

Regional Workshop on Responsible AI Governance in the Arab States

Human rights by design

Preventive safeguards for AI governance

Akram Khalifa
Middle East and North Africa, OHCHR
Manama, Bahrain | 16 September 2026

The Governance challenge

AI is moving faster than the institutions designed to evaluate, regulate and remedy its impacts.

>1 billion
people use
conversational AI
weekly

75% / 15%
top AI compute
concentrated in the
US / China

4–7 months
doubling of task
length some AI agents
can complete

Opportunities

- Health, education, public services and knowledge access
- Scientific discovery, accessibility and social goods
- Better local services when adapted to real needs

Risks

- Discrimination and exclusion at scale
- Surveillance, data extraction and cyber abuse
- Disinformation, deepfakes and erosion of public trust

Evidence dilemma

- Waiting for complete evidence may be too late
- Acting without evidence can be ineffective
- Governance must be precautionary and adaptive

OHCHR Framing

No separate “AI rights” are needed. Existing human rights obligations already apply.

International human rights law is the universal, legally grounded framework for AI governance.

The task is translation

- From abstract obligations to design choices
- From ethics principles to procurement rules
- From commitments to enforceable safeguards
- From accountability after harm to prevention before harm

What rights are often engaged

- Privacy and data protection
- Equality and non-discrimination
- Expression, opinion and access to information
- Due process, justice and effective remedy
- Work, health, education and social security

Message 1: Human rights law is the baseline

Public authorities cannot outsource their human rights duties to technology vendors.

State duties

- Respect: do not deploy AI that unlawfully restricts rights
- Protect: regulate private developers and deployers
- Fulfil: use AI to expand access without exclusion

Business responsibility

- Respect human rights under the UNGPs
- Conduct human rights due diligence
- Prevent and mitigate adverse impacts
- Provide or cooperate in remedy

Rule-of-law tests

- Legality and legitimate aim
- Necessity and proportionality
- Reasoned decisions and review
- Non-discrimination and equality in practice

Voluntary ethics principles may help, but they do not replace enforceable duties, oversight and remedies.



Arab-region lens

Language, representation and context determine whether AI inclusion is real or only formal.

Arabic is not one test case

- Modern Standard Arabic and dialects vary widely
- Systems may perform unevenly across countries and communities
- Arabic data quality, moderation and evaluation need local expertise

Data gaps become rights gaps

- Under-documented people may be absent from training data
- Refugees, migrants and persons with disabilities face compounding barriers
- Poor representation can undermine fair access to services

Who bears the risk?

- Women and children exposed to deepfakes and abuse
- Workers affected by automation and platform management
- Communities affected by data centres, energy and water use

Equality requires asking whose language, lived experience and risks are understood by the system.

Message 2: Prevention across the lifecycle

A human rights-based approach (HRBA) starts before approval and continues for as long as the system is used.



Key shift

- Do not wait until a person loses a benefit, job, service or liberty
- Risk assessment must consider less intrusive alternatives
- Rights-by-design must accompany secure-by-design

Operational test

- Who can stop the system?
- Who can explain the decision?
- Who can reverse the outcome?
- Who is accountable when harm occurs?

Human rights impact assessment (HRAI) is not a box-ticking exercise

It must be independent where appropriate, public to the greatest extent possible, and repeated when systems change.

Minimum content of an HRIA

- Intended purpose and affected rights
- Affected communities and differentiated impacts
- Foreseeable misuse and cumulative effects
- Error rates across demographic groups
- Consequences of failure and mitigation plan

Beyond technical accuracy

- Gender, child-rights and disability impacts
- Labour impacts and worker autonomy
- Environmental costs: water, energy, minerals and waste
- Supply-chain and data-labelling labour conditions
- Local language and context testing

Evidence for oversight

- Disaggregated testing results
- Documented data provenance and limitations
- Independent audit access
- Public summary and consultation record
- Triggers for suspension or redesign

Meaningful human oversight

A human merely clicking “approve” is not human oversight.

Oversight means competence, information, time and actual authority to question or reverse an AI-supported outcome.

Competence

- Trained officials understand system limits
- Staff know when to override or escalate
- Judges, regulators and NHRIs can interrogate evidence

Authority

- Clear legal responsibility
- Power to suspend, modify or withdraw systems
- No “black box” procurement that blocks inspection

Autonomy risk

- AI agents plan and act with limited oversight
- Control methods are still incomplete
- Greater autonomy requires stronger traceability and limits

Message 3: National law needs enforceable safeguards

Strategies set direction; regulation gives people rights, duties and remedies.

1. Technology-neutral scope across the AI value chain
2. Risk-based duties based on harm to people
3. Clear red lines and moratoria for rights-violating uses
4. Mandatory HRDD and HRIAs for high-risk systems
5. Transparency: registers, notices, explanations and public documentation
6. Independent oversight with expertise and access to information
7. Accessible complaints, human review and effective reparation
8. Procurement clauses: audit access, incident reporting and termination

Core idea: responsibility should not disappear between agencies, developers, vendors and subcontractors.

Procurement, remedy and accountability

The most practical safeguard may be the contract governments sign before deployment.

Procurement conditions

- Human rights compliance as a condition of purchase
- Audit and inspection rights
- Documentation of data, limits and model changes
- Incident reporting and cooperation with oversight bodies
- Suspension or termination when serious risks persist

Remedy by design

- Notice when AI affects a consequential outcome
- Understandable explanation of the basis of decision
- Prompt human review with authority to reverse
- Effective reparation when harm occurs
- Liability rules that do not burden victims with complex value chains

Public-sector test

- Do not buy what cannot be evaluated
- Do not deploy what cannot be explained
- Do not scale what cannot be independently audited
- Do not continue what cannot be made rights-compliant

Institutions worthy of trust

Capacity is a governance requirement, not an optional add-on.

Capacity gaps

- Many States lack compute, evaluation expertise and technical capacity
- Access to an AI tool is not the same as capacity to govern it
- Local evidence, datasets and expertise are essential

Institutional ecosystem

- Data protection authorities
- Sector regulators
- Courts and prosecutors
- Public administrations and procurement bodies
- Academia, workers and civil society

NHRIs as bridges

- Advise lawmakers and review strategies
- Scrutinize high-risk deployments and procurement
- Receive complaints and conduct inquiries
- Engage affected groups and connect to international mechanisms

Participation cannot be symbolic: those most likely to bear AI risks must help define the rules and test whether safeguards work.

A Pan-Arab regulatory toolkit: practical added value

Regional cooperation can pool scarce expertise while respecting national contexts.

Common minimum safeguards

- Shared rights-based definitions and risk tiers
- Model HRIA and procurement clauses
- Minimum transparency and public-register standards
- Incident reporting and remedy templates

Regional evidence base

- Arabic and dialect evaluation benchmarks
- Disaggregated performance testing
- Repository of public-sector AI use cases
- Shared lessons from complaints and audits

Implementation support

- Guidance for lawmakers and regulators
- Capacity-building for NHRIs and public institutions
- Civil-society participation channels
- Cross-border cooperation on high-risk systems

A useful toolkit should help governments ask the right questions at every stage — and give NHRIs and affected people means to enforce the answers.



UNITED NATIONS
HUMAN RIGHTS
OFFICE OF THE HIGH COMMISSIONER

Thank You