MarketImpact AidGPT

AI Workflow Card: Safe and Effective Use Guide

A Practical Framework for Applying AI Responsibly in Humanitarian and Development Work

Version 1.0 – November 2025

AidGPT | MarketImpact Digital Solutions Ltd
Responsible AI in Practice – Training & Advisory

Before You Start: Purpose, Structure and Limitations

Purpose of This Tool

This guide provides a humanitarian staff member with a structured method to design, test, and document safe AI-assisted workflows. It is designed to help you move from "experimenting" with AI to using it responsibly for **routine text-based tasks**, consistent with your organisation's data protection standards.

What Is In This Toolkit?

This document is divided into three practical sections:

1. The Approved Data & System Checklist (Decision Aid) A reflective set of questions to help you determine *before you start*:

- Is the AI tool I want to use approved?
- Is the data I want to process safe and appropriate?

2. How to Use the AI Workflow Card (The Method) A step-by-step guide on how to:

- Select the right task for automation.
- Build a "3-Step Prompt" (Role, Rules, Details) for consistent results.
- Verify the output using manual and automated checks.

3. The AI Workflow Card (The Template) A printable or digital form to document your prompt, your verification method, and your safety checks. This creates a repeatable "recipe" that you or your team can use safely in the future.

Scope and Limitations

This Tool Is Not Intended For High-Risk Workflows This framework is strictly for routine, low-risk work (e.g., drafting summaries, formatting reports, editing non-sensitive text).

Do NOT use this workflow for content classified as Restricted or High-Risk. Common examples of content that must never be entered into AI tools without explicit authorization include:

- Protection or survivor information (including case management details).
- Security or access incidents (locations, movements, or specific threats).
- Sensitive HR matters (performance reviews, complaints, or salaries).
- Politically sensitive communications (donor negotiations or advocacy strategy).
- Identifiable Vulnerability Data (details about specific beneficiaries or households).

Policy Compliance & "Default Stance"

Your organisation's Standard Operating Procedures (SOPs) and AI usage policies are the final authority on what content is appropriate.

⚠ Important: The "Default Stance" If your organisation does not yet have a specific AI policy or guidance, you must treat all non-public information as Restricted. Do not input internal data into any AI system without clear, written approval from leadership.

The ultimate responsibility for accuracy, appropriateness, and safe use of AI-generated content remains with the human user.

1. Approved Data and System Checklist

A reflective guide for humanitarian and development organisations

This checklist supports informed, responsible decision-making when using AI systems for work-related tasks. It helps staff assess two essential questions:

1. **Is the AI system I am using approved by my organisation?**
2. **Is the content I want to process appropriate to enter into that system?**

Both conditions must be met before proceeding. This is a reflective tool, not a rulebook; organisational Standard Operating Procedures (SOPs) remain the final authority.

Understand the Nature of the Information

Before engaging an AI tool, consider both the *category* and the *sensitivity* of the information.

A. Content Category

Ask yourself:

- What is the purpose of this information?
- Is it descriptive, operational, analytical, or administrative?
- Is this routine work content, or does it fall under a category requiring enhanced caution according to your organisation's SOPs?

Routine work content typically includes:

- Draft reports or programme updates
- Meeting notes and internal action points
- Templates, outlines, and presentation drafts
- General correspondence and email text
- Summaries and bullet points
- Non-sensitive contextual descriptions

Categories that often require additional consideration (refer to your organisation's specific guidance) include:

- Protection or safeguarding-related information
- Security, access, or operational risk content
- HR-sensitive information
- Strategic donor or partnership communications
- Case management information
- Content involving identifiable individuals or vulnerable groups

B. Sensitivity Level

Reflect on:

- Could any part of this information identify individuals or communities?
- Does this information involve confidential, restricted, or politically sensitive issues?
- How does your organisation classify this category of content?

Content classification varies across organisations. Always refer to your organisation's SOPs and information management policies.

Consider Who Could Be Affected

Assess who might be impacted if the AI misunderstands, simplifies, or reframes the information.

Consider whether:

- Misinterpretation could create risk or harm for beneficiaries, partners, or staff
- Omission of nuance might distort sensitive issues
- AI-generated phrasing could be misread by internal or external stakeholders
- The content involves trauma, power dynamics, cultural nuance, or contested narratives

This encourages ethical reflection. Many tasks involving safety, dignity, or sensitive decision-making require direct human judgement and should not be automated.

Assess Organisational Policy Requirements

Most organisations classify information into categories such as *internal*, *confidential*, *restricted*, or *sensitive*. Review your organisation's policies and ask:

- How is this type of content classified?
- What does our AI usage policy or SOP state about using AI systems with this information?
- Do data protection, safeguarding, information security, or donor requirements apply?
- Are additional approvals required for this type of content?

- Have there been recent updates to AI or data policies?

If uncertain, consult your Data Protection Officer, Digital/IT team, or programme leadership before proceeding.

Confirm That the AI System Is Approved

System approval and content appropriateness are separate considerations.

Approved enterprise systems (e.g., Copilot 365 Enterprise, ChatGPT Enterprise, Gemini for Workspace Enterprise, Claude for Teams) operate under contractual data protection agreements. This generally means:

- Data is processed within secure organisational environments
- Data is not used to train public models
- Routine, non-sensitive work content can typically be processed without anonymisation

Examples of routine content in approved systems (confirm with your organisation):

- Draft reports and programme updates
- Non-sensitive meeting notes
- Templates and outlines
- Email content and internal correspondence
- Summaries and bullet points
- Presentation drafts

Categories requiring confirmation with your organisation before use

These categories often carry specific restrictions, but policies vary. Confirm before proceeding:

- Protection, safeguarding, and case information
- Security or access-related incidents
- HR-sensitive matters
- Strategic communications, donor negotiations, or political analysis
- Content involving identifiable individuals or vulnerable communities

If you cannot confirm the policy for a given content type, pause and seek guidance.

Before using any AI system, confirm:

- The tool appears on your organisation’s approved list
- You are authenticated with your official account (not a personal one)
- You have checked your organisation’s SOPs or AI usage guidance for this content type
- You have consulted your supervisor or Data Protection/IT team if uncertain

If using a non-approved or personal system, organisational policies typically prohibit entering any work-related content.

Evaluate Whether the Task Can Be Completed with Less Information

Data minimisation is good practice. Ask:

- Is the AI able to help if I generalise the content?
- Can I remove specific details and still accomplish the task?
- Would a template, outline, or description be sufficient?

For approved systems handling routine content, full text may be appropriate. For sensitive content, minimisation or abstraction may be required.

Comparison of Approved and Non-Approved Systems

Category	Approved Enterprise Systems	Non-Approved or Personal Systems
Data Protection	Covered by organisational agreements	Not covered; data may be used for training
Suitable Content	General internal, non-sensitive text	Only anonymised or public text
Sensitive Data	Prohibited	Prohibited
Recommended Use	Routine workflows	Learning, experimentation only
Risk Level	Low to Moderate	High

Compliance Note

Work-related content must only be processed in systems that are approved under your organisation’s data protection and responsible AI frameworks.
This is a core requirement for safeguarding individuals, programmes, and institutional integrity.

2. How to Use the AI Workflow Card

This section explains **how to fill in the AI Workflow Card** and how to **safely use** the workflow in your day-to-day work.

Choose the Right Workflow

Pick **one** workflow where AI can realistically help and that is:

- Text-based (writing, summarising, structuring, translating, etc.)

- Repetitive (you do it regularly)
- Appropriate per your organisation's policies (check your SOPs)
- Easy for you to check and correct

Good examples:

- Summarising weekly or monthly updates
- Drafting or rewriting parts of donor reports
- Turning rough notes into bullet points
- Drafting simple email replies
- Drafting short programme or MEAL briefs
- Preparing a presentation outline

Check your organisation's SOPs before using AI for workflows involving:

- Protection or safeguarding-related information
- Security or access incidents
- Sensitive HR matters
- Strategic donor or political communications
- Case management information
- Any content your organisation classifies as restricted

Common examples that often require additional caution or are not appropriate for AI use (depending on your organisation's policy) include:

- Protection cases or survivor information
- Security incidents
- Performance reviews or HR complaints
- Political analysis or donor negotiations

When you cannot confirm whether content is appropriate, be safe and consult your supervisor or IT/Digital team before proceeding.

You'll write the name of this task and why it matters in **Section 1 of the card**.

Build Your Prompt Using the 3-Step Method

On the card, you'll design **one clear prompt** you can reuse for this task.

STEP 1 – ROLE
Describe who the AI should act as, so it behaves consistently. For example:

- "You are a clear-writing reporting assistant."
- "You are a MEAL analyst who summarises monitoring findings."
- "You are a communications assistant who rewrites text in simple English."

You'll write this in the **ROLE** field.

STEP 2 – RULES

Add a small number of copy-paste rules that tell the AI *how* to work.
You'll write these in the **RULES** field and later copy them into your prompt.

Examples you can use:

Format rules (choose 1–3):

- "Write in bullet points."
- "Keep the response under 150 words."
- "Use clear headings."
- "Write short, simple sentences."
- "Use a table format."

Tone rules (choose 1–2):

- "Use a neutral, professional tone."
- "Write in clear, simple English."
- "Remove jargon."

Safety & behaviour rules (choose 2–4):

- "Do not make assumptions or add information."
- "Highlight missing or unclear information."
- "Keep all content anonymised."
- "Ask clarifying questions if the input is incomplete."
- "List any uncertainties clearly."

STEP 3 – DETAILS

Prepare the input you will provide to the AI.

If using an approved system with routine content:

- You can use full content as needed for the task
- Data protection is covered by enterprise agreements

If your organisation requires extra caution for certain content:

- Use anonymised versions (placeholders like "Staff A", "Location X")
- Remove or redact specific identifiers

Partical Tip: Instead of just removing names, try to aggregate data (e.g., 'reporting on trends across three districts' rather than 'a specific incident in one village').

Verification (Your Quality Control)

Before using AI output, you **must verify** it.
This protects against assumptions, invented information, or tone issues.

You can verify in **three simple ways** — use one or combine all three.

A. Manual Review (Always Required)

Check the AI's output yourself, asking:

Accuracy Check

- Did the AI add anything that wasn't in your input?
- Did it remove something important?
- Does the text misrepresent any points?

Tone Check

- Is it neutral, professional, and appropriate for your intended audience?

Safety Check

- Is the content handled in line with your organisation's data protection rules (for example, anonymised where required)?
- Is anything too sensitive to include?

If something feels "off," trust your judgement — correct it manually.

B. Second-AI Cross-Check (Optional but Extremely Useful)

Ask a *different AI model* (or a new chat) to check the output.

Use this **copy-paste prompt**:

"Review the text below.
List any assumptions, inaccuracies, missing information, invented content, or places where the model may have guessed.
Do not rewrite — only analyse."

Paste your output beneath it.

This catches errors the first AI missed.

Example cross-check pairs:

- Draft in ChatGPT → Verify in Gemini
- Draft in Copilot → Verify in ChatGPT
- Draft in Gemini → Verify in Claude

C. Structured Verification Prompt (In the Same Chat)

Use this **copy-paste prompt** to verify inside the *same* tool:

"Check your previous output for assumptions, invented details, unclear statements, or anything not supported by the input.
List your findings clearly."

This forces the model to audit itself.

Combine Methods for Best Results

For low-risk tasks (summaries, rewriting, structuring):
→ Manual review + structured verification prompt

For medium-risk tasks (donor reports, MEAL summaries, SMT briefs):
→ Manual review + cross-check in second AI

For learning/testing workflows:
→ Cross-check + structured verification prompt

What You Should Document on the Workflow Card

In the **Verification & Human Oversight** section of the card, they should write:

- WHICH method(s) they will use
- WHY they chose those methods
- HOW they will check for assumptions
- WHO will review if the content is sensitive
- HOW they will finalise the output safely

Example entry on the card:

"I will manually review the output for tone & accuracy, then cross-check using Gemini to identify assumptions or errors. Final text must be validated by me before use."

WARNING: Only use the 'Second-AI' method if the second tool is ALSO on your organisation's approved/enterprise list. Never paste work content into a personal/free AI tool for verification.

Test & Review Your Workflow (Mandatory Step)

Before you use this workflow in your day-to-day work, you should run it once and check the output carefully.

Step 1 — Test Your Prompt Once

Take your prepared input text (in line with your organisation's policies, or a blank template) and run your full 3-Step Prompt.

This is to check if:

- The AI understands your instructions

- The format is correct
- The tone is appropriate
- The output is close to what you expected

This is **not extra work** — it is simply checking that your workflow actually works.

Step 2 — Review the Output Critically

Ask yourself:

Did the AI follow the *Role*?

Does it behave like the assistant you described (reporting assistant, MEAL analyst, etc.)?

Did it respect the Rules?

- Bullet points?
- Word limits?
- Neutral tone?
- No assumptions?
- No invented details?

Is anything unsafe?

- Did it add information you didn't provide?
- Did it infer locations or details?
- Did it include anything inappropriate for your organisation's context or intended audience?

Is the output useful?

- Clear and easy to edit?
- Saves time?
- Meaningful improvement over your usual workflow?

Step 3 — Adjust Your Workflow (If Needed)

If anything is off, update one or more of the following:

- The **Role** → make it clearer
- The **Rules** → add/remove instructions
- The **Details** → improve how you phrase your input
- The **Format** → headings, bullets, table, etc.

These refinements are simple and take 1–2 minutes.

You can bring your improved version to Session 2 — we'll refine them together.

Why This Matters

This quick test helps ensure your workflow is:

- Safe
- Predictable
- Useful
- Easy to repeat
- Ready for real work

A workflow that works once → will work every week.

3. Template AI Workflow Card

Team & Workflow Context

Team / Department: ____

Workflow Name: ____

Purpose of This Workflow:

Frequency:
□ Daily □ Weekly □ Monthly □ Quarterly □ As needed

Current Time Cost: ________ per task
Estimated Time Saved With AI: ________ per week / ________ per month

Target Audience:
□ SMT □ Donor □ Internal team □ Cluster/sector □ External partners

Expected Output Format:
□ Bullets □ Short paragraph □ Structured brief □ Table □ PPT outline □ Email □ Other: __________

Approved AI Tool

□ ChatGPT □ Copilot 365 □ Claude □ Gemini

□ Other (approved): ___________________

Build Your Prompt (3-Step Method)

STEP 1 — ROLE

"____"

STEP 2 — RULES (copy/paste the ones you will use)

Format Rules:
“ ”
“”

Tone Rules:
“____”

Safety & Behaviour Rules:
“ ”
“”

Your Own Rule:
“____”

STEP 3 — DETAILS (Safe, Anonymised Text)

Paste or describe your safe input:

Full Prompt (Copy–Paste Ready)

AI Output (Optional)

Usefulness:
□ Very useful □ Mostly useful □ Needs improvement □ Not useful

Accuracy:
□ Accurate □ Minor issues □ Missing key info □ Invented information

Verification & Human Oversight

Verification Method(s) I Will Use:

▢ Manual review
▢ Second-AI cross-check
▢ Structured verification prompt
▢ Supervisor/colleague review
▢ Other: ________________________________

My Final Human Check Before Use:

Risks & Safeguards

This Workflow Is NOT Suitable For (per my organisation's policies):

▢ Protection or safeguarding-related content

▢ Security or access information

▢ Sensitive HR matters

▢ Content my organisation classifies as restricted

▢ Other: ________________________________

I have confirmed this workflow is appropriate by:

▢ Checking my organisation's AI usage SOPs

▢ Consulting with my supervisor/IT team

avoid▢ Reviewing the Approved Data and System Checklist