Digital Convergence USP2030 | AI HUB

A TAXONOMY FOR AI IN SOCIAL PROTECTION:

Defining capabilities and application areas

AI HUB FOR SOCIAL PROTECTION

The AI Hub for Social Protection (AI Hub) is an initiative under the Digital Convergence Initiative (DCI), which helps social protection institutions to harness the benefits of responsible AI for the public good. It supports the development of strategies, frameworks, digital systems and institutional capacities to design, govern and implement AI solutions that are ethical, effective and uphold data sovereignty.

Visit our website at **spdci.org/ai-hub**.

ACKNOWLEDGMENTS

This report was authored by Thomas Byrnes, Avril James (Consultants, MarketImpact Digital Solutions Ltd) and Valentina Barca (independent consultant). We would like to thank the many colleagues who reviewed the various drafts: Juul Pinxten and Phillippe Leite (World Bank), Rob Worthington (Kwantu), Gabriele Erba (UNICEF), Franziska Clausius and Nour Barakat (GIZ). We would also like to thank Susan Sellars (copyeditor) for helping to bring the report into its final form.

REQUIRED CITATION

DCI AI Hub for Social Protection, *A Taxonomy for AI in Social Protection: Defining Capabilities and Application Areas*, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026.

AI HUB KNOWLEDGE ARCHITECTURE

The AI Hub is committed to advancing the global dialogue on the use of AI in social protection by creating and sharing knowledge on the responsible adoption of AI solutions. This Taxonomy is part of a broader portfolio of AI Hub knowledge products that support policymakers, practitioners, and partners in navigating the opportunities and challenges of AI in social protection.

Artificial intelligence is a fast-moving field, with new developments, evidence, and applications emerging continuously. While this report reflects the best available information at the time of publication, it may not include subsequent advances. Readers are encouraged to consult our AI Hub website for the most recent developments, publications, and resources: **spdci.org/ai-hub**.

PREFACE: WHY A TAXONOMY?

DEFINITION OF TAXONOMY:

the naming, describing, and classifying/categorising of concepts

Evidence on the use of AI in social protection has been accumulating rapidly across dozens of pilots and deployments (see *Global Evidence Review*[1] for details).

Yet a central challenge has become increasingly clear: the field lacks a shared, technically precise, and functionally grounded taxonomy of AI in social protection. Without such a taxonomy, tools are inconsistently classified, risks are misunderstood, and learning across country experiences is stifled. This taxonomy seeks to address this gap. It is the result of numerous discussions with DCI partners (including the World Bank, ISSA, ILO and the OECD), aiming to find common ground. It aims to unpack the broad and often ambiguous term 'artificial intelligence' by explaining key AI technologies, capabilities, and concepts in an accessible manner. Building on these foundations, the taxonomy explores how different AI capabilities can be translated into practical use cases within the social protection sector, and how these map to core social protection functions.

Successfully navigating AI innovation and realising the potential of these technologies in social protection requires a clear understanding of what AI is (and is not), how it works, and the value it can create. This taxonomy not only clarifies key AI terminologies and technological approaches, but also provides a structured overview of potential applications, making it a practical resource for practitioners and decision-makers – and a backbone for all the other AI Hub knowledge products.

Building organisational readiness for AI is not solely the responsibility of engineers. Managers, practitioners, and policymakers must also understand the opportunities, risks, limitations, and broader implications of AI systems. This taxonomy should be helpful to all of these categories to navigate a landscape they may not be familiar with.

HOW TO USE THIS TAXONOMY:

This document is not to be read back to back. We invite you to browse it as you would browse a dictionary, digging into key classification areas as relevant and useful.

IMPORTANT:

Given the rapid pace of technological development, the capabilities, applications, and value propositions of AI will continue to evolve. This taxonomy will continue to be updated as this happens.

1 DCI AI Hub for Social Protection, *AI Adoption in Social Protection: A Global Evidence Review (2020–2025)*, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026a

CONTENTS

ACRONYMS

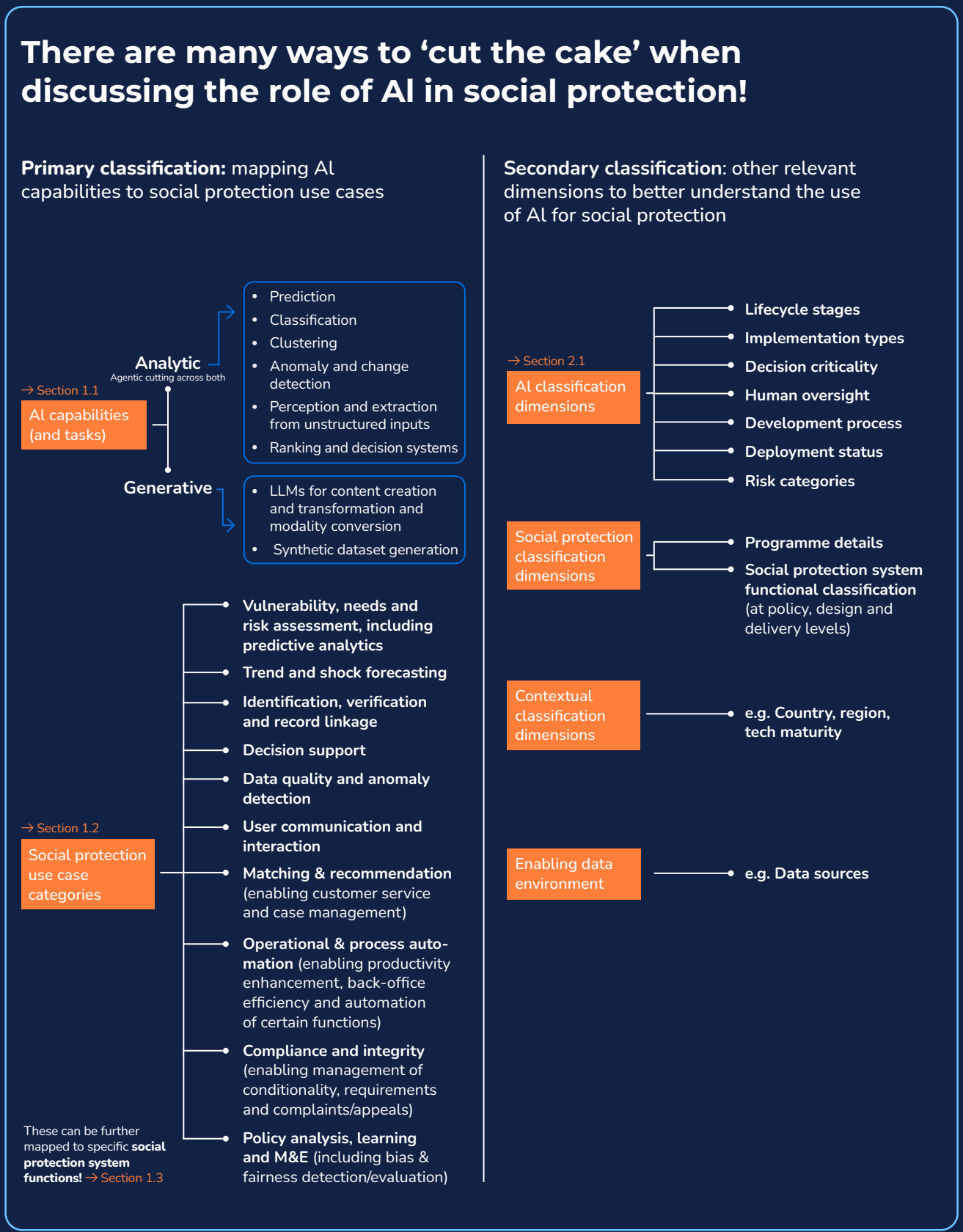
AI	artificial intelligence
CDR	call detail records
CNN	convolutional neural networks
DCI	Digital Convergence Initiative
DHS	demographic health survey
GIS	geographic information system
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
GTMI	GovTech Maturity Index
HITL	human-in-the-loop
HOTL	human-on-the-loop
HOOTL	human-out-of-the-loop
IPC	Integrated Food Security Phase Classification
ISSA	International Social Security Association
KYC	know your client
LLM	large language model
LMIC	low- and or middle-income country
LSMS	living standards measurement surveys
MICS	multiple indicator cluster surveys
MIS	management information systems
ML	machine learning
NLP	natural language processing
OCR	optical character recognition
OECD	Organisation for Economic Co-operation and Development
PMT	proxy means testing
RNN	recurrent neural networks

INTRODUCTION

This document serves as a foundational definitional framework for the use of artificial intelligence (AI) in social protection systems, drawing on insights shared in the **Workshop on Taxonomy** organised by the AI Hub for Social Protection of the Digital Convergence Initiative (DCI) and organising key concepts within all of the DCI foundational knowledge products. This taxonomy provides a complete and structured way of classifying AI systems used in social protection:

- It begins with a **primary classification** that links *what AI does to what AI is used for in the social protection sector* (Section 1).
- It proceeds with **secondary classification** dimensions, further classifying the use of AI, contextual factors and social protection characteristics.

FIGURE 1. CLASSIFYING THE ROLE OF AI IN SOCIAL PROTECTION



1. PRIMARY CLASSIFICATION:

mapping AI capabilities to social protection use cases

Core AI capabilities are the main technical functions and tasks AI models can perform on data, no matter what sector (Section 1.1). **These capabilities we combine with key social protection use cases** (Section 1.2).

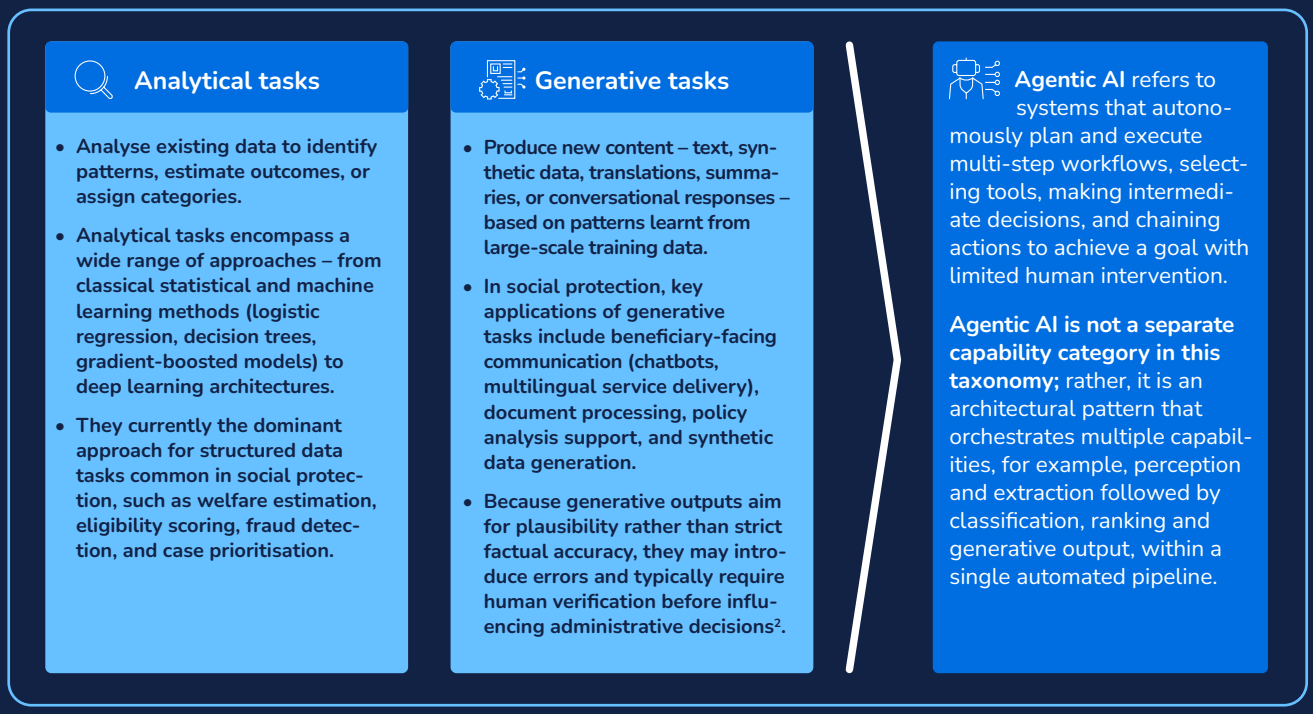
- These use case categories represent a clustering of the key roles AI can play when mapped against social protection functions. Each of the use cases has different practical value if applied to different stages of the social protection delivery chain and broader social protection system (Section 1.4), or to different types of programmes across key social protection pillars (social assistance, social insurance, labour market policies).
- Most of these use cases combine several AI capabilities, as shown in the matrix-view at the end of Section 1.2.
- This is not intended to be a single or exclusive lens, but a structure that can be cross-classified with other frameworks such as the OECD AI taxonomy or the ISSA typology used in the Techbytes. It has been chosen to focus attention on what the model technically does and how this maps to the key AI-supported functions that benefit the social protection sector.

1.1 AI capabilities (and tasks)

The taxonomy distinguishes between analytical AI tasks and generative AI tasks. While analytical AI analyses existing data to produce a score, label, grouping or flag, generative AI can produce new content such as text, synthetic data or translations. The distinction reflects the primary function of the output rather than the underlying model architecture, as modern systems, particularly large language models (LLMs), can perform both types of tasks. In practice, analytical tasks currently dominate social protection deployments for structured data (welfare estimation, eligibility scoring, fraud detection), while generative tasks are growing rapidly for text-based and communication applications.

Agentic AI systems are not a separate capability category in this taxonomy; rather, they represent an architectural pattern that orchestrates multiple capabilities, for example perception and extraction followed by classification, ranking and generative output, within a single automated pipeline. See Figure 2 for a complete picture.

FIGURE 2. THE TWO MAIN CATEGORIES OF AI CAPABILITIES



Analytical AI tasks

Key analytical AI tasks include the following:

PREDICTION (including forecasting)

What it is: A model that estimates outcomes based on patterns learnt from historical data. Prediction (also known as 'estimation') models output a numeric value or probability that quantifies an outcome of interest (e.g. expected cost, predicted demand, risk score). Forecasting is a subtype of estimation in which the value is explicitly tied to a future time horizon (next week/month/quarter). These models support planning, budgeting, and risk quantification by translating historical and contextual data into calibrated quantitative expectations.

Application to social protection: *Prediction outputs typically serve as advisory inputs – informing human judgment, programme design, or downstream analytical processes rather than directly determining entitlements or access to services.*

CLASSIFICATION

What it is: Classification is similarly based on patterns learnt from historical data and, in practice, the same modelling approaches as for prediction are frequently used. The distinction lies mainly in the form of the output and the governance implications. Classification produces a category or label (e.g. eligible/not eligible, priority tier), often with associated probabilities that can be thresholded into decisions. Classification is used when processes require consistent categorisation, routing, or tiering of cases.

Application to social protection: *The act of assigning a categorical label, particularly when it gates access to entitlements or services, typically carries higher decision criticality than producing a numeric estimate, and may require stronger human oversight, auditability of threshold-setting, and access to appeal. This governance distinction is the primary reason for treating classification as a separate capability from prediction in this taxonomy.*

2 International Social Security Association, *AI Applications in Social Security: Building Evidence and Insights from the TechByte*. Geneva: ISSA, 2025.

CLUSTERING (i.e. similarity and grouping)

What it is: Similarity and grouping capabilities organise data based on how alike items are and how they naturally form coherent segments. In practice this includes similarity scoring and nearest-neighbour matching (often via embeddings) as well as unsupervised clustering that produces group assignments and profiles. The primary outputs are similarity scores/nearest neighbours and/or cluster memberships, supporting matching, deduplication, segmentation, and exploratory understanding of populations or documents.

Application to social protection: *Clustering supports population segmentation for programme design, the deduplication of beneficiary records across registries, and the matching of individuals to similar profiles for case management or service referral.*

ANOMALY AND CHANGE DETECTION

What it is: These are models that identify rare, unusual, or suspicious observations that deviate from expected patterns, often when labelled examples of problematic cases are scarce or constantly evolving. Outputs are anomaly scores or alerts that prioritise items for human review, commonly used for fraud and integrity monitoring, data quality checks, and detection of systemic shifts or model drift. This capability supports early warning, integrity monitoring, and quality control by flagging cases that merit attention because they deviate from normal behaviour.

Application to social protection: *Anomaly detection is widely used for fraud and error detection in payment systems, flagging irregular transaction patterns or duplicate claims for human review, as well as monitoring data quality across registries and detecting systemic shifts that may signal emerging needs or programme drift.*

PERCEPTION AND EXTRACTION FROM UNSTRUCTURED INPUTS (i.e. video, audio and text analysis)

What it is: These models transform unstructured data – text, images, audio, or video – into structured observations that can be searched, counted, and acted upon.

Outputs include:

- Extracted fields from documents (optical character recognition, OCR)
- Entities and relations from text (information extraction)
- Objects/regions in images and video (detection/segmentation)
- Transcripts from speech (speech recognition)

This category encompasses optimisation, reinforcement learning and resource allocation tasks, which represent a distinct capability from both pattern recognition and content generation.

Application to social protection: *Extraction capabilities enable the processing of identity documents, application forms and case notes into structured data that can feed eligibility assessments, referrals, complaints and appeals triage or compliance checks, and are often the first step in automating manual document-heavy workflows.*

RANKING AND DECISION SYSTEMS (including optimisation and AI augmented rule-constrained decision support)

What it is: These models use reinforcement and personalised learning, acting as ranking and decision systems, to produce a prioritised list of options or a selected action/policy designed to achieve an objective under constraints. This includes learning-to-rank and recommender approaches (ordering cases, items, or actions) as well as optimisation-oriented decision models that allocate scarce resources or choose interventions. While these systems may use predictions or classifications as inputs, they are distinct because their primary output is what to do (or what to do first), rather than a descriptive measurement. This capability often serves as the bridge that makes raw content usable by downstream estimation, classification, or decision workflows.

Application to social protection: *Ranking and decision systems support case prioritisation across large caseloads, resource allocation across geographic areas or programme components, and the recommendation of tailored service or benefit packages based on individual profiles and constraints.*

For many social protection use cases these would be AI-augmented capabilities, not AI-led. For example, eligibility, compliance, and benefit determination in social protection are often rules-first, model-assisted and legally constrained.

Generative AI tasks

LLMS FOR CONTENT CREATION AND TRANSFORMATION (intra-modal generation) AND MODALITY CONVERSION (cross-modal generation)

What it is: LLMs for content creation/transformation and modality conversion can be classified in two main sub-categories:

- **Intra-modal generation:** These are generative models that produce a new artifact or an edited version of an existing artifact, while keeping the same modality. For text, this includes drafting, rewriting, summarising, translating, and correcting (text→text). For images and audio, it includes editing operations such as enhancing, inpainting, style modification, denoising, or controlled modification of an existing file (image→image; audio→audio). The defining feature is that the output remains in the original modality, enabling creation and refinement of content within a single medium.
- **Cross-modal generation:** These are generative models that express content in a different modality from the input, effectively translating information across media. Typical examples include text-to-speech (text→audio), generating images from text prompts (text→image), describing or captioning visual content (image→text), and related cross-modal transformations. The defining feature is that the model's output is in a different modality than the input, supporting accessibility, multimodal interfaces, and communication across formats.

Application to social protection: *Generative capabilities support beneficiary-facing chatbots and voicebots, multilingual translation of programme information, drafting of decision notices and case summaries for staff, and accessibility tools such as text-to-speech for beneficiaries with low literacy.*

SYNTHETIC DATASET GENERATION (data-level generation)

What it is: Synthetic dataset generation refers to models that produce datasets or collections of records – tabular data, document corpora, image sets, or audio samples – designed to approximate relevant statistical properties of real data. The goal is typically utility for testing, simulation, augmentation, or privacy-aware sharing, rather than generating a single human-facing artifact. Outputs may be used to test pipelines end-to-end, stress-test systems under varied scenarios, augment training data in low-resource settings, or create sandbox environments where direct use of sensitive data is restricted. The generation of synthetic datasets can be achieved via machine learning (with higher performance) and via LLMs with lower performance.

Application to social protection: *Synthetic data generation supports privacy-preserving sharing of beneficiary data for research or cross-agency collaboration, the stress-testing of eligibility models under varied demographic scenarios, and the augmentation of training datasets where real data is scarce or restricted.*

1.2 Social protection use case categories

In what ways can the AI tasks discussed above improve the execution of core functions of social protection systems? To answer this question, we developed ten broad use case categories, grouping key roles which AI can play when supporting social protection functions[3]. For each we provide a brief description, select AI-driven examples and a comment on governance sensitivity (linked to decision criticality, as discussed in Section 2.1).

Vulnerability, needs and risk assessment, including predictive analytics

Description: Identifies for whom and where support is most likely to be needed (and may be needed in the future) by estimating poverty, vulnerability, or risk levels and segmenting populations to guide policy, design and operational decisions (e.g. programmatic adjustments), as well as resource allocation.

AI-driven examples: Tailoring intervention design (e.g. eligibility criteria) to population needs, geospatial poverty mapping, predictive vulnerability and risk scoring to target prevention efforts and to facilitate planning, non-uptake detection, outreach/registration segmentation and optimisation

Governance sensitivity, linked to decision criticality: *Indirectly affects rights and entitlements.* Outputs inform policy, programme design and resource allocation at population level rather than individual-level decisions. Governance sensitivity escalates to high risk where vulnerability scores are used as direct inputs to individual eligibility determination, at which point they effectively become part of the decision support function.

Trend and shock forecasting

Description: Anticipates future trends, shocks or crises to enable early planning and preventive or adaptive measures, including anticipatory action.

Note: *Forecasting is about predicting what will happen (events, trends, shocks), whereas vulnerability assessment (above) is about estimating who is or will be affected.*

AI-driven examples: Climate and hazard forecasting/early warning systems (EWSs), conflict and displacement early warning, demographic, socio-economic and health trend/shock prediction/simulation (potentially at micro, meso or macro-level)

Governance sensitivity, linked to decision criticality: *Primarily informational.* Outputs support planning, preparedness and early warning rather than individual-level decisions on rights and entitlements. Governance sensitivity escalates where forecasts directly trigger anticipatory transfers or automatic benefit adjustments, as these may affect entitlements without individual-level assessment.

Identification, verification and record linkage

Description: Verifies identity, validates claims or credentials, and links records across various systems to create a unified and reliable view of individuals or households.

AI-driven examples: Biometric/ID verification (e.g. face/fingerprint matching), deduplication, automated evidence validation, cross-registry linkage, entity resolution, conditionality monitoring, attendance verification

Governance sensitivity, linked to decision criticality: *May directly affect rights and entitlements.* Errors in identity verification or record linkage can result in wrongful exclusion, denial of benefits or false fraud flags. Governance sensitivity is highest for biometric systems, given documented cases of exclusion among the elderly, disabled and manual-labourers.

3 In other words, this is an organic clustering of the key roles AI can play vis-a-vis social protection functions. This means functions are grouped in a way that will be less familiar for social protection stakeholders, but that ensures mutually exclusive categories that are AI-specific and that can be subsequently mapped against standard social protection functional classifications (Section 1.3). This choice was prioritised because classifying solely against social protection functions would mean each AI use case would appear multiple times against different core functions, with different permutations (as Section 1.3 shows). As an example, decision support is an AI function that can be applied at different levels, e.g. to inform eligibility determination or to inform the right benefit and services package responding to different needs.

Decision support

Description: Assesses whether individuals or households qualify for programmes and/or what level or type of support they should receive, typically through rule-based or model-assisted decision support.

Note: *This category only encompasses cases where AI directly supports the elegibility or benefit level determination (e.g. model-assisted scoring, automated benefit calculation), not cases where AI-generated inputs from other categories (e.g. a vulnerability score) feed into a non-AI eligibility process.*

AI-driven examples: Scoring, benefit package configuration (amount and type), service tailoring, enrolment decision support

Governance sensitivity, linked to decision criticality: *Directly affects rights and entitlements.* This is the highest criticality use case category, as outputs directly gate access to programmes, determine benefit levels or trigger entitlement changes. Requires strongest human oversight (HITL), auditability of threshold-setting, and accessible appeal mechanisms. Read the *AI Hub Risk Assessment Framework for Social Protection: A Practical Guide to Responsible AI Deployment* for guidance.

Data quality and anomaly detection

Description: Prepares, validates and enhances data, while also detecting unusual or suspicious patterns, prioritising cases for data quality checks or investigation to protect programme integrity and manage risk.

Note: *This use case (vis-a-vis the compliance and integrity one below) focuses on data and system integrity, i.e. answering questions such as: is this record accurate? is this transaction pattern unusual? is this dataset complete? The primary purpose here is data or system integrity.*

AI-driven examples: Enhancing the data environment (e.g. structuring unstructured data), transaction anomaly detection, data consistency checks and quality audits, predictive integrity risk scoring, network or collusion analytics, and audit prioritisation – many of these enabling detections of error and fraud (via AI analysing data patterns to flag irregularities)

Governance sensitivity, linked to decision criticality: *May directly affect rights and entitlements.* Outputs are typically flags for human review, but governance sensitivity is high where anomaly scores trigger automatic actions such as payment suspension or fraud investigation. Documented cases of harmful outcomes from automated systems in this category (e.g. the Dutch childcare benefits scandal) underscore the need for human review before adverse action.

User communication and interaction

Description: Facilitates two-way information exchange with beneficiaries/applicants or even staff by providing guidance, explanations, and personalised messages through automated or assisted channels.

Note: *For this use case the primary function is staff or user-facing dialogue/interaction.*

AI driven examples: Chatbots or voicebots, translation, explainable decision notices, personalised messaging (including for behaviour change)

Governance sensitivity, linked to decision criticality: *Low decision criticality but high visibility.* Outputs do not determine entitlements, but errors in chatbot guidance, mistranslation or misleading information can affect beneficiary behaviour and trust in the system. Governance sensitivity increases where automated communication replaces rather than supplements human service channels.

Matching and recommendation (enabling customer service and case management)

Description: Connects individuals to suitable benefits, services, support opportunities, jobs, training (etc.) by analysing profiles, preferences, and information from other relevant data sources (e.g. labour-market).

AI driven examples: Case management, job matching, skills profiling, career pathway guidance, programme or opportunity recommendation

Governance sensitivity, linked to decision criticality: *Moderate criticality.* Recommendations influence access to services, jobs and opportunities, but typically

require human confirmation. Governance sensitivity escalates when matching outputs directly determine programme referrals, job placements or service access without caseworker review.

Operational and process automation (enabling productivity enhancement, back-office efficiency and automation of certain functions)

Description: Streamlines administrative workflows through two distinct sub-categories:

- **Document processing and generative staff assistance:** extracting, digitising and routing information from documents (e.g. OCR on application forms, medical invoices), and generating drafts, summaries or case file syntheses to augment staff productivity – primarily leverages perception/extraction and generative AI capabilities.
- **Workload and resource forecasting:** predicting operational demand, caseload volumes and staffing requirements to support planning and resource allocation – primarily leverages prediction capabilities and is closer in nature to forecasting functions, applied to internal operational contexts rather than external trends or shocks.

AI driven examples: Supported/automated processes (e.g. registration), document processing, generative drafting or summarisation for:

- Supported/automated processes (e.g. registration), document OCR and field extraction, generative drafting or summarisation for staff, workflow routing
- Caseload forecasting, staffing demand prediction, resource allocation planning

Governance sensitivity, linked to decision criticality: *Primarily internal and operational*. The two sub-types carry different governance profiles: document processing and generative assistance raise data quality and accuracy risks (particularly where extracted or generated content feeds downstream decisions), while workload forecasting carries minimal direct rights impact. Governance sensitivity escalates when automated workflow routing affects processing times in ways that materially delay access to benefits for specific groups.

Compliance and integrity (enabling management of conditionality, requirements and complaints/ appeals)

Description: Monitors adherence to programme rules, requirements and conditions, manages complaints and appeals.

Note: *This use case (vis-a-vis the data quality and anomaly detection one above) focuses on beneficiary and programme-rule adherence, i.e. answering questions such as: are conditions being met? Are requirements fulfilled? Here the primary purpose is assessing individual compliance behaviour.*

AI driven examples: Conditionality verification, behavioural compliance analytics (e.g. monitoring conditionality adherence, detecting non-compliance patterns), complaint triage and pattern detection

Governance sensitivity, linked to decision criticality: *May directly affect rights and entitlements.* Outputs monitor beneficiary adherence to programme rules and can trigger consequences including benefit reduction or suspension. Governance sensitivity is high, with documented cases of harmful outcomes from automated compliance systems (e.g. Robodebt in Australia). Requires human review before any adverse action affecting entitlements.

Policy analysis, learning and M&E (including bias and fairness detection/evaluation)

Description: Generates insights to improve policy and programme design and implementation, improving system performance over time.

AI driven examples: Policy microsimulation, counterfactual analysis and synthetic evaluation, model performance and bias/fairness monitoring, trend monitoring for service improvements, improved analytics/dashboards, system stress-test under varied scenarios

Governance sensitivity, linked to decision criticality: *Primarily internal and operational*. Outputs can inform communication to the general public for transparency and accountability purposes, with important implications for trust in public institutions. Outputs inform programme improvement and system performance rather than individual decisions. Governance sensitivity is generally low, but increases when bias and fairness evaluations reveal systemic discrimination, requiring transparent reporting and corrective action.

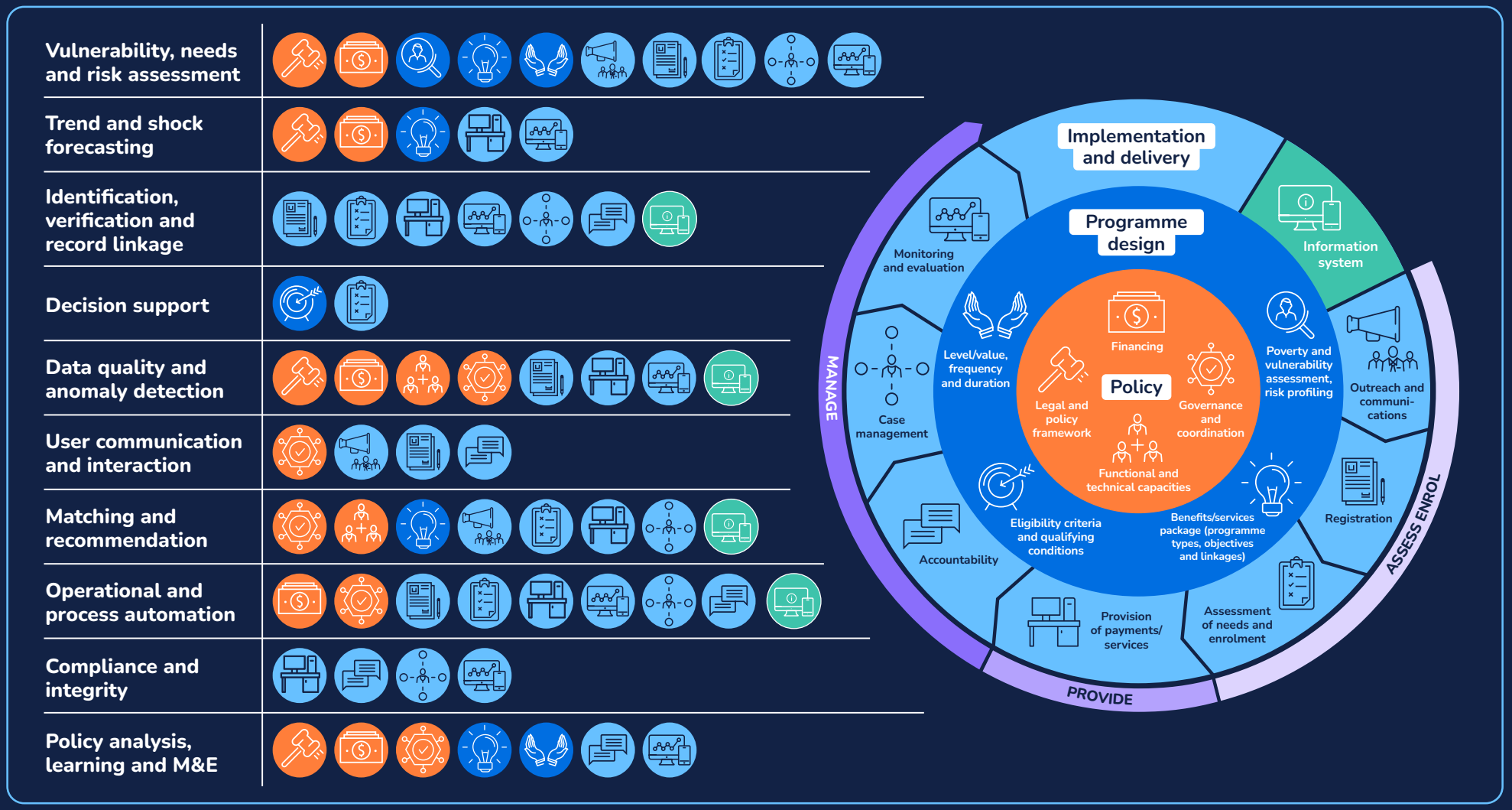
Table 1 shows how AI use cases (Section 1.2) serving the social protection sector leverage AI capabilities and tasks (Section 1.1).

TABLE 1. MATRIX VIEW ACROSS AI CAPABILITIES AND SOCIAL PROTECTION AI USE CASES

USE-CASE / AI CAPABILITY	VULNERABILITY, NEEDS & RISK ASSESSMENT, INCLUDING PREDICTIVE PROFILING	TREND & SHOCK FORECASTING	IDENTIFICATION, VERIFICATION AND RECORD LINKAGE	DECISION SUPPORT (e.g. for determination of eligibility, and benefits and services packages)	DATA QUALITY AND ANOMALY DETECTION	USER COMMUNICATION & INTERACTION	MATCHING & RECOMMENDATION	OPERATIONAL & PROCESS AUTOMATION	COMPLIANCE AND INTEGRITY (conditions, requirements, complaints/ appeals)	POLICY ANALYSIS, LEARNING AND M&E (including bias & fairness evaluation)
	ANALYTICAL AI TASKS									
Prediction (including forecasting)	✓	✓		✓				✓		✓
Classification	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Clustering (i.e. similarity and grouping)	✓	✓			✓		✓		✓	✓
Anomaly and change detection			✓		✓			✓	✓	
Perception and extraction from unstructured inputs (i.e. video, audio and text analysis)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ranking and decision systems	✓	✓		✓	✓		✓	✓	✓	✓
	GENERATIVE AI TASKS									
LLMs for content creation/transformation and modality conversion						✓		✓	✓	✓
Synthetic dataset generation (data-level generation)	✓	✓								✓

1.3 Mapping AI use cases against specific social protection system functions

FIGURE 3. THE TOP 10 AI USE CASE CATEGORIES SERVE DIFFERENT SOCIAL PROTECTION FUNCTIONS AT THE POLICY, DESIGN AND IMPLEMENTATION LEVELS OF THE DELIVERY CHAIN



This mapping enables a further **matrix view of the ten generic AI use cases (Section 1.2) against specific social protection functions at different levels** (as categorised in depth in Section 2.2). **It illustrates how each of the ten AI use cases play slightly different functions at different levels** (policy, design, implementation), across different pillars (e.g. social assistance, social insurance, social services, active labour market policies), and how they can produce different outcomes depending on exactly what they are applied to.

Table 2 should also serve as inspiration to broaden the current scope of AI use cases de-facto prioritised by many countries (see the DCI AI Hub's *Global Evidence Review*[4] for details), towards a broader set of functions that can help achieve the overarching objectives of universal social protection as embedded in ILO Recommendation 202 (providing income security and affordable access to needed services along the life cycle).

Policy and governance level

The following table contains the key dimensions of social protection institutionalisation within government.

TABLE 2. POLICY LEVEL: MATRIX VIEW OF GENERIC AI USE CASES AGAINST SOCIAL PROTECTION FUNCTIONS

KEY FUNCTIONAL AREA (see also the 'Social Protection as a Solar System' publication)	RELEVANT AI USE CASE CATEGORIES (Only the most relevant for each dimension are shared here)	HOW THE AI USE CASE CATEGORY APPLIES TO EACH SOCIAL PROTECTION FUNCTION
 **LEGAL AND POLICY FRAMEWORKS** *Clarity of strategic vision in terms of social protection objectives, functions, internal coherence and linkages with other sectors, ideally embedded in legal and policy frameworks*	**Vulnerability, needs and risk assessment, including predictive analytics**	AI-generated population segmentation and (predictive) poverty/risk scoring to identify unmet needs, highlight coverage gaps (non-take-up) and guide legal framework and policy design, including anticipatory legal provisions for adaptive/risk-informed social protection.
	Trend and shock forecasting	AI-supported demographic, socio-economic and climate/shock forecasting to stress-test the adequacy of legal frameworks and anticipate the need for expanded statutory entitlements or new programme/benefit categories (e.g. ahead of emerging trends and shocks).
	Policy analysis, learning and M&E	AI-assisted microsimulation and counterfactual analysis e.g. to assess the distributional impacts of proposed policy, legal or regulatory changes before adoption.
	Data quality and anomaly detection	AI-driven analysis of registry coverage gaps, data completeness and cross-system inconsistencies to inform policy decisions on data collection requirements, interoperability mandates and data governance frameworks.

4 DCI AI Hub for Social Protection, *AI Adoption in Social Protection: A Global Evidence Review (2020–2025)*, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026a

 **FINANCING** *Sustainable and domestic financing sources, including secure future budget commitments in place*	**Vulnerability, needs and risk assessment, including predictive analytics**	AI-informed needs and risk assessment to estimate and forecast the size, needs and composition of target populations, supporting cost modelling for budget planning and fiscal sustainability analysis.
	Trend and shock forecasting	AI-informed actuarial forecasting based on demographic/economic/hazard data to model contingent liability from anticipated trends and shocks, supporting financial planning and the development of risk-informed financing such as parametric financing instruments, reserve fund sizing and anticipatory financing trigger design.
	Policy analysis, learning and M&E	AI-led policy microsimulation to model the fiscal impact of eligibility and benefit level changes, enabling evidence-based budget planning and expenditure tracking.
	Data quality and anomaly detection	AI anomaly detection applied to financial flows and administrative data to improve budget forecasting accuracy, detect leakage and improve the integrity of expenditure reporting used for financing decisions.
	Operational and process automation	Caseload demand forecasting and workload projection to support budget planning, staffing models and fiscal sustainability analysis for programme operations.
 **COORDINATION AND GOVERNANCE** *Established governance arrangements and coordination mechanisms* +  **TECHNICAL AND FUNCTIONAL CAPACITIES**	**User communication and interaction**	Supporting staff in their workflows – including AI-powered tools for staff-facing communication, including internal knowledge management assistants, chatbots, and automated translation of guidance materials across languages.
	Matching & recommendation	AI tools that match staff profiles to training opportunities, recommend tailored learning journeys, and support career pathway development for the social worker workforce – directly supporting technical and functional capacity building.
	Operational & process automation	Automating processes to enable staff to focus on high-value work, e.g. automating reporting, meeting note-taking and inter-agency information exchange.
	Policy analysis, learning and M&E	AI-generated dashboards, performance analytics and bias/fairness evaluation to support governance oversight bodies and technical capacity building functions.
	Data quality and anomaly detection	Monitoring data quality and consistency across agencies participating in coordination arrangements, flagging mismatches that may indicate gaps in inter-agency data-sharing or process alignment.

Programme design level

Programme design choices not only inform individual programme design, but also how the range of programmes within a system can best work together cohesively to maximise outcome. These are inherently linked to the universal social protection aspiration of enhancing the comprehensiveness, coverage and adequacy of social protection systems, via a system of programmes that address different lifecycle contingencies and risks. These are agreed in international standards (including Convention 102 and Recommendation 202), meaning that many country have obligations under international law to comply with certain benefit levels, durations and eligibility criteria.

TABLE 3. PROGRAMME DESIGN LEVEL: MATRIX VIEW OF GENERIC AI USE CASES AGAINST SOCIAL PROTECTION FUNCTIONS

KEY FUNCTIONAL AREA (see also the '**Social Protection as a Solar System**' publication)	RELEVANT AI USE CASE CATEGORIES (Only the most relevant for each dimension are shared here)	HOW THE AI USE CASE CATEGORY APPLIES TO EACH SOCIAL PROTECTION FUNCTION
Programmatic decisions on the most appropriate  **BENEFITS AND SERVICE PACKAGE:** *Choice of programmes that best address different vulnerabilities and needs of different groups*	**Vulnerability, needs and risk assessment, including predictive analytics**	Tailoring intervention design to population needs, including predictive vulnerability and risk scoring to design prevention efforts.
	Trend and shock forecasting	Informing programme-specific design choices, including anticipatory actions.
	Policy analysis, learning and M&E	AI-assisted analysis of programme complementarity and gap coverage – e.g. microsimulation of how different combinations of benefits and services (across different pillars) reach different vulnerability profiles – to inform decisions on optimal package composition and encouraging transitions to social insurance where appropriate.
	Matching and recommendation	Recommendation systems that analyse which combinations of benefits and services are most effective for different vulnerability profiles, informing programme design decisions on package composition and cross-pillar linkages.

Programmatic decisions on the most appropriate  **ELIGIBILITY CRITERIA AND QUALIFYING CONDITIONS:** *Targeting design decisions, which differ across programmes and should focus on the inclusion of those facing the highest barriers and lifecycle risks*	**Vulnerability, needs and risk assessment, including predictive analytics**	Informing benefit-specific eligibility design choices such as tailored eligibility criteria (e.g. via a combination of geographic, household and individual level dimensions).
	Decision support	AI-augmented simulation tools that test how different eligibility threshold combinations affect coverage rates, inclusion/exclusion errors and equity outcomes across population subgroups, enabling evidence-based calibrations.
Programmatic decisions on the most appropriate  **LEVEL, VALUE, FREQUENCY AND DURATION OF SUPPORT**	**Vulnerability, needs and risk assessment, including predictive analytics**	Informing programme-specific design choices on level/value (depending on if delivered in cash or kind), frequency and duration of support, based on an understanding of individual and household needs and circumstances (e.g. disability status, etc.).
	Trend and shock forecasting	Projections of how inflation, labour market shifts or climate shocks will affect the real value of a fixed transfer level over time, informing indexation decisions.
	Policy analysis, learning and M&E	Microsimulation of how changes to benefit levels, payment frequency or programme duration affect poverty outcomes, fiscal costs and adequacy across different population subgroups.

Implementation level ('delivery chain')

Table 4 presents the administrative functions necessary to deliver benefits or services to those in need, when in need. Most of these functions can be performed in an integrated way, serving multiple programmes.

TABLE 4. IMPLEMENTATION LEVEL (DELIVERY CHAIN): MATRIX VIEW OF GENERIC AI USE CASES AGAINST SOCIAL PROTECTION FUNCTIONS

KEY FUNCTIONAL AREA (see also the 'Social Protection as a Solar System' publication)	RELEVANT AI USE CASE CATEGORIES (Only the most relevant for each dimension are shared here)	HOW THE AI USE CASE CATEGORY APPLIES TO EACH SOCIAL PROTECTION FUNCTION
 **OUTREACH / COMMUNICATIONS / SENSITISATION ACTIVITIES:** *To promote awareness of entitlements and procedures, on an ongoing basis and inclusive manner*	**User communication and interaction**	Supporting two-way information exchange, including responses to queries, navigation of the user journey (application, enrolment, payments, etc.) and provision of information (e.g. on rights and responsibilities). Note: ***This is valid across all client-facing dimensions below, especially rights-critical stages and accountability.***
	Vulnerability, needs and risk assessment, including predictive analytics	Geospatial poverty mapping and non-take-up detection to identify individuals, communities and population groups not yet reached, informing targeted outreach prioritisation and tailored communication channel selection.
	Matching & recommendation	Recommendation systems that match communication channels, languages, messengers and messaging content to the profiles of specific under-reached population segments, improving the inclusivity and effectiveness of sensitisation campaigns.

 **REGISTRATION:** *To gather information on people's characteristics, and sometimes on needs, and conditions, ideally via continuous and on-demand processes*	**Vulnerability, needs and risk assessment, including predictive analytics**	Supporting prioritisation and deployment decisions for different types of registration processes (including via geospatial mapping), ensuring non-take-up detection, etc.
	Identification, verification and record linkage	Supporting interoperability across existing databases to support registration function.
	Operational & process automation	Methods for data collection optimisation (e.g. adaptive questionnaires) and for forecasting workloads or resource needs.
	Data quality and anomaly detection	Predictive methods for data imputation, data quality checks, deduplication etc. adopted at registration stage to ensure data on potential beneficiaries is correct and complete.
	User communication and interaction	Chatbots and guided digital interfaces that support applicants through the registration process, answering questions, explaining documentation requirements and providing multilingual assistance.
 **MANAGEMENT OF CONTRIBUTIONS, CLAIMS AND WITHDRAWALS** (SOCIAL INSURANCE)	**Operational & process automation**	Automated processing of contribution records, employment status updates, claims and withdrawal requests; AI-assisted reconciliation of payroll data with contribution histories; and workload forecasting for contribution management operations.
	Identification, verification and record linkage	Verification and cross-registry linkage to authenticate identity at point of contribution or withdrawal, detect ghost contributors, and reconcile records across employer payroll systems, tax registries and beneficiary databases.
	Data quality and anomaly detection	Anomaly detection applied to contribution patterns to flag irregular or potentially fraudulent employer or worker records, and to ensure contribution histories are complete and accurate for entitlement calculations.
	Compliance and integrity	Monitoring adherence to mandatory contribution requirements, detecting systematic non-compliance by employers, and identifying patterns of under-reporting or evasion that affect long-term entitlement adequacy.
	Trend and shock forecasting	Forecasting contribution revenue based on labour market trends, employment projections and economic conditions, supporting financial planning and reserve management for social insurance schemes.

 **ASSESSMENT OF NEEDS/ CONDITIONS:** *To determine eligibility for different programmes based on each programme's different qualifying conditions and eligibility criteria* **+** **ENROLMENT:** *To formalise the enrolment of eligible beneficiaries via authentication of identity, notification and onboarding*	**Identification, verification and record linkage**	Confirms identity and relevant documentation.
	Decision support for determination of eligibility, and benefits and services packages	AI-assisted scoring and model-augmented decision support that translates verified applicant data into eligibility recommendations or benefit package configurations, supporting caseworkers or automated systems in determining entitlement while maintaining appropriate human oversight given the high decision criticality of this function.
	Data quality and anomaly detection	Predictive methods for data imputation, data quality checks, deduplication etc. adopted at enrolment to ensure beneficiary data is correct and complete.
	Operational & process automation	OCR methods for document processing.
	Matching and recommendation	Recommendation systems that suggest the most appropriate combination of programmes and services for newly enrolled beneficiaries based on assessed needs, vulnerability profile and available provision, supporting joined-up enrolment across multiple programmes simultaneously and reducing the risk of eligible individuals missing complementary entitlements.
	Vulnerability, needs and risk assessment	Predictive profiling of applicant needs beyond binary eligibility, supporting holistic assessment of multi-dimensional vulnerability to inform appropriate package assignment and referral at the point of enrolment.
 **PROVISION OF PAYMENTS AND / OR SERVICES:** *To deliver benefits or services to beneficiaries via channels and processes that cater to their different needs and preferences, while curbing the potential for fraud*	**Identification, verification and record linkage**	Confirms identity and relevant documentation to comply with know your client (KYC) to enable payments while enabling payment reconciliation.
	Data quality and anomaly detection	Predictive methods for data imputation, data quality checks, deduplication etc. adopted at payments stage to ensure payment integrity.
	Operational & process automation	Methods for forecasting workloads or resource needs to optimise the payment delivery process.
	Matching & recommendation	Recommendation systems: flagging the most appropriate payment channel, modality or service delivery mechanism based on beneficiary characteristics and preferences (e.g. mobile money vs agent networks vs bank transfer), improving accessibility and reducing failure rates.
	Compliance and integrity	Payment-stage compliance verification, such as confirming conditionality adherence or programme requirement fulfilment before disbursement is authorised, including flagging cases where conditions have not been met for caseworker review.

 **ACCOUNTABILITY MECHANISMS (e.g. ADDRESSING COMPLAINTS AND APPEALS):** *To enable people to hold the state and service/benefit providers accountable – may include channels to complain, provide feedback, or appeal service/benefit provider's decisions, as well as processes for these to be responded to and addressed*	**Compliance and integrity**	AI-driven triage and classification of incoming complaints and appeals by type, urgency and prior case history, combined with pattern detection across complaints and appeal data to identify systemic non-compliance issues, recurring programme-rule failures or clusters of similar complaints that may indicate broader delivery chain problems requiring escalation.
	Operational & process automation	Automated routing and triage of incoming complaints/appeals based on complaint type, urgency and prior case history, combined with workload forecasting to ensure adequate staffing of redress mechanisms.
	Policy analysis, learning and M&E	Systematic analysis of grievance patterns to detect systemic programme design or delivery failures, informing iterative improvements to accountability mechanisms and feeding back into programme design.
	User communication and interaction	Chatbots and automated channels for submission of complaints and appeals, complaint status tracking and multilingual support, improving accessibility of redress mechanisms particularly for populations with low literacy or in remote areas.
	Identification, verification and record linkage	Verifying the identity of complainants and linking complaints and appeals records to beneficiary case files across systems, enabling tracking of complaint resolution and preventing duplicate or fraudulent submissions.
 **CASE MANAGEMENT** *To manage the complexity of individuals' multidimensional needs, while also supporting beneficiaries in their interactions with programmes/services and facilitating referrals and linkages across programmes*	**Vulnerability, needs and risk assessment, including predictive analytics**	Predictive risk profiling to identify complex or multi-dimensional need cases that require specialist case management referral; this also includes ongoing predictive risk scoring applied during case management to identify beneficiaries at elevated risk of adverse outcomes (e.g. exit from programme, deteriorating compliance, emerging crisis), enabling proactive caseworker intervention before situations escalate.
	Matching & recommendation	Connecting individuals to programmes providing suitable benefits, services, support opportunities, jobs, training.
	Identification, verification and record linkage	Confirms identity, enables linkages and case management across different databases and enables compliance monitoring.
	Compliance and integrity	Conditionality verification (where this is appropriate) (e.g. monitoring conditionality adherence, detecting non-compliance patterns).
	Operational & process automation	AI-assisted case file management, automated data updates from linked registries, and document processing tools that reduce the administrative burden on caseworkers and enable them to focus on high-value direct support interactions.

 **PROFILING, JOB MATCHING AND SUPPORT SERVICES** FOR LABOUR MARKET PROGRAMMES	**Identification, verification and record linkage**	Cross-registry linkage between labour market programme registries, employment service databases, social registries and tax/payroll records to build a unified and continuously updated employment profile for each participant, supporting accurate skills assessment, job matching and benefits.
	Matching & recommendation:	AI-powered job matching that analyses participant skills profiles, labour market vacancy data, geographic accessibility and employer preferences to generate ranked job recommendations; combined with career pathway guidance tools that suggest training sequences aligned with realistic employment trajectories.
	Vulnerability, needs and risk assessment, including predictive analytics	Predictive profiling of labour market programme participants to segment by distance-to-employment and identify those at highest risk of long-term unemployment, enabling differentiated support intensity and early intervention.
	Trend and shock forecasting	Labour market forecasting to anticipate demand for different skills and occupations, informing training programme.
	Operational and process automation	Automated processing of skills assessment documentation, CV parsing, certification verification and administrative workflow automation in employment service operations.
 **MONITORING AND EVALUATION:** *To measure – and iteratively improve – performance against agreed delivery standards and intended outcomes*	**Data quality and anomaly detection**	Automated data quality monitoring across programme management information systems (MISs), with anomaly detection applied to programme performance data to flag implausible trends, data entry errors or potential manipulation of performance metrics, thereby ensuring the integrity of the evidence base for M&E.
	Policy analysis, learning and M&E	Trend monitoring for service improvements, improved analytics/dashboards, user satisfaction, system stress-test under varied scenarios, model performance and bias/fairness monitoring, policy microsimulation, counterfactual analysis and synthetic evaluation.
	Vulnerability, needs and risk assessment	Tracking how population vulnerability and needs profiles change over time, detecting emerging coverage gaps or shifts in target population characteristics that feed back into programme redesign and recalibration.

2. SECONDARY CLASSIFICATION:

other relevant dimensions to better understand the use of AI for social protection

2.1 AI classification dimensions

Beyond the primary dimensions set out in Section 1, the following dimensions will also be used to classify AI.

AI lifecycle stages

The AI lifecycle is adapted from established frameworks[5, 6]. The AI lifecycle recognises that risks accumulate over time, rather than appear at a single stage.

TABLE 5. AI LIFECYCLE STAGES

LIFECYCLE STAGE	DESCRIPTION
Problem identification and design	Defining the problem the AI system is meant to solve, the intended users and context, and the success criteria, while considering feasibility, risks, and ethical implications from the outset.
Data gathering and preparation	Collecting, cleaning, labelling, and managing data needed to train the model, ensuring data quality, representativeness, and compliance with legal and ethical standards.
Model selection and training	Choosing appropriate algorithms and model architectures, then training and validating the model to optimise performance while addressing issues such as bias, overfitting, and explainability.
Integration and deployment	Embedding the trained model into existing systems or workflows and releasing it into real-world use, including testing for reliability, security, and user interaction.
Monitoring, maintenance, and decommissioning	Continuously tracking performance and impacts over time, updating or retraining the model as conditions change, and safely retiring the system when it is no longer effective or appropriate.

5 Organisation for Economic Co-operation and Development (OECD), *Recommendation of the Council on Artificial Intelligence, Organisation for Economic Co-operation and Development*, Paris, 2019 (revised 2024), https://oecd.ai/en/ai-principles

6 National Institute of Standards and Technology (NIST), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (NIST AI 100-1), U.S. Department of Commerce, Gaithersburg, Maryland, 2023, https://doi.org/10.6028/NIST.AI.100-1

AI implementation types

The primary classification (Section 1) distinguished AI capabilities by task type – what the output does. This sub-section classifies the implementation approach – how the output is produced. The two dimensions are independent: the same analytical task (e.g. classification) can be implemented via classical machine learning (ML), deep learning, or a foundation model, and the choice of implementation carries distinct implications for validation, operational feasibility and governance.

This distinction matters for social protection practitioners, because the implementation approach significantly affects three practical dimensions:

- **Validation and reliability.** Classical ML approaches for structured data (logistic regression, decision trees, gradient-boosted models) are generally easier to validate, more reproducible, and more stable across runs. Foundation model approaches can be more sensitive to prompting, context and configuration, and may require additional guardrails to ensure consistent outputs in operational settings.

- **Compute and operational sustainability.** Classical ML models are typically lightweight and can be trained and deployed in resource-constrained environments. Foundation models often require substantially greater compute for both training and inference, affecting total cost of ownership, infrastructure requirements and environmental footprint, considerations that are particularly relevant in low- and middle-income country (LMIC) contexts with limited technical infrastructure.

- **Risk and governance profile.** Both implementation types raise fairness and due-process risks, but foundation model deployments introduce additional concerns including plausible-but-incorrect outputs (hallucination), prompt-related vulnerabilities, reduced interpretability, and dependency on third-party model providers. These typically require stronger human review protocols, output logging and use boundaries, particularly in higher-stakes workflows.

In practice, many deployments are hybrid – for example, a foundation model performing extraction and structuring of unstructured inputs, feeding into a classical ML model for scoring and routing; or retrieval-augmented generation with rule-based guardrails.

For these reasons, analytical (traditional) approaches – particularly classical ML on structured administrative data – are likely to remain the default implementation for core social protection functions, such as welfare estimation, eligibility scoring, fraud detection and case prioritisation in LMIC contexts, for the foreseeable future. Foundation model approaches are most immediately applicable when generative tasks are required or unstructured data processing is the primary function.

TABLE 6. AI IMPLEMENTATION TYPES

IMPLEMENTATION TYPE	DESCRIPTION	CHARACTERISTICS	COMMON SOCIAL PROTECTION APPLICATIONS
Classical machine learning	Statistical and ML methods including logistic regression, decision trees, random forests, gradient-boosted models (e.g. XGBoost), support vector machines, and k-nearest neighbours. Models are trained on structured, tabular datasets with explicitly engineered features.	Lightweight compute requirements; high interpretability (particularly linear and tree-based models); well-established validation methods; strong reproducibility; limited to structured inputs without additional preprocessing	Proxy means testing (PMT)/welfare estimation, eligibility scoring, fraud detection on transaction data, workload forecasting, recertification risk scoring
Deep learning	Neural network architectures including convolutional neural networks (CNNs), recurrent neural networks (RNNs/LSTMs), and transformer-based models trained for specific tasks on domain data. Requires larger datasets and more compute than classical ML, but can learn directly from unstructured inputs.	Moderate-to-high compute requirements; reduced interpretability compared to classical ML; strong performance on image, audio and sequential data; requires larger labelled training datasets; task-specific training	Satellite imagery analysis for poverty mapping, biometric matching (fingerprint, facial, iris), document OCR and field extraction, speech recognition for voice-based services
Foundation models	Large pre-trained models (including LLMs and multimodal models) that can be adapted to downstream tasks through fine-tuning, prompting, or retrieval-augmented generation. Trained on broad data at scale, not task-specific by default.	High compute requirements for training (typically accessed via API or pre-trained); adaptable to multiple tasks without retraining; risk of plausible-but-incorrect outputs; prompt sensitivity; dependency on external model providers in most LMIC deployments; rapidly evolving capabilities	Chatbots and conversational assistants, document summarisation, translation and multilingual support, policy document analysis, complaint/appeal triage, synthetic data generation, caseworker drafting support
Hybrid	Systems combining multiple implementation types in a pipeline – for example, a foundation model for extraction feeding a classical ML model for scoring, or a classical model with an LLM-generated explanation layer.	Governance profile determined by the highest-risk component; requires clear documentation of which component produces which output; integration complexity	Extract-and-score pipelines (e.g. NLP extraction from case notes followed by ML-based prioritisation), retrieval-augmented generation with rule-based guardrails, AI-assisted decision support combining predictive scoring with generated explanations

AI decision criticality

Decision criticality determines the required level of human oversight based on the rights impact of the decision the system supports.

TABLE 7. LEVELS OF DECISION CRITICALITY

LEVEL	DESCRIPTION
High criticality	Systems in this category require human-in-the-loop (HITL) oversight, because errors may directly affect rights, livelihoods, or entitlements. Examples include eligibility recommendation or payment-suspension triggers based on anomaly detection.
Moderate criticality	Moderate criticality decisions affect service quality, but not entitlements. Errors may delay services, but are generally reversible. These systems require human-on-the-loop (HOTL) oversight to ensure early detection of any issues.
Low criticality	Low criticality decisions support internal administration with minimal rights impact. Examples include workload forecasting or automated form processing, under the operational and process automation category. These systems may operate with human-out-of-the-loop (HOOTL) oversight, supported by periodic audits.

Human oversight requirements for AI

TABLE 8. HUMAN OVERSIGHT REQUIREMENTS (HITL, HOTL, HOOTL)

TYPE	WHAT IT IS	DESCRIPTION
Human-in-the-loop (HITL)	HITL requires active human review before a system influences an outcome.	Officers must confirm or override AI recommendations before rights-affecting decisions are finalised. This oversight is essential for eligibility decisions and benefit adjustments, for example.
Human-on-the-loop (HOTL)	HOTL requires human monitoring of system performance.	Staff observe the tool during operation, intervene when required, and escalate complex cases. This is suitable for routing of complaints and appeals or forecasting tools.
Human-out-of-the-loop (HOOTL)	HOOTL permits autonomous operation with periodic audit.	Routine administrative tasks, such as sorting documents or processing optical character recognition, may operate autonomously where the rights impact is low and the task is reversible.

AI development process

TABLE 9. AI DEVELOPMENT PROCESS (AND POTENTIAL SOCIAL PROTECTION IMPLICATIONS)

TYPE	DESCRIPTION	POTENTIAL SOCIAL PROTECTION IMPLICATIONS
Fully in-house development	AI systems designed, built, and maintained entirely by government entities	**Advantages:** stronger alignment with public interest and user-centred design; greater accountability to oversight bodies; sensitive beneficiary data (income, health, biometrics) remains under government control, reducing privacy risks; institutional knowledge stays within government, supporting long-term sustainability **Risks:** • Likely capacity constraints/bottlenecks, leading to slower deployment • Limited access to cutting-edge AI expertise
Mix of in-house and third party development	AI systems co-developed through collaboration between government teams and external vendors or partners	**Advantages:** can help balance private sector innovation with public accountability, supporting knowledge transfer that builds lasting government capability Suffers some of the **risks** described below, while possibly leading to: • Unclear accountability, i.e. harder to attribute responsibility between government and vendor • Risk of dependency creep: hybrid arrangements can gradually shift towards over-reliance on the vendor, eroding internal capacity despite original intent • Requires strong government capacity to define requirements, evaluate model outputs, and maintain oversight – this is often underestimated and underfunded
Fully third-party developed (commissioned)	AI systems fully designed and delivered by external providers under government contract	**Advantages** centre around accessing cutting-edge AI expertise, **risks** include: • Lack of capacity for setting clear objectives, guardrails and managing complex and costly AI-focused contracts • Less likely alignment with ultimate social protection objectives (private sector optimises for efficiency, not equity, reaching the hardest-to-reach and impact) • Potential liability issues – requires strong scrutiny and safeguards, noting issues of algorithmic opacity (government may have limited ability to audit AI models for bias, error, or discriminatory patterns, undermining due process and the right to contest decisions) • Vendor lock-in and external dependencies • Private sector has limited incentives to transfer knowledge • Data protection risks: sensitive data processed by private entities may be subject to weaker safeguards, secondary use, or cross-border transfer inconsistent with national frameworks and sovereignty • Service continuity risks as well as limitations to adaptability: policy changes may require costly contract renegotiations

AI deployment status (and potential social protection implications)

Lifecycle stages describe the technical development process a system passes through (from problem definition through to monitoring and decommissioning), while deployment status captures how far the system has progressed into real-world operational use and institutional embedding. Both dimensions are independent, as a system may be technically mature but not yet scaled, or widely deployed but inadequately monitored.

TABLE 10. AI DEPLOYMENT STATUS

TYPE	DESCRIPTION	POTENTIAL SOCIAL PROTECTION IMPLICATIONS
Concept / exploration phase	• Problem definition, feasibility assessment, stakeholder consultation, or initial scoping underway • No functional system yet	• Opportunity to embed rights-based principles (data protection, accountability, explainability) from the outset, far harder to retrofit later • Readiness assessment crucial, especially from a governance perspective (assessing legal frameworks, oversight structures, and data availability early can prevent costly redesigns) • Early stakeholder engagement, including with intended beneficiaries, frontline workers, civil society organisations and social partners, is critical to avoid designing systems that replicate existing exclusions or biases, or simply do not address real needs
Design & development phase	• Model design, data preparation, and system development in progress • Internal testing may be underway, but not yet used in live operations	• Data readiness (ensuring training data quality) and risk assessment leveraging • Risk Toolkit • Explainability and contestability designed in from the start • Involving frontline workers, social partners and potential beneficiaries in early user testing • Ensuring system is not optimising for efficiency metrics (processing speed, cost reduction) at the expense of equity and inclusion outcomes
Pilot / controlled trial phase	• Limited deployment in a defined setting (e.g. specific region, department, or use case) • Performance, risks, and operational viability being evaluated	• Ensuring complaints and appeals mechanisms are operational even at pilot stage to pick up any early issues • Detecting for bias in the choice of pilot populations • Considering independent evaluation of pilot results with inclusion/exclusion rates disaggregated by population subgroups

Operational deployment (limited rollout)	• Actively used in real workflows, but not yet scaled across the full intended scope • May still include monitoring and refinement	• Rights-critical phase as consequences of AI decisions affect people – complaints and appeals and adaptive re-programming crucial • Staff will require substantial change management support • Integration with live delivery systems will surface further data quality and interoperability issues not caught in testing that will need addressing • Monitoring must track distributional impacts across vulnerable subgroups (e.g. by gender, disability, geography), not only system-level performance averages • Importance of resisting institutional pressure to scale before the system has been adequately stress-tested
Full production deployment	• Fully integrated into government operations at intended scale • Embedded in business processes with governance and maintenance structures in place	• Errors at scale will have large aggregate human impact. These need to be closely monitored and addressed at speed, also meaning that accountability infrastructure (audit trails, explainability, complaints and appeals mechanisms) must remain accessible and functional • Model drift risk, needs continuous monitoring too • Risk of complacency as AI decisions become routine: human oversight and meaningful review of algorithmic outputs can erode over time
Scaled & institutionalised	• System embedded in policy frameworks, budgets, and long-term operational planning • Includes formal oversight, auditing, and lifecycle management	• Risks of over-reliance and un-learning among staff, while knowledge of AI system's limitations, assumptions, and failures modes may erode, making it harder to challenge or correct • Periodic independent auditing essential to detect performance degradation, emergent biases, or gaps between system design and current policy intent • Long-term sustainability requires dedicated budget lines, maintenance capacity, and governance structures • Many points in previous rows also hold true

AI risk categories and societal impact dimensions

The AI Hub *Risk Assessment Framework for Social Protection*[7] identifies five structural sources of risk that accumulate across the AI lifecycle, together with a downstream layer capturing how these risks materialise as human and societal impacts. Failures in one domain frequently amplify weaknesses in another: a model trained on biased data will produce unfair outcomes, even if technically robust; strong governance cannot compensate for poor integration; and secure data practices can be undermined if market structures restrict auditability or migration. Risk assessment must, therefore, consider all five sources in combination, not in isolation. Full definitions, mitigation strategies and assessment tools are provided in the Risk Assessment Framework for Social Protection and the companion Toolkit.

TABLE 11. STRUCTURAL RISK DIMENSIONS

RISK CATEGORY	DESCRIPTION	RISK DIMENSIONS
Data-related risks	Risks arising when the information used to train, validate or operate an AI system does not accurately or fairly reflect the real world. In social protection, systematic gaps in registries and administrative data tend to under-represent populations most in need.	Representation bias, data quality failure, cross-dataset inconsistency, data or concept drift, weak provenance or lineage, consent or lawful basis gap
Model-related risks	Risks originating in the architecture, objectives and internal behaviour of the algorithm itself. These are particularly consequential in social protection where outputs directly influence eligibility, sanctions or payment decisions.	Opacity or limited explainability, objective misalignment, model misspecification, shortcut learning and proxy reliance, subgroup bias, reliability or generalisation failure, hallucination or misinformation, model security vulnerability, behavioural drift
Operational and system integration risks	Risks arising when an AI model is deployed within real infrastructures, workflows and data pipelines. In social protection, these risks are compounded by staff tendency to defer to algorithmic outputs without critical scrutiny.	Legacy system integration failure, pipeline fragility, threshold or rule misconfiguration, monitoring gap, inadequate real-world validation, scalability or latency constraint, weak operational management of deployed models, automation complacency
Governance and institutional oversight risks	Risks emerging when institutions lack accountability, capacity, oversight mechanisms, legal safeguards or meaningful control over AI systems. These factors determine how much de facto autonomy AI systems acquire and whether errors are detected, corrected or ignored.	Unclear accountability, insufficient human oversight, weak monitoring ownership, weak documentation or auditability, inadequate avenues for complaints and appeals or redress, regulatory non-compliance, purpose limitation failure, insufficient institutional capacity, inadequate resourcing
Market, sovereignty and industry structure risks	Risks arising from external dependencies on vendors, infrastructure providers and global technology ecosystems. These risks are extensively addressed in the companion report on data sovereignty considerations published by the AI Hub for Social Protection of the DCI[8].	Vendor lock-in, interoperability constraints, jurisdictional hosting risk, restricted audit access, upstream model or API dependency, absent exit or migration pathway, market concentration, opaque supply chain, LMIC power asymmetry, systemic or cascading fragility

7 DCI AI Hub for Social Protection, *AI Hub Risk Assessment Framework for Social Protection: A Practical Guide to Responsible AI Deployment*, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026.

8 DCI AI Hub for Social Protection, *AI in Social Protection: Data Sovereignty Considerations*, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026.

The five structural risk sources are upstream conditions; the human and societal impact layer captures how they materialise in people's lives and in institutional outcomes. These downstream impacts are organised around the four guiding principles established in the Risk Assessment Framework for Social Protection.

TABLE 12. HUMAN AND SOCIETAL IMPACT DMENIONS

IMPACT DOMAIN	DESCRIPTION	IMPACT DIMENSIONS
Privacy and data security	Harms arising from excessive data collection, unlawful linkage, breaches of sensitive information, or loss of individual control over personal data	Privacy violation or data breach, disproportionate surveillance or profiling, loss of individual control over personal data
Equality, non-discrimination, fairness and inclusion	Discriminatory outcomes, systematic exclusion of marginalised groups, adverse inclusion, disparate error rates, and reinforcement of structural inequities	Discriminatory outcomes, systematic exclusion from benefits or services, disparate error rates across groups, reinforcement of structural inequity
Autonomy, human dignity and due process	Loss of agency, opaque or unexplained decisions, inability to contest outcomes, erosion of procedural justice through rigid automation, and psychological stress or stigma from intrusive profiling	Loss of individual agency or autonomy, opaque or unexplained decisions, inability to contest or appeal outcomes, psychological stress stigma or dignity harm
Accountability, transparency and redress	Absence of a named decision owner, missing audit trails, inaccessible complaint and appeals mechanisms, and inability to trace, explain or correct AI-assisted decisions	No identifiable decision owner, untraceable decision pathway, no accessible or effective remedy
Systemic and societal	Impacts that erode public trust, reinforce structural inequities, trigger political backlash, increase administrative burden on frontline staff, and deepen digital deepen digital divides, while also weakening social solidarity	Erosion of public trust in social protection system, political backlash litigation or controversy, increased administrative burden on frontline staff, deepened digital divide

2.2 Social protection classification dimensions

Beyond the primary AI use cases (developed to showcase core social protection functions that can be supported by AI), the following dimensions are also relevant for classification.

Programme details

Where the use cases relate to specific programmes, rather than the social protection system as a whole, the following dimensions will be relevant.

TABLE 13. PROGRAMME DETAILS

CATEGORY	SUB-CATEGORY	DETAILS
Social protection branch / pillar and programme type	Non-contributory	• Social assistance cash transfers (conditional or unconditional) • Child grants/benefits (universal or poverty-targeted) • Non-contributory disability grants (universal or poverty-targeted) • Non-contributory social pensions (universal or poverty-targeted) • In-kind transfers • Public works programmes • School feeding programmes • Fee waivers and targeted subsidies • Other (including ad-hoc emergency transfers)
	Contributory	• Old age, survivors and disability pensions (contributory) • Unemployment insurance • Health insurance • Work injury and occupational insurance • Maternity and paternity benefits • Other
	Labour market programmes	• Skilling and training programmes • Entrepreneurship support • Intermediation services • Employment incentives/wage subsidies • Job search assistance and placement services • Public employment services • Productive inclusion programmes • Other
Programme name	Name	
Programme description	Brief description	

Note: *Where a programme combines elements across pillars (e.g. a conditional cash transfer with a graduation/skills training component), assign one primary pillar based on the programme's core function and record the secondary pillar as a linkage.*

2.3 Contextual classification dimensions

The dimensions in Table 14 provide contextual information.

TABLE 14. CLASSIFYING CONTEXT

CATEGORY	WHAT IT IS
Country information	• Country: name of the country • Region: region as per World Bank Classification FY26 • Income level: as per World Bank Classification FY26 • Fragile and conflict affected state (FCAS) classification: yes/no
Tech maturity as per available index	• GovTech Maturity Index (GTMI) – World Bank
AI maturity as per available index	• Oxford Insights – Government AI Readiness Index
Respondent details (where relevant)	• Name of the respondent • Institution the respondent is affiliated with • Email ID of the respondent

2.4 Enabling data environment classification

Data sources, key categories

For each documented AI-enabled use case, it is important to document structured attributes for every data source that may be used, capturing not only which data sources are involved, but how they are used in the AI pipeline, their sensitivity and governance status, and any constraints on availability or quality. Understanding the data quality attributes and de-facto status of these different data sources supports assessment of whether applications documented in one setting are transferable to others with different data environments.

TABLE 15. ENABLING DATA ENVIRONMENT – DATA SOURCE CATEGORIES FOR SOCIAL PROTECTION AI

DATA SOURCE CATEGORY	DESCRIPTION	KEY SOCIAL PROTECTION USE CASES
Social registries	Information systems containing potential beneficiaries' socio-economic information, collected through census-sweep or on-demand registration processes. Social registries hold data on a wider population than those currently enrolled in programmes, enabling cross-programme targeting and coordination. Note: these are not always centralised – in many LMIC contexts they are federated, fragmented or programme-led, but used across programmes. Common limitations include reliance on self-reported data, infrequent updates, and under-registration of marginalised groups. These vary widely across countries in terms of their coverage, the data they collect, their level of institutionalisation, etc.	Vulnerability and needs assessment (poverty scoring, geographic targeting), eligibility decision support (PMT inputs), identification and record linkage (deduplication across programmes), policy analysis (coverage and exclusion analysis)
Beneficiary registries/ integrated beneficiary registries & management information systems (MIS)/beneficiary operations management systems (BOMS)	Programme-specific records of enrolled beneficiaries, including entitlement details, payment histories, compliance records, case notes and status changes. These are the operational backbone of programme delivery. Note: in practice, benefit administration records, payment and reconciliation records, and complaints and appeals records may be stored in separate systems with different governance and quality profiles, even when nominally part of a single MIS. Not all countries invest in integration across programme-specific MISs (enabling an overview of who is receiving what across programmes and social protection pillars).	Compliance and integrity monitoring, data quality and anomaly detection (transaction-level fraud flagging), operational automation (workflow routing, caseload management), recertification and exit decision support
Civil registration and vital statistics (CRVS)	Records of births, deaths, marriages and other vital events maintained by civil authorities. Completeness and timeliness vary significantly across LMICs, which directly affects the feasibility of AI-enabled applications that depend on verified identity or life-event data. Late registration and delays in death reporting can be operationally decisive for proof-of-life verification, survivor benefits and fraud prevention. Not all countries have one, or have high coverage.	Identification and record linkage (verifying identity, confirming vital status), eligibility decision support (age-based entitlements, survivor benefits), data quality (cross-verification of registry records against vital events)

National ID and biometric databases	Identity infrastructure including foundational ID systems (e.g. national ID numbers) and associated biometric templates (fingerprint, iris, facial). Coverage varies – some LMICs have extensive foundational ID systems (India Aadhaar, Pakistan NADRA), others have significant gaps, particularly among migrants, remote populations or people lacking documentation. Note: high nominal coverage can mask the exclusion of marginalised groups. Foundational ID registers and biometric authentication services should be distinguished, as some systems provide IDs widely, but have limited biometric matching service coverage, uptime or integration. Many countries lack these, prioritising functional IDs over foundational ID systems.	Identification and verification (biometric matching, proof-of-life), record linkage (entity resolution across fragmented systems), data quality (deduplication, ghost beneficiary detection)
Administrative data from other sectors	Records held by government agencies outside social protection. Three broad sub-types are relevant: (i) economic and labour data – tax, payroll, formal employment records, enterprise-level data; (ii) human capital and service utilisation data – education enrolment and attendance, health system records; (iii) asset and residence proxies – land and property registries, vehicle registration, utility connections. These records are often the richest data source for welfare estimation models (e.g. PMTs), but availability depends heavily on interoperability infrastructure and data-sharing agreements between agencies.	Vulnerability assessment (income and welfare estimation using cross-sectoral proxies), eligibility decision support (cross-verification of declared income/assets), data quality and anomaly detection (consistency checks across systems), non-take-up detection
Financial and payments data: programme operations	Transaction records from programme payment operations – disbursement records, reconciliation data, failed payment logs, duplicate payment flags. These are generated within the programme's own payment infrastructure and are usually justifiable within programme mandate without additional consent requirements.	Data quality and anomaly detection (duplicate disbursements, payment reconciliation errors), operational automation (payment processing, exception handling)
Financial and payments data: beneficiary financial behaviour	Broader financial data including mobile money transaction patterns, banking records, credit histories and financial inclusion indicators. This data is accessed through payment service providers, mobile money platforms or banking systems. It has materially different governance requirements from programme payment data, typically requiring an explicit legal basis, informed consent and proportionality assessment, and raising significant privacy concerns.	Vulnerability assessment (financial inclusion proxies, consumption estimation from mobile money patterns), compliance and integrity (cross-verification of reported income) **Note: use of this data for social protection targeting or compliance monitoring requires careful proportionality assessment given the sensitivity of personal financial information.**

Geospatial and remote sensing data	Satellite imagery (optical and radar), nighttime lights data, land use and land cover classifications, climate and weather data (historical and forecast), digital elevation models and GIS infrastructure layers (roads, markets, health facilities). This data is increasingly accessible through open data platforms and from commercial providers. Note: remote sensing models typically require ground truth data from household surveys or administrative sources for calibration and periodic revalidation – this creates a dependency on survey data recency.	Vulnerability assessment (geospatial poverty mapping, asset wealth estimation from satellite features), trend and shock forecasting (climate hazard prediction, drought monitoring, flood risk), policy analysis (geographic coverage gaps, service accessibility modelling)
Telecommunications and mobile data	Call detail records (CDR, aggregates or individual-level metadata), mobile usage patterns, connectivity and coverage data, and mobility aggregates. These are used as alternative data sources for welfare estimation and mobility analysis, particularly where traditional survey or administrative data is sparse. Typically accessed through partnerships with mobile network operators. Note: representativeness is a documented concern that should be assumed unless demonstrated otherwise in a given context – mobile data tends to under-represent the poorest and most marginalised populations. The privacy-preserving access modality matters significantly: on-operator processing, secure enclaves and de-identified aggregates carry different risk profiles from individual-level data extraction.	Vulnerability assessment (alternative welfare proxies where survey data is scarce), trend and shock forecasting (population displacement and mobility tracking), outreach optimisation (identifying connectivity patterns for digital service delivery)
Survey and census data	Household surveys (e.g. demographic health surveys [DHS], living standards measurement surveys [LSMS], multiple indicator cluster surveys [MICS]), population and housing censuses, labour force surveys and other nationally representative data collection instruments. These are typically the ground truth against which AI-enabled models are trained and validated, but are expensive and conducted at multi-year intervals (commonly every decade for censuses and every few years for major surveys, subject to country practice and financing). In humanitarian and fragile settings, programme-specific rapid assessments, post-distribution monitoring data and Integrated Food Security Phase Classification (IPC)/food security classifications serve similar functions, but with different coverage, frequency and institutional ownership. Model staleness risk: models trained on survey data with a multi-year lag may not reflect current conditions, particularly in contexts affected by rapid economic change, displacement or climate shocks. Concept and policy drift may also invalidate learnt relationships even when population characteristics remain stable.	Vulnerability assessment (training and validating poverty prediction models), trend and shock forecasting (baseline demographic and socio-economic parameters), policy analysis (microsimulation, impact evaluation, distributional analysis)

Unstructured and text-based content	Case notes and caseworker assessments, beneficiary correspondence and complaint/appeal submissions, policy documents and programme guidelines, media and news feeds. These are increasingly processable through NLP and LLM capabilities. Note: governance requirements vary significantly by source type within this category. Complaints, appeals and feedback data has high relevance to rights and redress, distinctive retention rules and frequent NLP applications (triage, routing, summarisation). Social media and news feeds are usually weakly justified for operational social protection decisions and should be treated as high-risk, low-proportionality inputs, unless used strictly for aggregate situational awareness. Language coverage and low-resource language support are often the primary determinants of NLP feasibility in LMIC contexts.	User communication and interaction (chatbot training, translation, explainable notices), compliance and integrity (complaint/appeal triage and pattern detection), policy analysis (document synthesis, evidence review), operational automation (case note summarisation, correspondence drafting)

Data source coding template

For each data source used by a documented AI-enabled use case, the following structured fields should be recorded. This template operationalises the data source classification for consistent coding and enables users to filter the tracker by data environment characteristics.

TABLE 16. DATA SOURCE CODING TEMPLATE

TRACKER FIELD	VALUES
Source category	From Table 14 (select one per row)
Data sensitivity tier	Non-personal | Personal | Sensitive | Special category (biometrics, health, financial)
Cross-system data linkage	Single source (no linkage) | Links data across multiple systems
Data availability status	Currently available and used | Available but not yet used | Precondition that would need to be established
Key constraints	Short text

CONCLUSION:

From taxonomy to practice

This taxonomy is not an end in itself – it is rather a tool for design and application. It provides a structured lens for understanding the diverse landscape of AI technologies, capabilities, and application areas relevant to social protection. By unpacking AI into its constituent capabilities rather than treating it as a single technology, this report aims to replace the ambiguity surrounding AI with a shared vocabulary that practitioners, managers, and policymakers can use to ask better questions and make more deliberate choices. In doing so, it can foster **informed discussions about where, how, and why AI can create value within social protection systems.**

The taxonomy should not be viewed as a blueprint for implementing AI, nor as a ranking of technologies or use cases. Rather, it is a practical tool for speaking a common language, identifying opportunities (e.g. against social protection applications in Section 1.3), assessing feasibility, and facilitating dialogue between policymakers, practitioners, technical experts, and service providers. Understanding the relationship between AI capabilities and social protection functions helps organisations move beyond technology-driven approaches and instead focus on solving concrete operational and policy challenges.

In practice, the adoption of AI does not start with choosing a technology, it starts by clearly defining the problem to be addressed. Whether the objective is improving service delivery, enhancing administrative efficiency, detecting non-take up of benefits, strengthening fraud detection, supporting case management, or generating policy insights, organisations must first understand the desired outcome and then identify which AI capabilities are most appropriate. The taxonomy can support this process by helping stakeholders map organisational needs to specific AI capabilities and application areas. Social protection institutions are best served by identifying a specific operational challenge; eligibility verification, detecting unmet needs or fraud detection, complaints/appeal triage, demand forecasting, case management – and using the taxonomy as a tool to identify appropriate technical approaches, and necessary considerations.

The rapid evolution of AI technologies means that both capabilities and applications will continue to expand. New models, tools, and approaches are likely to emerge, while existing technologies become more accessible and integrated into everyday workflows. For exactly these reasons, this taxonomy should be understood as a living framework; a foundation for continuous learning and adaptation rather than a fixed classification.

The value of AI in social protection systems is not determined by the sophistication of the technology, but by its capacity to help institutions improve service delivery to people. By providing a common language and a structured understanding of AI capabilities and applications, this taxonomy seeks to support evidence-based decision-making and responsible innovation, enabling organisations to move from exploration and experimentation toward meaningful and sustainable impact.

Next to this Taxonomy, the AI Hub for Social Protection of the DCI has developed a broader set of tools to support the process of AI adoption for social protection. The portfolio includes a report on Data Sovereignty Considerations, a Risk Assessment Framework and accompanying toolkit, and a Global Evidence Review of AI adoption from 2020–2025, all of which are available on spdci.org/ai-hub.

IMPRINT

Published by
Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Registered offices
Bonn and Eschborn

Friedrich-Ebert-Allee 32 + 36
53113 Bonn, Germany
T +49 228 44 60-0
F +49 228 44 60-17 66
info@giz.de

Dag-Hammarskjöld-Weg 1–5
65760 Eschborn, Germany
T +49 6196 79-0
F +49 6196 79-11 15

Required citation
DCI AI Hub for Social Protection, *A Taxonomy for AI in Social Protection: Defining Capabilities and Application Areas*, Deutsche Gesellschaft für Internationale Zusammenarbeit, Germany, 2026.

Authors
Thomas Byrnes
Avril James
Valentina Barca

Responsible for the content
Bernd Schramm, GIZ

Editing
Susan Sellars

Layout
EYES-OPEN, Berlin

Bonn, August 2026

AI Hub for Social Protection

The AI Hub is a global partnership, with close cooperation among nine core implementing partners: German Social Accident Insurance Institution for the Construction Sector (BG BAU), Government of Canada – Employment and Social Development Canada (ESDC), Dataprev (the Social Security Information and Technology Enterprise of the Brazilian government), Federation of German Pension Insurance Institutions (DRV Bund), Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), International Labour Organization (ILO), International Social Security Association (ISSA), Organisation for Economic Co-operation and Development (OECD) and the World Bank.

The AI Hub is implemented by:

BG BAU Berufsgenossenschaft der Bauwirtschaft

Government of Canada / Gouvernement du Canada

DATAPREV

Deutsche Rentenversicherung

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

issa

OECD

THE WORLD BANK

International Labour Organization