

# Vendor Questions for AI in Contractor & Permit Software

Two pages. Put these to any vendor, get the answers in writing and file them with the contract. Use with the guide How AI Should Be Used in Contractor & Permit Software.

Vendor and product

Feature (for example, document review or method statement review)

Date asked (DD/MM/YYYY)

## 1. Scope: what the AI does, and who answers for it

- ☐ The AI reads, cross-checks and flags against criteria we write. It does not approve, waive or close anything on its own.
- ☐ Anything it fills in, such as an expiry date read off a document, is shown to a person to confirm or correct.
- ☐ The contractor chooses whether to run it, and can submit without it or despite a low score. Where one ran, the approver sees it beside the document; where not, the approver works without it.
- ☐ A named person, by role, is accountable for each permit type and document type: \_\_\_\_\_
- ☐ The tool that judges a document is not the one that prepared it.

## 2. The eight data and control questions

| # | Ask the vendor                                                                      | A good answer sounds like                                                        | Their answer, in writing | OK? |
|---|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|--------------------------|-----|
| 1 | Which AI model and provider reads our documents, and in which country?              | Names the provider and the location. Matches their published sub-processor list. |                          |     |
| 2 | Is our data used to train or improve any model?                                     | A plain no, for inputs and outputs, in the contract and not only on a web page.  |                          |     |
| 3 | How long are documents, inputs and outputs kept, and can we delete them?            | A stated period, a deletion route and what happens at the end of the contract.   |                          |     |
| 4 | Who can see an AI review, and can it show anyone more than they could already open? | Reviews follow the user's own permissions and sites.                             |                          |     |
| 5 | Is the AI optional for the contractor, and does it block submission?                | Optional, and it does not block. Any required step is our choice, not theirs.    |                          |     |
| 6 | Is every review logged with the criteria used and the reviewer's decision?          | Assessment, criteria, comment and decision sit together and can be exported.     |                          |     |
| 7 | What testing can you show us?                                                       | Evidence on documents like ours, not a general claim of accuracy.                |                          |     |
| 8 | How will you tell us when the model or the criteria library changes?                | Advance notice, a change record and a way to re-test before it reaches us.       |                          |     |

A vague answer to 2 or 3 is a reason to stop. A document with a licence number and a date of birth on it is personal information.

### 3. Write criteria two people would mark the same way

- ☐ One observable test per criterion. 'Adequate insurance' fails; 'public liability of at least [your minimum] for the whole period of the work' passes.
- ☐ Pass, Partial and Fail defined in words.
- ☐ The criteria that must pass are marked. A Fail on any of them stops the review, whatever the total score.
- ☐ Local rules are in: state, site and client standards.
- ☐ Contractors can see the criteria before they submit.

### 4. Test before you rely on it

| Step                                                                                                                                                                | Pass mark, set before you run it                                       | Result |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------|--------|
| Run 20 to 30 past submissions you have already decided, including a few broken on purpose (expired yesterday, wrong company, blank template, poor photo)            | For example, no misses on any must-pass criterion                      |        |
| Tally: AI passed and you passed; AI flagged and you rejected; AI flagged what you approved (noise); AI passed what you rejected (a miss)                            | Noise low enough that reviewers still act on flags                     |        |
| Run live in shadow for 2 to 4 weeks. On submissions where the contractor ran the review, reviewers still read everything and record each disagreement with a reason | A clean record on a criterion before anyone reads the assessment first |        |
| Run the test set again after any change to a criterion or to the model                                                                                              | Same pass mark                                                         |        |

The numbers 20 to 30 and 2 to 4 weeks are a rule of thumb, not a published figure.

### 5. Measure before you switch it on, and again after 30 days

| Number (take the baseline from 10 recent reviews)                                  | Before | Day 30 | Notes |
|------------------------------------------------------------------------------------|--------|--------|-------|
| Minutes a reviewer spends on a full read                                           |        |        |       |
| Rounds between first submission and approval                                       |        |        |       |
| Days from submission to approval                                                   |        |        |       |
| Share sent back for something a reviewer could have caught earlier                 |        |        |       |
| Share of submissions where the contractor ran the AI review                        | n/a    |        |       |
| How often a reviewer disagreed with the AI                                         | n/a    |        |       |
| What the AI missed that surfaced later (audit finding, lapsed certificate on site) | n/a    |        |       |

Time saved alongside more misses is a loss moved to the audit.

### 6. Keep with a person, not the AI

- ☐ Approving a contractor or issuing a permit.
- ☐ Waiving or overriding a failed criterion, with a reason on the record.
- ☐ How an incident is classified, who is told and what the investigation concludes.
- ☐ Who may change the criteria.