

# Probing the rules on automated decision making

***Duc Tran, Senior Associate at Herbert Smith Freehills LLP, applies the rules and guidance available in relation to automated decision making to real-world decision making processes***

As organisations increasingly seek to employ new and disruptive technologies such as machine learning to automate and optimise the effectiveness and efficiency of their operations and decision making processes, they are having to pay ever closer attention to the rules set out in Article 22 of the General Data Protection Regulation ('GDPR') which govern the use of automated decision making.

Given that the purpose behind these rules is to protect individuals from important decisions being made about them by solely automated means, the GDPR only permits this type of processing if one of the limited exceptions set out in Article 22 applies (see Table 1 on page 8), which means that it can be more difficult to legitimise than most other types of processing — there is no option to rely on the legitimate interests condition available under Article 6 of the GDPR, for instance.

It follows that if an organisation wishes to implement a process which could amount to automated decision making, it needs to ensure that one of the relevant exceptions applies or, if this is not possible, change the process so that it no longer amounts to automated decision making (by, for example, involving a human in the processing) and therefore falls outside the scope of Article 22.

The degree of ambiguity which currently surrounds the application of the rules regarding automated decision making, however, means that this can pose a challenge to any organisation in this position. Given that the GDPR is still near the beginning of its lifecycle, there is an absence of relevant case law and relatively limited official guidance.

In addition, the age-old problem of technological innovation outpacing the ability of data protection legislation and guidance to keep up has served to compound this ambiguity, resulting in organisations having to apply the sections of the GDPR and guidance issued by the UK Information Commissioner's Office ('ICO') and European Data Protection Board ('EDPB') relating to automated decision making to technologies and techniques that simply weren't envisaged at the time of writing.

In recognition of this, the ICO is part way through releasing up-to-date guidance on this topic as part of its AI Framework blog series, which focuses on the meaningfulness of human review in non-solely automated AI applications.

In this article, we seek to consider some of these areas of ambiguity which arise when attempting to apply the text of the GDPR and official guidance to real-world decision making processes.

## What is automated decision making?

Article 22(1) of the GDPR defines automated decision making as 'a decision based solely on automated processing [...] which produces legal effects concerning [a data subject] or similarly affects [the data subject].'

While automated decision making is sometimes confused with profiling, it should be noted that the two are distinct concepts. Profiling is a type of processing activity which involves evaluating, analysing or predicting a data subject's characteristics or behaviour, whereas automated decision making is making a decision based on a processing activity (which could, but doesn't have to, include profiling).

In order for a given activity to constitute automated decision making, two criteria must be satisfied: there must be a decision which produces a legal effect or similarly significant effect; and the decision must be based solely on automated processing.

## Legal effect or similarly significant effect

A legal effect is anything that impacts a person's legal status or rights, such as access to social welfare benefits or a person's right to work.

A similarly significant effect is more difficult to define. It refers to any effect on an individual that has an equivalent impact to one of the aforementioned legal effects.

*[\(Continued on page 8\)](#)*

[\(Continued from page 7\)](#)

The Article 29 Working Party Guidelines ('the Guidelines'), which were developed prior to the GDPR but have since been endorsed by the EDPB, suggest that in order for an effect to be 'sufficiently significant' for the purposes of Article 22, the relevant decision must have the potential to significantly affect the circumstances, behaviour or choices of individuals; have a prolonged or permanent impact on the data subject; or 'at its most extreme, lead to the exclusion or discrimination of individuals'. The Guidelines indicate that such decisions might include those which affect individuals' financial circumstances, or access to health services, education, or employment.

Clearly some decisions, particularly those made in the context of certain industries (recruitment, education, healthcare), will always potentially have a 'significant effect' on individuals for the purposes of Article 22.

However, identifying whether other decisions, for example in the context of targeted advertising, will have a significant effect on individuals can prove more challenging.

While a social media post advertising a new clothing range to a profiled demographic representing a brand's target audience might seem unlikely to have a significant effect on those targeted, an online casino sending multiple intrusive advertisements to an individual with a gambling addiction may well be deemed to have a significant effect on that individual and their financial circumstances. As such, organisations seeking to employ decision making processes in this context must take account of a range of factors such as the content, intrusiveness, level of tracking, method of delivery of a targeted advert, and the expectations, wishes and vulnerabilities of the individuals targeted when assessing whether or not the relevant processing activity falls within scope of Article 22.

## Solely automated means

A decision is deemed to be based on 'solely automated means' if there is no human intervention involved. The ICO provides two scenarios (one which involves human intervention, and one which does not) to help illustrate the nature of human involvement required.

### Table 1: When an organisation may conduct automated decision making

If it is either:

- necessary for entering into, or the performance of, a contract between the individual and a controller;
- authorised by EU or Member State law; or
- based on the individual's explicit consent (freely given, specific, informed and unambiguous affirmative indication of the individual's wishes, which must be an express form of consent).

**Scenario 1:** An employee is issued with a warning about late attendance at work. The warning was issued because the employer's automated clocking-in system flagged the fact that the employee had been late on a defined number of occasions.

However, although the warning was issued on the basis of the data collected by the automated system, the decision to issue it was taken by the employer's HR manager following a review of these data.

In this scenario the decision was not taken solely by automated means.

**Scenario 2:** A factory worker's pay is linked to their productivity, which is monitored automatically. The decision about how much pay the worker receives for each shift they work is made automatically by referring to the data collected about their productivity. This is an example of solely automated decision making.

Whilst neither of these examples serve to explore the nuances of human involvement in the decision-making process, they do appear to indicate that the human intervention which serves to take an activity outside the scope of Article 22 needs to occur after an automated process has been carried out (in Scenario 1, the clocking-in system automatically flagging the employees' lateness and in Scenario 2, the monitoring of workers' productivity), but before a final decision is made, interrupting the causal chain between automated processing and output.

This message is reinforced by the accompanying guidance which explains that in order for human intervention to be deemed to have occurred, a human needs to have reviewed a decision made by an automated system and needs to have had the authority and discretion to alter that decision before it is applied, rather than just apply a decision that has been made by automated means.

Whilst this all seems straightforward in principle, applying this concept to advanced or black box processing involving vast amounts of data and which might rely on algorithms and/or machine learning techniques can prove challenging. It is unclear whether a human reviewing a decision generated by such processing would ever have the ability to determine whether that decision should be overturned, and therefore whether their involvement at the output stage could ever amount to sufficient human intervention for the purposes of Article 22. This issue is explored in the first of the ICO's entries in its AI Framework blog series (copy at [www.pdpjournals.com/docs/88796](http://www.pdpjournals.com/docs/88796)) which identifies the lack of interpretability in complex systems, such as those which use deep learning, and automation bias, whereby humans routinely rely on a system's outputs without critical evaluation or using their own judgment, as risk factors which may preclude meaningful human review, potentially rendering a system solely automated.

## Human intervention at input stage?

The focus on human involvement at the output stage of the decision-making process in the guidance raises the question of whether human involvement at the input stage of the decision-making process can also be interpreted to amount to human intervention for the purposes of Article 22.

On the face of it, this does not appear to be the case. The guidance explicitly states that a decision might be considered to be solely automated even 'if a human inputs the data' and any involvement of a human at the input stage of both Scenario 1 (a human determining the value of 'x' late clock-ins which would trigger a warning) and Scenario 2 (a human determining the criteria against which a worker's productivity is measured and the corresponding pay bands) does not appear to be determinative for the purposes of these examples.

The ICO guidance set out in the AI Framework blog does not say anything explicit about the input stage, but does make reference to human intervention being required before a 'final decision' is made. However, this may be an overly literal interpretation of the guidance.

Consider the example of a receptionist working at a busy GP surgery. When a patient arrives and checks in for their appointment, the receptionist checks whether the patient is 'on time' (i.e. they have arrived less than 5 minutes after the appointment start time), or 'late' (i.e. they have arrived 5 or more minutes after the appointment start time). If the patient is late, then the receptionist politely informs them that their appointment has been cancelled, and that they must reschedule. However, the receptionist is spending a lot of her time dealing with check-ins and rescheduling latecomers, so she orders a check-in machine, and programs the machine such that it allows on-time patients to check-in, and informs latecomers that they must reschedule, by reference to the same 5-minute rule above.

Based on the ICO examples and guidance set out above, this system

would appear to constitute automated decision making, since denying access to a healthcare-related appointment would have a significant effect on the individual, and the check-in machine's decision-making process involves no human intervention. However, it could be argued that it defies common sense to classify this system as automated decision-making because the machine is simply applying exactly the same criteria, and making exactly the same binary decisions, as the receptionist would have done.

This example is, of course, simplistic. Even allowing for a more sophisticated machine (for example, one which could identify and take into account whether the relevant clinician was running late), as long as the machine is replicating the receptionist exactly, it is difficult to see why simply using a machine to improve efficiencies should be deemed automated decision making. This view would seem to be supported by a statement in the ICO blog that "AI systems that only support or enhance human decision-making are not subject to [Article 22]". Indeed, if such a literal interpretation of the guidance were to be adhered to, this would result in almost all automated processes carried out in a healthcare context amounting to automated decision making for the purposes of Article 22.

Applying this reasoning to Scenario 2 provided by the ICO, if the automated system is measuring the factory worker's 'productivity' by somehow assessing their efficiency or prioritisation skills, then it seems fairly obvious that in the absence of human intervention — and we would query whether a system capable of making this kind of assessment would be sufficiently interpretable to accommodate meaningful human review — this would constitute automated decision making.

However, if the worker's productivity is measured by simply counting how many widgets the worker produces, then the system is merely applying a rule and generating a pre-determined output (e.g. the worker will be paid £x per widget). In this case, assuming a human selects the amount £x, then there is no real distinction between this example and the GP surgery ex-

ample, which it arguably does not make sense to classify as automated decision making.

As such, in the absence of explicit guidance to the contrary, there is an argument that human involvement which takes place solely at the input stage of a decision-making process could amount to human intervention for the purposes of Article 22. The strength of any such argument would depend on the nature of the human involvement at the input stage and the extent to which it has an impact on the decision-making process.

While there may be a case for arguing that the GP surgery example and Scenario 2 (to the extent that the worker's productivity is simply measured by counting how many widgets they produce) involve sufficient human intervention to take the processing outside the remit of Article 22, given that the automated process which follows the human's input simply entails the binary application of a rule, and merely supports or enhances human decision making, it would be more difficult to argue that a human's involvement in programming a self-adjusting algorithm capable of weighing up different factors in order to make decisions to move away from instructions initially given to it would amount to human intervention in the same way. Indeed, if one was to argue that a human's involvement in programming any type of algorithm would be sufficient to constitute human intervention, arguably nothing would ever amount to automated decision making.

## Conclusion

Given the current ambiguity around the meaning of concepts such as 'significant effect', 'solely automated' and 'human intervention', there are only a limited number of real-world decision-making processes which can be labelled with certainty as amounting to automated decision making or not.

Until the application of Article 22 to real-world decision-making processes is further clarified by way of case

*(Continued on page 10)*

[\(Continued from page 9\)](#)

law and/or further guidance (and we recommend keeping a watching brief on the ICO's AI Framework blog series, as it is updated every few weeks for the next six months or so), we consider that different decision-making processes will likely fall along a scale ranging from those which do not amount to automated decision making to those which do, with the potential effect of the decision on individuals and the nature and degree of human involvement being determining factors.

It is up to organisations to identify how far along this scale a decision-making process that they are seeking to implement falls. This will likely involve a cost-benefit analysis balancing the efficiency gains and other benefits associated with implementing new technology against the potential impact on individuals and associated regulatory risk. Depending on risk appetite, some organisations may be prepared to stray further along the scale than others in determining whether to treat their decision-making process as automated decision

making or not. However, as with all things GDPR, we would recommend that any such organisation takes a detailed note of their thought processes and conclusions (taking into account the available guidance) to keep on file in order to provide an audit trail if ever required.

#### Links to the ICO's blog posts on AI:

Known security risks exacerbated by AI, 26th March 2019.  
[www.pdpjournals.com/docs/88796](http://www.pdpjournals.com/docs/88796)

Accuracy of AI system outputs and performance measures, 12th April 2019.  
[www.pdpjournals.com/docs/88797](http://www.pdpjournals.com/docs/88797)

An overview of the Auditing Framework for Artificial Intelligence and its core components, 23rd May 2019.  
[www.pdpjournals.com/docs/88798](http://www.pdpjournals.com/docs/88798)

---

**Duc Tran**

Herbert Smith Freehills LLP  
Duc.Tran@hsf.com

---

**pdp** TRAINING

# Practitioner Certificate in Data Protection



The UK's leading qualification for those working in the data protection field

**Qualify as a GDPR Data Protection Practitioner - course dates available throughout the year around the UK**

Can also be taken as an **Intensive Residential Programme** or on a **Distance Learning Basis**

*"The course content was informative and well presented, with very knowledgeable trainers. The exam was challenging, so I feel a real sense of achievement in having gained this qualification."*  
Caroline Chalk, Civil Aviation Authority

The Programme modules can be taken on an intensive basis or at different times throughout the year.

For more information, go online or contract our training team on +44 (0)207 014 3399

[www.dataprotectionqualification.com](http://www.dataprotectionqualification.com)