AI for non-profits.

helping non-profits use AI well

Before You Give AI Permission to Act.

A simple guide to AI agents for non-profits.

An AI assistant can help with a task when you ask it. An AI agent can take a series of steps towards a goal. It may search files, use connected tools, update records or prepare and send something.

That can save time. It can also create a different kind of risk.

The question is not just, “Can this tool do the task?” It is, “What are we giving it permission to read, decide, change or send?”

Start with the work.

Be clear about the job you want the agent to do. Write it in one sentence.

Start with work around the work. For example, sorting non-sensitive enquiries into themes or preparing a draft meeting summary. Do not begin with decisions about people, money, contracts, safeguarding or external communications.

Name a human owner.

One named person should be responsible for the agent. They need to understand what it is for, what it can access, when it needs checking and what happens if something goes wrong.

An agent does not take responsibility away from the charity.

Keep its access narrow.

Give the agent the smallest amount of access it needs to do the agreed task.

Do not give it access to whole drives, inboxes or systems simply because that is easier. Consider what it can read, what it can change and which connected tools it can use.

Treat personal, confidential and safeguarding information with particular care. Check your organisation's policy, lawful basis and any data protection requirements before connecting a system to this kind of information.

Set a clear approval point.

Decide which actions always need a person to approve before they happen.

For most charities, that should include sending external messages, publishing content, changing records, making commitments, handling money, accepting or declining people, or making decisions that could affect someone's rights, safety or access to support.

Test it before you trust it.

Try the agent in a small, controlled setting first. Use test material or permitted non-sensitive information where possible.

Check not only whether it completes the task, but how it behaves when the information is unclear, incomplete or unusual. Look for confident mistakes, missed context, actions outside the brief or information it should not have used.

Know when it must stop.

Write down the situations in which the agent must pause and ask for help.

This could include uncertainty, missing information, a request involving personal or sensitive data, a complaint, a safeguarding concern, an unexpected cost, a conflict with your policy or a task outside its agreed role.

Make sure someone can stop or disconnect it quickly if needed.

Keep a record.

For each agent, keep a short record of:

- What it is for

- Who owns it
- What systems and information it can access
- What it can do without approval
- What always needs human approval
- What it must not do
- When it should pause or escalate
- How you will check its actions
- When you will review whether it is still useful and safe

Review it regularly.

AI tools, connected systems and the work itself can change. Check whether the agent still has the right permissions, whether it is making useful mistakes or harmful ones, and whether the time it saves is worth the risk and oversight it requires.

A sensible place to start.

Start small. Keep the task low-risk. Keep access limited. Keep people involved.

This guide is a practical starting point, not legal, data protection or cyber security advice. Seek specialist advice where your organisation's circumstances require it.