



PHARMA-AI Ethical Guardrail Builder

Documentation & User Guide

Release Date: August 2026 | Application Version: 1.0.1

Platform: Universal HTML5 / JavaScript - Client-Side Decision Support

Central principle

The learning outcome should determine the role of AI. The application helps educators make that reasoning visible, consistent and documentable; it does not replace institutional authority or professional judgment.

1. Introduction

The PHARMA-AI Ethical Guardrail Builder is a browser-based decision-support application developed for pharmacy educators, preceptors, assessment leaders and academic governance teams. It converts ethical principles into a structured review of a proposed educational use of artificial intelligence and generates practical, editable policy language.

Unlike a static checklist, the application recalculates its findings whenever the user changes the educational setting, learning outcome, stakes, AI function, data classification or safeguard. This allows users to test alternative designs and understand why a recommendation changes.

Key features

1. **Dynamic PHARMA risk engine:** Assesses Privacy, Human judgment, Academic integrity, Reliability, Maintenance of competence, and Access and equity.
2. **Quick and Full review modes:** Supports an 11-minute workshop activity as well as detailed institutional review.
3. **Critical override rules:** Prevents serious incompatibilities from being hidden by a reassuring average score.
4. **Live decision trace:** Explains which inputs raised or reduced each ethical-risk dimension.
5. **Scenario exploration:** Includes 12 preloaded pharmacy-education scenarios and side-by-side comparison.
6. **Policy Studio:** Produces four editable outputs for faculty, students and experiential settings.
7. **Local profiles and outcomes:** Saves work in the browser and supports portable JSON backup and restoration.
8. **Professional reporting:** Provides copyable language, A4 print layouts and browser-based PDF saving.

Professional review required

The Guardrail Index is a structured decision-support indicator, not a psychometrically validated risk instrument. The application does not provide legal, regulatory, accreditation or institutional approval.



2. The PHARMA Ethical AI Lens

The application organizes ethical review around six pharmacy-specific domains. Each domain is scored separately so that a low concern in one area cannot obscure a major concern in another.

Letter	Domain	Guiding question
P	Privacy and confidentiality	Are patient, student, research or institutional data being exposed?
H	Human judgment and accountability	Who remains responsible when the AI output is wrong, unsafe or misleading?
A	Academic integrity and attribution	Is AI use permitted, transparent and appropriately acknowledged?
R	Reliability, bias and verification	Could the output be inaccurate, fabricated, outdated, incomplete or biased?
M	Maintenance of competence	Is AI supporting development, or performing the learner's thinking?
A	Access and equity	Can all learners participate fairly, safely and accessibly?

The three cognitive risks

1. **Deskilling:** A previously developed ability weakens through repeated dependence on AI.
2. **Never-skilling:** Learners outsource cognitive work before the underlying competence is developed.
3. **Mis-skilling:** Learners internalize and repeat inaccurate or biased AI-generated reasoning.

Recommended sequence

Human first -> AI second -> Human verification

3. Decision Methodology

The Guardrail Index

Each response contributes to one or more PHARMA dimensions. The engine combines baseline risk, contextual modifiers, selected AI functions and safeguards, then normalizes each dimension and the overall Guardrail Index to a 0-100 scale.

The overall score is intentionally influenced by both the average risk and the highest individual PHARMA risk. Critical override rules can set a minimum decision level when an incompatibility should not be averaged away.

Index	Decision category	Operational meaning
0-19	Suitable for purposeful integration	The proposed use aligns with the learning purpose under standard safeguards.
20-34	Permitted with documented safeguards	Use is reasonable if the specified controls are implemented and monitored.
35-54	Restricted use only	Only selected AI functions should be permitted under documented limitations.
55-74	Redesign and approval required	Major risks remain and require redesign plus appropriate institutional review.
75-100 or critical override	Do not proceed in the proposed form	A critical incompatibility must be removed before proceeding.

Suggested AI involvement levels

Level	Label	Interpretation
Level 0	No AI	AI use is incompatible with the proposed activity in its current form.



Level	Label	Interpretation
Level 1	Mechanical support only	Limit use to low-cognition administrative or language functions.
Level 2	Limited assistive use	Permit defined functions while protecting substantive independent work.
Level 3	Declared and supervised collaboration	Use AI transparently with human-first work and verification.
Level 4	Purposefully AI-integrated learning	AI use is itself part of the learning outcome and is intentionally taught.

Examples of critical override rules

1. Identifiable patient or sensitive information paired with a public or unevaluated AI environment.
2. Patient-facing or progression decisions without mandatory qualified human review.
3. AI performing the same substantive competence a high-stakes assessment is intended to judge, without independent evidence of performance.
4. Required paid AI use without an equivalent alternative, triggering an equity redesign requirement.

4. Getting Started

Access and browser requirements

1. Open the live application in a current version of Chrome, Edge, Safari or Firefox.
2. Use a laptop or tablet for the most detailed review. Mobile phones are fully supported for Quick Mode and scenario exploration.
3. Do not enter patient-identifiable, confidential, unpublished or otherwise sensitive information.
4. Choose Quick Mode for workshop use or Full Review for institutional documentation.

Home screen

1. **Start a quick review:** Opens the five-step Builder with only high-impact decisions visible.
2. **Open full institutional review:** Adds activity identity, purpose, review date and extended safeguards.
3. **Explore scenarios:** Opens preconfigured cases that demonstrate contextual differences.
4. **Saved work:** Displays locally stored assessments and institutional profiles.

Local data notice

Saved work belongs to the current browser and website origin. Export a JSON backup before changing devices, clearing browser data or moving the application to another domain.

5. Quick Workshop Mode

Quick Mode is designed for rapid group use during the workshop. It retains the complete decision engine while reducing the visible inputs to the factors most likely to change the ethical recommendation.

Step	Focus	What the user decides
1	Purpose and setting	Select the educational setting, stakes and intended learning outcomes.
2	Proposed role of AI	Identify every function AI will perform.
3	Data and environment	Classify the information involved and the type of AI environment.
4	Human judgment and competence	Specify human review, human-first reasoning, independent demonstration and verification.
5	Integrity, disclosure and equity	Clarify AI-use rules, disclosure and equivalent access.



The live panel recalculates after every selection. On mobile devices, tap the floating result button to open the same information in a touch-friendly drawer.

6. Full Institutional Review

Full Review preserves the same five-stage structure but adds fields needed for governance, record keeping and later reassessment.

1. **Activity identity:** Assessment title, institution, course or rotation, faculty owner and review date.
2. **Educational purpose:** A brief activity description that must exclude names or sensitive information.
3. **Required use:** Whether the AI tool is mandatory, optional or not yet decided.
4. **Oral defence:** Whether learners must explain and defend the submitted work.
5. **Supporting controls:** Prompt retention, user training, accessibility arrangements and expert review.

Good review practice

1. Assess the activity design, not an individual student's conduct.
2. Select the most sensitive information that could realistically be involved.
3. Describe what AI will do, not merely the name of the tool.
4. Treat 'not confirmed' as an unresolved control requiring follow-up.
5. Save a completed assessment only after reviewing the generated decision trace.

7. Understanding Live Results and the Decision Trace

Live results panel

1. **Decision category:** The current operational recommendation.
2. **Guardrail Index:** A 0-100 comparison indicator based on the current responses.
3. **Suggested involvement:** The recommended Level 0-4 role for AI.
4. **PHARMA bars:** Separate scores for all six ethical domains.
5. **Priority alert:** The most important hard stop or unresolved safeguard.
6. **What changed?:** An explanation of how the most recent input changed the risk profile.

Full Results view

The Results view expands the live summary into four operational components:

1. Executive decision and suggested AI involvement.
2. Dimension-by-dimension PHARMA profile with primary reasons.
3. Critical findings and required safeguards.
4. Transparent decision trace listing the strongest activated risk and control rules.

Interpretation rule

A lower overall score does not authorize use when a critical incompatibility remains. Read the critical findings and required safeguards before relying on the category label.

8. Scenario Library and Comparison

The Scenario Library contains 12 examples spanning didactic, assessment, experiential, research and faculty-development settings. Each scenario can be loaded into the Builder and modified.

Scenario	Setting	Primary issue demonstrated
----------	---------	----------------------------



Scenario	Setting	Primary issue demonstrated
Real patient case in a public chatbot	Experiential	Confidentiality and patient-data exposure
Synthetic case with supervised AI critique	Formative	Purposeful supervised integration
AI-assisted pharmaceutical care plan	Summative	Integrity and competence replacement
Forty overnight AI-generated MCQs	Faculty	Unvalidated keys, references and bias
Biased patient-case generation	Didactic	Representational bias
Confidential manuscript in public AI	Research	Confidentiality and intellectual property
Required premium AI platform	Graded	Access and equity
Grammar refinement after a first draft	Summative	Limited assistive use
Formative drug-information exercise	Formative	Evidence verification
AI patient-counselling simulation	Simulation	Supervised communication practice
Faculty-reviewed OSCE case generation	Faculty	Expert validation
Reflective writing outsourced to AI	Experiential	Authenticity and reflective competence

Using Compare

1. Open Compare from the navigation or a scenario card.
2. Choose two preloaded or saved assessments.
3. Review differences in decision category, overall index, data environment and all PHARMA dimensions.
4. Identify the input or safeguard responsible for the difference.



9. Policy Studio

Policy Studio converts the current ethical decision into editable language using deterministic templates. It does not send prompts or information to an external AI service.

Generated output	Purpose
Faculty decision statement	Documents the decision, suggested AI involvement, principal concerns, safeguards and retained accountability.
Student-facing syllabus clause	States permitted and prohibited functions, required process, disclosure, privacy and consequences.
Experiential-rotation clause	Appears when the setting is experiential, simulated or patient-facing and emphasizes confidentiality and preceptor review.
Student AI-use disclosure statement	Provides a structured record of the tool, purpose, verification and learner responsibility.

Editing and copying

1. Select an output from the Policy Studio menu.
2. Edit the text directly. Changes are preserved in the current draft.
3. Use Copy this section for one output or Copy all outputs for a complete policy pack.
4. Use Regenerate from current answers to discard manual edits and rebuild the language.

Important

Generated language remains a draft. Review terminology, institutional policy references, misconduct procedures and local data-protection requirements before adoption.

10. Saving Profiles, Outcomes and Backups

Saved assessments

Save outcome stores the current responses, PHARMA scores, decision, generated language, application version and methodology version in browser storage. Saved assessments can be opened, duplicated or deleted.

Institutional profiles

An institutional profile stores reusable defaults such as institution, course or program, faculty owner, approved AI environment, disclosure expectations, equivalent-access arrangements and supporting controls.

Backup and restoration

1. Open Saved work.
2. Choose Export backup to download a JSON file containing profiles, assessments and the current draft.
3. Store the file in an institutionally appropriate location.
4. Choose Import backup on another browser or device. Importing replaces the saved work in that browser after confirmation.



Do not use browser storage as a clinical repository

Browser storage is convenient local persistence, not an approved location for patient-identifiable, confidential or sensitive information.

11. Copying, Printing and PDF Reports

Available report formats

1. Executive decision and PHARMA risk profile.
2. Detailed ethical audit with critical findings, safeguards and methodology.
3. Policy and syllabus pack containing the applicable generated statements.

Print or save as PDF

1. Open Results or Policy Studio.
2. Select Print / Save PDF or Print policy pack.
3. Choose A4 or Letter in the browser print dialog if required.
4. Select Save as PDF or a physical printer.
5. Review the PDF before distribution, particularly after extensive manual edits.

The print layout automatically hides navigation, touch controls and accessibility skip links. Reports remain readable in colour or grayscale and preserve genuine paragraph breaks in all generated statements.

Standalone report export

The application can also produce a self-contained HTML report. The exported report contains the current assessment output and can be opened or printed independently.

12. Mobile and Tablet Operation

The application is designed for touch use and adapts its information architecture rather than simply shrinking the desktop interface.

Device or orientation	Interface behaviour
Mobile portrait	Single-column Builder, bottom navigation and full-height live-results drawer.
Mobile landscape	Compact single-column or split view where screen width permits.
Tablet portrait	Questionnaire with collapsible live-results drawer.
Tablet landscape	Side-by-side questionnaire and live PHARMA results.
Desktop	Full two-panel workspace with persistent navigation and insights.

Touch and accessibility features

1. Large selection cards and touch targets of approximately 44-48 pixels.
2. Sticky Back, Next and Results actions during an assessment.
3. Readable body text and responsive headings without horizontal page scrolling.
4. Keyboard navigation, visible focus indicators and screen-reader labels.
5. Text, icons and labels used alongside colour for risk status.
6. Reduced-motion support and no essential information conveyed only through animation.



13. Privacy, Security and Data Handling

Privacy-by-design architecture

1. **No external processing:** All calculations and policy generation occur within the browser.
2. **No external AI API:** The application does not send assessment content to a language model.
3. **No analytics by default:** The self-contained application includes no tracking service.
4. **Local persistence only:** Drafts, profiles and outcomes use browser storage.
5. **Portable backup:** Users control transfer through explicit JSON export and import.
6. **Safe text rendering:** User-entered content is treated as text rather than executable HTML.

Data-entry rule

Describe the educational activity without entering names, medical-record numbers, dates of birth, case identifiers, unpublished results or any other information that could identify a patient, student, colleague or confidential project.

The first ethical guardrail

A governance tool does not need sensitive data to evaluate whether sensitive data would be involved.

14. Troubleshooting

Issue	Likely cause	Recommended action
The live app looks unchanged after an update	The browser may be using a cached copy.	Refresh the page. If needed, use a hard refresh and reopen the live link.
Saved work is missing on another device	Browser storage is local to a device, browser and domain.	Export a JSON backup from the original browser and import it on the new device.
Copy button does not work	Clipboard access may be restricted outside a secure connection.	Use the live HTTPS site, or select and copy the editable text manually.
Print dialog does not open	Pop-ups or printing may be blocked by the browser.	Allow printing, then retry from Results or Policy Studio.
PDF contains interactive controls	The page was printed without using the app's report action or from an old cached version.	Refresh the app and print through Print / Save PDF.
Literal backslash-n text appears	An older application version or old manually edited statement is open.	Refresh to version 1.0.1. Previously saved text is cleaned automatically when reopened.
Risk remains high after adding disclosure	Disclosure does not resolve privacy, safety or competence incompatibilities.	Review Critical findings and address the specific hard stop or required safeguard.
Imported backup is rejected	The file is not a valid PHARMA-AI backup or exceeds the size limit.	Use an unmodified JSON backup exported by the application and keep it below 5 MB.

15. Technical Specifications and Deployment

Specification	Implementation
Architecture	Single-page, self-contained browser application.
Core technologies	HTML5, embedded CSS3 and vanilla JavaScript.
Application size	Approximately 100 KB; no external runtime dependencies.
Processing	Deterministic client-side rules engine; no server calculation.
Persistence	Browser localStorage with JSON backup and restore.
Reporting	Clipboard API with fallback, Blob downloads and browser printing.
Responsive design	Mobile, tablet and desktop layouts with touch and keyboard support.
Hosting	Compatible with Cloudflare Pages or any static HTTPS hosting service.
Network behaviour	No application data transmission after page load.
Current version	1.0.1, August 2026.



Cloudflare Pages deployment

1. Save the self-contained application as index.html, or retain the supplied filename and route to it.
2. Upload the file through Cloudflare Pages Direct Upload or connect a Git repository.
3. No build command, server function, database or environment variable is required for the standalone edition.
4. Use HTTPS so clipboard operations and browser security controls function consistently.
5. After changing domains, remind users that saved browser data does not automatically move to the new origin.

Version 1.0.1 release notes

1. Removed the accessibility skip-link element from printed and PDF reports.
2. Converted statement separators to genuine paragraph breaks across all Policy Studio outputs.
3. Added automatic cleanup for previously saved statements containing literal backslash-n characters.

16. Methodological Limitations and Governance

The application structures educational judgment but cannot determine whether an activity complies with every institutional policy, law, regulator, accreditation standard or contractual tool term. Those requirements vary by jurisdiction, institution and time.

The scores are designed for consistent comparison and explanation, not for psychometric interpretation. Institutions adopting the tool should review the rule set, document any local modifications, identify accountable decision makers and schedule periodic review.

Recommended governance use

1. Use the application before approving an AI-enabled activity, not only after an incident.
2. Attach the report to course, assessment or experiential-education documentation where appropriate.
3. Record the responsible reviewer and the institutional profile used.
4. Reassess when the AI tool, data environment, assessment stakes or learning outcome changes.
5. Investigate critical findings through the relevant privacy, assessment, clinical or research governance route.

17. Glossary

1. **AI environment:** The technical and contractual setting in which an AI system processes information.
2. **Critical override:** A rule that prevents a serious incompatibility from being reduced by averaging or minor safeguards.
3. **Guardrail Index:** The 0-100 decision-support indicator generated from the current responses.
4. **Hard stop:** A condition requiring removal or redesign before the activity proceeds.
5. **Human-first workflow:** Learners attempt or document reasoning before consulting AI.
6. **Inherent risk:** Risk present before safeguards are considered.
7. **Residual risk:** Risk remaining after selected safeguards are applied.
8. **Synthetic information:** Invented data that do not represent a real identifiable person or confidential record.



18. Development and Attribution

1. Concept and ethical logic: Prof. Muhammad Al-Shorbagy, Dean, College of Pharmacy, Gulf Medical University.
2. Development approach: AI-assisted VIBE coding, with faculty ownership of the educational and ethical decision logic.
3. Application family: Acabuild AI LabHub - applications by academicians for academicians.
4. Methodology: PHARMA Ethical AI Lens, Rules Version 1.0.

Innovation is not measured by how much AI is inserted into a process, but by how appropriately it is used - or intentionally not used.