AI tools in recruitment

Audit outcomes report

ico.
Information Commissioner's Office

November 2024

Contents

Executive summary

The ICO have carried out consensual audit engagements with developers and providers of artificial intelligence (AI) powered sourcing, screening, and selection tools used in recruitment. We recognise that the use of AI tools in recruitment processes can offer benefits to employers, but their use can also lead to risks for people and their privacy and information rights. We undertook this work as part of our upstream monitoring of the wider AI ecosystem to understand how the development and provision of AI recruitment tools complies with UK data protection law.

Our audits found areas for improvement in data protection compliance and management of privacy risks in AI as well as areas of good practice. We recommended actions both to improve compliance with data protection law and promote the good practices in our published guidance.

Many providers monitored the accuracy and bias of their AI tools and took action to improve them. However we did witness instances where there was a lack of accuracy testing. Additionally, features in some tools could lead to discrimination by having a search functionality that allowed recruiters to filter out candidates with certain protected characteristics. Others estimated or inferred people's gender, ethnicity, and other characteristics from their job application or even just their name, rather than asking candidates directly. This inferred information is not accurate enough to monitor bias effectively. It was often processed without a lawful basis and without the candidate's knowledge.

We were concerned to find tools that collected far more personal information than was needed. In some cases, personal information was scraped and combined with other information from millions of peoples' profiles on job networking sites and social media. This was then used to build databases that recruiters could use to market their vacancies to potential candidates. Recruiters and candidates were rarely aware that information was being repurposed in this way.

We found several instances where AI providers incorrectly defined themselves as processors rather than controllers, and subsequently had not complied with the data protection principles. Some had attempted to pass all responsibility for compliance to recruiters using their tool. In these cases the arrangements were usually subject to vague or unclear contracts, that appeared to be deliberately broad or left recruiters in the dark.

However, we also noted many encouraging practices. Some providers gave recruiters their own bespoke AI model, that they could tailor to their

own needs and which avoided collecting unnecessary personal information. Others worked to be as transparent as possible, and shared detailed information online about the AI and how it worked in order to build people's trust.

During the course of our work we made almost 300 recommendations to improve compliance, all of which were accepted. These recommendations covered a number of requirements under the law ranging from;

- processing personal information fairly in the AI;
- explaining the processing clearly;
- keeping personal information collected to a minimum;
- not repurposing or processing personal information unlawfully; and
- conducting risk assessments to understand the impact to people's privacy.

Both AI providers and recruiters should follow the recommendations in this report.

By having high standards of data protection compliance, organisations developing and using AI in recruitment can innovate and deliver great services, while building trust with the public.

Introduction

We have carried out a programme of consensual audit engagements with organisations that develop or provide AI tools used in recruitment. Recruitment tools audited were broadly used for sourcing, screening, and selection.

Sourcing tools included:

- suggesting potential candidates that match or best fit a recruiter's job vacancy from a database of potential candidate profiles; and
- finding candidates that may increase the recruiter's workforce diversity, based on their predicted or inferred gender, ethnicity, age, or other diversity characteristics.

Screening tools included:

- scoring candidate competencies and skills from written applications and CVs;
- predicting a candidate's 'interest' in a job vacancy based on their interactions with recruiters; and
- predicting the likelihood of a candidate being successful in the recruiter's selection process.

Selection tools included:

- assessing a candidate's skills and fit to a role based on performance in AI-powered behaviour games or psychometric assessments;
- scoring candidate competencies and skills from written responses to interview questions and text transcriptions of in-person or video interviews; and
- evaluating a candidate's language, tone, and content in video interviews to predict their personality type.

This work covered a range of AI use cases such as machine learning, including natural language processing. We did not include AI used to process biometric data, such as emotion detection in video interviews, as we have reviewed and are producing separate guidance on biometric data and neurotech. We also did not include tools using generative AI in this work, such as for chatbots and drafting job adverts or role descriptions. Although, we are aware of the increasing use of generative AI models in recruitment and are exploring risks to people's privacy in other work.

We undertook this work as part of our upstream engagement and monitoring of the wider AI ecosystem. This helped us to understand the privacy risks and potential non-compliance with UK data protection law in the development, provision, and use of AI recruitment tools.

We recognise that AI offers opportunities that could bring improvements for society, such as efficiency, scalability, consistency and process simplification. When used in recruitment processes, AI can enable organisations to handle potentially high volumes of applications and process them consistently and in a timely manner.

However, shifting the processing of personal information to these complex and sometimes opaque systems comes with inherent risks to people and their privacy. Human recruiters may be influenced and make recruitment decisions based on AI outputs, scores, or predictions that might have limited scientific validity[1]. As detailed by the UK government in their Responsible AI in Recruitment Guide, AI recruitment algorithms can be unfair, learn to emulate human bias, and perpetuate digital exclusion of minorities[2]. The Centre for Data Ethics and Innovation noted in their Industry Temperature Check in December 2022 that AI systems holding vast amounts of personal information can be targets for cyber-attacks and

[1] REC. *REC responds to report showing risk to UK jobs from AI* (27 March 2024). https://www.rec.uk.com/our-view/news/press-releases/rec-responds-report-showing-risk-uk-jobs-ai

[2] Department for Science, Innovation, and Technology. *Responsible AI in Recruitment guide* (25 March 2024). https://www.gov.uk/government/publications/responsible-ai-in-recruitment-guide

interference[3], especially if information is kept and stored for longer than necessary. AI can process personal information in an untransparent and unexplainable way, or rely on consent that is not valid and informed.

Further to the National AI Strategy published in September 2021, the UK government published an AI regulation policy paper in March 2023. This sets out plans to implement a pro-innovation approach to AI regulation, based on the principles of:

- safety, security, and robustness;
- appropriate transparency and explainability;
- fairness;
- accountability and governance; and
- contestability and redress.

These principles are closely linked to the data protection principles in the UK GDPR. By having high standards of data protection compliance, organisations developing and using AI in recruitment can innovate and deliver great services, while building trust with the public.

Key recommendations

Our audits found some considerable areas for improvement in data protection compliance and management of privacy risks in AI. We recommended actions both to improve compliance with data protection law and promote the good practices in our published guidance.

Our recommendations were tailored to the AI use case, the personal information processed, and the context of the organisation. However we have summarised the most common areas into seven key recommendations, which are crucial to all organisations when designing and using AI recruitment tools.

These key recommendations are relevant for organisations that develop or provide AI recruitment tools (AI providers), and organisations that use or are thinking of using an AI tool in their recruitment (recruiters).

AI providers and recruiters should follow our recommendations, to ensure AI recruitment tools comply with UK data protection law.

Recommendation: Fairness

[3] Centre for Data Ethics and Innovation. *Industry Temperature Check: Barriers and Enablers to AI Assurance* (December 2022). https://assets.publishing.service.gov.uk/media/638f3af78fa8f569f7745ab5/Industry_Temperature_Check_-_Barriers_and_Enablers_to_AI_Assurance.pdf

AI providers and **recruiters** must ensure that they process personal information fairly by AI. This includes monitoring for potential or actual fairness, accuracy, or bias issues in the AI and its outputs, and taking appropriate action to address these. Depending on the decisions made and the level of human involvement as a result, the accuracy being better than random is not enough to demonstrate that AI is processing personal information fairly.

Additionally, **AI providers** and **recruiters** must also ensure any special category data processed to monitor for bias and discriminatory outputs is adequate and accurate enough to effectively fulfil this purpose. They must also ensure this processing complies with data protection law. Inferred or estimated data will not be adequate and accurate enough, and will therefore not comply with data protection law.

Recommendation: Transparency and explainability

Recruiters must ensure that they inform their candidates how they will process their personal information by AI. They should do this by providing detailed privacy information, or ensuring this is provided by the AI provider. This should clearly explain:

- what personal information is processed by AI and how;
- the logic involved in making predictions or producing outputs; and
- how they use personal information for training, testing, or otherwise developing the AI.

AI providers should support the transparency and explainability of their AI by proactively providing relevant AI technical information or details about the AI logic to the recruiter.

AI providers and **recruiters** must ensure that contracts clearly define which party is responsible for providing privacy information to candidates.

Recommendation: Data minimisation and purpose limitation

AI providers should comprehensively assess:

- the minimum personal information they require to develop, train, test, and operate each element of the AI;
- the purpose for processing and compatibility with the original purpose for processing; and
- how long the require the personal information for.

Recruiters should:

- ensure that they collect only the minimum personal information necessary to achieve the AI's purpose; and
- confirm that they only process this personal information for that specific limited purpose and they do not store, share, or reprocess it for an alternative incompatible purpose.

Recommendation: Data protection impact assessments (DPIA)

AI providers and **recruiters** must:

- complete a DPIA early in AI development and prior to processing, where processing is likely to result in a high risk to people; and
- update the DPIA as AI develops and when processing changes.

The DPIA must include:

- a comprehensive assessment of privacy risks to people as a result of personal information processing;
- appropriate mitigating controls to reduce these risks; and
- an analysis of trade-offs between people's privacy and other competing interests.

Even when acting exclusively as processors, **AI providers** should consider completing a DPIA to assess and mitigate privacy risks and evidence technical and organisational controls in place.

Recommendation: Data controller and processor roles

AI providers and **recruiters** must:

- define whether the AI provider is the controller, joint controller, or a processor for each specific processing of personal information; and
- record this clearly in contracts and privacy information.

The **AI provider** is the controller if it exercises overall control of the means and purpose of processing in practice. For example, if it uses the personal information it processes on the recruiter's behalf to develop a central AI model that they deploy to all recruiters.

Recommendation: Explicit processing instructions

Recruiters must set explicit and comprehensive written processing instructions for the AI provider to follow when processing personal information on its behalf as a processor. This includes deciding the:

- specific data fields required;
- means and purposes of processing;
- output required; and
- minimum safeguards to protect personal information.

Recruiters should periodically check that AI providers are complying with these instructions and not sharing or processing personal information for additional alternative purposes.

AI providers must only follow the recruiters' explicit instructions when they process personal information as a processor for the recruiter. The AI provider must not retain personal information, share it without permission, or process it for their own purposes beyond their instructions.

Recommendation: Lawful basis and additional condition

AI providers and **recruiters** must:

- identify the lawful basis they relied on for each instance of personal information processing where they are the controller, before processing any personal information;
- identify an additional condition, where they are processing special category data;
- document, describe in privacy information, and record in contracts the lawful basis and condition;
- when relying on legitimate interests, complete a legitimate interests assessment; and
- when relying on consent, ensure that consent is specific, granular, has a clear opt-in, appropriately logged and refreshed at periodic intervals, and as easy to withdraw as it was to give.

Methodology

From August 2023 to May 2024, we conducted consensual audit engagements with organisations that develop or provide AI-powered recruitment tools.

The scope of the audits covered these key areas:

- **Privacy management framework** – to review the management framework supporting privacy in AI systems, including:
  - comprehensive privacy policies and procedures;
  - compliance mechanisms and KPIs;
  - specialised privacy and AI training for key staff; and

    - identification of appropriate lawful bases and additional conditions for processing personal information.
- **Data minimisation and purpose limitation** – to ensure that personal information is not repurposed for AI development or provision, and personal information processed is minimal, adequate, and not retained longer than necessary.
- **Third party relationships** – to ensure that AI providers and recruiters understand and fulfil their controller and processor responsibilities and have formalised these in contracts.
- **Information security and integrity** – to confirm that technical security measures and access controls are in place and effectively protecting personal information during collection, in transit, and at rest.
- **Transparency** – to ensure that people are informed how their personal information is processed in AI recruitment tools.
- **DPIAs and risk management** – to ensure that data protection impact assessments (DPIAs) have been completed and include a comprehensive assessment of the privacy risks to people, and effective mitigations to reduce these risks.
- **Privacy trade-offs within AI** – to confirm that potential and existing trade-offs in AI systems between people's privacy and other competing values or interests have been assessed and navigated carefully.
- **Using personal information to train and test AI** – to review how personal information has been used fairly and transparently to develop AI.
- **Accuracy, fairness, and bias mitigation in AI** – to assess how potential and actual fairness, accuracy, and bias issues have been mitigated in AI development and are monitored effectively through the lifecycle of AI.
- **Human reviews in AI** – to ensure that AI, its processing, and its outputs are subject to meaningful human checks and formalised reviews, and issues addressed in a timely manner.

The audits were conducted following our data protection audit methodology. The key elements of this were:

- desk-based reviews of relevant policies and procedures;
- interviews with key privacy compliance and AI technical staff; and

- reviews of evidential documentation, including AI design documents, system specifications, and management information.

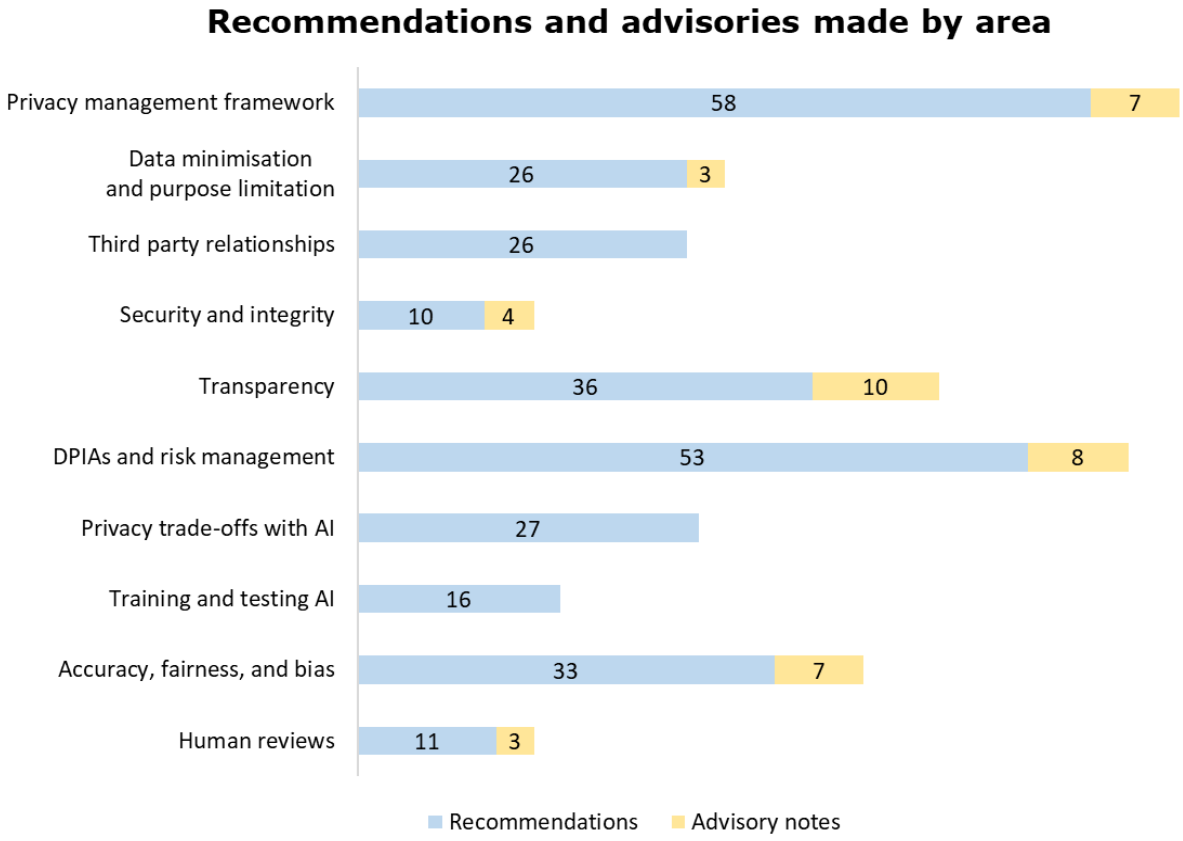
We reviewed the same focus areas for each organisation, so that we could identify key themes.

The findings from our work were taken as a 'snapshot in time' and are based on what we found at the time of each engagement. Organisations may have taken actions since to improve compliance and mitigate risks.

Each organisation received an individual audit report. Where we identified weaknesses or opportunities, we made recommendations to improve compliance with data protection law and enhance existing processes.

Impact

ICO auditors made 296 recommendations and 42 advisory notes across all engagements. These were broken down by area as follows:



Following the initial audit engagement, we asked all organisations to respond to our recommendations with appropriate actions. Organisations responded positively and were willing to take swift action to improve compliance on a voluntary basis, as follows:

- **97%** of recommendations were accepted, and actions set.
- **3%** of recommendations were partially accepted, and actions set.
- No recommendations were rejected.

We also asked for feedback on the audit experience and value added to the organisation. Respondents scored areas out of 10 as follows:

- **9.3** for improving their understanding of the requirements of UK data protection law.
- **9.7** for improving their understanding of key privacy risks in their AI tool.
- **9.0** for helping them to mitigate privacy risks in their AI tool.

- **9.3** for helping them to raise awareness of information privacy with senior leaders.

Organisations also provided the following comments about their engagements with us:

> "The process is easy to follow and efficient."
>
> "It was well managed and very professional."
>
> "Very useful and encouraging."
>
> "The audit confirmed some of our positioning around controller and processor relationships and encouraged our own thinking and research."
>
> "The audit definitely prompted us to consider our DPIAs and any gaps we might have."

Finally, after the initial audit we followed up with certain organisations where there were significant outstanding risks or areas of non-compliance. We reviewed progress and supporting evidence in these key risk areas and confirmed that these organisations had undertaken work towards implementing the recommendations we made.

Summary of findings

The findings below summarise the key observations, opportunities for improvement, and good practice we've seen during our programme of audits.

Data minimisation and purpose limitation

Developing AI systems generally requires large amounts of personal information to train AI models to reliably reproduce tasks or produce outputs. These can conflict with the data protection principles, particularly data minimisation and purpose limitation. We reviewed:

- what personal information they were processing;
- whether this was limited to what was necessary; and
- whether they stored it only for as long as needed, and did not repurpose it for other incompatible uses.

This is to comply with UK GDPR articles 5(1)(a)-(e).

The majority of AI providers had considered data minimisation in their approach to developing their AI tool. Generally, AI providers limited the information collected from people to:

- the person's name;
- contact information;
- career experience;
- relevant skills; and
- relevant qualifications or certifications.

Many AI providers also processed additional information, if instructed to do so by the recruiter.

Consider: Design AI-powered games or assessment tools to only collect the candidate's name and email address, where possible.

Example: AI providers maintaining databases of potential candidate profiles from public job networking sites generally only collected and stored the person's name, contact information, career experience, relevant skills, and relevant qualifications or certifications.

A small number also collected and stored less essential information, such as photos of the person. We recommended that these providers assess the minimum personal information needed to operate each AI element.

Most AI providers had assessed the minimum personal information needed to operate their AI tool effectively. In particular, for training and testing the AI before launch and maintaining it after launch. Some of these had recorded a minimum data profile in the DPIA or policies, with clear justification for why each data field was essential or not.

Consider: Develop AI using only pseudonymised personal information, or only aggregated information, where possible. This minimises the risk of people being identified or AI learning from irrelevant information.

Consider: Train and test AI tools using minimised datasets and techniques such as k-fold cross-validation. This allows you to use datasets several times and improve accuracy without needing large amounts of information.

The majority of AI providers had repurposed candidate personal information in their system to train, test, and maintain their AI tool. In several cases, they used it to develop other products too, usually by pseudonymising or anonymising candidate profiles. In many cases, the providers could not demonstrate that this secondary use of candidate personal information was compatible with the purpose for processing that they originally collected the information for.

Consider: Check personal information is effectively anonymised for the processing to fall outside UK data protection law. De-identified or pseudonymised information is still subject to UK data protection law.

Example: AI providers maintaining databases of potential candidate profiles typically pulled this information in bulk from public profiles on job networking sites, social media, and other open-source web content. When scraping large amounts of information this way, or purchasing scraped information from data vendors, not all providers:

- could demonstrate that the new use of information was compatible with the original purpose for processing; and
- always had a contract or written agreement from job networking or social media sites confirming that information had been collected lawfully and protected from privacy risks and potential harms.

We recommended that providers not process personal information for a new purpose and lawful basis that is incompatible with the original purpose and lawful basis it was collected for. We also recommended that these arrangements were documented in a contract or written agreement.

Consider: Assess purpose compatibility throughout the information supply chain, and build this into contracts, due diligence, and ongoing assurance checks completed with data vendors, to comply with the purpose limitation principle.

Most AI providers relied on the recruiter to set the retention period for their candidate information. This was usually one or two years after the job requisition was closed and often documented in the contract. Contracts also generally included a provision for candidate information to be retained for a short period after termination, in order to allow some time for the AI provider to stop processing and transmit the information back to the recruiter.

Consider: Check that automated retention mechanisms are deleting personal information at the end of the retention period as expected.

Example: Several AI providers maintaining a large database of potential candidate profiles had recorded their intention to retain personal information in their database indefinitely. They did not periodically 'weed' that information to remove any that might be out-of-date, inaccurate, or no longer necessary. Retaining information for longer than necessary, or

indefinitely, is unlikely to comply with the UK GDPR data minimisation and storage limitation principles.

We recommended that personal information was only retained as long as necessary to fulfil the intended purpose for processing, and that retention periods were recorded clearly and transparently.

Consider: Look for opportunities to 'weed' or delete personal information that is no longer needed, likely inaccurate, or out-of-date.

Recommendations to **AI providers** include:

- Assess the minimum personal information required to operate each element of the AI, and consider alternatives that achieve the same or a similar outcome using less or no personal information.
- Ensure all personal information processed is clearly adequate and accurate to fulfil the intended purpose.
- Document the approach to data minimisation, purpose limitation, and the other data protection principles in relevant policies and AI development documents, to promote a pro-privacy culture.
- Do not process personal information for a new purpose and lawful basis that is incompatible with the original purpose and lawful basis it was collected for. This includes retained information and information sourced from third parties, such as public job networking sites, data vendors, or recruiters.
- Retain personal information only as long as necessary to fulfil the intended purpose for processing, and record retention periods in contracts and privacy information. Do not retain personal information indefinitely, or just in case it is useful in the future.

Recommendations for **recruiters** include:

- Review the personal information collected by the AI tool and ensure this is the minimum necessary to fulfil your purpose for processing.
- Check that the personal information is not processed by the AI provider for a new purpose and under a different lawful basis that is incompatible with your original purpose and lawful basis.
- Record retention periods consistently and in detail in contracts, privacy information, and a retention schedule. This includes how long the AI provider keeps each category of personal information and why, and what action they take at the end of the retention period.

☑ **Good practice:** An AI provider gave recruiters an indicative grade for each element of a candidate's written application. They had embedded a

data minimisation approach in their AI tool and comprehensively assessed the minimum personal information needed to operate the AI. This resulted in decisions such as only collecting up to 10 years of job experience. They also embedded purpose limitation in their AI design, by:

- providing each recruiter with a separate version of the AI tool that was trained and tested using only that recruiter's candidate information; and
- optionally collecting candidate's demographic information directly from them exclusively for the purpose of monitoring bias in the AI tool.

Case study: Organisation A sourced candidate profiles from public job networking sites and data vendors to build a large database of potential candidates. They used an AI search tool to identify candidates with relevant skills or experience for a recruiter's job vacancy. Their published privacy information stated that candidate profiles in its database were retained for one year. However, the retention period was restarted every time profiles were updated with new information, which meant in practice the majority of their database was retained indefinitely. When procuring personal information in bulk from data vendors, they had not checked where it had been sourced from or considered whether they were processing it for a new purpose that was incompatible with the original purpose for collection.

We recommended that the organisation review their retention periods and how they are applied, and not hold information indefinitely by restarting the retention period regularly. We also recommended that purpose compatibility is assessed carefully when procuring information in bulk from data vendors, to ensure information is not repurposed unlawfully.

Recruiters should also assess purpose compatibility when using similar services, to ensure they aren't also repurposing information when using these services to identify possible candidates.

Find out more:

Purpose limitation - Principle (b)

Data minimisation - Principle (c)

Storage limitation - Principle (e)

What considerations about the data minimisation principle do we need to make?

Using personal information to train and test AI

We reviewed whether AI had been adequately trained using quality and representative datasets, and tested using separate datasets to ensure it produces consistent and reliable outputs, to comply with UK GDPR articles 5(1)(a) and (b), and 5(2).

Most AI providers had developed a single AI model, which they trained and tested centrally using information from all recruiters before deploying it to all recruiters in the same way. AI providers tested changes to AI before rolling them out to recruiters, such as changes to scoring or grading algorithms.

Almost all AI providers had trained and tested their tools using candidate information that they had already collected from recruiters. They had usually pseudonymised, de-identified, or anonymised the information before using it to train or test the AI.

Consider: You are the controller when using personal information from multiple recruiters to train, test, or otherwise develop your own AI tool or products. This is because you are likely to exercise control over the means and purpose of this processing in practice.

Consider: Develop the base AI model without using personal information, and train and test AI tool separately for each recruiter using only that recruiter's candidates. This also allows you to tailor AI to each recruiter.

Consider: Train and test AI tools using k-fold cross validation. This allows you to train and test AI multiple times using a minimised dataset.

AI providers had generally separated training data and testing data, to ensure they were not testing AI with the same information they used to train it. They used a range of techniques to ensure clear separation, such as data labelling, assigning a 'train' or 'test' key, or storing information in separate databases.

Most AI providers were aware of the risk of bias in AI caused by imbalances in training and testing data. However, not all had used sampling techniques to ensure datasets were diverse and representative of the relevant population. Other AI providers had either attempted to mitigate bias by 'cleaning' datasets of demographic information and proxies, or had accepted the risk of bias without mitigating it.

Recommendations to **AI providers** include:

- Ensure data labels and data labelling processes are clear and applied accurately, particularly in 'edge cases' or unusual situations, This avoids you misinterpreting or mislabelling information.
- Document the process for training and testing AI tools, including specific criteria or targets that must be met in order to progress to the next stage of AI development. This ensures there is a consistent approach to training and testing AI.
- Inform people clearly when you are using their personal information for training or testing AI. This includes the lawful basis you rely on to process it for this purpose.
- Delete training and testing datasets once you no longer need them.
- Monitor the demographic characteristics of information you use to train and test AI. Minimise imbalances such as over or under representation of a characteristic or group to ensure datasets are representative of the population or different groups of people.

Recommendations for **recruiters** include:

- Seek assurance or obtain evidence that the AI provider:
    - monitors the demographic characteristics of information used to train and test the AI tool; and
    - has identified and minimised any imbalances in the training and testing datasets, such as over or under representation of a characteristic or group.
- Be clear whether the AI provider is using your candidates' personal information to train, test, or develop their AI tool or products. They may be a processor, if they train a bespoke AI tool or algorithm with your information, just for your use, and are acting on your explicit instructions. If the AI provider trains a central AI tool or algorithm with your information which is used by all or several of its recruiters, they are the controller for that processing.

☑ **Good practice:** An AI provider makes game-based assessments to predict the candidates' skills and behaviours. They used an optional survey, issued after assessments to collect demographic characteristics directly from candidates, which they added to their anonymous profile. They used this to ensure training and testing datasets were representative of the relevant population. Where candidates did not provide their demographic information, AI providers added anonymised profiles to a third dataset. This was used to validate the tool again after initial training and testing, with a dataset to compare accuracy and bias metrics, where the diversity of the dataset was unknown.

Case study: Organisation A sourced candidate profiles from public job networking sites and data vendors to build a large database of potential candidates. They used an AI search tool to identify candidates with relevant skills or experience for a recruiter's job vacancy. They had attempted to split their entire database between training, testing, and validation. However, it wasn't clear how they kept these separate in practice. Where AI is trained with information, then tested with the same information, accuracy or bias issues may remain undetected and not addressed before 'go live'. They also hadn't ensured datasets were representative, as they had split information randomly between training, testing, and validation.

We recommended that the organisation keep training and testing data separate, to avoid testing AI with the same information it was trained with. We also recommended that the organisation ensure testing and training datasets are representative of the demographics and minimise imbalances, before beginning training and testing.

Recruiters should seek assurance or obtain evidence that AI tools are trained and tested with representative datasets, before using them.

Find out more:

How should we distinguish purposes between AI development and deployment?

Accuracy, fairness, and bias mitigation in AI

Fairness must be a key consideration throughout the design and development of AI. We assessed whether AI providers had regularly monitored accuracy and bias and swiftly addressed issues throughout the AI lifecycle, to comply with UK GDPR articles 5(1)(a) to (e) and 25.

AI providers had usually considered the accuracy of their AI tools during development, and assessed it before launch to test that it reliably generated accurate outputs. They measured accuracy using a combination of precision, recall, area under curve, and other similar metrics, to test for positive correlation between the AI and expected results.

Example: Providers tended to set their own tolerances or minimum targets for accuracy when developing their AI tools. Some providers only included data points that produced near-perfect positive correlation in their AI tool. Others accepted a wider range of data points ranging from those with slight positive correlation, relying on the indicative nature of

their tool. AI providers usually had a process to exclude or reduce the weighting of data points that did not positively correlate, in order to improve overall accuracy of the tool.

The majority of AI providers repeated accuracy tests at least periodically after launch and especially before implementing changes or updates. This ensured accuracy remained within tolerance and did not erode over time.

Example: An AI provider had not formally assessed the accuracy of their AI tool using a planned testing methodology, before launching. Instead they relied on their AI tool being 'at least better than random'.

In this case, the AI produced only indicative grades in very limited areas and was only designed to support recruiting managers. Recruiting managers also received training and resources that clearly stated grades were indicative only. Recruiters changed inaccurate grades, which the AI learned from to improve accuracy over time.

However, being 'at least better than random' would usually not be sufficient to comply with data protection law, where AI actively makes recruitment decisions without human intervention. In these cases we would recommend that providers must assess and monitor the accuracy of their AI before launching, and take action to address accuracy issues. The AI should reach the target accuracy level before processing personal information.

We recommend recruiters consider whether an AI tool is sufficiently accurate before using it, based on how they intend to use it and what influence the decision will have in their recruitment process. Recruiters should not rely on inaccurate AI to make decisions alone.

Consider: Assess the validity and accuracy of all data points prior to launching the tool. Restrict creating further data points without testing. This will prevent accuracy erosion after launch.

Consider: Engage cognitive behavioural and psychometric experts to regularly test and review AI logic, scoring, and outputs for potential accuracy or bias issues.

AI providers had also considered the potential for bias in their AI tools, usually by measuring potential bias using an adverse impact analysis methodology. In many cases they used the 'four fifths rule' as a minimum threshold. This means the selection rate for any group must be at least four fifths or 80% of the selection rate of the group with the highest rate.

AI providers could generally demonstrate actions they took to improve the AI tool where they identified bias. For example, reducing the weightings of data points or excluding data points that they considered to be negatively impacting bias or causing an adverse impact on groups. They typically repeated bias tests periodically at the least, or before launching changes to AI tools.

Example: Several AI providers chose to estimate or infer people's characteristics from their personal information or other parts of their candidate profile, rather than collecting it directly. This usually involved predicting the person's gender and ethnicity – often from their name but sometimes also from elements of their candidate profile or application. They added this to the information they already held about them.

Information intentionally inferred in this way is still special category data. AI providers did not always treat inferred information as special category data or identify an additional condition for processing. Processing special category data without a lawful basis and additional condition is unlikely to be lawful under the UK GDPR. Creating additional personal information in this way was often invisible and not transparent. This was because candidates weren't always informed clearly that this was happening and were unable to access additional information created.

We recommended that these providers identify a lawful basis and additional condition before processing special category data, and not use inferred information instead of collecting accurate information lawfully.

Using inferred or estimated information to measure, monitor, and address bias in AI tools, had several limitations. Bias monitoring was limited to only gender, ethnicity, and age. They could not estimate other protected characteristics in the UK Equality Act 2010 from the information available. They could only estimate gender, ethnicity, and age accurately as large categories, for example "male" or "female", or "white", "black", and "Asian" for ethnicity. Smaller categories were too difficult to estimate reliably, and would likely be too small a sample to effectively measure.

AI providers using inferred information were generally unable to demonstrate that it was reliable and accurate enough to mitigate bias effectively in their AI tools. This means that if bias creeps into the AI tool, there is a significant risk that it won't be highlighted by bias testing. Very few AI providers had adequately assessed this risk.

Consider: It is more accurate to measure bias by repurposing demographic information you collect from successful candidates for equality monitoring of new hires, than inferring demographic information. However, repurposing information in this way is unlikely to comply with

the UK GDPR purpose limitation principle and you will need an appropriate lawful basis for this purpose.

Consider: Collect demographic information directly from candidates, via an optional survey issued after the recruitment process. You can use this to more accurately monitor bias where provided, or for additional validation tests with unknown samples where not provided.

Where AI providers did not measure potential bias using an adverse impact analysis methodology, they usually relied on blinding AI to personal characteristics in input information. They did this by removing demographic information and proxies from input information.

Consider: Monitor for bias where AI is processing video interview recordings, and reduce bias by removing proxies for demographic characteristics or modifying the AI learning process or model.

Consider: Engage external audits of AI tools, such as adverse impact assessments or reviews of AI source code for errors.

Consider: Engage regularly with national projects, industry partnership groups, and stakeholder networks to develop new ideas and share good practices in AI development and provision.

Recommendations to **AI providers** include:

- Demonstrate that AI is operating fairly and not discriminating against minority groups by:
    - testing regularly for fairness, accuracy, or bias issues in AI tools or outputs;
    - addressing any issues you identify effectively;
    - reporting key performance indicators for accuracy and bias regularly to senior managers and key stakeholders; and
    - retaining test results or reports and evidence of actions you've taken to address issues.
- Assess and mitigate potential or actual fairness, accuracy, and bias risks in AI tools being unintentionally taken into account by the AI, including:
    - human bias present in AI development;
    - sampling bias in training and testing information;
    - incorrect or inadequate information labelling; and
    - demographic information present in text or video.

- Consider a wide range of characteristics when monitoring fairness and bias, including:
    - gender and gender identity;
    - racial or ethnic origin, disability; and
    - other characteristics listed in UK GDPR recital 71 or protected characteristics in the UK Equality Act 2010.
- Evaluate algorithmic fairness limitations and how you can navigate them, such as:
    - unequal distribution of protected characteristics;
    - people with multiple protected characteristics; and
    - proxies of protected characteristics present in training and testing data.
- Document how staff should log and respond to candidates challenging AI outputs, particularly complaints about unfair, inaccurate, or biased outputs.

Recommendations for **recruiters** include:

- Check how the AI provider monitors and mitigates fairness, accuracy, and bias in the AI tool, particularly what personal information they use to do this and where they sourced it from.
- Review potential or actual fairness, accuracy, and bias risks of processing personal information or operating the AI tool, and the measures in place to mitigate them. Record these risks in a DPIA.
- Request test results or reports and evidence of actions AI providers have taken to address fairness, accuracy, or bias issues in AI tools or outputs. Ensure these demonstrate that AI is operating fairly and not discriminating against minority groups.

☑ **Good practice:** Some AI providers gave game-based assessments to predict the candidates' skills and behaviours. They had built in accuracy and bias assessments at each stage of tool development. They tested the accuracy and validity of each data point prior to launching the tool, and measured bias and adverse impact by each demographic group for each data point. They repeated assessments periodically after launch. Tool outputs were overseen by a team of cognitive behaviour and psychometric subject matter experts who compared distributions of scores for each demographic group after each recruitment campaign.

Case study: Organisation A sourced candidate profiles from public job networking sites and data vendors to build a large database of potential candidates. They used an AI search tool to identify candidates with relevant skills or experience for a recruiter's job vacancy. They tested the accuracy of the search tool, however, they had not internally tested for

bias. Instead they used an external organisation that tested for adverse impact using only artificial datasets made for this purpose. They inferred the gender and ethnicity candidates from their name and other profile information. However, instead of bias mitigation, they used this to allow recruiters to filter demographics in or out of the list of suggested candidates, which was unlikely to be fair or have an appropriate lawful basis for processing.

We recommended that the organisation test and monitor the potential for bias and discrimination, using accurate and adequate information to do this effectively. We also recommended that the organisation assess and mitigate fairness, accuracy, and bias risks of their AI and processing.

Recruiters should check that the AI is operating fairly and not discriminating against minority groups before using it and regularly after, by reviewing test results and evidence of issues being addressed.

Find out more:

What do we need to know about accuracy and statistical accuracy?

What about fairness, bias and discrimination?

What are technical approaches to mitigate discrimination in ML models?

Using AI to make inferences

Fairness in the AI lifecycle

Good Work Algorithmic Impact Assessment (IFOW)

Transparency

It is important that AI developers and users are open and transparent about how they process personal information using AI to make recruitment decisions or produce outputs. We checked that people were actively provided with clear non-technical explanations that they could understand, to comply with UK GDPR articles 5(1)(a), 13, and 14.

AI providers had published privacy information in a privacy policy on their website, which was text-based and structured into sections with sub-headings. Most privacy policies we reviewed contained at least:

- an overview of the personal information fields they processed;
- the primary purpose for processing;
- safeguards to protect information;
- people's rights under UK GDPR; and

- contact information for the AI provider and the supervisory authority.

Example: Some AI providers had published one privacy policy that covered several different instances of personal information processing, multiple different categories of people, or multiple different jurisdictions or laws. In most cases, this was misleading and might confuse people.

We recommended that these providers produce clear privacy information that gives people relevant information. For example by structuring privacy information in clear sections by activity or producing tailored privacy information for each processing or category of people.

Consider: Produce a privacy policy specifically for candidates on your AI platform and relevant to the UK GDPR requirements. This ensures candidates understand which information is relevant to them and are informed correctly.

Consider: Supplement text-based privacy information with informative pop-up messages or bite-sized information at the point that processing takes place, or visual aids such as data flow maps.

Several privacy policies did not contain sufficient detail about:

- each specific instance of personal information processing;
- the lawful basis and additional conditions specifically they were relying on; or
- how long they would retain the information for.

Some privacy policies did not state or incorrectly stated whether the AI provider was a controller or processor.

Consider: Review privacy information or resources regularly to ensure they are accurate, particularly before implementing changes to the information processing or AI functionality.

Example: Where AI providers processed personal information for secondary or alternative purposes, they did not always mention it in the privacy policy. For example, when they used it to infer demographic information for bias mitigation or to train and test AI, candidates were not aware this processing was taking place.

We recommended that providers inform people transparently and fully how their information is processed. Personal information processing that is not transparent is unlikely to be lawful under UK GDPR article 5(1)a.

Consider: Provide model privacy information or text that explains technical AI processing in a clear understandable way. This helps recruiters understand the AI system and accurately inform candidates.

Consider: Check people actually understand how you process their information by:

- testing privacy information with users;
- conducting focus groups or surveys; or
- tracking when candidates open privacy information.

Several privacy policies we reviewed only referenced AI very broadly, if at all. The processing is effectively invisible if people are not informed specifically how their personal information is processed within AI tools eg the logic involved in making predictions or producing outputs, or how personal information is used to train and test the AI.

Consider: Proactively publish information and resources and avoid overly complex explanations, or technical or legalistic language. This will build trust in AI products.

We expected AI providers and recruiters to have contracts in place that clearly set out which party was responsible for informing candidates how they were processing their personal information. Most AI providers relied on the recruiters as controllers to inform candidates how they were processing their personal information. However, contracts between AI providers and the recruiters were often unclear about which party was responsible for informing people or providing privacy information.

Recommendations to **AI providers** include:

- Provide detailed privacy information to inform people how you are processing their personal information, if you are the controller or contractually responsible for informing people.
- Inform people clearly how their personal information is processed within AI tools, including the logic involved in making predictions or producing outputs, and how you use personal information to train, test, or otherwise develop the AI.
- Inform people clearly when creating additional personal information or special category data about them eg inferring their gender or ethnicity from their name. Identify an appropriate lawful basis and additional condition for this processing.
- Provide privacy information to candidates within one month of obtaining their information where you do not collect this directly from them eg from job networking sites, social media, other public

sites, or third party data vendors. If a valid exemption applies, document this assessment and justification in sufficient detail and keep it under regular review.

Recommendations for **recruiters** include:

- Ensure contracts clearly define how you provide privacy information to candidates, and which party is responsible for this.
- Provide detailed privacy information to inform candidates how you process their personal information. Where you instruct the AI provider to do this, check the privacy information is clear, accurate, and detailed.

☑ **Good practice:** Some AI providers made game-based assessments to predict the candidates' skills and behaviours. Although they were processors, they published several resources on their website explaining:

- how the assessment tool processed personal information;
- the science involved in predicting skills and behaviours; and
- how they trained and tested the tool using this information.

Resources included both text and graphics, and they signposted candidates to them with the assessment invite.

Case study: Organisation A provided an AI tool that scored candidates' written responses to interview questions. Their privacy policy did not contain sufficient detail about processing and directed candidates to the recruiters' privacy policies. However, the recruiters' privacy policies directed candidates back to Organisation A's privacy policy. Contracts with recruiters were unclear which party was responsible for informing candidates, and as a result people were not sufficiently informed by either party. Organisation A was the controller when anonymising candidate applications for training and testing the central AI tool, and inferring demographic characteristics from candidate names to monitor for bias. Their privacy policy contained very limited information about this processing. Therefore, they did not inform candidates and processing was effectively 'invisible', which is likely to breach UK GDPR article 5(1)(a).

We recommended that the organisation inform candidates before processing their personal information as a controller. We also recommended that the organisation specify clearly in their template contracts whether they or the recruiter is responsible for providing privacy information, to avoid not informing candidates at all.

Recruiters should also check which party is responsible for informing candidates how their information is processed, and check this is done, before using the AI tool.

Case study: Organisation B sourced candidate profiles from public job networking sites and data vendors to build a large database of potential candidates. They used an AI search tool to identify candidates with relevant skills or experience for a recruiter's job vacancy. They published a privacy policy on their website. However, they did not actively inform people that they were processing and storing their personal information, relying on the 'disproportionate effort' exemption in UK GDPR article 14(5)(b). Organisation B had each candidate's name and email address and hadn't considered available options. They therefore couldn't justify why providing privacy information would involve disproportionate effort.

We recommended that the organisation actively provide privacy information to people within one month, assess available options before relying on exemptions, and keep exemption decisions under review.

Recruiters should check that potential candidates are informed how their information is processed, before using similar services.

Find out more:

Principle (a): Lawfulness, fairness and transparency

How do we ensure transparency in AI?

Explaining decisions made with AI

What methods can we use to provide privacy information?

Right to be informed checklist

Privacy trade-offs within AI

When developing or using an AI tool, there can be different values and interests to consider that may pull in different directions. We reviewed how organisations had identified and navigated trade-offs between privacy and other competing values or interests when developing their AI tools, to comply with UK GDPR articles 5(2) and 24-25.

AI providers consistently identified the following key trade-offs between privacy and other competing values or interests:

- Accuracy versus explainability – including how more data points improves output accuracy, but makes it harder to explain how the AI works to people.
- Data minimisation versus statistical accuracy and validity – collecting and processing more personal information can improve the accuracy and validity of outputs, but collecting more information than needed is unlikely to comply with the data minimisation principle.
- Transparency versus understandability – explaining AI in granular technical detail in privacy information may seem more transparent, but impacts how understandable the privacy information really is.

Consider: Assess the benefits and limitations of available AI methodologies, including complex machine learning models that improve output accuracy but are less transparent and explainable.

Example: Potential trade-offs in AI tools were typically recorded in a DPIA or an AI product specification document and signed off by the product manager or legal counsel.

However, several AI providers had not recorded the assessment of trade-offs anywhere, and others could not demonstrate that they had considered or navigated trade-offs at all. We recommended that these providers identify and assess all trade-offs in their AI tool, and document the chosen approach and reasons, such as in a DPIA.

Consider: Record design considerations, decisions, and justifications in internal wikis or technical documents. These can be reviewed and used as a guide by relevant staff, senior leaders, and legal counsel.

Recommendations to **AI providers** include:

- Identify and assess all potential and existing trade-offs in AI tools between the person's information privacy and other competing values or interests, as part of your DPIA. For each trade-off you should:
    - consider the options available;
    - assess the impact on people's rights and freedoms; and
    - record the chosen approach and justification.
- Document the process for identifying and assessing trade-offs during the design and development of AI, including:
    - how you will assess the impact on people's privacy rights;
    - who you will consult; and
    - who will sign off trade-off decisions at a senior level.

- Review trade-off analysis and decisions regularly, particularly before making changes to the AI or processing. Consider new or emerging trade-offs or new technical approaches that are available.

Recommendations for **recruiters** include:

- Identify and assess all potential and existing trade-offs in AI tools between a person's information privacy and other competing values or interests. Do this as part of your DPIA and wider approach to privacy.

☑ **Good practice:** An AI provider predicted the likelihood of a candidate being progressed positively to the next selection stage by a recruiter. They had assessed trade-offs when developing their AI tool, such as using more complex models that increased accuracy but decreased transparency and explainability. They recorded the considerations and decisions on product tickets that stakeholders could add comments to.

Case study: Organisation A suggested candidates who matched or best fit a recruiter's job vacancy from a large database of potential candidates. They had recorded their key design decisions in a DPIA. However, they had not recorded the benefits and risks of available approaches, or how they had balanced privacy and the competing interests. They had also not reviewed or updated decisions regularly, as in some cases they had taken a different approach in practice than what they recorded.

We recommended that the organisation:

- record the available approaches and reasons for their decision;
- update and review this regularly to consider new trade-offs or new approaches available; and
- document this process in development roadmaps to ensure it happens in future.

Recruiters should also assess trade-offs in AI tools, as part of a DPIA.

Find out more:

How should we manage competing interests when assessing AI-related risks?

How should we balance data minimisation and statistical accuracy?

How should we assess security and data minimisation in AI?

What about fairness, bias and discrimination?

Human reviews in AI

We reviewed how AI outputs or decisions have been meaningfully reviewed and quality checked, to comply with UK GDPR articles 5(1)(a) to (e). We also reviewed if AI made automated decisions with legal or similarly significant effects, and if this complied with article 22.

Most AI providers included human intervention at some point in the AI process, such as a human review or sample of AI outputs.

Consider: Engage cognitive behaviour and psychometric subject matter experts, who are closely involved in the AI tool operation. Randomly sample AI outputs to ensure they are fair, valid, and accurate.

Consider: Complete both random and risk-based human reviews of AI outputs, where risk-based reviews are triggered by:

- uncertain or ambiguous inputs;
- unexpected or ungraded outputs; or
- where performance metrics highlight potential bias.

Example: Where human reviews were a formal stage of the process, staff completing reviews were trained on the review methodology, including what to check, how to identify and record issues, and what action to take.

Where human reviews were not formalised, staff were typically not trained and they did not complete reviews consistently and thoroughly. We recommended that these providers formalise and document the human review process and provide relevant training to reviewers, to ensure reviews are consistent.

AI tools reviewed in this work were designed and intended only to support human recruiters to make decisions, rather than to make automated recruitment decisions without human intervention. Most AI tools provided only indicative grades or fit scores, or suggested a candidate's behaviour traits or skills which a human recruiter could consider in their decisions.

Consider: Prevent using AI outputs to make automated recruitment decisions. Prevent human recruiters from progressing or rejecting candidates based solely on indicative grades or fit scores produced by the AI. Clearly record the intended use in contracts, marketing materials, or training and resources you provide to recruiters.

Recommendations to **AI providers** include:

- Subject AI outputs to robust and meaningful human reviews or quality checks, so you effectively address output accuracy or bias issues at an early stage.
- Complete sampling checks on changes to AI algorithms, to prevent introducing errors or bias into AI unintentionally.
- Keep records of completed human reviews or quality checks, including any actions you've taken, changes you've made, and feedback given to development teams, as well as reasons or justification.
- Document in your policies the process for human reviews of the AI outputs, including when humans may override the algorithm, and how managers will sample and check human reviews.

Recommendations for **recruiters** include:

- Ensure that recruiting managers do not use AI outputs (particularly 'fit' or suitability scores) to make automated recruitment decisions, where AI tools are not designed for this purpose.
- Offer a simple way for candidates to object to or challenge automated decisions, where AI tools make automated decisions.

☑ **Good practice:** Some AI providers gave recruiters an indicative grade for each element of a candidate's written application. They reviewed a random sample of grades and how the AI scored and weighted each element, to ensure grades were fair and accurate. They also reviewed grades that had been changed by recruiters, to identify issues or trends in AI scoring. Finally, internal stakeholders monitored numbers of grades changed by recruiters and other relevant performance metrics.

Case study: Organisation A sourced candidate profiles from public job networking sites and data vendors to build a large database of potential candidates. They also used an AI search tool to identify candidates with relevant skills or experience for a recruiter's job vacancy. They did not review outputs of the AI search tool to check it was working as intended. Instead they relied on recruiters to flag if the AI suggested candidates that did not fit the vacancy. However, they had not made this clear to recruiters, provided a mechanism for recruiters to provide this feedback or highlight errors, and couldn't evidence how else AI outputs were reviewed.

We recommended that the organisation introduce robust and meaningful human reviews or quality checks of AI outputs, so issues are addressed at

an early stage. We also recommended that the organisation implement a feedback mechanism for recruiters to report errors for review.

Recruiters should not use AI tools to make automated recruitment decisions, where the AI is not designed for this purpose.

Find out more:

What is the role of human oversight in AI decisions?

DPIAs and risk management

DPIAs are likely to be required by law for AI tools, as they almost always involve processing or innovative technology that is likely to result in a high risk to people's rights and freedoms. We checked whether organisations had completed a DPIA for their AI tool, and reviewed DPIAs to ensure they were meaningful and detailed. We also checked that organisations had identified and mitigated the risks to people before processing their personal information, to comply with UK GDPR articles 5(2), 24-25, and 35-36.

The majority of AI providers had completed a DPIA for their AI tool before they used it to process personal information. However, in some cases, AI providers had completed DPIAs retrospectively or just prior to the audit. Some DPIAs did not include dates so it was unclear when they were completed or due for review.

DPIAs we reviewed usually included at least:

- an overview of the purpose and scope of processing;
- the personal information fields that would be collected; and
- a summary of the safeguards in place.

However, in many cases DPIAs were not sufficiently detailed, and often did not include key elements such as:

- a detailed map of data flows through the AI system;
- consideration of how to meet the data protection principles;
- meaningful assessment of the necessity and proportionality of processing; and
- consideration of alternative approaches that might use less personal information to achieve the same outcomes.

Example: Several DPIAs included an assessment of key risks to or potential impacts on people, and proposed measures to reduce these risks to an acceptable level.

It was often unclear how new risks or changes to risks were captured, or who had checked that mitigating controls were fully implemented and effective before processing started. We recommended that DPIAs and risk mitigation measures are reviewed regularly, including checks that controls are working effectively.

Consider: Assess the risks to people's rights and freedoms of processing their information, rather than risks to the organisation. Identify and implement measures to mitigate each risk.

Key risks to people that had been consistently assessed in DPIAs were:

- potential anomalies in the AI tool or processing operations, resulting in inaccurate processing or outputs;
- potential bias in the AI tool or in training data, resulting in biased processing or outputs;
- inappropriate staff or third party access to personal information or AI source code;
- personal data breaches, cyber-attacks, or other interference to the AI system; and
- accidental collection of unnecessary personal information in written or video responses, without a lawful basis and purpose for collection.

Example: Most AI providers suggested that relevant internal stakeholders were involved in the DPIA. However, very few had clearly recorded any feedback they received from internal experts, how they had considered this, and what they had changed as a result. Almost no AI providers had sought the views of the wider public on the intended processing, particularly to find out whether this use of their personal information was reasonably expected and transparent to them.

We recommended that providers consult meaningfully with relevant internal and external stakeholders, consider results and feedback received, and clearly record changes made following consultation.

Consider: Use a privacy compliance tool to complete DPIAs, record stakeholder comments, track changes, and automatically prompt reviews.

Consider: Document the DPIA process and when you are required to do a DPIA in relevant policies, product development roadmaps, and project flow charts.

DPIAs usually included at least some advice from the internal privacy lead or staff member acting as Data Protection Officer, but in many cases had not formally been approved by a senior manager.

AI providers had not regularly reviewed or updated most DPIAs after the AI tool was made live, or did not clearly record when reviews had taken place or what had changed. As a result, it was difficult to determine if:

- identified risks had changed;
- they had checked mitigating controls and whether they were still effective; or
- new privacy risks had arisen that they hadn't assessed and mitigated at all.

Consider: Support recruiters to complete a DPIA by providing relevant technical information about AI or proposed processing, or evidence of controls. Do not charge a fee for this as it might discourage recruiters or prevent them from robustly assessing the risks.

Consider: Provide training to new recruiters, demonstrate the AI tool and how to interpret outputs, or transparently publish reference guides and resources on your website. This will support recruiters to use AI tools in the way intended and understand potential privacy risks.

Recommendations to **AI providers** include:

- Complete a DPIA before commencing processing that is likely to result in a high risk to the people's rights and freedoms, from early in development and before you process information.
- Consider completing a DPIA for proposed processing activities using AI or other innovative technology, even when acting as a processor.
- Ensure DPIAs are comprehensive and detailed, including:
    - the scope and purpose of the processing;
    - a clear explanation of relationships and data flows between each party;
    - how processing will comply with UK GDPR principles; and
    - consideration of alternative approaches.
- Consult with us at the ICO on DPIAs where there is still a high residual risk to the rights and freedoms of people in the UK after mitigation.
- Ensure you robustly review DPIAs and an appropriate senior manager formally approves them (such as one responsible for privacy and data protection).

- Review DPIAs and risk mitigation measures regularly, and check controls are working effectively. Carry out more frequent reviews when there is a system change or change to information processing.

Recommendations for **recruiters** include:

- Complete a DPIA before commencing processing that is likely to result in a high risk to the people's rights and freedoms such as procuring an AI recruitment tool or other innovative technology.
- Ensure DPIAs are comprehensive and detailed, including:
    - the scope and purpose of the processing;
    - a clear explanation of relationships and data flows between each party;
    - how processing will comply with UK GDPR principles; and
    - consideration of alternative approaches.
- Assess the risks to people's rights and freedoms clearly in a DPIA, and identify and implement measures to mitigate each risk.
- Follow a clear DPIA process that follows the recommendations above.

Case study: Organisation A provided game-based assessments to predict the candidates' skills and behaviours. They had completed DPIAs for each component of their AI tool. However, these were very light in detail, including:

- they did not clarify whether they were a controller or processor;
- they did not describe technical or organisational measures to protect information;
- they contained several other anomalies or errors;
- several questions were marked 'N/A' or left unfinished;
- they did not show any input from internal stakeholders, but did include some high-level advice from the external DPO;
- risks related to the organisation rather than people, and relevant staff were not fully aware of mitigating controls that should be in place; and
- they had not reviewed them since 2019, despite several significant changes to the AI tool in that time.

We recommended that the organisation ensure DPIAs are comprehensive and detailed, and include consultation with stakeholders and regular robust reviews of risks to people. We also recommended that mitigating controls are communicated to relevant staff and checked regularly.

Recruiters should also complete detailed DPIAs, identify risks to people, and regularly check that mitigating controls are in place and effective.

Find out more:

Data Protection Impact Assessments (DPIAs)

When do we need to do a DPIA?

How do we do a DPIA?

What do we need to consider when undertaking DPIAs for AI?

Information security and integrity

AI systems are likely to be integrated with several other software components and third party systems, and involve more complex data flows. We assessed how effectively information security, integrity, and access risks had been managed, and whether appropriate measures were in place to protect the AI tool and personal information in it, to comply with UK GDPR articles 5(1)(f) and 32 to 34.

AI providers generally hosted their AI tools on third party infrastructure, in many cases using 'elastic' cloud servers with no fixed capacity in order to minimise availability risks. Many providers were able to offer servers in specific countries or states to recruiters. This maintained data sovereignty from collection and while in transit.

AI providers had implemented automated monitoring systems to monitor their infrastructure, including:

- scanning for vulnerabilities;
- detecting and analysing real-time security threats;
- taking limited remedial action to automatically restrict the potential impact of a threat; and
- reporting threat alerts to relevant staff and senior leaders.

Consider: Implement several monitoring and identification systems operating in 'layers', to provide assurance that systems are working as expected. This also increases the likelihood that at least one system identifies the threats.

Consider: Run a 'bug bounty' and incentivise reports of bugs and possible vulnerabilities, so you can resolve them before they are exploited.

AI providers had implemented a range of technical controls to protect their AI tool and personal information. In most cases they had clearly

documented these in internal information security policies and system operating documents, including:

- encrypting information to minimum AES 256-bit symmetric or equivalent asymmetric standard, at collection, transit, and rest;
- malware, anti-virus protection, and organisation-configured security software on all workstations and devices connected to the network;
- network access restrictions, firewalls, intrusion detection alerts, and automated real-time traffic monitoring and filtering;
- a robust patching process with priority for external-facing assets and critical or emergency patches;
- practicing secure development by using 'sandbox' or a separate test environment for product development and testing AI code changes;
- independent line-by-line review of AI code changes and robust authorisation processes before deploying code changes;
- logging or tagging information assets, and secure asset disposal;
- business continuity plans and fallback processes; and
- automated full and partial back-ups and restoration processes.

Consider: Undertake annual external assessments of information security management systems, including vulnerability and penetration tests, This ensures AI and wider systems meet relevant ISO and SOC standards. Also consider rotating providers periodically so findings are independent and unbiased.

Most AI providers had a data breach policy or response plan which set out how they would investigate, manage or report data breaches or near misses.

Consider: Document personal data breach processes in detail, including:

- key staff responsibilities;
- the statutory requirement to report relevant breaches to the ICO within 72 hours;
- processes for notifying affected people; and
- differentiate processes as a controller or processor.

AI providers usually assigned access permissions to joiners, movers, and leavers based on a role map. This set out the minimum access required to systems and personal information for specific job roles. Access permissions to the AI tool and personal information in it were typically restricted to only a small number of senior leaders, and subject to additional controls, such as time-limited connection restrictions.

Example: In some cases it was unclear how AI providers handled changes in access for staff moving role internally, and how they regularly

reviewed role maps or existing access permissions. In other cases, they automatically logged user activities, including access, read, edit, and delete. However, they did not meaningfully review logs or subject them to automated monitoring to prevent inappropriate access going undetected.

We recommended that access is granted or changed in a timely manner, and access management processes are formalised. We also recommended that access activity logs are periodically reviewed to identify instances of inappropriate access or trends.

Consider: Review all access permissions assigned in AI systems regularly, including privileged permissions to access AI code and personal information.

Recommendations to **AI providers** include:

- Record technical and organisational controls in policies and contracts, including how you monitor these, and regularly review them to ensure the information is up to date and accurate.
- Assess security risks or vulnerabilities, record findings and risk treatment actions in a risk register. Review regularly to ensure mitigating controls are fully in place and effective.
- Clearly document key decision-making processes and staff information security responsibilities in relevant policies and staff guidance. Ensure staff with security responsibilities are sufficiently trained.
- Implement equivalent technical security controls on staff devices and monitor that controls remain fully in place, to protect personal information to at least the same level as on company devices.
- Test the effectiveness of your data breach management processes in practice. For example, by holding periodic walkthrough exercises, desktop scenarios, or simulations with key staff.

Recommendations for **recruiters** include:

- Undertake meaningful due diligence that includes obtaining evidence that technical and organisational controls are in place and personal information is secure during collection, transit, and at rest.
- Complete regular compliance checks throughout the contract lifecycle to get assurance that technical and organisational controls remain in place and effective.
- Document required technical and organisational controls clearly in the contract, including:
    - access management controls;
    - change management processes; and

- clear responsibilities for each party in the event of a data breach or near miss.

☑ **Good practice:** One AI provider had engaged a third party that conducted 20-25 security assessments of its AI tool and system infrastructure every year. This ensured security measures remained in place and effective, and they continuously improved their security.

Find out more:

Security, including cyber security

Information security checklist

UK GDPR data breach reporting and self-assessment

Personal data breaches: a guide

Management frameworks

It is vital that AI systems are developed within an embedded management framework with clear accountability for privacy and data protection. In our audits, we reviewed:

- how organisations fulfilled their responsibilities as controllers or processors;
- how they identified an appropriate lawful basis for processing and an additional condition, where relevant, to comply with UK GDPR articles 6, 7, and 9;
- processes to handle individual rights requests, to comply with UK GDPR articles 12-22; and
- whether there was effective senior leadership oversight, adequate and relevant staff training, and robust policies, to comply with UK GDPR article 5(1)(a).

The majority of AI providers could evidence clear privacy management frameworks that supported oversight of data protection compliance from senior leadership throughout the organisation, particularly in developing and providing AI tools. AI providers had either appointed a Data Protection Officer or nominated a senior manager responsible for privacy, who regularly checked compliance, monitored KPIs or performance metrics, and reported risks to senior leaders. Most AI providers had also formalised data protection responsibilities in staff contracts and job descriptions, and recorded their agreed approach to data protection compliance in policies, so staff were aware what was expected of them.

Consider: Do periodic staff surveys or tests to check privacy knowledge, measure awareness of policies, and identify gaps or relevant training.

Consider: Provide additional mandatory training for staff with key privacy responsibilities, such as 'privacy by design' training for product teams and 'AI fairness' training for AI technicians. Ensure staff refresh training regularly and are equipped with the relevant knowledge.

Consider: Implement regular process reviews, internal or external privacy compliance assessments, and reviews of privacy risks. This will improve the AI control environment and ensure that AI complies with privacy policies.

In UK GDPR, controllers are the main decision-makers and exercise overall control over the purposes and means of the processing. Processors process personal information on the behalf of, and only on the instructions of, the controller. Controllers and processors have different responsibilities and obligations under UK GDPR. Organisations can simultaneously be a controller for some processing activities and a processor for others. Or they can be both a controller and a processor of the same personal information, if processing it for different purposes.

AI providers had considered whether they were a controller or processor when processing personal information in their AI tool. Several AI providers had determined their role as a controller or processor correctly. However, others could not demonstrate that they had determined their role correctly or at all.

Example: Providers of AI screening or selection tools determined that they were the controller when processing personal information to develop a single central AI tool, and a processor when processing candidate information through the AI tool on the recruiter's instructions. Some providers created a separate version of the AI tool for each recruiter that they developed with and used for only that recruiter's candidate personal information. This demonstrated that they were a processor acting on the recruiters' instructions, and that the recruiter could exercise control over the purpose and means of processing.

Other providers developed a single, central AI tool. They were controllers as they decided how and why personal information was processed in practice when developing the AI. Future recruiters can't feasibly be controllers of processing that took place before they procured the AI tool.

Example: Providers of AI sourcing tools generally determined they were:

- the controller when processing personal information to build a database of candidate profiles and developing search algorithms; and
- a processor when processing candidate information to source relevant candidates on the recruiter's instructions.

Consider: You are the controller if you:

- exercise overall control of the means and purpose of processing of the personal information; or
- process the personal information again for your own purposes.

Consider: You are a processor only if:

- recruiters are able to exercise meaningful control of the means and purpose of processing; and
- you don't process the personal information again for your own purposes.

AI providers acting as controllers generally relied on legitimate interests as a lawful basis for processing. However, they had not always completed a legitimate interests assessment to balance their interests with a person's interests and privacy rights. They also hadn't always informed people that they were processing their personal information.

Generally, most AI providers did not rely on consent as a lawful basis for processing personal information in their AI tools.

Where processing special category data as controllers (such as to measure and monitor potential bias in AI tools), AI providers had not always treated inferred information as special category data. This requires an additional condition for processing, so they were processing it without an additional condition. Processing personal information without a lawful basis, or processing special category data without a lawful basis and additional condition, is unlikely to be lawful under the UK GDPR.

Example: Some AI providers with sourcing tools and candidate databases were processing special category data on the additional condition that it was inferred from information that people had manifestly made public on social media or job networking sites. The providers could not clearly describe or demonstrate how this additional condition was appropriate in practice.

We recommended that these providers reconsider their lawful basis and additional condition for processing inferred special category data, and

inform people before processing their information. We also recommended that they cease processing and permanently delete the information, if they cannot identify an appropriate lawful basis and additional condition.

Consider: When using your AI tool for your own recruitment, you are the controller and responsible for meeting UK GDPR requirements. This includes identifying a lawful basis, and an additional condition if you are processing special category data.

Consider: If you issue an optional survey after assessments to collect demographic characteristics directly from candidates and process based on their clear explicit consent, ensure people can withdraw consent as easily as they can give it to comply with the law.

Recommendations to **AI providers** include:

- Identify whether you are a controller or processor for each specific instance where you are processing personal information. Record this in privacy information, contracts, DPIAs, and other documents.
- Before you start processing, identify a lawful basis for each instance where you are processing personal information as the controller. Also identify an additional condition where you are processing special category data. Record this in privacy information, contracts, DPIAs, and records of processing activities (RoPAs).
- Do not process personal information if you cannot identify an appropriate lawful basis. Do not process special category data if you cannot identify both an appropriate lawful basis and additional condition.
- Produce a RoPA based on regular data flow mapping that records every processing activity in detail. This should include the purpose, the lawful basis and additional condition, and who you share information with.
- Record data protection and AI privacy processes in detail in policies, so staff can find information and understand their responsibilities.
- Document and implement processes to comply with individual rights requests. This should include how each individual right will be handled within the AI tool, and how you will communicate requests to recruiters or other relevant third parties.

Recommendations for **recruiters** include:

- Ensure the AI provider's role as a controller or processor has been correctly identified for each instance where they process personal

information. Record this clearly and consistently in privacy information, contracts, the DPIA, and other documents.

- Check you can fully control the means and purpose of processing as the controller and tailor processing to your requirements. If not, the AI provider may be the controller or a joint controller.
- Before you start processing, identify an appropriate lawful basis for each instance where you are processing personal information. Also identify an additional condition where you are processing special category data. Record this clearly in privacy information, contracts, DPIAs, and the RoPA.
- Do not process personal information if you cannot identify an appropriate lawful basis. Do not process special category data if you cannot identify both an appropriate lawful basis and additional condition.
- Produce a RoPA based on regular data flow mapping that records every processing activity in detail. This should include the purpose, the lawful basis and additional condition, and who you are sharing information with.
- Seek assurance that the AI provider is complying with their privacy obligations, by requesting evidence of periodic internal privacy compliance checks and KPIs or compliance metrics.
- Consider and document how individual rights requests will be handled within the AI tool, and how you will communicate requests to AI providers or other relevant third parties.

☑ **Good practice:** Some AI providers gave recruiters an indicative grade for each element of a candidate's written application through their screening tool. They initially developed their AI tool without using personal information and provided each recruiter with a separate unique version of the AI tool. They trained and tested it using only that recruiter's candidate personal information following their instructions. By taking this approach, they did not process personal information from all recruiters to train or test a single central AI tool or develop further products, for which they would likely be the controller.

Case study: Organisation A also provided recruiters with an indicative grade for each element of a candidate's written application. They were the controller when processing personal information to develop their AI tool, and had identified legitimate interests as the lawful basis but hadn't completed a legitimate interests assessment. They also were a processor when processing candidate personal information through the AI tool on recruiters' instructions. When using the AI tool for their own recruitment, Organisation A first relied on consent as a lawful basis for processing.

However, they reverted to relying on legitimate interests if candidates did not consent, which does not comply with UK GDPR. They had not identified an additional condition for processing special category data.

We recommended that the organisation identify an appropriate lawful basis for each processing activity, and an additional condition for processing special category data. Where relying on consent, we recommended that the organisation ensure consent mechanisms comply with UK GDPR article 7 requirements, and not switch to an alternative lawful basis if consent is not given freely. We also recommended that they cease processing and permanently delete the information, if they cannot identify an appropriate lawful basis and additional condition.

Case study: Organisation B sourced candidate profiles from public job networking sites and data vendors to build a large database of potential candidates. They used an AI search tool to identify candidates with relevant skills or experience for a recruiter's job vacancy. They processed personal information to create the database of potential candidates before engaging clients, and they trained and tested the AI search tool independently from clients. They determined they were a processor and therefore hadn't identified a lawful basis for this processing. Both Organisation B and the recruiters had agreed their roles as processor and controller respectively in the contract. However, in practice Organisation B exercised control over the means and purpose of processing and so was the controller in practice, and had not ensured processing complied with UK GDPR.

We recommended that the organisation consider that they are the controller when producing the database, as recruiters cannot feasibly control processing that took place before they were engaged. As a result, we recommended that the organisation assess their compliance with all aspects of UK GDPR, including identifying a lawful basis.

Recruiters should check they can fully control the means and purpose of processing in practice, before agreeing to be the controller in a contract.

Find out more:

Principle (a): Lawfulness, fairness and transparency

Controllers and processors

A guide to lawful basis

Special category data

How do we ensure lawfulness in AI?

What do we need to document under UK GDPR article 30?

Accountability Framework

What are the accountability and governance implications of AI?

How do we ensure individual rights in our AI systems?

Third party relationships

AI systems can involve potentially complex data supply chains. We checked that there were written contracts with clear data protection responsibilities in place and being followed by all parties, to comply with UK GDPR articles 5(1)(e) and (f), and 24 to 29.

AI providers usually had contracts or data processing agreements in place with recruiters. Explicit processing instructions from recruiters were included within contracts where AI providers were processors, but were often worded broadly and covered only principles to follow.

Example: Some contracts we saw were too broad and did not include enough specific detail. For example, they did not include sufficient information about:

- what personal information they would process and how;
- the responsibilities of each party;
- technical and organisational measures for each party to implement; or
- how they would handle information in AI models if the contract ended.

We recommended that these providers revise contracts to include all the required details and data protection clauses above.

Consider: Use plain language data protection clauses in contracts that clearly set out the controller and processor obligations of each party and explain your proposed processing of personal information transparently.

Contracts we saw were often based on a template produced by the AI provider, and were agreed as part of the client onboarding process with recruiters. Contracts were typically in force until actively terminated by either party. Many AI providers and recruiters had built-in scheduled reviews of contracts (usually annually) to check that contractual terms and explicit processing instructions were adequate and fit for purpose.

Consider: Ensure recruiters can add to, change, or remove contractual terms or processing instructions to meet their needs. In particular when providing a standard template contract for recruiters to agree to. Where recruiters cannot meaningfully control the means and purpose of processing, or if the AI provider determines this in practice, the AI provider is the controller rather than a processor.

Example: AI providers engaged their own processors for their AI tool, such as for infrastructure and security, customer support platforms, and messaging services. AI providers usually had written contracts with at least equivalent protections in place. However, they could not always show that they received written authorisation from recruiters before engaging additional sub-processors after agreeing contracts.

We recommended that providers get clear written authorisation before engaging additional sub-processors, and add this process to contracts.

AI providers had generally undertaken at least some due diligence before engaging processors. They continued to receive ongoing evidence of compliance to demonstrate that technical measures remained in place.

Consider: Proactively publish due diligence and ongoing compliance evidence in a portal, or on your website, for recruiters to review.

Consider: Agree contracts with third parties or data vendors where you are obtaining personal information in bulk from public sources (eg job sites, social media, or networking sites). These should clearly set out the legality and transparency of processing and compatibility with the original purpose you collected it for. Consider completing due diligence or ongoing checks to verify that large datasets comply with data protection law.

Recommendations to **AI providers** include:

- Agree a timebound written contract or a data processing agreement that clearly sets out the responsibilities of each party as controller or processor, and granular detail of the proposed processing.
- Ensure explicit processing instructions cover:
    - the specific personal information you are processing;
    - how and why you are processing it;
    - what the output will be;
    - how you will store it;
    - how long you will retain it for;
    - who you will share it with; and
    - what safeguards will be in place.

- Review contracts with recruiters and sub-processors periodically to ensure they are accurate, sufficient, and fit for purpose.
- Seek assurance that sub-processors are complying with contracts. Do this by completing routine compliance checks or requesting evidence that they are following contract terms and processing instructions.
- Document the requirement for a data processing agreement or data protection clauses in contracts, client onboarding processes, project management processes, and system procurement policies.

Recommendations for **recruiters** include:

- Agree a timebound written contract or a data processing agreement that clearly sets out the responsibilities of each party as controller or processor, and granular detail of the proposed processing.
- Ensure explicit processing instructions cover:
    - the specific personal information you are processing;
    - how and why you are processing it;
    - what the output will be;
    - how you will store it;
    - how long you will retain it for;
    - who you will share it with; and
    - what safeguards will be in place.
- Ensure you can fully exercise control over the means and purpose of processing as the controller. Ensure you can meaningfully alter standard contracts or explicit processing instructions to your needs.
- Review contracts with AI providers periodically to ensure they are accurate, sufficient, and fit for purpose.
- Seek assurance that the AI provider is complying with contracts. Do this by completing routine compliance checks or requesting evidence that they are following contract terms and processing instructions.

☑ **Good practice:** Some AI providers used an AI tool that automatically scored candidates' written responses to interview questions. They used a template contract that recruiters could tailor to their particular needs. This included detailed data protection clauses and responsibilities. They retained signed copies of contracts and could check terms quickly.

Case study: Organisation A provided game-based assessments to predict the candidates' skills and behaviours. They had a template contract but this included only very broad information about the AI tool and processing and did not include explicit instructions to follow. Contracts also did not mention additional processing for Organisation A's

own purposes, such as reusing candidate information to develop their AI tool and other products. They had also not periodically reviewed contracts to check that terms were appropriate and being followed.

We recommended that the organisation include detailed information and specific processing instructions in contracts, and review these periodically.

Recruiters should also only agree to detailed and specific contracts, and seek assurance that AI providers are complying with contracts.

Find out more:

Controller and processor relationships in AI

Contracts and liabilities between controllers and processors

What needs to be included in the contract?

What responsibilities and liabilities do processors have in their own right?