Digital Convergence USP 2030 | AI HUB

Tool 4 - Final Go / Pilot / No-Go Checklist

Tool 4 is the mandatory decision gate that determines whether an AI system may:

- Go → proceed to deployment with standard safeguards
- Pilot → proceed under controlled, limited, and closely supervised conditions
- No-Go → pause for redesign, capacity strengthening, or legal clarification

This tool ensures that no AI system affecting rights, eligibility, exclusion, sanctions, or appeals is deployed without documented due diligence. It consolidates all prior assessments—intrinsic risk (Tool 1), contextual risk (Tool 2), and the Risk Register (Tool 3)—into a single, auditable decision point.

While developed for social assistance use cases, this tool is adaptable for broader use across social protection pillars.

Sources and Methodology

This checklist synthesises minimum safeguard requirements from multiple authoritative frameworks:

- **Rights against automated decision-making**: The checklist requirements for human oversight, contestability, and the prohibition on fully automated rights-critical decisions reflect the Handbook on Data Protection in Humanitarian Action (ICRC, 2024, Chapter 17.3.1), which establishes that "Data Subjects have the right to not be subjected to solely automated decision making… when such decisions produce legal effects or similarly significantly affect the individual." The Handbook specifies that human oversight must be "meaningful, rather than just a token gesture" and that decision-makers must have "the ability to refute the machine's decision."
- **Data Protection Impact Assessment (DPIA) and transparency requirements**: The requirements for published DPIAs, plain-language explanations, and notice to affected individuals align with ICRC (2024, Chapter 17.6) guidance that "publicly available DPIAs can also help increase beneficiaries' acceptance" and with the Handbook's emphasis on transparency as "crucial" for AI governance.
- **EU AI Act and GDPR:** The checklist's structure reflects the tiered risk approach in the EU Artificial Intelligence Act (European Union, 2024) and data protection requirements under GDPR (Regulation (EU) 2016/679), particularly regarding high-risk AI systems and automated decision-making.
- **AI ethics frameworks:** Requirements for documentation, monitoring, and accountability reflect the OECD AI Principles (OECD, 2019) and UNESCO Recommendation on the Ethics of AI (UNESCO, 2021).

- **Humanitarian data responsibility:** The OCHA Data Responsibility Guidelines (OCHA, 2025) inform the checklist's emphasis on documented justification, grievance mechanisms, and the principle that deployment decisions must be "proportionate, transparent, and evidence-based."

A note on rating scales: This tool draws on the risk tier produced by combining Tool 1 (intrinsic risk, rated Low / Medium / High) and Tool 2 (institutional readiness, rated Strong / Moderate / Weak). These scales measure different things and are combined using the decision matrix described in the companion report (Byrnes, Dobre and van der Erve, 2026, Section 5.4). The Risk Register (Tool 3) serves as the single source of truth, with each risk tagged as Blocking, Pilot Condition, or Monitor to directly inform the decision below.

Mandatory Use

Use of Tool 4 is **mandatory** before any pilot, deployment, or scale-up, and whenever the model, data, workflow, or legal basis changes. It is the final safeguard to confirm that minimum requirements are met and that governance capacity aligns with the level of risk.

System Information

Field	Details
AI Component / Module	
Brief Description	
Version Number	
Date Assessed	
Date for Next Review	

Step 1 – Preliminary Decision Based on Risk Tier

Use of Tool 4 is **mandatory** before any pilot, deployment, or scale-up, and whenever the model, data, workflow, or legal basis changes. It is the final safeguard to confirm that minimum requirements are met and that governance capacity aligns with the level of risk.

Requirement	Confirmed
Pre-Assessment Screening (Tool 1) completed and no unresolved red flags	☐ Yes ☐ No

If No → Do not proceed. Return to Tool 1 and address fundamental concerns first.

Summary Ratings

Intrinsic Risk Tier (from Tool 1)	☐ Low ☐ Medium ☐ High
CRM Tier (from Tool 2)	☐ Low ☐ Medium ☐ High
Final Overall Risk Tier (from Tool 2)	**☐ Low ☐ Medium ☐ High**

Preliminary Decision Table

Final Overall Risk Tier	Preliminary Guidance	Tick
Low	Go – likely safe to deploy with standard safeguards, monitoring, and annual audit. *Proceed to Step 2 to confirm minimum requirements.*	☐
Medium	Pilot – may proceed under controlled, time-bound conditions with enhanced oversight. *Proceed to Step 2 to confirm minimum requirements.*	☐
High	No-Go / Pause – defer deployment until risks are reduced or governance capacity strengthened. *Step 2 may identify specific gaps to address.*	☐

Note: This preliminary guidance is based on risk tier only.

Step 2 confirms whether minimum safeguards are in place. A "Low" risk tier does not guarantee a "Go" decision—all checklist items must be satisfied.

Step 2 – Confirm Minimum Safeguards

Before any deployment or pilot begins, all essential safeguards must be verified. This step ensures that rights protections, transparency measures, and governance structures are in place. Even if the risk tier suggests "Go" or "Pilot," failure to meet these requirements results in an automatic **No-Go** decision.

All items must be "Yes" to proceed. If any item = No → Automatic No-Go.

#	Requirement	Yes/No
	Pre-assessment & Documentation	
1	Pre-Assessment Screening (Tool 1) completed with no unresolved red flags (proportionality, do-no-harm, accountability confirmed)	
2	Tool 1 & Tool 2 justifications fully documented with evidence cited	
3	Risk Register (Tool 3) complete with all identified risks, owners, mitigations, deadlines, and decision implications assigned. All risks tagged as *Blocking* have been resolved/mitigated to an acceptable level.	
	Legal & Data Protection	
4	Legal basis for data processing verified & documented (not assumed)	
5	Data Protection Impact Assessment (DPIA) or AI Impact Assessment (AIA) completed and public summary available	
	Human Oversight	
6	Eligibility, exclusion, sanction, and appeal decisions require human review and approval before action—not automated	
7	Named human decision owner assigned for each critical decision point (a specific person, not a team or system)	

8	Staff have authority and mechanism to override or reject AI recommendations	
	Transparency & User Rights	
9	Beneficiaries are informed when AI contributes to decisions affecting them	
10	Plain-language explanations provided for adverse or high-impact AI-influenced outcomes	
11	Users can request a fully human-led decision process (opt-out pathway)	
	Grievance & Redress	
12	Appeal and remedy channel available, with tracked response times and clear escalation paths	
	Monitoring & Boundaries	
13	Monitoring plan defined, including thresholds for errors or bias and clear responses when exceeded	
14	Logging and model documentation complete (model card, assumptions, versioning, known limitations)	
15	Pre-deployment validation completed, including bias testing across relevant subgroups and, for high-risk systems, adversarial or red-team testing. Results documented with evidence that identified issues have been addressed.	
16	Data hosting arrangements confirmed (jurisdiction, localisation compliance, encryption key ownership, prohibition on model reuse, and contractual audit rights verified)	

17	Deployment scope explicit—covering geography, population, and timeframe	
18	Stop conditions defined—criteria that would trigger pause or rollback	
	Stakeholder Involvement	
19	Affected communities or their representatives consulted during design, review, or oversight	
	Final Confirmation	
	Are all items marked Yes?	

Step 3 – Final Deployment Decision

This step consolidates all evidence and confirms whether the system can proceed. The decision is based on the **risk tier AND the safeguard checklist**. If any safeguard is missing, the decision defaults to **No-Go**, regardless of risk tier.

Overall Risk Tier & Checklist Status	Decision	Safeguards & Review Schedule
Low **AND** Yes on all items from checklist	Go	Full deployment with standard safeguards as defined in the companion report (Byrnes, Dobre and van der Erve, 2026). Human-on-the-Loop (HOTL) oversight is acceptable for advisory and operational functions, meaning staff monitor system outputs and can intervene when anomalies or errors are detected. All rights critical processes, including eligibility determination, exclusion, sanctions, and appeals, require Human-in-the-Loop (HITL) review, meaning a named human decision maker reviews and approves each decision before it takes effect. Beneficiaries must be informed when AI contributes to decisions affecting them,

		and plain language explanations must be available for adverse or high-impact outcomes. A monitoring plan must be active before deployment, covering review cadence, performance thresholds, and drift detection. An incident response protocol must be documented with clear escalation paths. The Risk Register (Tool 3) must be updated quarterly, and an independent audit conducted annually. Model cards and supporting documentation must be published. Stop conditions and decommissioning criteria must be defined in advance.
Medium **AND** Yes on all items from checklist	Pilot	Controlled, time-bound deployment under enhanced oversight. Human-in-the-Loop (HITL) review is mandatory for all decision points, not only rights-critical ones, meaning every AI output is reviewed and approved by a named individual before action is taken. Pilot boundaries must be defined, including geographic scope, population coverage, and time limit. Beneficiaries must be informed that they are participating in a pilot involving AI, and must have access to a fully human-led decision pathway if they choose. Enhanced monitoring is required, including quarterly reporting on accuracy, error rates, subgroup performance, and grievance volumes. Automatic stop conditions must be defined, specifying thresholds that trigger immediate pause or rollback. The Risk Register (Tool 3) must be reviewed monthly during the pilot period. Tool 4 must be reapplied before any scale-up from pilot to full deployment.
High **OR** No on **any** item from checklist	No-Go	Deployment prohibited until safeguards, legal basis, or governance foundations are strengthened: • Legal, procedural, and independent audit safeguards required before any live use • Redesign may be required • Tool 4 repeated after improvements

Narrative Justification

Summarise: • Key risks and residual risks • CRM considerations and governance capacity • Rights implications • Safeguards required • Rationale for final decision

Decision Sign-Off

Role	Name	Signature	Date
Programme Manager			
Data Protection Officer			
Oversight/Audit Representative			
MIS/IT Lead (optional)			
Legal Advisor (optional)			
National Data Protection Authority (if applicable)			
Vendor Representative (acknowledgement only)			