

PROCUREMENT TEMPLATE

AI Procurement Requirements and Supplier Question Set

Requirements schedule, tender question set and scoring matrix for buying an AI system or software containing AI features

Document reference	PROC-AI-01	Version	1.0 (draft)
Document owner	[Name, job title]	Approved by	[Board / SLT / named director]
Date of issue	[DD Month YYYY]	Next review	[DD Month YYYY — 12 months]
Applies to	Procurement, legal, information security and the business owner of the requirement	Classification	Internal

How to use this template. Text in square brackets must be replaced. Delete any clause that does not apply to your organisation, and add clauses where your circumstances require them. This is a drafting aid, not legal advice, and it is not tailored to your organisation. Legal and regulatory positions change; confirm the current position before you rely on any statement here.

1. How to use this document

This document is issued to bidders as part of an invitation to tender, a request for proposal, or a direct negotiation. Section 3 sets out mandatory requirements. Section 4 is the question set bidders must answer. Section 5 is the scoring matrix, which the buyer completes and does not normally issue.

Proportionality matters. For a low-value tool that does not process personal data and does not influence decisions about people, issue sections 4A, 4B and 4G only. Issue the full document where the system will inform decisions about individuals, process personal or confidential data, operate at scale, or be difficult to replace.

Drafting note. The single most common procurement failure is buying an AI capability against a specification written for conventional software. Conventional specifications assume deterministic behaviour and stable performance. Neither assumption holds, so the specification must instead capture accuracy under stated conditions, behaviour on failure, and notice of change.

2. Statement of requirement

Buying organisation	
Reference number	
Business outcome sought	[Expressed as an outcome, not a technology]
Current process to be replaced or supported	
Population affected by outputs	
Estimated volume of use	
Contract term and extension options	
Indicative budget	
Deadline for clarification questions	
Deadline for responses	
Format of response	

3. Mandatory requirements

A bidder that cannot meet a mandatory requirement will be excluded. Mark each requirement as mandatory or desirable before issuing; do not issue this list unamended.

Ref	Requirement	M / D
M1	The supplier states, in writing, its role in respect of the system under applicable AI legislation, and identifies the party that bears provider obligations.	M
M2	The supplier names the underlying model or models and their versions, including third-party models.	M
M3	The supplier gives not less than thirty days' notice before replacing or materially changing an underlying model.	M
M4	The buyer's data is not used to train, fine-tune or evaluate any model, whether the supplier's or a third party's, without the buyer's specific written consent.	M
M5	The supplier provides written instructions for use covering capabilities, limitations, accuracy metrics, and the conditions under which the system should not be relied upon.	M
M6	The supplier provides logs sufficient to reconstruct any individual output, including the model version and inputs used, and retains them for not less than twelve months.	M

Ref	Requirement	M / D
M7	The system permits a human operator to override, disregard or reverse any output, and the supplier does not restrict that ability.	M
M8	The supplier discloses all sub-processors and downstream services to which buyer data is transmitted, including any invoked through retrieval or tool use.	M
M9	The supplier states the countries in which processing occurs and notifies the buyer before adding a country.	M
M10	On termination, the supplier exports buyer data, configuration and output history in a documented machine-readable format and certifies deletion.	M
M11	The supplier notifies the buyer of any incident affecting the accuracy, availability or security of the system within the period stated in the contract.	M
M12	The supplier holds, and maintains for the term, professional indemnity and cyber insurance at not less than the levels stated.	M

4. Question set for bidders

Answer every question. Where a question does not apply, state why rather than leaving it blank. Where an answer is commercially sensitive, say so and provide it under separate cover; do not omit it.

4A. Supplier and product

#	Question	Word limit
A1	Legal entity, registered number, country of establishment and ultimate parent.	100
A2	Product name and version offered, and the date it became generally available.	100
A3	Which parts of the product use AI, and can those parts be disabled without losing the remaining functionality?	300
A4	How many customers use the AI functionality in production, and in which sectors?	200
A5	Provide two references from customers using the AI functionality for a comparable purpose.	200

4B. Fit to requirement

#	Question	Word limit
B1	Describe how the product meets the business outcome in section 2, and identify any part of the requirement it does not meet.	600
B2	What does the product produce — a recommendation, a score, a classification, a draft, or a decision?	200

#	Question	Word limit
B3	Describe the intended role of the human operator, and what the product expects them to check.	400
B4	Name at least three uses for which you consider the product unsuitable.	300
B5	What configuration or tuning is required before the product performs as described, and who performs it?	400

Assessment note. Question B4 is the most informative in the set. A bidder unable to name any unsuitable use has either not examined the product or is unwilling to be candid, and both are material.

4C. Models, data and supply chain

#	Question	Word limit
C1	Name each model used, its version, its developer, and where it is hosted.	300
C2	What is your dependency on any single upstream model provider, and what happens if that provider withdraws or restricts access?	400
C3	How and when will you notify us of a model change, and what testing will you have performed before it?	300
C4	State your default position on using customer inputs and outputs for training, and whether it can be contractually excluded.	300
C5	List every sub-processor and third-party service that receives customer data, including through retrieval or tool use.	400
C6	What retention applies to prompts, outputs and logs, and can the buyer configure it?	200
C7	Where personal data is processed, do you act as processor or controller, and for which activities?	300

4D. Performance, testing and fairness

#	Question	Word limit
D1	State your accuracy or quality metrics, the test set used, and the date of measurement.	400
D2	How does performance vary across languages, accents, dialects, age groups and other relevant populations?	400
D3	Have you tested for disparate outcomes across groups? Describe the methodology and provide the results.	500
D4	Describe the documented failure modes and give three realistic examples of incorrect output.	500

#	Question	Word limit
D5	How do you detect performance degradation over time, and what do you do when it is detected?	300
D6	Will you support a proof of concept on the buyer's own data before contract award, and on what terms?	200

Assessment note. Treat a headline accuracy figure quoted without test conditions as no answer. Ask instead for the metric, the population, the date, and whether the figure was produced by the supplier or independently.

4E. Regulatory and assurance position

#	Question	Word limit
E1	How have you classified this system under applicable AI legislation, and on what reasoning?	400
E2	What documentation will you supply to help the buyer meet its own obligations as a user of the system?	400
E3	Does the product generate content requiring disclosure or machine-readable marking, and do you apply it?	300
E4	Which management system or security certifications do you hold, and what is their scope?	300
E5	Have you been subject to a regulatory enquiry, investigation or enforcement action concerning AI or data protection?	300

4F. Security and resilience

#	Question	Word limit
F1	How do you defend against prompt injection, and against data exfiltration through model outputs?	400
F2	Have the AI components been subject to adversarial testing? By whom, and when?	300
F3	State your availability commitment, how it is measured, and the remedies for failure.	300
F4	State your incident notification period and the information the buyer will receive.	300
F5	What continuity arrangements apply if a model provider suspends service?	300

4G. Commercial, intellectual property and exit

#	Question	Word limit
G1	Set out the pricing basis and how cost changes with volume, including any consumption-based element.	400
G2	Who owns the outputs generated from the buyer's inputs?	200
G3	Do you indemnify the buyer against third-party intellectual property claims arising from outputs? State the limits, exclusions and conditions.	400
G4	What is your position on the provenance of the data used to train the models the product relies on?	400
G5	Describe the realistic effort required to migrate to an alternative supplier, and what you will provide to assist.	400

5. Scoring matrix

Complete internally. Weightings must be set before responses are opened and must not be adjusted afterwards.

Criterion	Weighting	Scored from	Score	Weighted
Fit to requirement	[25]%	4B		
Demonstrated performance and fairness testing	[20]%	4D		
Transparency of models and supply chain	[10]%	4C		
Regulatory and assurance position	[10]%	4E		
Security and resilience	[10]%	4F		
Exit and portability	[5]%	4G		
Price and total cost of ownership	[20]%	4G		
Total	100%			

Score	Meaning
0	No response, or a response that fails to address the question.
1	Addresses the question but with material gaps or unverified assertions.
2	Adequate. Meets the requirement with evidence, but without distinction.
3	Good. Meets the requirement with specific, verifiable evidence.

Score	Meaning
4	Excellent. Exceeds the requirement and demonstrates candour about limitations.

Drafting note on scoring. Award credit for candour. A bidder that discloses failure modes and unsuitable uses is a lower risk than one that presents the product as universally reliable, and a scoring scheme that rewards only positive claims will systematically select the latter.

6. Evaluation record

Bidders invited	
Responses received	
Bidders excluded, and on which mandatory requirement	
Highest scoring bidder	
Matters to be resolved in contract negotiation	
Residual risks accepted, and by whom	
Evaluation panel	
Date of decision	