government, London.
[124] Roberto Ulloa, Mykola Makhortykh, and Aleksandra Urman. 2022. Scaling up Search Engine Audits: Practical Insights for Algorithm Auditing. *Journal of Information Science* (02 May 2022), 01655515221093029. https://doi.org/10.1177/01655515221093029
[125] Michael Veale. forthcoming. Denied by Design? Data Access Rights in Encrypted Infrastructures. In *Researcher Access to Digital Infrastructures*, Jef Ausloos and Siddharth de Souza (Eds.). Cambridge University Press, Cambridge. https://doi.org/k5mk
[126] Michael Veale and Irina Brass. 2019. Administration by Algorithm? Public Management Meets Public Sector Machine Learning. In *Algorithmic Regulation*, Karen Yeung and Martin Lodge (Eds.). Oxford University Press, Oxford, 121–149. https://doi.org/10/gfzvz8
[127] Michael Veale and Frederik Zuiderveen Borgesius. 2021. Demystifying the Draft EU Artificial Intelligence Act. *Computer Law Review International* 22, 4 (2021), 97–112. https://doi.org/10/gns2s9
[128] Kaiser Wolfram and Johan Schot. 2014. *Writing the Rules for Europe.* Palgrave Macmillan.
[129] Richmond Y. Wong, Michael A. Madaio, and Nick Merrill. 2023. Seeing Like a Toolkit: How Toolkits Envision the Work of AI Ethics. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1 (16 Apr 2023), 145:1–145:27. https://doi.org/10.1145/3579621
[130] Lorna Woods and Will Perrin. 2021. Obliging Platforms to Accept a Duty of Care. In *Regulating Big Tech*, Martin Moore and Damian Tambini (Eds.). Oxford University Press, Oxford, 93–109. https://doi.org/10.1093/oso/9780197616093.003.0006
[131] Lorna Woods, Will Perrin, and Maeve Walsh. 2019. Draft Online Harm Reduction Bill. https://carnegieuktrust.org.uk/publications/draft-online-harm-bill/
[132] Meg Young, Michael Katell, and P. M. Krafft. 2022. Confronting Power and Corporate Capture at the FAccT Conference. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Seoul, 1375–1386. https://doi.org/10.1145/3531146.3533194
[133] Skye Zhu and Soon-Yeow Phang. 2024. How Do Lead Auditor Instructions Influence Component Auditors' Evidence Collection Decisions? The Joint Influence of Construal Interpretations and Responsibility. *Contemporary Accounting Research* 41, 1 (2024), 591–619. https://doi.org/10.1111/1911-3846.12911