The Information Commissioner's Office's (ICO's) response to the Science and Technology Committee's call for evidence on algorithms in decision-making

The Information Commissioner has responsibility for promoting and enforcing the Data Protection Act 1998 (DPA), the Freedom of Information Act 2000, the Environmental Information Regulations and the Privacy and Electronic Communications Regulations 2003. She is independent from government and upholds information rights in the public interest, promoting openness by public bodies and data privacy for individuals. The Commissioner does this by providing guidance to individuals and organisations, solving problems where she can, and taking appropriate action where the law is broken.

The Commissioner welcomes the opportunity to respond to the Science and Technology Committee's call for evidence on algorithms in decision-making. She recognises the timeliness of the inquiry, given the increasing use of artificial intelligence (AI) and machine learning across all sectors, and the imminent implementation of the General Data Protection Regulation (GDPR), which gives individuals powerful rights relating to automated decisions and profiling.

The DPA and the GDPR are concerned with the collection and use of personal data. Personal data is information that by itself, or in conjunction with other information, identifies a living individual. Algorithmic decision-making can be used in a number of scenarios, many of which do not involve the use of personal data. The Commissioner's regulatory remit does not extend to such decisions and therefore, her response is focused on those decisions that do involve the use of personal data and can have an impact on individuals.

The Commissioner recognises the benefits that can flow from the use of algorithms in decision-making and also the implications that they can have for privacy and data protection. However, she sees these implications as opportunities rather than barriers; opportunities to encourage creativity and innovation, and to foster trust with and empower individuals. This view is reflected in her response to the points below.

Point 1 – The extent of current and future use of algorithms in decision-making in Government and public bodies, businesses and others, and the corresponding risks and opportunities

Current use

In her foreword to the ICO's recently updated paper, 'big data, AI, machine learning and data protection'[1], the Commissioner makes the point that big data analytics has spread throughout both the public and private sectors. While algorithmic decision-making is just one aspect of this, the Commissioner's statement would still stand if it only related to algorithmic decisions. Some current examples from the public and private sectors are provided below:

Public sector

Criminal justice – A number of states in the USA (such as New Jersey[2] and Wisconsin[3]) use algorithms to make criminal justice decisions on matters including sentencing, bail and recidivism. Private companies like Northpointe[4] make algorithmic tools available to public officials for this purpose.

Fraud detection – Government departments (such as HMRC[5] in the UK and the Financial Intelligence Unit[6] in the Netherlands) use algorithms to make decisions on which cases or unusual transactions to investigate for potential fraudulent activity.

Education – In France, central government uses an algorithm when dealing with teacher assignments in public schools[7]. Teachers are matched with positions in schools across the country based on decisions made by the algorithm.

Private sector

[1] https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf
[2] https://qz.com/920196/criminal-court-judges-in-new-jersey-now-use-algorithms-to-guide-decisions-on-bail/
[3] http://www.bbc.co.uk/news/magazine-37658374
[4] http://www.northpointeinc.com/
[5] https://sage-exchange.co.uk/news/industry-news/how-does-hmrc-select-tax-returns-for-enquiry
[6] http://link.springer.com/article/10.1007/s41125-017-0012-x
[7] http://www.matching-in-practice.eu/matching-practices-of-teachers-to-schools-france/

Employment – Companies such as Predictive Hire[8] offer their services to businesses around the world for the purposes of candidate selection. Based on existing business data, predictive models are built, from which algorithmic decisions are then made about which candidates to hire.

Credit – For a long time, credit reference agencies such as Equifax[9] and CallCredit[10] in the UK have used algorithms to decide people's credit scores. Lenders also use their own algorithms to decide on which applicants to accept and what interest rates to set.

Insurance – Similarly, insurers have for some time been using algorithms to make decisions about potential customers. The algorithms profile people and place them into risk categories. Decisions can then be made on whether to accept applications and if so, what level of insurance premium to set.

Future use

The 'fuel' powering the algorithms in all these examples is personal data. The increasing variety of sources from which personal data can be obtained, observed, inferred and derived will only lead to the use of algorithms in decision-making in more and more contexts. In the near future it's possible that such scenarios will include:

Public sector

Healthcare – In the UK, a company called Your.MD has been in talks with the NHS about using its personal health assistant to streamline and reduce patient-doctor interactions[11]. Your.MD's engine uses algorithms that make decisions on the medical conditions that people may have based on their reported symptoms.

Private sector

Transport – In the UK, companies including Nissan and Volvo are starting to introduce driverless cars to London[12]. This will allow data to be collected in a real-life city environment; data that will in turn be

[8] http://www.predictivehire.com/how-it-works/
[9] https://www.equifax.com/personal/education/credit/score/how-is-credit-score-calculated
[10] http://www.callcredit.co.uk/press-office/news/2013/08/callcredit-adopts-featurespaces-aric-technology
[11] How AI can transform the NHS, BusinessCloud Magazine, December 2016, p32
[12] http://www.independent.co.uk/life-style/motoring/nissan-to-trial-autonomous-cars-in-london-next-month-a7533101.html

used to train and refine the algorithms that underpin self-driving technology and make autonomous driving decisions.

Risks

The increasing use of algorithms in scenarios like those detailed in the examples above can create several risks for both organisations and individuals alike. The risks are often referred to using terms such as 'opacity', 'discrimination', 'loss of control', 'bias', and 'responsibility'. While the language used may differ, each of these matters is intrinsically linked to the principles of data protection. Some of the key principles and corresponding risks for algorithmic decision-making are as below:

Fairness

Under the DPA and GDPR, the use of personal data must be fair[13]. Fairness is about transparency, effects and expectations.

Transparency – In data protection terms, transparency means that people should be given some basic information about the use of their personal data, such as the purpose for its use and the identity of the organisation using it. In the context of algorithmic decision-making, the provision of such information can be challenging due to the complexity of the algorithms, the unforeseen uses of data and the variety of data sources.

Effects – Algorithmic decisions made about people will always have some sort of effect on them. Whether that effect is fair or not will depend on how intrusive the decision is and what its justifications are. There is a risk that algorithmic decisions may be intrusive and unjustified where they have discriminatory effects on people. For instance, it was reported in 2015 that a female doctor was locked out of a gym changing room in the UK because the automated security system had decided she was male due to associating the title 'Dr' with men[14].

Expectations – To be fair, the use of personal data should be within the reasonable expectations of the people concerned. Expectations are shaped by factors such as trust and what people are told about the use of their data. It's possible that algorithmic decisions may be unfair if they use data that people wouldn't reasonably expect to form part of a decision. For example, would people expect what they post on social media to be used to assess their insurance risk?[15]

[13] DPA Schedule 1, Part I, 1 & GDPR Article 5(1)(a)
[14] http://www.mirror.co.uk/news/uk-news/doctor-locked-out-womens-changing-5358594
[15] http://www.bbc.co.uk/news/business-37847647

Lawfulness

Under the DPA and the GDPR organisations must satisfy a 'condition'[16] or have a lawful basis[17] for the use of personal data. 'Consent' is only one of several lawful bases, but where organisations do choose to rely on it there is a risk that it may not be valid.

For consent to be valid it must be (among other things) informed. This means there should be a clear explanation of what people are actually consenting to in terms that they can easily understand. As with transparency above, this may be difficult due to the complex nature of algorithmic decision-making.

Accuracy

The DPA and GDPR both say that personal data shall be accurate[18]. The data used to train algorithms may be inaccurate, or it may be classified and labelled inaccurately. Where this is the case, there is a risk that the resulting algorithmic decisions may also be inaccurate and perpetuate bias. For instance, algorithmic risk scores used in some US states inaccurately classified black defendants as future criminals at almost twice the rate as white defendants[19], perpetuating a bias that already existed in the training data.

However, even when raw data is recorded accurately, a dataset may not be representative of the relevant population. Where such datasets are used to profile people, there is a risk that some derived or inferred data may be inaccurate (e.g. for minority groups not represented in the dataset). In turn this may lead to unreliable algorithmic decisions that are based on that data.

Algorithms are often applied to datasets in order to find unexpected patterns in the data, but correlation does not necessarily equal causation. Where algorithmic decisions are made based on such patterns, there is a risk that they may be biased or inaccurate if there isn't actually any causality in the discovered associations.

Accountability

[16] DPA Schedule 2
[17] GDPR Article 6
[18] DPA Schedule 1, Part I, 4 & GDPR Article 5(1)(d)
[19] https://www.nytimes.com/2016/08/01/opinion/make-algorithms-accountable.html?_r=2

While only an implicit requirement in the DPA, the principle of accountability is made explicit in the GDPR[20]. It requires organisations to demonstrate their compliance with the data protection principles. Among other things, this means being able to show that algorithmic decisions are fair and accurate; demonstrating this may be difficult to achieve in practice.

In order to know whether a decision is fair or accurate, it is necessary to have an understanding of its underlying reasons, i.e. the factors that influenced the decision and their weightings. Machine learning algorithms used in decision-making are often referred to using terms such as 'opaque' and 'black box'. This is because, even to the data scientists that originally developed them, their inner workings become incomprehensible as they learn and evolve based on new data and on their own output.

There is a risk therefore that organisations may not be able to hold their machine learning algorithms to account. This risk also extends to the new right that individuals have under the GDPR to obtain an explanation of an automated decision[21] (see response to point 3 below). If, for instance, an insurance company does not understand the reasons why the algorithm for its online application system turns some people away but accepts others, it cannot explain this to the individuals affected.

Opportunities

While there are a number of risks associated with the use of algorithms in decision-making, it would be reductive to only see these risks as barriers. Data protection is not simply a legal requirement to be ticked off by the compliance department; rather it is an opportunity to be seized. The key areas of opportunity are as below:

Creativity & innovation

The creation and development of algorithms is not a simple process. It requires expertise, skill and ultimately an investment of time and resources. The use of algorithmic decision-making involving personal data challenges organisations to be as innovative in their application of data protection measures as they are in developing and using the algorithms.

It is an opportunity for organisations to foster innovation in new business areas by encouraging and investing in their compliance

[20] GDPR Article 5(2)
[21] GDPR Recital 71

professionals and data scientists alike, to help develop new creative technical and organisational approaches to privacy.

This is linked with the concept of privacy by design, which is included in the GDPR under the heading 'data protection by design and default'[22]. It will help to encourage such creativity by obliging organisations to take appropriate measures to implement data protection principles throughout both the design and application of a processing operation.

Trust & competitive advantage

Despite the assertion by some that people are becoming less concerned about the use of their personal data, there are numerous studies that show privacy is still a key issue for many[23]. As such, getting data protection right is a clear opportunity to build trust with customers and citizens.

Being able to demonstrate that an algorithmic decision is fair, accurate and explainable will give people the confidence to provide their data for certain products and services, and empower them to exercise their rights. Not only that, but in a commercial context it can also set an organisation apart from others and help to deliver competitive advantage.

Data quality & information governance

Organisations need to know that they can trust the validity of the decisions that are being made. If a particular algorithm is producing unreliable decisions, this can have serious repercussions for both the organisation and individuals. For instance, if a loan application algorithm is offering inappropriately high-value loans to certain customers, both the lender and its customers are likely to suffer financially as a result.

Data quality and information governance are crucial issues for organisations looking to ensure the validity of their algorithmic decisions. This involves looking at matters such as data sources, accuracy, age, relevancy and completeness; all issues that fit neatly with the principles of data protection. Addressing the data protection risks highlighted above is therefore a key opportunity to tackle these

[22] GDPR Article 25
[23] E.g. https://www.bcgperspectives.com/content/articles/information_technology_strategy_consumer_products_trust_advantage_win_big_data/, https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2016/11/crossing-the-line.pdf and http://www.sciencewise-erc.org.uk/cms/public-views-on-big-data/

data quality issues. In short, getting data protection right will help to ensure data quality.

Point 2 – Whether 'good practice' in algorithmic decision-making can be identified and spread, including in terms of:

- **The scope for algorithmic decision-making to eliminate, introduce or amplify biases or discrimination, and how any such bias can be detected and overcome;**

- **Whether and how algorithmic decision-making can be conducted in a 'transparent' or 'accountable' way, and the scope for decisions made by an algorithm to be fully understood and challenged;**

- **The implications of increased transparency in terms of copyright and commercial sensitivity, and protection of an individual's data.**

Good practice

As detailed in response to the previous point, the use of algorithms in decision-making can create data protection risks. However, this does not mean that algorithmic decisions cannot be fair, accurate or accountable. There are several existing and developing approaches that can help to address these risks. While the below is by no means a complete list, it does detail some of the key methods that can help organisations conduct algorithmic decision-making in a fair, lawful, accurate and accountable manner.

Algorithmic auditing

Auditing can help to make algorithmic decision-making more accountable[24]; certain auditing techniques for instance have been found to be successful at identifying the discrete factors that influence algorithmic decisions[25]. This can help organisations to determine whether those factors mean that a decision would be discriminatory or biased and would therefore fall foul of the fairness and accuracy principles of data protection.

When developing new algorithms, organisations should encourage their data scientists to find innovative ways of building in auditability, to allow an on-going internal review of algorithmic behaviour. While, currently, this may be too technically difficult or resource intensive for

[24] https://www.technologyreview.com/s/602933/how-to-hold-algorithms-accountable/
[25] http://sorelle.friedler.net/papers/auditing_icdm_2016.pdf

some organisations, third parties could be enlisted to provide this service (organisations offering algorithmic audits are already being set up[26]). With regards to the protection of copyright or proprietary information, a basic overview of such audits could be made publicly available, while the audits themselves could be carried out in confidence.

Ethics boards

Further to the Science and Technology Committee's call for a national Council of Data Ethics[27], internal organisational ethics boards can also be setup to apply ethical principles and assess difficult issues that can arise in the creation and use of algorithms in decision-making[28]. Such boards, or committees, can raise relevant questions about matters of fairness and accuracy in order that any potential issues can be identified and then addressed by the data scientists responsible for the algorithm.

As well as addressing these risks internally, this approach can also help to increase the transparency around the development of algorithms externally. For instance, the ethics boards could publish their meeting minutes or written reports to the public so that the development of the algorithm is openly documented. This is approach is linked with 'data protection impact assessments', which are discussed in response to point 3 below.

Privacy notices

Although privacy notices are certainly not restricted to the use of personal data in this context, they are a vital and necessary tool for helping to ensure the fairness of algorithmic decisions, particularly in terms of transparency and expectations. At the point of collection of personal data, as well as some basic information about its use, the GDPR says that information must also be provided about (among other things) the existence of, and logic involved in, automated decision-making[29]. The discoveries made from auditing, reviewing and documenting the development of an algorithm can help to inform the information provided here.

Organisations can use layered privacy notices to provide information in increasing detail at each layer. This can help to reduce the burden of supplying lots of complex information at once, while still allowing

[26] http://www.oneilrisk.com/
[27] https://www.publications.parliament.uk/pa/cm201516/cmselect/cmsctech/468/468.pdf
[28] https://research.gold.ac.uk/13195/1/Neyland%20Final%203.pdf
[29] GDPR Article 13(2)(f)

people to access the detailed information about the areas that matter to them. Where data is collected ad hoc throughout an individual's relationship with an organisation, just-in-time notices can be useful to help spread the provision of information as and when it is relevant. This can be used to remind people about the logic of an algorithmic decision, and their right to challenge it (see response to point 3 below) at the moment before such a decision is taken, e.g. when applying for a loan online. This can also help people to understand what they are consenting to when an organisation is relying on consent as the lawful basis for the algorithmic decision.

In order to effectively convey privacy notice information, it can be useful to facilitate cooperation between a compliance department and service design / user experience professionals[30]. This can help to ensure that not only is the right information provided, but it is done in a way that people find engaging and intuitive. In turn, this will help people to absorb and understand what will happen with their data when an algorithmic decision is made.

Interactive outputs

The use of visual and interactive interfaces can help people to understand and challenge the reasons behind algorithmic decisions and improve their accuracy moving forward[31]. The idea is that graphs and charts can display the factors taken into account by the algorithm, and adjustable sliders can allow people to change the weight given to certain factors.

The presentation of, and interaction with, such information can also allow people to spot and correct any inaccuracies in the data. If an individual can see that an algorithmic decision was heavily influenced by an inaccurate piece of data (that may have been derived or inferred through profiling for instance), they would have the opportunity to correct this and seek a new decision.

[30] https://iapp.org/news/a/design-jam-provides-new-approach-to-data-transparency-and-control/
[31] E.g. http://d-scholarship.pitt.edu/19382/1/IUI-aduna-revised-0.5.4.pdf and http://www.christophtrattner.info/pubs/iui2014.pdf

Point 3 – Methods for providing regulatory oversight of algorithmic decision-making, such as the rights described in the EU General Data Protection Regulation 2016.

Regulatory oversight

The risks associated with algorithmic decision-making have led some to call for new regulation in this space. In the Commissioner's view however, although the means of decision-making are changing, the underlying issues remain the same. Are people being treated fairly? Are decisions accurate and free from bias? Is there a lawful basis for the decision? These are issues that data protection regulators like the ICO have been addressing for many years and under the GDPR, the regulatory toolkit will be further sharpened. Some of the key rights and provisions for regulatory oversight of algorithmic decisions are detailed below:

Rights related to automated decision-making

The DPA already contains provisions on automated decision-making[32], but the more detailed provisions and more powerful rights in the GDPR[33] reflect their increasing use. People will have the right not to be subject to an automated decision producing legal or significant effects (e.g. credit and job applications) unless it is necessary for a contract, authorised by law or based on explicit consent.

This is a powerful right which gives people greater control over automated decisions made about them. Unless it's a trivial decision, necessary for a contract or authorised by law, organisations will need to obtain explicit consent be able to use algorithms in decision making. Furthermore, subsequent to an automated decision being made, people will also have the rights to obtain an explanation and challenge the decision.

Right to complain to the regulator

On an individual basis, regulatory oversight of these rights is provided in the GDPR by the right to lodge a complaint with a regulator such as the ICO[34]. So should an individual feel that an organisation has failed to provide an appropriate explanation of an algorithmic decision for

[32] DPA Part II, Section 12
[33] GDPR Article 22 and Recital 71
[34] GDPR Article 77

instance, they can complain to the ICO so that the matter can be investigated.

Powers of the regulator

While the precise mechanisms for complaints and investigations are still being developed, there are a number of investigative powers available to the ICO under the GDPR. For instance, we will have the power to conduct investigations in the form of data protection audits[35], obtain access to the personal data necessary for an investigation[36], and obtain access to the premises and data processing equipment of an organisation[37].

Furthermore, there are also a number of corrective powers available to the ICO under the GDPR. These include (among others) the power to impose a ban on processing operations[38] and the ability to issue an administrative fine[39] of up to €20,000,000 or 4% of annual worldwide turnover, whichever is higher.

Data protection impact assessments and prior consultation

In addition to the areas of investigation and enforcement, the GDPR will also provide for regulatory oversight of algorithmic decision-making in other ways. For instance, it is likely that organisations will often need to conduct data protection impact assessments (DPIAs) when developing systems that use algorithms in decision-making. This is because the GDPR makes this a requirement in situations where there is likely to be high risk to the rights and freedoms of individuals, particularly when new technologies are being used[40].

A DPIA is a tool to assess the data protection risks of, and identify the mitigation measures for, a particular processing operation. The ICO has long been a champion of this type of risk assessment and has encouraged organisations to undertake and publish material relating to them[41].

Following a DPIA, if an organisation is not able to identify a way of mitigating any high risks, it will need to consult with the ICO prior to implementing a proposed processing operation[42]. Regulatory oversight

[35] GDPR Article 58(1)(b)
[36] GDPR Article 58(1)(e)
[37] GDPR Article 58(1)(f)
[38] GDPR Article 58(2)(f)
[39] GDPR Article 58(2)(i)
[40] GDPR Article 35(1)
[41] https://ico.org.uk/media/for-organisations/documents/1595/pia-code-of-practice.pdf
[42] GDPR Article 36(1)

via this mechanism will provide opportunity for the organisation to receive written advice on the data protection aspects of the proposed operation (e.g. a new algorithmic decision-making system). At this stage, the ICO will also be able to use the power to prohibit the processing operation where necessary.

Codes of conduct and certification

The increasing emphasis on accountability in the GDPR is reflected in new provisions relating to codes of conduct[43] and certification[44]. While neither is mandatory, they can still help to provide regulatory oversight of the use of personal data in a number of contexts, including algorithmic decision-making.

Codes of conduct may be drawn up by trade associations or bodies representing specific sectors in order to assist the proper application of the GDPR. For instance, a trade association for a sector where algorithmic decisions are particularly prevalent may want to prepare a code of conduct that focuses on the exercise of the rights of individuals in relation to automated decision-making. The code would first need to be approved by the ICO and could then be monitored by an ICO-accredited body.

Similarly, certification schemes may be established (and certification issued), either by the ICO or an accredited body. This would allow compliant processing operations to be checked, tested and certified as so. For example, if an algorithmic decision-making system is demonstrated as having appropriate safeguards to prevent discriminatory and inaccurate decisions, this could receive a certification 'seal'. The seals can help to inform people about the data protection compliance of a particular product or service.

The ICO is currently looking into how certification schemes can be set up and managed in practice.

[43] GDPR Article 40 and 41
[44] GDPR Article 42 and 43