Digital Regulation Cooperation Forum

CMA — Competition & Markets Authority · Ofcom — making communications work for everyone · ICO. — Information Commissioner's Office · FCA — FINANCIAL CONDUCT AUTHORITY

DRCF Algorithmic Processing workstream: Response to Call for Input Submissions

1. SUMMARY

In April 2022, the Digital Regulation Cooperation Forum (DRCF) published two papers related to algorithmic processing. One looked at the specific **harms and benefits** posed by algorithmic systems, including those powered by machine learning methods. The other looked at the merits of **algorithmic auditing**, as a means of documenting the risks associated with these systems and assuring stakeholders that they are safe to use.

The purpose of this research was to reveal areas of common interest among the DRCF members, ensuring that any future interventions we make as individual regulators are aligned and complementary. Our intention was also to highlight emerging ideas for how to mitigate the risks posed by algorithms, and to signal to industry those solutions that appear the most promising, and which we would like to explore further. This includes interventions that promote transparency for users and developers (e.g. via model and system cards), that give people more control over the systems that shape their daily lives, and that enable people to seek redress where they have been harmed.

At the same time as publishing these papers, we launched a **"call for input"** where we invited views from stakeholders on our findings and conclusions. In the case of the paper on algorithmic auditing, we specifically asked for views on several hypotheses, which set out a vision for the role that could be played by regulators in using algorithmic audits and supporting the development of a third party auditing market. We also asked stakeholders for their opinions on where the DRCF should focus its energy in the future.

In total, we received 31 responses to our call for input. Annex A lists the names of those organisations who have given their permission to be cited. We received submissions from stakeholders in the public, private and third sectors, and from both large and small organisations. Among them were social media platforms, vendors of AI systems, consumer groups, management consultants, industry trade bodies and chartered professional institutes. Some respondents provided views on both papers, while others focused their feedback on one or the other, occasionally concentrating on a single issue.

The majority of respondents were supportive of the two papers and welcomed the four digital regulators collaborating on this topic. There was widespread agreement that algorithmic processing creates opportunities and risks across the entire economy, and that as such there is value in regulators pooling knowledge and working together to promote best practice.

With regard to our first paper on the **benefits and harms of algorithmic processing**, most respondents felt that this had captured the majority of pressing issues, among them risks to fairness, transparency and competition. However, some stakeholders felt that our analysis could have paid closer attention to the impact of algorithms on the environment, as well as to workers and their experiences of the labour market. While some respondents disagreed with several claims made in the paper, for example that the use of algorithms on social media may be driving a polarisation of views in society, in most cases the feedback we received sought to build on our analysis, as well as to highlight some of the challenges in addressing the risks we identified. To take one example, several stakeholders cautioned that efforts to improve algorithmic transparency could prove counterproductive if they allow systems to be more easily gamed.

We received a similar set of responses to our paper on **algorithmic auditing.** Stakeholders were generally supportive of the hypotheses outlined by the paper, including (H1) that there would be value in regulators clarifying how third party audits could support the regulatory process; and (H2) that there may be a role for some regulators in producing guidance on how third parties should conduct audits. However, not all the ideas raised within this paper drew support. For instance, a number of stakeholders cautioned against the idea of creating mechanisms for information sharing between third party auditors and regulators due primarily to security and privacy concerns. Several also underscored the challenges involved in accrediting auditors, which could be significantly costly for regulators and create disproportionate burdens for auditing firms.

In addition to sharing views on the contents of each paper, respondents provided comments on the wider DRCF workplan, as well as who we should be working with to maximise our impact. Some stakeholders stressed the importance of the DRCF engaging with all types of organisation, small as well as large. Several called for greater collaboration with other regulators, particularly the Equality and Human Rights Commission who the DRCF already engages with via its regulatory round table. Respondents' ideas for future DRCF projects varied considerably, however a recurring theme throughout the submissions was a request for the DRCF to help clarify what organisations need to do in order to comply with regulation when they build or use algorithmic systems.

We will use the feedback gathered through this exercise to inform the next stage of the DRCF project on algorithmic processing and relevant work across the individual regulators. Over the coming work programme year (2022/23), we plan to undertake several new activities, including a stream of work looking at how to improve the procurement of AI systems (such as content moderation or age verification technology), and another stream of work where the four regulators will share lessons learnt and best practice in how to conduct algorithmic audits.

To keep updated on progress, please continue to check the [DRCF website](), or email drcf.algorithms@drcf.org.uk with any specific enquiries.

Box 1: Make up of respondents

- In total we received 31 responses to our Call for Input exercise (please see the Annex for a list of the organisations who provided a substantive response and who gave permission to be named within this document).

- 48% provided commentary on both papers, while 52% provided a response to one of the two.

- 16% of these were from private individuals, 26% were from private firms, and 58% were from other types of organisation, including civil society groups.

- 50% of responses from organisations came from organisations that operate globally, while 50% were from organisations that are predominantly active in the UK.

Box 2: Definition of algorithmic processing

We define algorithmic processing as the processing of data (both personal and non-personal) by automated systems. This includes artificial intelligence (AI) applications, such as those powered by machine learning (ML) techniques, but also simpler statistical models. Our interest covers the processing of data, as well as the context in which that processing occurs, such as the means used to collect and store that data, and the ways humans interact with the results of any processing. Algorithmic processing can be used both to produce an output (e.g. video or text content) and to make or inform decisions that have a direct bearing on individuals.

Real world examples of algorithmic processing include:

- Detecting fraudulent activity in someone's bank transactions

- Translating a foreign news site into English language

- Recommending content on streaming services or social media news feeds

- Estimating someone's age where they seek to buy products for over 18s

- Making a prediction about the likelihood of a customer defaulting on a loan

---

2. PAPER #1 - THE BENEFITS AND HARMS OF ALGORITHMIC PROCESSING

Box 3: A reminder of 'Benefits and Harms' paper's key messages:

- Algorithmic processing has the potential to deliver significant benefits, from detecting fraudulent activity in financial services to translating languages at the click of a button

- However, it also poses several risks. When deployed without due care, for example, algorithmic processing can amplify bias in decision-making, and lead to the collection of large volumes of data that could put people's privacy in jeopardy
- The paper highlights six areas of concern and provides examples of risks in relation to each. These are: transparency; fairness; access to information and goods; resilience of infrastructure (inc. cyber security issues); individual autonomy; and competition
- In each of these areas, algorithmic processing can generate harms as well as benefits. With regard to resilience, for example, algorithms can be weaponised for the purpose of launching cyber attacks, yet they can also be used to bolster cyber defence systems
- The paper finishes by reflecting on the factors that lie behind the misuse of algorithms, and what might be required to ensure they are used more responsibly
- One key takeaway is that many of the harms caused by algorithmic processing are likely to be inadvertent, driven less by malicious acts and more by a lack of awareness among those using them. We also flag that those purchasing algorithmic systems often know little about their origins and limitations, making it more difficult for them to mitigate risks

Issues highlighted in the paper that resonated with respondents

Respondents agreed with several of the findings and arguments presented in the paper, among them:

1. The importance of considering context when determining risk

In our paper we stressed the importance of considering context when evaluating the risks posed by algorithmic processing. We cautioned against viewing individual technologies as inherently high risk, arguing that what matters is how those technologies are used, by whom and for what purpose. Many of our respondents shared this sentiment emphasising that it was the use case, or combination of algorithm and use case that should be assessed, potentially on a case by case basis, rather than the algorithm (or type of algorithm) alone. One stakeholder, for example, noted that the use of algorithms to generate synthetic media is not always done for malicious reasons; while algorithms can be used to generate visual forms of disinformation (e.g. by depicting a politician saying or doing something incendiary in a "deepfake" video), they can also be used to create special visual effects in film and tv production. Another example given was the use of facial recognition software. One stakeholder claimed that while this technology is often seen as high risk due to its use in law enforcement and surveillance settings, it can also be used for more benign purposes, for example to allow people to verify their identity when logging into a banking app.

2. *The false belief that human oversight is a foolproof way of mitigating risks*

We argued in our paper that the impact of algorithmic systems cannot be understood without taking into account their users. While the outputs they generate are important, we also need to reflect on whether they can be properly interpreted and acted upon. This view was shared by some of our respondents, including one large firm that said algorithmic systems should always be viewed through a "socio-technical lens". Related to this point, many respondents agreed with our claim that having a "human in the loop" is not a foolproof safeguard to protect against harm. One stakeholder noted their belief that human involvement might even increase risks of bias and inaccuracy, for example in the context of identity verification, as well as slow down service delivery. As this

stakeholder put it, "we must not introduce rules that may appear to be strengthening safeguards, but actually serve only to increase the cost and complexity of solutions without value…".

Our suggestion that further research be carried out on the topic of "human computer interaction" was welcomed by stakeholders. It was noted in one response that the National Institute for Standards and Technology in the US is seeking to develop formal guidance about how to implement human-in-the-loop processes that do not amplify or perpetuate bias.

3. Transparency is critical but there is no one-size-fits-all solution to the disclosure of information

There was widespread agreement about the importance of transparency in algorithmic processing, for example that consumers using online platforms should be informed of why they are seeing a certain type of content, since low transparency means consumers are less discerning than they should be. It was suggested that the DRCF should consider steps to help improve standards of transparency and to educate people about their rights surrounding algorithmic processing. However, some respondents also pointed out that more information does not always lead to more clarity and could create significant risks such as enabling bad actors to game systems.

4. Fairness is required for public trust and to help protect vulnerable groups.

The topic of fairness was also raised by several respondents since, as well as being important in its own right, fair outcomes also help to foster public trust in algorithmic systems. A respondent commented that the papers provided a thorough account of the different ways algorithms can shape fair outcomes. Others agreed with our argument that incomplete or skewed datasets are not the only reason for biased outcomes, and that bias can emerge at other points in the production-use chain. The issue of data was also highlighted in respect of protecting vulnerable users by ensuring diversity and inclusion in source data. It was pointed out that the harms generated by algorithmic systems are often felt more keenly by vulnerable populations, such as children or the elderly, for example via targeted advertising.

5. Buying algorithms from 3rd parties can cause many issues including poor performance and confusion over accountability and responsibility

Finally, there was concern from a large number of stakeholders about the problems that can arise when algorithms are bought "off the shelf". It was pointed out that risk identification requires action on behalf of both the developer of the system, and the deployer, and that for AI systems this requires greater interaction than is typically required for classic software. In addition, the involvement of multiple parties in the development and use of algorithms can lead to confusion over where accountability lies and who is responsible for any harm. The lack of transparency of AI models can also lead to them being built or sold for use cases where they either should not be used or where they do not function as expected. It was requested that the DRCF issue guidance on outsourcing and "*encourage the development and adoption of innovative digital technologies that will help organisations in (possibly international) supply chains work together responsibly.*" This feedback shows the importance of the work on algorithmic procurement which the DRCF is conducting this year.

Additional issues raised by the respondents
1.   There are risks of algorithms distorting or exploiting consumer's views of the world

Several respondents drew attention to the role of algorithms in generating fake content and portraying false impressions of the world. Attention was drawn to algorithms that attempt to detect people's emotional state (including vulnerable customers), and research indicating that this could lead to a polarisation of views. Similarly, there was concern over too much emotive content on platforms, as well as the practice of platforms limiting people's exposure to alternative viewpoints and/or political microtargeting. One respondent suggested that platforms could audit accounts on their sites to determine whether the users behind them are misrepresenting themselves.

For some respondents, concerns over false impressions went beyond polarisation. There was particular concern that algorithms (particularly price personalisation algorithms) could harm consumers of financial services, such as vulnerable groups who may end up paying more for insurance and other banking products.

2. *Other effects of algorithms need to be considered*

Several respondents requested that the six "shared priorities" laid out in the paper be expanded to cover additional issues, including the effect of algorithms on employment, health and the environment. The idea was raised of the DRCF working more closely with stakeholders in these areas, including for instance the Health and Safety Executive.

Respondents also called for competition and innovation that leads to better outcomes for all those affected by algorithms, not just those who are consumers of algorithmic products or services. Stakeholders argued that technology has the potential to improve society in many ways, such as enhancing agency for citizens, facilitating collaboration throughout globalised supply chains, and encouraging productivity and growth that is sustainable.

3. *A healthy online ecosystem is highly dependent on how algorithms are used*

The role of algorithms in facilitating the spread of harmful and illegal content online was singled out by several respondents as an issue worthy of more consideration. Emotionally triggering content was flagged as particularly dangerous, and something that could lead to the polarisation of views, hatred and social division.

One stakeholder, however, was resistant to the idea that algorithms were an unambiguous driver of polarisation, stressing other media sources as a probable cause.[1]

Several respondents highlighted that algorithms could also be used to moderate content. This varied from using algorithms for 'Just-in-Time' regulation (one based on the use of algorithms to monitor a platform's own decision algorithms), to practical examples of how algorithms were currently being used in online platforms to spot and remove misinformation that could lead to harm. Several stakeholders argued that exposing source code could lead to malicious actors 'gaming' algorithms, allowing them to further amplify the dissemination of harmful and illegal content.

There was concern from multiple respondents about online platforms leveraging their perceived dominant market position for their advantage, leading to reduced innovation. Several respondents, for example, stated that they did not have full access to data on how their media products behaved

---

[1] Our Algorithmic Auditing paper does, however, draw attention to a study by Ribeiro et al., which presents a more nuanced picture of how recommender systems influence the likelihood of internet users viewing extreme content online. See Ribeiro et al. (2019) *Auditing radicalisation pathways on YouTube*.

on so-called gatekeeper platforms. One stakeholder argued that tech companies should "not be allowed to expand exponentially, without constraint or proper regulatory oversight".

4. The public may not be able to fully understand the risks posed by algorithmic systems

A number of stakeholders highlighted that the public's ability to learn about and protect themselves from the risks posed by algorithmic systems may be overstated. Transparency guidelines, they argued, would have to be adapted to the diverse range of platforms and their business models and features. Volume and detail of information does not necessarily translate into clarity for end users. It was suggested that the strategic delivery of bite-sized, digestible information at particular points in people's digital journeys may be more impactful.

In contrast, some respondents highlighted the importance of giving users more control over their data and of the algorithmic systems they directly interact with, for example by choosing what they wish to see on algorithmically-driven news feeds. Separately, several stakeholders were concerned about 'dark patterns', whereby services and products are designed to encourage consumers to share more data than they are comfortable with (via complex T&Cs, and multi-step processes for opting out of data sharing).

3. PAPER #2 - AUDITING ALGORITHMS: THE EXISTING LANDSCAPE, ROLE OF REGULATORS AND FUTURE OUTLOOK

Box 4: A reminder of the 'Auditing Algorithms' paper's key messages:

- Algorithmic auditing refers to a range of tools and methods that can be used to reveal information about the use and impact of an algorithmic system. It includes assessments of an organisation's governance arrangements, as well as deep technical reviews.
- Audits can be used either to prove compliance with a regulatory regime, or to highlight general risks that an organisation should attend to. The findings of audits could be of interest to regulators, the public, as well as the organisations using algorithms.
- The paper looks in particular at the merits and future prospects of a "third party auditing ecosystem", where private firms offer auditing services.
- A market for algorithmic auditing has already begun to emerge, for example with some consultancies and technology firms providing tools to test algorithms for bias. However, our research suggests the growth of this market is being held back by a lack of common standards, which mean audits are applied inconsistently and vary in quality.
- The paper considers what role DRCF members could play in supporting this auditing ecosystem. One option would be for regulators to set standards for audits, and/or to accredit auditors. Another option would be to allow industry to take the lead, with DRCF members providing input as and when required.
- We finish by setting out a number of hypotheses (based on these options), which readers were invited to reflect on and respond to. These related to:
    - o H1 – Clarifying how external audits support the regulatory process

> o H2 – Providing guidance on how third parties should conduct audits
> o H3 – Shaping the development of auditing standards
> o H4 – Sharing information between auditors and regulators
> o H5 – Accrediting third party auditors
> o H6 – Establishing sandboxes

Respondents highlighted several key themes that should be taken into consideration across all the following hypotheses posed. They were:

1. *The importance of considering the type of industry/product*

 Several respondents asserted that the type of industry/product context should play a large role in informing which regulatory role is most appropriate in that setting – therefore stating that there will be no 'one size fits all' regulatory approach across industries and across product/service applications. Respondents noted that one way in which regulatory approaches could be adapted was to choose an 'appropriate' level of audit related to the level of perceived 'risk'. As an example, one respondent noted that the hypothesis chosen would vary largely from a 'low risk' scenario like that of an individual decision with low impact, to a conversely 'high risk' scenario where the outcome will have a large impact on many individuals (with the example provided of nationwide educational outcomes).

2. *The importance of the regulatory approach being adaptable to change*

 Several respondents noted that any approach taken by regulators will need to be adaptable to change, with respondents noting that the power of algorithmic systems and the ways in which they can be applied will continue to grow. This will likely be a pertinent challenge for regulators seeking to set guidelines or standards. The rapid speed of change will also be a relevant consideration in determining the frequency of any proposed potential audits.

Respondents to the call for input aligned their responses to the hypotheses presented, providing a brief summary of their position, followed by their views on any advantages or disadvantages of each proposed regulatory action.

> **H1 Clarifying how external audit supports the regulatory process**
>
> **There may be a role for some regulators in clarifying how external audits could support the regulatory process, for example, as a means for those developing and deploying algorithms to demonstrate compliance with regulation, under conditions approved by the regulator.**

Respondents generally agreed that there would be value in regulators issuing guidance on the need and purpose of external audits, including how this approach can be used to demonstrate compliance. Stakeholders highlighted several **advantages** of this approach, among them:

1. *Creating the incentive for firms to audit*

A respondent questioned the incentive for firms to audit if not for regulatory oversight. They indicated that, in the absence of regulatory clarity around what an audit entails and why it may be needed, firms are unlikely to commence auditing practices.

2. *Setting firms' audit expectations*

Respondents suggested that an advantage of this approach would be that the guidance sets firms' expectations of what an audit entails, and that it would also provide greater clarity around the benefits of such audits to firms.[2] One respondent claimed that clearer guidance may also lead to the creation of more consistency in practices across organisations. Moreover, respondents noted that there was value in regulators not only providing information on the benefits of external audits, but also the 'haphazard risks' of auditing.

Respondents also highlighted **risks, potential issues** and **key considerations** for regulators pursuing an approach aligned to clarifying the purpose of external audits, among them:

3. *Caution against guidance being too narrow or not adaptable to different contexts*

Several respondents were cautious that an external audit would be specific to a single use of an algorithmic system at a specific moment in time. An audit may not lead to a long term shift in the culture of the organisation using that system.

Numerous respondents said that any guidance would need to strike the right balance between supporting consistency in the use of audits, while also allowing for different intensities of audit that would match the level of risk posed by the algorithmic system in question (e.g. auditing the use of a facial recognition system for proving one's identity on a banking app would necessarily need to look different to an audit of a facial recognition system used for law enforcement purposes). Regulators should be aware that there may need to be different applications of auditing depending on the 'risk' of the system.

It was also noted by one respondent that firms should have some choice over whether to undertake external or internal auditing, including in relation to demonstrating compliance.

One respondent highlighted the concern that audit adoption would likely be affected by firm 'centralization' (interpreted as firm size and whether domestic or international). It may be that any guidance could incorporate considerations relating to firm size and/or guidance on multi-national corporation best practice where there may be conflicting guidance across countries.

H2 Guidance on how third parties conduct audits

There may be a role for some regulators in producing guidance on how third parties should conduct audits and how they should communicate their results to demonstrate compliance with our respective regimes.

---

[2] One stakeholder noted a distinction between 1) assessments for general due diligence; 2) assessments that validate specific claims made about a system (e.g. that it does not produce outputs or predictions that are biased against a particular demographic group), and 3) assessments that assess the overall risk profile of a system, regardless of whether any claims have been made about it. This stakeholder suggested that only the first two could be described as audits, in the specific sense of the word.

Respondents generally agreed that there would be value in regulators issuing guidance on how third parties should conduct external audits. Stakeholders highlighted several **advantages** of this approach, among them:

1. *Cross-regulatory guidance would support consistency*

As with the responses relating to hypothesis 1, there was general agreement that there would be value in guidance for third party audits, as the involvement of all regulators would support consistency (whilst also recognising that there may be some differences in approach due to varying remits). Standardised evaluation mechanisms, cross-cutting themes and minimum 'red lines' were identified as potential areas for alignment across regulators.

2. *Where respondents thought such guidance would create most value*

Respondents set out what they would like the guidance to cover, with suggestions that it should also cover "how audits should be completed across the supply chain" and "how stakeholders should be consulted and notified". One respondent commented that guidance would create the largest value as a "consumer-centered best practice guidance".

Respondents also highlighted a potential **risk** of the approach:

3. *Such guidance may be overly burdensome*

Multiple respondents highlighted the burdens that such guidance can create, commenting that guidance can be labour intensive to produce and update. One respondent advocated for updates at "reasonable intervals" so as to avoid overly burdening audited organisations. A respondent also cautioned against the risk of scope-creep, advocating for "clear scoping and guidance" to avoid the guidance becoming too burdensome.

> **H3 Shaping the development of auditing standards**
>
> **There may be a role for some regulators in assisting standards-setting authorities to convert regulatory requirements into testable criteria for audit.**

In general, the respondents supported a role for regulators in standards, with comments suggesting regulatory action in this space would provide much-needed clarity for firms. Respondents provided advice on how regulators could **add the most value to standards**, among them:

1. *Regulators should work closely with relevant standards setting bodies*

Most respondents detailed the need for regulators to work closely with professional standards setting bodies, as well as working to incorporate international standards (such as through ISO/IEC) wherever possible. One respondent noted the requirements for documentation and impact assessment (pre and post market deployments) within the EU AI Act may have implications for the DRCF regulators. Another respondent stressed the importance of standards as a way of preventing fragmentation in approaches to assessing and governing the use of algorithms.

2. *Clear regulatory requirements*

A respondent highlighted the need for regulators to provide clarity on what testable criteria are meaningful for regulatory requirements and what are not. The respondent also commented that to support this it may be helpful to clarify each body's responsibility.

However, some respondents **cautioned against this approach**:

3. *Difficulty in creating non-circumventable testable criteria*

Multiple respondents argued that converting regulatory requirements into testable criteria would be both difficult and potentially circumventable, claiming that there would be a high risk of manipulation. One respondent cited the example of audit recipients being able to tweak data set selection. Therefore, regulators should be particularly aware of any data requirements needed to facilitate third party auditing.

> **H4 Sharing information between auditors and regulators**
>
> **Some regulators may have a role to provide mechanisms through which internal and external auditors, the public and civil society bodies can securely share information with regulators to create an evidence base for emerging harms. Such mechanisms could include a confidential database for voluntary information sharing with regulators.**

Some respondents welcomed regulatory action in the context of sharing information, with one respondent describing the access to relevant information with regulators as "essential". Some positive responses related to:

1. *Beneficial information sharing between firms and regulators*

One respondent supported the DRCF exploring this hypothesis further, commenting that it may allow firms to derive insights into the types of risk events being recorded in other firms. A different respondent noted that the sharing of information may benefit regulators to better understand the pace of AI change within industries.

2. *Expanding to regulators sharing information with the public*

One respondent was keen to suggest the importance of 'reverse information' - that whilst the description of the hypothesis proposed information sharing from the public (public, civil society bodies and researchers) to regulators, there was likely to be value in regulators sharing information with the public too. Examples were provided, with mention to the proposed reporting on the findings of impact assessments in the US Accountability for Algorithms Act 2022 by the Federal Trade Commission.

Many respondents, however, argued **against** pursuing this form of regulatory involvement, chiefly due to:

3. *The risk to security and potential for data breaches*

Many respondents asserted that such information sharing was a significant risk to security and a liability for potential data breaches. Respondents commented that particular care must be taken to protect commercially sensitive information, whilst also avoiding the compromise of intellectual property rights. Respondents also posited that information sharing done inappropriately could lead to a significant negative impact in the relationship between regulators and industry.

H5 Accrediting third party auditors

There may be a role for some regulators in accrediting organisations to carry out audits, and in some cases these organisations may certify that systems are being used in an appropriate way (for example, through a bias audit) in order to demonstrate compliance with the law to a regulator.

Respondents were generally positive towards the suggestion that regulators could have a role in accrediting organisations:

1. *Enhanced confidence and trust in the processes and individuals conducting the auditing*

Multiple respondents commented that accreditation may help to provide firms with validation and confidence around their systems usage and output. One respondent went further, commenting that for an auditing regime to be trusted, there must be some form of accreditation assigned to the auditors, as well as help for organisations to identify which audit firms would be best placed to use the audit tools in an effective and responsible way.

Respondents suggested the following **considerations** for any regulatory involvement in the accreditation space:

2. *The need to work with professional bodies*

One respondent emphasised the need for regulators to work collaboratively with professional bodies to develop professional practice, akin to respondents' earlier comments relating to working collaboratively with standards setting bodies in relation to any creation of auditing standards.

H6 Establishing sandboxes

For some regulators there may be a further role to play in expanding the use of regulatory sandboxes (where a regulator has power to do so) to test algorithmic systems in a controlled environment.

Respondents were positive about the use of regulatory sandboxes, seeing sandboxes as an effective way for organisations to understand the application of regulation and the impact of algorithms, whether as a real-world trial or as a simulation. A key **positive** of sandboxes and regulatory action to expand their use was the following:

1. *Effective collaboration between regulators and industry (SMEs in particular)*

Respondents also recognised this as an opportunity for regulators and industry to work together, helping provide consistency and safeguards in the ways regulators access algorithmic systems. One respondent particularly encouraged regulators to support SMEs to participate in sandboxing exercises, as regulatory burdens may be felt most by smaller businesses.

The feedback we received did, however, highlight a key limitation of sandboxes:

2. Prioritise making sandboxes inexpensive and easy of use

Lastly, one respondent cautioned that if sandboxes are not inexpensive and easy to use, their practical utility may be limited. The respondent also highlighted that such expense is likely to be most burdensome on smaller businesses or new entrants.

4. VIEWS ON HOW THE DRCF SHOULD OPERATE AND WHERE IT COULD ADD THE MOST VALUE

As well as asking stakeholders for their views on the findings of our two papers, we also canvassed opinion on where the DRCF could make the greatest impact in future. A number of respondents highlighted specific issues that they would like to see DRCF members collectively address. One respondent saw the use of algorithms to power harmful "choice architecture" as a priority concern. A body representing consumer interests, meanwhile, thought the DRCF should focus on addressing the lack of accountability for the outcomes of algorithmic systems, particularly where those systems are procured from third parties. A number of respondents also stressed that future DRCF work should focus on the use cases and applications that pose the highest risks to consumers and the public.

In addition to highlighting issues for the DRCF to explore, some respondents set out ideas for specific interventions and projects. Some stakeholders suggested that the four digital regulators should develop a single route through which the public can flag concerns about the misuse of algorithms, which could then be directed to the appropriate regulator. The rationale would be to save the public and consumers having to determine for themselves which regulatory regime their complaint falls under. Other respondents thought the DRCF could run activities to educate the public on the fundamentals of algorithmic processing, including what it means, how it shapes their lives, and ways they can protect themselves from harm.

In all, respondents' ideas for future DRCF projects varied considerably. Yet a recurring theme throughout the submissions was a request for the DRCF to help organisations understand their obligations under the different regulatory regimes. More than one respondent suggested the DRCF could develop joint guidance or position statements on algorithmic processing where regulatory responsibilities overlap, and in particular where regimes appear to be in conflict.

Although it was a not specific question set to respondents, many of those who engaged with our call for input also included views on how the DRCF should operate more generally. Many of these views related to who we should we be engaging with as part of our research. Some stakeholders felt strongly that the DRCF should be proactive in seeking out the opinions of smaller organisations who build or use algorithmic systems, noting that they may have different perspectives than the "big tech" firms that tend to dominate conversations about the responsible use of AI and algorithms.

In addition, several respondents called for the DRCF to collaborate with other regulators that have an interest in algorithmic systems, particularly the Equality and Human Rights Commission, which the DRCF already engages with closely via its regulatory round table (as well as other bodies on a bilateral basis). Others pointed to the need to collaborate with regulators operating in the field of health and social care, which have valuable experience in assessing and auditing algorithms. Some

also saw value in the DRCF working with professional bodies that represent the interests of those working to build, use or oversee the use of algorithms, such as the British Computer Society.

A more general request from some respondents was for the DRCF not to lose sight of the many benefits that algorithms can bring to society, and to avoid falling into a situation where our work only highlights the risks of this technology. These stakeholders did not want to see the DRCF making or encouraging interventions that would undermine innovation or result in valuable algorithmically-driven products and services being more difficult to access.

5. NEXT STEPS FOR THE DRCF'S WORK ON ALGORITHMIC PROCESSING

The feedback that we have collected through our call for input will be used to shape and inform the next phase of our work on algorithmic processing. In particular, we are conscious of the need to retain a clear focus on high risk use cases and applications, to engage with a range of stakeholders through our work (we plan to engage with both the Centre for Data Ethics and Innovation and the Alan Turing Institute in our planned next steps), and to ensure that any interventions we make in future do not create a disproportionate burden on the firms we regulate. While we have already committed to a series of activities on algorithmic processing over 2022/23, we intend to use the feedback from our Call for Input exercise to steer and guide these initiatives as they get underway.

Our 2022/23 project on algorithmic processing includes three workstreams:

1. **Improving the capability of regulators to undertake algorithmic audits** – This will involve a series of internal workshops where the four regulators will share lessons in how to conduct effective algorithmic audits. These workshops will consider which auditing methods are effective in different contexts, as well as which types of skills and resources are required to undertake such assessments. We will publish a summary of these discussions in early 2023.

2. **Understanding the role of third party algorithmic audit in supporting the regulatory process** – This workstream will explore how third party audits of algorithmic systems could complement the regulatory process. We will follow up the hypotheses laid out in our algorithmic auditing paper to develop a firmer view around the role audits can play in assisting firms with compliance and improving quality and transparency in the use of algorithmic systems. As we are developing our views on the hypotheses, we will draw upon the comments made within this paper. We will also continue to consider what role regulators can play in the emerging ecosystem of algorithmic audit.

3. **Promoting transparency in the procurement of algorithmic systems** – This workstream will look at how to improve transparency for buyers of algorithmic systems, such that they have a better understanding of what they are purchasing, its strengths and limitations, and how it can be used appropriately. The work will focus on several use cases, which may include for example the procurement of content moderation systems and age verification technology. We will publish the findings of this research in early 2023.

Stakeholder feedback on our algorithmic auditing paper will be particularly useful to inform the first and second of these workstreams, while feedback on the algorithmic harms and benefits paper will be valuable for the third of these.

Annex: List of organisations which provided a substantive response to the Call for Input and gave permission to be named within this document.

5Rights Foundation
Association of British Insurers (ABI)
BBC
BCS, The Chartered Institute for IT
BT
Deloitte
Direct Line Group
Equality and Human Rights Commission
Financial Services Consumer Panel
Future PLC
Gambling Commission
Institute for the Future of Work
ITI, Information Technology Industry Council
Meta
Microsoft
MPA, Motion Picture Association
Onfido
Reset
techUK
Trustpilot
TUC
UK Finance
Which?

(Please note that this list does not include the names of private individuals who responded to our call.)