



TEMPLATE

Decision Record and Vendor Question Set

Document every AI use in your hiring process, and test each supplier on evidence, not the demo.

WHERE BETTER BEGINS

September 2026 | Version 1.0

Editable template. Not legal advice.

READ THIS FIRST

How to use this template

Two tools in one document.

Part A: Decision Record is a form to complete once for every AI or automated tool used in your hiring process, before it goes live and again at each review. It forces the questions most organisations skip: what problem is this solving, what did we measure first, who can overrule it, and when will we stop using it. Copy Part A for each use case.

Part B: Vendor Question Set is a scored list of questions to put to suppliers. It is built to compare up to three vendors side by side on the same evidence. Score the answers, not the demonstration.

Good practice

- Complete Part A **before** you sign a contract, not after. A decision record written after the event is a justification, not a decision.
- Involve the person responsible for data protection from the start. The ICO expects a data protection impact assessment where automated decision-making is used in recruitment.
- Keep the completed records together. They are your evidence trail if a candidate, regulator or auditor asks how a decision was reached.
- Review at least every six months, and whenever the vendor changes the model or a feature is switched on.

Companion documents

AI in Hiring Assurance Checklist to audit your whole hiring process. The Signal Problem for the evidence and regulatory background behind these questions.

This template is a practical tool and not legal advice. Take advice on your own position.

Part A: Decision Record

One record per tool or feature. Duplicate this section as needed.

1. The basics

Tool or feature name	
Vendor	
Hiring stage it operates in	
Decision owner (name and role)	
Date of this record	
Next review date	

2. The problem

The problem this is meant to solve, in one sentence	
What happens if we do nothing?	
Did we consider fixing or removing the step instead of automating it?	

3. What the tool does

☐ Drafts content ☐ Summarises ☐ Schedules ☐ Verifies identity ☐ Recommends ☐ Scores ☐ Ranks ☐ Filters out ☐ Rejects ☐ Other (describe):

4. Data

Candidate data the tool uses (including anything inferred)	
Does it use or infer special category data, for example from voice, face or behaviour?	☐ Yes ☐ No ☐ Unsure
Does our candidate data train the vendor's models, or anyone else's?	☐ Yes ☐ No ☐ Unsure
Where is data stored and processed? Sub-processors?	

5. Baseline and target

Record the position before the tool goes live. If you cannot, note why, and treat any later claim of improvement with caution.

Measure	Baseline (and date)	Target	How measured
Time to fill			
Cost per hire			
Conversion at this stage			
Offer acceptance rate			
Candidate drop-out at this stage			
Candidate satisfaction			
Hiring manager satisfaction			
Quality of hire measure (state which)			
Outcomes by group, such as adverse impact ratio			

6. Decision effect and human involvement

Does the tool make, or materially shape, decisions with a significant effect on candidates (for example sifting out, ranking or rejecting)?	☐ Yes ☐ No ☐ Unsure
Our reasoning	
Who reviews the tool's output, and at what point?	
What authority, information and time does the reviewer have to change the outcome?	
Is review applied to every candidate at this stage?	☐ Yes ☐ No
How are overrides recorded and reviewed?	

7. Legal and governance checks

Data protection impact assessment: reference and date	
---	--

Lawful basis for processing	
Where candidates are told about the tool, and what they are told	
How a candidate asks for human review (process, owner, timescale)	
How a candidate asks for a reasonable adjustment	
Bias and adverse impact monitoring: method and schedule	
EU AI Act relevance (hiring into the EU, or vendor sells there)	

8. Risks and mitigations

Risk	Likelihood	Impact	Mitigation	Owner
Bias or adverse impact on a protected group				
Candidate fraud, or gaming the tool				
Data protection or transparency failure				
Vendor dependency, or a model change we do not notice				
Candidate distrust and drop-out				
Other:				

9. Stop conditions

Agree these before go-live. We will pause or switch off this tool if:

1.	
2.	
3.	

10. Decision

Decision	☐ Approve ☐ Approve with conditions ☐ Pilot only ☐ Decline
----------	--

Conditions or scope of pilot	
Approver (name and role)	
Date approved	

11. Review log

Date	Reviewer	Results against baseline	Findings	Actions

PART B

Vendor Question Set

Put the same questions to every supplier, in writing, and score the answers on the evidence you have actually seen. Ask for documents, not reassurance.

Score 0	No answer, or an evasive one.
Score 1	An answer, but without evidence, or only a marketing claim.
Score 2	A clear answer, backed by evidence we have seen and kept.

#	Question	What a good answer looks like	Red flags	Vendor A	Vendor B	Vendor C
	Vendor name and product					
A. What it does and what it uses						
1	What exactly does the tool do at each step: draft, score, rank, filter or reject? Which decisions does it influence?	A plain description of inputs, outputs and the decisions it affects. Clear about where a person takes over.	"AI-powered matching" with no explanation of how or what it decides.			
2	Which AI features are switched on in our account by default, and can we turn each one off?	A feature-by-feature list with admin controls we can see in the product.	Cannot say what is on. Features cannot be disabled.			
3	What data does the tool use, including anything inferred from video, voice, text or behaviour?	A data map. States whether special category data is used or inferred, and the safeguards.	Refuses to disclose. Vague about inference.			
4	Does our candidate data train your models, or anyone else's?	A contractual no, or an opt-in only.	A broad right to use our data for product improvement.			
5	Where is data stored and processed? Who are your sub-processors? How are international transfers covered?	A current list. UK or EU hosting, or a valid transfer mechanism.	No list. Cannot say where data goes.			

#	Question	What a good answer looks like	Red flags	Vendor A	Vendor B	Vendor C
B. Fairness and evidence						
6	How have you tested for bias: what data, which measures, when and by whom? Can we see results by protected characteristic?	A documented method, some independent involvement, results shared, tested within the last 12 months.	"Our AI is unbiased." Marketing statement only. No results.			
7	How do you retest after model or feature updates, and how do you tell us?	Defined retest triggers and change notifications.	No retest after updates. Changes announced after the event.			
8	Can the tool explain why a candidate was scored or ranked, in terms a reviewer and a candidate can understand?	Reason codes or factors shown to the reviewer.	An opaque score with no explanation.			
9	What evidence links the tool's output to job performance or quality of hire?	A validation study with method and data, ideally not only the vendor's own.	Speed and volume metrics and testimonials only.			
10	How does it perform for disabled and neurodivergent candidates, older candidates and non-native speakers? What adjustments exist?	Tested. A clear route to an alternative or adjustment.	Never tested. "Talk to us if there is a problem."			
C. Legal and governance						
11	Will you support our data protection impact assessment and automated decision-making assessment with documentation?	Provides a documentation pack and named support.	None. Tells us it is our responsibility alone.			
12	How does the product support candidate notices, requests for human review and challenges?	Built-in features or a documented workflow.	Nothing built in. Left to us.			
13	How do you classify the product under the EU AI Act, and what is your plan for December 2027?	A clear classification and a roadmap.	Blank looks. "It does not apply to us" with no reasoning.			

#	Question	What a good answer looks like	Red flags	Vendor A	Vendor B	Vendor C
14	Who is responsible when the product produces an unlawful or discriminatory outcome? What does the contract say?	Clear allocation, reasonable cooperation on regulatory queries.	All liability disclaimed. No cooperation commitment.			
15	Have you been subject to regulatory enquiry, audit or litigation about AI in hiring? What did you change?	A candid answer and evidence of remediation.	Evasion. "Not that we can discuss."			
D. Human oversight						
16	How does the product support meaningful human review: what does the reviewer see, and can they override?	The reviewer sees the basis for the output, can override, and overrides are logged.	Automatic rejection with no override.			
17	Can we configure thresholds and rules, or are they fixed?	Configurable, with audit history of changes.	Fixed and hidden.			
18	Is there a mode in which a person reviews every case before any rejection?	Yes, and we can enforce it.	No such mode.			
E. Integrity and security						
19	What fraud and identity verification does the product offer? What are the false positive rates and how can a candidate challenge a flag?	Layered checks, published error rates, a route for candidates to appeal.	Claims of near-perfect deepfake detection. No appeal route.			
20	What security certifications do you hold, and how quickly do you tell customers about a breach?	Current ISO 27001 or SOC 2, a stated notification time.	None, or cannot state a timescale.			
21	How do you cope with applicants using AI to generate applications or answers?	Acknowledges the limits and explains its design response.	Says the product cannot be gamed.			

#	Question	What a good answer looks like	Red flags	Vendor A	Vendor B	Vendor C
F. Commercial and exit						
22	What is the total cost, including implementation, per-seat and per-candidate charges, volume overage and premium AI features?	One clear, itemised price for our expected volume.	A headline price with the AI features priced separately.			
23	What happens to our data and configuration when we leave?	Export in a usable format and certified deletion.	Data locked in. No exit terms.			
24	Can we speak to customers of similar size and sector about outcomes? Tell us about an implementation that did not work and why.	References offered. A candid account of a failure.	Only case studies. "We have never had a failed implementation."			
Total (maximum 48)						

Reading the scores

A total is a summary, not a decision. Any score of 0 in sections B, C or D is a reason to stop and ask again, whatever the total. A vendor who cannot explain how their tool has been tested for bias, who is responsible when it goes wrong, or how a person can overrule it, has told you what you need to know.

Record the outcome in Part A. Keep the completed question set with the vendor's written answers and documents.

If you would like an independent view of a supplier's claims, or of whether your hiring passes the Beta Test, speak to MβB through the contact page of the site you downloaded this from.