HARM BLINDNESS FRAMEWORK

ENFORCEMENT TEMPLATES FOR POLICY IMPLEMENTATION

Version: 2.0
Release Date: November 19, 2025
Created by: Real Safety AI Foundation (Travis Gilly, Executive Director)
Contact: t.gilly@ai-literacy-labs.org
Website: realsafetyai.org

Legal Disclaimer: These templates are provided for informational purposes. Organizations should review with appropriate legal counsel before implementation. The author and Real Safety AI Foundation provide no legal representation and assume no liability for use of these templates.

Collaboration Welcome: Modifications require collaborative involvement with author to maintain systematic rigor. Contact above email to discuss applications, adaptations, or integrations.

ABOUT THESE TEMPLATES

These templates provide ready-to-use policy language for mandating systematic stakeholder analysis in various contexts. Policy makers, regulators, and institutional leaders can adapt this language for:

- Government procurement requirements
- Regulatory compliance frameworks
- Industry certification standards
- Institutional governance policies
- Professional licensing requirements
- International development standards

Each template is structured to be:

- Concrete and enforceable (not abstract principles)
- Auditable (clear documentation requirements)
- Scalable (works for organizations of different sizes)

- Legally defensible (based on documented methodology)

IMPORTANT: These are PROPOSED/MODEL templates for potential adoption. No jurisdiction currently mandates these requirements. Policymakers and institutional leaders can adapt this language when creating:

- Government procurement requirements
- Regulatory compliance frameworks
- Industry certification standards
- Institutional governance policies

TEMPLATE 1: FOR Government Procurement Officers - Model RFP Language

For Federal/State/Local Government AI Procurement

Short Form (for RFPs):

STAKEHOLDER ANALYSIS REQUIREMENT

All AI systems [purchased by / deployed by / sold to] [Agency Name] must demonstrate completion of systematic stakeholder analysis during development.

Vendors must provide documentation showing:

1. **Checkpoint 1 (Ideation):** Identification of all affected stakeholder groups beyond direct users
2. **Checkpoint 2 (Design):** Analysis of incentive structures and potential harms
3. **Checkpoint 3 (Testing):** Documentation of testing coverage and gaps
4. **Checkpoint 4 (Launch):** Complete stakeholder impact assessment with executive sign-off

Documentation must be produced for review prior to contract award and updated annually for duration of contract.

Non-compliance will result in contract ineligibility.

Long Form (for Detailed Procurement Standards):

COMPREHENSIVE STAKEHOLDER ANALYSIS MANDATE

1. REQUIREMENT SCOPE

This requirement applies to all technology systems that:

- a) Process data about individuals
- b) Make or influence decisions affecting people
- c) Automate workflows previously performed by humans
- d) Scale to affect more than 1,000 individuals

2. MANDATORY CHECKPOINTS

Vendors must demonstrate completion of four stakeholder analysis checkpoints during development:

CHECKPOINT 1 - IDEATION PHASE Required Documentation:

- Complete list of stakeholder groups affected (direct and indirect)
- Assessment of impacts at scale (1000x current usage)
- Analysis of non-adopters and disadvantaged populations
- Initial risk identification
- Project owner sign-off acknowledging stakeholder identification

CHECKPOINT 2 - DESIGN PHASE Required Documentation:

- Technical architecture overview
- Incentive structure analysis (behaviors rewarded/punished)
- Identification of potential perverse incentives
- Mitigation strategies for identified risks
- Technical lead sign-off on design implications

CHECKPOINT 3 - TESTING PHASE Required Documentation:

- Testing methodology and population coverage
- Identified gaps in testing (who was NOT tested)
- Results from diverse user testing
- Accessibility compliance verification
- QA lead sign-off acknowledging limitations

CHECKPOINT 4 - LAUNCH PHASE Required Documentation:

- Complete stakeholder analysis table (benefits vs harms for each group)

- Net outcome assessment
- Long-term precedent consideration
- Public defense statement ("can we defend this publicly?")
- Executive sign-off accepting responsibility for impacts

3. DOCUMENTATION STANDARDS

All checkpoint documentation must include:

- Date of checkpoint completion
- Names and roles of participants
- Complete stakeholder analysis
- Identified risks with probability and impact ratings
- Mitigation strategies with assigned owners
- Final decision (proceed/modify/cancel) with reasoning
- Sign-off from designated decision-maker

Documentation must be:

- Searchable and accessible for audit
- Preserved for contract duration + 5 years
- Produced upon government request within 5 business days

4. AUDIT REQUIREMENTS

Government reserves right to:

- Audit documentation at any time
- Interview checkpoint participants
- Review methodology and quality of analysis
- Verify stakeholder representation was adequate
- Assess whether identified harms were properly mitigated

5. ENFORCEMENT

Non-compliance will result in:

- Contract ineligibility for current procurement

- Requirement to remediate before future eligibility
- Notification to other government entities
- Potential contract termination for existing agreements

Falsification of documentation will result in:

- Immediate contract termination
- Debarment from government contracts
- Referral for legal action

6. COMPLIANCE TIMELINE

- For new contracts: All checkpoints completed before contract execution
- For existing systems: Retroactive analysis due within 180 days
- For updates/changes: New checkpoints required for major modifications

7. DEATH GATE PROTOCOL REQUIREMENTS

Systems identified with death risk must:

- Immediately activate Death Gate Protocol (Framework Part 2)
- Complete Stage 1: Public warning labels on all interfaces
- Complete Stage 2: Regulatory authorization before deployment
- Complete Stage 3: Independent coalition validation
- Cannot proceed without completing all three stages
- Bypassing protocol triggers criminal liability for executives

Death risk includes:

- Direct causation of preventable death
- Suicide facilitation or encouragement
- Violence enabling capabilities
- Life-critical system failures

Legitimate exceptions (medical devices, emergency systems) must complete full three-stage approval before deployment.

8. MIT AI RISK REPOSITORY VERIFICATION

As supplementary verification, covered entities should review:

- All 7 primary domains from MIT AI Risk Repository
- 24 subdomains for comprehensive coverage
- Use as checklist to catch overlooked risk categories

Reference: Framework Appendix C

TEMPLATE 2: FOR Regulatory Agencies - Proposed Compliance Framework

STAKEHOLDER HARM PREVENTION REGULATION

[Regulation Number]: Systematic Stakeholder Analysis for [Industry Sector]

1. REGULATORY AUTHORITY

Under authority of [Enabling Statute], [Agency Name] requires systematic stakeholder analysis for all [regulated products/services] that [trigger conditions: process personal data / affect public health / influence financial decisions / etc.].

2. COVERED ENTITIES

This regulation applies to:

- Companies developing [AI systems / automated decision-making / etc.]
- Organizations deploying systems affecting more than [threshold] people
- Any entity seeking [regulatory approval / market authorization]

3. MANDATORY ANALYSIS FRAMEWORK

Covered entities must implement stakeholder analysis at four development checkpoints:

- a) Before development begins (Ideation)
- b) Before implementation starts (Design)
- c) Before preparation for launch (Testing)
- d) Before deployment to market (Launch)

At each checkpoint, entities must:

- Identify ALL affected stakeholder groups
- Assess potential benefits and harms for each group

- Prioritize risks by probability and impact
- Define mitigation strategies with assigned responsibility
- Obtain executive sign-off on decisions

4. DOCUMENTATION REQUIREMENTS

Entities must maintain checkpoint documentation including:

- Stakeholder identification and analysis
- Risk assessments with mitigation plans
- Decision rationales
- Executive accountability (signatures)
- Follow-up actions and monitoring plans

Documentation must be preserved for [10 years / product lifetime] and produced upon regulatory request.

5. AUDIT AND INSPECTION

[Agency] may:

- Request documentation at any time
- Conduct on-site inspections
- Interview personnel involved in checkpoints
- Assess quality and completeness of analysis
- Verify implementation of mitigation strategies

6. ENFORCEMENT ACTIONS

Failure to comply will result in:

- Warning letter with 90-day remediation period
- Fines: $[amount] per violation per day
- Product recall or market withdrawal
- Injunction against future violations
- Criminal referral for willful violations

7. SAFE HARBOR

Entities demonstrating good faith compliance with this regulation:

- Are presumed to have exercised reasonable care
- Have affirmative defense in private lawsuits
- May qualify for reduced penalties in enforcement actions

8. EFFECTIVE DATE

- This regulation takes effect [date]
- Existing products must be brought into compliance within [180 days]
- New products must comply before market entry

9. GUIDANCE DOCUMENTS

[Agency] will publish additional guidance on:

- Checkpoint implementation best practices
- Documentation standards and templates
- Industry-specific considerations
- Small business compliance alternatives

10. DEATH GATE PROTOCOL REQUIREMENTS

Systems identified with death risk must:

- Immediately activate Death Gate Protocol (Framework Part 2)
- Complete Stage 1: Public warning labels on all interfaces
- Complete Stage 2: Regulatory authorization before deployment
- Complete Stage 3: Independent coalition validation
- Cannot proceed without completing all three stages
- Bypassing protocol triggers criminal liability for executives

Death risk includes:

- Direct causation of preventable death
- Suicide facilitation or encouragement
- Violence enabling capabilities
- Life-critical system failures

Legitimate exceptions (medical devices, emergency systems) must complete full three-stage approval before deployment.

11. MIT AI RISK REPOSITORY VERIFICATION

As supplementary verification, covered entities should review:

- All 7 primary domains from MIT AI Risk Repository
- 24 subdomains for comprehensive coverage
- Use as checklist to catch overlooked risk categories

Reference: Framework Appendix C

TEMPLATE 3: FOR Professional Associations - Sample Certification Standard

ETHICAL AI CERTIFICATION STANDARD

Certification: [Name] Ethical AI Development Certification

1. CERTIFICATION PURPOSE

This certification verifies that organizations implement systematic stakeholder analysis to prevent harm during AI development.

2. ELIGIBILITY

Organizations developing or deploying AI systems that:

- Make decisions affecting individuals
- Process personal data
- Automate human workflows
- Impact stakeholder wellbeing

3. CERTIFICATION REQUIREMENTS

Organizations must demonstrate:

PROCESS IMPLEMENTATION

- Designated checkpoint facilitators (trained)
- Integration into development workflows
- Stakeholder representation mechanisms
- Executive accountability structure

- Documentation systems

CHECKPOINT COMPLETION

- Evidence of all four checkpoints for sample projects
- Quality documentation meeting standards
- Stakeholder representation in analysis
- Appropriate risk mitigation
- Executive sign-offs on decisions

ORGANIZATIONAL CULTURE

- Training programs for staff
- Psychological safety to raise concerns
- Authority to delay/cancel based on checkpoints
- Success metrics and monitoring
- Continuous improvement processes

4. CERTIFICATION PROCESS

- Step 1: Application and self-assessment
- Step 2: Documentation review (sample of 3-5 projects)
- Step 3: Interviews with staff and facilitators
- Step 4: Stakeholder feedback verification
- Step 5: Audit of processes and culture
- Step 6: Certification decision

5. MAINTENANCE

Certification valid for [2 years] with:

- Annual checkpoint documentation submission
- Bi-annual process audits
- Ongoing monitoring of public incidents
- Stakeholder complaint review

6. PUBLIC DISCLOSURE

Certified organizations may:

- Use certification mark in marketing
- List on public registry
- Reference in procurement bids
- Display on website and materials

7. REVOCATION

Certification may be revoked for:

- Falsified documentation
- Significant harm to stakeholders
- Non-compliance with requirements
- Failure to remediate identified issues

Revocation will be:

- Publicly announced
- Published on registry
- Reported to relevant regulators

8. RECERTIFICATION

Following revocation, organizations may apply for recertification after:

- Minimum [1 year] waiting period
- Demonstration of remediation
- Additional auditing and oversight

9. DEATH GATE PROTOCOL REQUIREMENTS

Systems identified with death risk must:

- Immediately activate Death Gate Protocol (Framework Part 2)
- Complete Stage 1: Public warning labels on all interfaces
- Complete Stage 2: Regulatory authorization before deployment
- Complete Stage 3: Independent coalition validation
- Cannot proceed without completing all three stages

- Bypassing protocol triggers criminal liability for executives

Death risk includes:

- Direct causation of preventable death
- Suicide facilitation or encouragement
- Violence enabling capabilities
- Life-critical system failures

Legitimate exceptions (medical devices, emergency systems) must complete full three-stage approval before deployment.

10. MIT AI RISK REPOSITORY VERIFICATION

As supplementary verification, covered entities should review:

- All 7 primary domains from MIT AI Risk Repository
- 24 subdomains for comprehensive coverage
- Use as checklist to catch overlooked risk categories

Reference: Framework Appendix C

TEMPLATE 4: FOR Institutional Leaders - Model Governance Policy

ORGANIZATIONAL POLICY: STAKEHOLDER HARM PREVENTION

Policy Number: [XXX]
Effective Date: [Date]
Review Frequency: Annual
Policy Owner: [Title]

1. PURPOSE

This policy establishes mandatory stakeholder analysis for [organization name] projects and initiatives that affect external stakeholders or deploy technology systems.

2. SCOPE

This policy applies to:

- All technology development projects
- Product launches and major feature releases

- Policy changes affecting stakeholders
- Research that impacts communities
- Partnerships involving data or automation

3. REQUIREMENTS

All covered projects must complete four stakeholder analysis checkpoints:

CHECKPOINT 1 - PROJECT KICKOFF

- Before: Development, implementation, or deployment begins
- Required: Stakeholder identification, scale analysis, initial risks
- Sign-off: Project owner

CHECKPOINT 2 - DESIGN APPROVAL

- Before: Implementation or development work starts
- Required: Incentive analysis, harm mitigation strategies
- Sign-off: Technical lead

CHECKPOINT 3 - PRE-LAUNCH REVIEW

- Before: Launch preparation begins
- Required: Testing gaps, stakeholder representation validation
- Sign-off: QA or research lead

CHECKPOINT 4 - LAUNCH AUTHORIZATION

- Before: Public deployment or implementation
- Required: Complete stakeholder analysis, net outcome assessment
- Sign-off: Executive sponsor

4. ROLES AND RESPONSIBILITIES

CHECKPOINT FACILITATORS

- Designated staff trained in framework methodology
- Authority to pause projects for analysis
- Not members of project teams (independence)
- Report to [governance body]

PROJECT OWNERS

- Ultimately accountable for stakeholder analysis
- Must address identified harms or document risk acceptance
- Cannot proceed without checkpoint completion
- Subject to accountability measures

EXECUTIVE SPONSORS

- Must sign off on Checkpoint 4 before launch
- Accept responsibility for stakeholder impacts
- Can be held accountable for ignored warnings

5. DOCUMENTATION

All checkpoints must produce:

- Stakeholder analysis tables
- Risk assessments with ratings
- Mitigation strategies with owners
- Decision rationales
- Executive sign-offs

Documentation must be:

- Stored in [designated system]
- Accessible for audit
- Retained for [duration]

6. MONITORING AND COMPLIANCE

[Governance body] will:

- Track checkpoint completion rates
- Review documentation quality
- Monitor post-launch outcomes
- Report annually to [leadership]

Non-compliance will result in:

- Project suspension until compliance
- Escalation to executive leadership
- Performance impact for responsible parties
- Policy violation consequences per HR

7. EXCEPTIONS

Exceptions require:

- Written justification
- [Level] executive approval
- Enhanced documentation of risk acceptance
- Increased post-launch monitoring

8. TRAINING

Required training:

- All project owners: 4-hour framework training
- Checkpoint facilitators: 16-hour facilitator certification
- Executive sponsors: 2-hour briefing on accountability

9. REVIEW AND UPDATES

This policy will be reviewed annually and updated based on:

- Effectiveness metrics
- Lessons learned
- Industry best practices
- Regulatory changes

10. DEATH GATE PROTOCOL REQUIREMENTS

Systems identified with death risk must:

- Immediately activate Death Gate Protocol (Framework Part 2)
- Complete Stage 1: Public warning labels on all interfaces
- Complete Stage 2: Regulatory authorization before deployment
- Complete Stage 3: Independent coalition validation

- Cannot proceed without completing all three stages
- Bypassing protocol triggers criminal liability for executives

Death risk includes:

- Direct causation of preventable death
- Suicide facilitation or encouragement
- Violence enabling capabilities
- Life-critical system failures

Legitimate exceptions (medical devices, emergency systems) must complete full three-stage approval before deployment.

11. MIT AI RISK REPOSITORY VERIFICATION

As supplementary verification, covered entities should review:

- All 7 primary domains from MIT AI Risk Repository
- 24 subdomains for comprehensive coverage
- Use as checklist to catch overlooked risk categories

Reference: Framework Appendix C

TEMPLATE 5: FOR Development Organizations - Proposed Technology Standard

DEVELOPMENT AID TECHNOLOGY DEPLOYMENT STANDARD

Standard: Stakeholder Harm Prevention for Technology Interventions

1. STANDARD PURPOSE

Technology deployed through development aid must undergo systematic stakeholder analysis to prevent harm to vulnerable populations.

2. APPLICABILITY

This standard applies to:

- Digital identity systems
- Financial technology (mobile money, digital payments)
- Agricultural technology
- Health technology

- Educational technology
- Government service digitization
- Data collection and management systems

3. CULTURAL CONTEXT REQUIREMENTS

Standard must be adapted for local context:

- Stakeholder identification includes traditional and modern structures
- Power dynamics reflect local realities
- Consent mechanisms respect cultural norms
- Documentation in appropriate languages
- Local facilitators trained and empowered

4. ENHANCED CHECKPOINT REQUIREMENTS

Given vulnerable populations, checkpoints must include:

CHECKPOINT 1 - COMMUNITY CONSULTATION

- Extensive stakeholder mapping beyond direct users
- Understanding of power structures and marginalized groups
- Assessment of gender dynamics and impacts
- Analysis of digital divide and exclusion risks
- Free, prior, and informed consent processes

CHECKPOINT 2 - CULTURAL ADAPTATION

- Design tested with diverse community members
- Literacy, language, and accessibility considerations
- Infrastructure requirements (connectivity, power, devices)
- Traditional systems integration or displacement analysis
- Economic impacts on existing livelihoods

CHECKPOINT 3 - INCLUSIVE TESTING

- Testing with marginalized populations specifically
- Gender-disaggregated testing results

- Accessibility testing for disabilities
- Testing under realistic infrastructure conditions
- Community feedback integration

CHECKPOINT 4 - SUSTAINABILITY AND EXIT

- Long-term community capacity building
- Local ownership and governance
- Exit strategy that doesn't leave harm
- Ongoing monitoring and support commitments
- Mechanisms for community to raise concerns

5. STAKEHOLDER PROTECTION

Special protections for:

- Women and girls
- Ethnic and religious minorities
- People with disabilities
- Elderly populations
- Informal workers
- Rural and remote communities
- Refugees and displaced persons

6. DOCUMENTATION STANDARDS

Checkpoint documentation must:

- Be accessible to local communities (language, format)
- Include community representatives as co-authors
- Demonstrate genuine consultation, not extraction
- Show how community feedback shaped decisions
- Be shared with communities (transparency)

7. ENFORCEMENT

Technology deployments without proper stakeholder analysis:

- Will not receive development aid funding
- Must be remediated before deployment
- May trigger aid suspension for implementing organization
- Require external audit and community redress

8. ACCOUNTABILITY

Implementing organizations must:

- Establish community grievance mechanisms
- Conduct post-deployment harm monitoring
- Provide remedy for identified harms
- Report annually on stakeholder impacts
- Face evaluation in future funding decisions

9. PARTNERSHIP REQUIREMENTS

Local partners must:

- Participate in checkpoint facilitation
- Have authority to halt deployment
- Receive capacity building in framework
- Lead community engagement
- Co-own decisions and accountability

10. DEATH GATE PROTOCOL REQUIREMENTS

Systems identified with death risk must:

- Immediately activate Death Gate Protocol (Framework Part 2)
- Complete Stage 1: Public warning labels on all interfaces
- Complete Stage 2: Regulatory authorization before deployment
- Complete Stage 3: Independent coalition validation
- Cannot proceed without completing all three stages
- Bypassing protocol triggers criminal liability for executives

Death risk includes:

- Direct causation of preventable death
- Suicide facilitation or encouragement
- Violence enabling capabilities
- Life-critical system failures

Legitimate exceptions (medical devices, emergency systems) must complete full three-stage approval before deployment.

11. MIT AI RISK REPOSITORY VERIFICATION

As supplementary verification, covered entities should review:

- All 7 primary domains from MIT AI Risk Repository
- 24 subdomains for comprehensive coverage
- Use as checklist to catch overlooked risk categories

Reference: Framework Appendix C

IMPLEMENTATION GUIDANCE

Adapting Templates to Your Context

1. Identify your authority

- Regulatory: Citation of enabling statute
- Procurement: Contracting authority
- Institutional: Governance structure
- Professional: Certification or licensing power

2. Define scope clearly

- Which products/services covered
- Thresholds for applicability
- Exceptions if any

3. Specify enforcement

- Audit mechanisms
- Penalties for non-compliance

- Timeline for remediation
- Escalation procedures

4. Ensure auditability

- Documentation standards
- Retention requirements
- Production timelines
- Review procedures

5. Build in accountability

- Sign-off requirements
- Personal responsibility
- Consequences for violations
- Protection for those who raise concerns

Legal Review Considerations

Before implementing these templates:

- Review with legal counsel for your jurisdiction
- Ensure authority is clear and defensible
- Verify enforcement mechanisms are appropriate
- Confirm documentation requirements are achievable
- Check alignment with existing regulations

Pilot Program Approach

Recommended implementation:

1. Start with voluntary pilot (6 months)
2. Mandate for subset of projects/products
3. Evaluate effectiveness and refine
4. Expand to full mandate
5. Monitor and continuously improve

SUPPORT

For policy implementation questions:
Email: t.gilly@ai-literacy-labs.org

For legal review support:
Contact for referrals to specialized counsel

For training materials:
Available at realsafetyai.org

Legal Disclaimer: These templates are provided for informational purposes. Organizations should review with appropriate legal counsel before implementation. The author and Real Safety AI Foundation provide no legal representation and assume no liability for use of these templates.

Collaboration Welcome: Modifications require collaborative involvement with author to maintain systematic rigor. Contact above email to discuss applications, adaptations, or integrations.

Version History:

- Version 1.0: Initial templates (November 10, 2025)
- Version 2.0: Added Death Gate Protocol requirements, MIT AI Risk Repository verification, clarified proposed/model status (November 19, 2025)