AssessReview AI (ARAI)

Documentation & User Guide

Release Date: October 2026 | Application Version: 3.1
Platform: Universal HTML5 / JavaScript — Client-Side · Gemini AI-Powered

Central principle

Exam quality review should be consistent, transparent, and evidence-based. ARAI helps assessment leads apply a structured, documented quality check to every examination — it does not replace professional academic judgement or institutional authority.

1. Introduction

AssessReview AI (ARAI) is a browser-based, AI-powered quality assurance tool developed for the College of Pharmacy at Gulf Medical University. It automates the peer-review process for ExamSoft examination files — Multiple Choice Question (MCQ), Short Answer Question (SAQ) and mixed examinations — applying standardised assessment-quality criteria aligned with the QF Emirates National Qualifications Framework.

Unlike a manual checklist, ARAI re-evaluates its findings in real time as you upload different files. This allows reviewers to check alternative versions and understand exactly why a defect is reported.

Key features

1. **MCQ and SAQ review:** Detects the question format of every exam file (or uses the format the reviewer selects) and applies MCQ-specific or SAQ-specific criteria — scientific accuracy, spelling, grammar, task clarity, option design and more.
2. **Reliable item identification:** Every question is read from "Question #" to "Item Psychometrics", so each defect, similarity match and blueprint row carries the correct ExamSoft Item ID.
3. **Word-by-word similarity index:** Compares the stem and options of every current question with every previous-exam question, shows the index and the changed words for each question, and flags ≥ 95.00% as reused.
4. **Blueprint Alignment:** The AI reads the blueprint, places every question under its blueprint topic and question type, and compares the expected and found numbers of MCQs and SAQs.
5. **Model cascade with time limits:** Tries up to 8 Gemini text models in order. Every request is time-limited, so a slow or unavailable model never stalls the review, and a Stop and show results button is always available.
6. **Structured Word export:** Produces a formatted Assessment Review Report (.docx) that mirrors the on-screen report, including the Similarity and Blueprint Alignment tables.
7. **Session restore:** Completed reviews are saved in the browser and can be restored on the same device.
8. **Zero server storage:** All processing happens in the user's browser — no exam data is stored on any server.
9. **Responsive design:** Fully functional on desktop, tablet, and mobile screens.

Professional review required

The ARAI Defect Report is a structured decision-support output, not a validated psychometric instrument. All AI-generated defects must be reviewed and confirmed by a qualified academic before any action is taken.

2. System Requirements

Browser and device

- Modern desktop browser: Chrome 110+, Edge 110+, Firefox 115+, Safari 16.4+
- Mobile / tablet: Chrome for Android, Safari for iOS (tablet recommended)
- JavaScript must be enabled; minimum screen width 320 px
- Active internet connection required during analysis (API calls and CDN libraries)

Google Gemini API key

Required: Gemini API key

Obtain a free key from **Google AI Studio** at aistudio.google.com/app/apikey. ARAI does not include an API key. The free tier supports most review sessions.

Supported file formats

FILE TYPE	EXTENSION	NOTES
ExamSoft Export PDF	.pdf	**Recommended** for current and previous exams. Questions are identified from "Question #" to "Item Psychometrics"; the ✓ marking the correct MCQ option is read.
Word Document	.docx / .doc	Exams and blueprints. A previous exam that is not an ExamSoft export is split on its question numbers (1., 2., 3. …).
Excel Spreadsheet	.xlsx / .xls	Supported, mainly for blueprint files.

Scanned PDFs

A PDF that contains only scanned images has no readable text. Export the exam from ExamSoft as a text PDF before uploading it.

3. Getting Started — Step-by-Step Workflow

Step 1 · API key setup

1	**Go to Google AI Studio** Navigate to aistudio.google.com/app/apikey and sign in with a Google account.
2	**Create API Key** Click **Create API Key**. Copy the key. Keep it confidential — treat it like a password.
3	**Paste into ARAI** On the AI Configuration screen, paste the key and click **Test Connection**, then **Connect & Continue**. ARAI lists the Gemini text models available to the key and starts with the newest Flash model.

Step 2 · Assessment information

Fill in all required fields: Course Name, Course Code, Program, Semester, Coordinator, Reviewer, Assessment Element, QF Emirates Level, and Offering Type (First or Repeated). **Question Format** is optional: *Auto-detect* (recommended) recognises SAQ items by their answer lines and handles MCQ, SAQ and mixed files; choose *MCQ* or *SAQ* to override it.

Offering type matters

Selecting *First Offering* disables the Previous Exams upload zone and marks Similarity as N/A. Select *Repeated Offering* when previous exam files are available.

Step 3 · Document upload

ZONE	FILES	PURPOSE
A. Exam Blueprint	Single file (required)	Read by the AI to list the blueprint topics, the expected number of items per question type (MCQ / SAQ) and their weights for Blueprint Alignment
B. Current Examination	Multiple files (required)	The ExamSoft export(s) to be reviewed — MCQ, SAQ or mixed
C. Previous Exams	Multiple files (optional)	Used for the word-by-word similarity (item reuse) analysis

Step 4 · Run analysis

Click **Start Assessment Review**. The analysis runs in four phases:

1. **Reading files (0–45%)** — Text is extracted in page order, each exam is split into items and the question format of each file is detected.
2. **AI quality review (45–90%)** — Questions are sent to Gemini in groups of 10, three groups at a time; the similarity index is calculated in the browser.
3. **Blueprint Alignment (91–96%)** — The AI reads the blueprint, then matches the questions to it in groups of 25.
4. **Report (97–100%)** — Results are combined, displayed and saved in the browser.

Time limits and Stop and show results

Each AI request has a time limit (150 seconds per model, 7 minutes per group of questions). If a model does not answer, the next model is tried automatically. **Stop and show results** ends the wait at any time: everything already checked is kept, and questions the AI has not finished are listed in the report.

How ARAI reads ExamSoft files

- Each item runs from "Question #" to "Item Psychometrics"; the Item ID, description, weight and attachment are taken only from inside that item.
- Item Categories tables, psychometrics, page headers and page numbers are excluded from the question text.
- The ✓ that marks the correct MCQ option is kept with its option, e.g. "D. Variance ✓".
- SAQ items are recognised by their answer lines; each file is reported as MCQ, SAQ or Mixed.

4. Assessment Review Criteria

ARAI applies one set of criteria to MCQ items and another to SAQ items; in a mixed exam each question is checked against the criteria for its own format. Each criterion carries a severity level.

Severity definitions

Critical — Must be resolved before the exam can be administered. **Major** — Significant quality concern; resolution strongly recommended. **Minor** — Advisory; low risk but worth improving.

MCQ items

#	CRITERION	SEVERITY	WHAT IS CHECKED
1	Scientific Correctness	CRITICAL	Solves the question independently; flags an incorrect marked answer, more than one correct option, or no marked answer
2	Stem Validity	MAJOR	Stem is clear, unambiguous and self-contained; flags vague or double-barrelled stems

3	Option Appropriateness	MAJOR	Every option is a plausible distractor; flags obvious decoys and options that repeat the stem
4	None of the Above	MAJOR	Reported only when the exact phrase appears in an option, which is quoted
5	All of the Above	MAJOR	Reported only when the exact phrase appears in an option, which is quoted
6	All Except	MAJOR	Reported only when the exact phrase appears in an option, which is quoted
7	Grammatical Errors	MINOR	Grammar mistakes, quoted with the corrected wording
8	Spelling Mistakes	MINOR	Misspelled words, quoted with the correct spelling
9	Equations for Calculations	MAJOR	A calculation question provides the equation it requires
10	QF Emirates Alignment	MAJOR	Cognitive demand matches the QF Emirates level selected for the course
11	INN / Generic Drug Names	CRITICAL	Brand/trade names detected; correct INN equivalent suggested
12	Parallel Grammatical Structure	MINOR	All options use the same grammatical form
13	Absolute Terms	MINOR	Flags "always", "never", "only", "must" in options as potential cueing words
14	Negative Stem Format	MINOR	Flags "NOT", "EXCEPT", "LEAST likely" stems (advisory)
15	Item Reuse / Similarity	CRITICAL	Similarity index ≥ 95.00% vs. previous exams; reports the exact Item ID of the match (see Section 5)

SAQ items

#	CRITERION	SEVERITY	WHAT IS CHECKED
1	Scientific Accuracy	CRITICAL	Every statement, premise, mechanism, drug name, value and unit is factually correct; flags questions that cannot be answered as written
2	Spelling Mistakes	MINOR	Every word of the question, its sub-parts and attachment names; quoted with the correct spelling
3	Grammatical Errors	MINOR	Malformed questions, tense, agreement, punctuation and sub-part labelling; quoted with the corrected wording
4	Task Clarity (Stem Validity)	MAJOR	The command verb and expected scope are clear; flags vague or double-barrelled tasks
5	Mark Allocation	MINOR	Multi-part items show marks per part, and the required answer is proportionate to the item weight
6	Referenced Material	MAJOR	An article, figure, table or compound referred to in the question is attached
7	Item Overlap / Cueing	MINOR	Two items assess the same content, or one item reveals the answer to another
8	QF Emirates Alignment	MAJOR	Cognitive demand matches the QF Emirates level selected for the course
9	INN / Generic Drug Names	CRITICAL	Brand/trade names detected; correct INN equivalent suggested
10	Equations for Calculations	MAJOR	A calculation question provides the equation or constants it requires
11	Negative Stem Format	MINOR	Flags "NOT", "EXCEPT", "LEAST likely" wording (advisory)

12	Item Reuse / Similarity	CRITICAL	Similarity index ≥ 95.00% vs. previous exams; reports the exact Item ID of the match (see Section 5)

Not listed in the Defect Report

CLO / PLO / Bloom's mapping — excluded from the review at the request of the Assessment Office. **Blueprint alignment** — reported in its own Blueprint Alignment section (Section 6). **Item reuse** — reported in the Similarity Report (Section 5). **Question-order randomisation** — not reviewed.

5. Similarity Analysis

When previous exam files are uploaded, ARAI compares every current question, word by word, with every question in all previous exam files. The comparison runs in the browser — it does not depend on the AI — and only the **question stem and options** are compared. The Item ID, description, categories, psychometrics, attachment names, the ✓ answer mark and the option letters are excluded.

Similarity index formula

```
Similarity Index = 2 × M ÷ (Nc + Np) × 100%
M = words that match in order (exact word = 1, one-letter difference = 0.5)
Nc, Np = number of words in the current / previous question
```

Text is compared after the same clean-up on both sides: letters are lower-cased, punctuation and option / sub-part labels (A., b., 1.) are removed, and hyphenated or line-broken words are joined. MCQ options are compared in alphabetical order, so a re-ordered option list does not lower the index. A word one letter away from its counterpart (a misspelling, singular / plural) counts as half a match; numbers must match exactly. An index ≥ 95.00% marks the question as **Reused** and triggers a **[Critical]** Item Reuse finding.

Similarity Report

COLUMN	CONTENT
Current Question ID	Item ID of the question in the current exam
Previous Exam Reference	Closest match: file name and **Item ID** of the previous question (e.g. EndSemesterJanCohortMCQ.pdf — Item ID: 632375 / 1), not a sequential question number
Similarity Index	Index of the closest match, to two decimals
Status	**Reused** (≥ 95.00%) or **New**

Differences	Words that differ, current → previous (e.g. "110, 6000 → 120, 240"); "Identical" at 100%

Every current question is listed — not only reused ones — sorted from the highest index. A summary line gives the number of questions compared and reused, and the number of questions read from each previous file. A previous file in which no questions can be identified is named in the report.

6. Blueprint Alignment

Blueprint Alignment runs as a separate AI step after the quality review:

1. **Read the blueprint:** The AI lists every topic row of the blueprint with its learning outcome, the expected number of items per question type (e.g. MCQ – 8, SAQ – 1) and its % weightage.
2. **Match the questions:** The AI reads each question's text and options (with its ExamSoft topic label as a hint) and decides which blueprint topic it assesses and whether it is an MCQ or an SAQ.
3. **Fill the table:** ARAI counts the matched questions per topic and per question type and compares them with the blueprint.

Blueprint Alignment table

COLUMN	CONTENT
Blueprint Topic	Topic as written in the blueprint (with its section heading, if any)
Expected (blueprint)	Expected number of items per question type, with the % weightage
Found (exam)	Number of exam questions matched to the topic, per question type
Exam Item IDs	The matched questions; hover over an Item ID to see why it was placed under that topic
Status	**Aligned**, **Under-represented (−n)**, **Over-represented (+n)**, **Not assessed**, or **Type mismatch** (e.g. the blueprint expects an SAQ but the question is an MCQ)

The last row, **Total number of questions**, always shows the MCQ, SAQ and overall totals — Expected (blueprint) against Found (exam) — with the difference per type (e.g. MCQ −7 · SAQ −1). Questions that fit no blueprint topic are listed with the AI's reason; if the blueprint cannot be read or some questions could not be matched, the section says so.

7. Review Results & Report

Review Summary strip

Four statistics are shown at the top of the results page: **Questions** · **Total Defects** · **Similar Items** · **Similarity %**. The report heading shows the course details and the question format of the exam (MCQ, SAQ or MCQ + SAQ).

Report sections

SECTION	CONTENT
Exam Items Detected	Collapsible list of every item found: file, Question #, Item ID, format and the start of the question text — use it to check that Item IDs are paired with the right questions
AI review not completed	Shown only when an AI request failed or was stopped; names the questions that were not checked
Similarity Report	Every question with its closest previous match, similarity index, status and differences (Section 5)
Blueprint Alignment	Expected and found items per blueprint topic and question type, with totals (Section 6)
Defect Report	One row per defect, with editable remarks
Overall Review Comments	Free-text comments from the reviewer

Defect Report table

COLUMN	CONTENT	EDITABLE?
Item ID / File	Question Item ID, source exam filename and defect category	No
Remarks	AI-generated defect description with quoted evidence	Yes
Actions	Save · Reset · Delete per defect	—

Filters

Category filter (drop-down) and free-text search filter the defect table in real time.

8. Export to Word

Click **Export Word** to download a .docx file. The exported document mirrors the on-screen report: title, subtitle with the question format, per-file similarity line, Course Information, Review Summary, Similarity Analysis table (with similarity index, status and differences), Blueprint Alignment table with totals, any AI-review-not-completed notes, Defect Report table and Overall Review Comments.

File naming: ProgramName_CourseCode_Semester.docx — e.g. BPharm_BIS101_Fall_2026.docx

Download in iframe / Wix

When embedded in a Wix widget or iframe the Word file may open in a new browser tab rather than downloading directly. Use the browser's Save As option from that tab.

9. Session Save & Restore

ARAI automatically saves the session to browser local storage on every edit. On re-opening, a banner offers to restore the last session. Autosave is silently disabled in private/incognito windows or when local storage is blocked by the browser.

10. Technical Specifications and Deployment

Technology stack

COMPONENT	TECHNOLOGY
AI engine	Google Gemini API — up to 8 text models (Flash models prioritised); 150 s per model attempt, 7 min per group of questions
AI batching	Quality review: 10 questions per request; Blueprint matching: 25 per request; 3 requests in parallel
File reading	PDF.js (Mozilla, v2.16.105) with page-order text rebuilding; Mammoth 1.5.1 (Word); SheetJS 0.18.5 (Excel)
Similarity engine	JavaScript in the browser — word-by-word index on stem and options
Word export	docx.js v7.8.2 (client-side)

Deployment unit	Single self-contained HTML file (~154 KB) — no installation or backend
Styling	Tailwind CSS (play CDN) + Font Awesome 6.4
Storage	Browser localStorage only — no server-side data

Cloudflare Pages deployment

1	**Create a Cloudflare Pages project** Log in to dash.cloudflare.com → Workers & Pages → Create → Direct Upload.
2	**Upload the HTML file** Rename ARAI_AssessReview_AI_v3.1_MCQ_SAQ.html to index.html and upload it.
3	**Share the URL** The tool is instantly accessible at the generated *.pages.dev URL.

Wix HTML widget embedding

1	**Add an HTML Component** In the Wix Editor → + → Embed Code → HTML iframe.
2	**Paste the full source code** Open ARAI_AssessReview_AI_v3.1_MCQ_SAQ.html, select all, paste into the HTML widget editor.
3	**Set size** Set component height ≥ 900 px, width 100%, scrolling enabled.

11. Mobile and Tablet Operation

ARAI v3.1 includes responsive CSS for mobile and tablet screens:

- Nav subtitle and model name hidden on screens < 640 px to prevent overflow
- Stat cards wrap to a 2-per-row grid on mobile and tablet
- Tables (Similarity, Blueprint Alignment, Exam Items, Defects) use horizontal scroll containers — the page body never scrolls sideways

- Minimum touch target size ensured for all interactive controls

12. Privacy, Security and Data Handling

Privacy-by-design architecture

- Exam question text and the blueprint are processed transiently by the Gemini API per Google's usage policies — they are not stored permanently by Google or by ARAI.
- Previous exams are compared in the browser and are not sent to the AI, unless the current exam cannot be split into questions.
- API keys are held in browser session memory only and are cleared when the tab is closed.
- Session data (defects, comments) is stored in browser local storage on the reviewer's device only.
- No data is transmitted to any GMU, ARAI, or third-party server beyond the Gemini API call.

Data-entry rule

Reviewers should treat exam content as confidential and ensure their Google account and API key usage comply with institutional data governance and exam security policies.

13. Methodological Limitations and Governance

- AI models may occasionally produce false-positive or false-negative defect reports.
- Subjective criteria (task clarity, mark allocation, option plausibility) require human judgement for final determination.
- The model cascade selects the best available model; different models may produce slightly different results for the same exam.
- The similarity index measures wording: a question rewritten in different words is not detected as reused. Similarity detection is a flagging mechanism — intent and context must be assessed by the reviewer.
- Blueprint matching is an AI judgement — check the topic assigned to each question (hover over its Item ID for the reason).
- Item identification relies on the ExamSoft export structure ("Question #" … "Item Psychometrics") — check the Exam Items Detected panel when a result looks wrong.

Recommended governance use

Use ARAI as a first-pass automated screen to identify potential issues. All flagged defects should be independently verified by the reviewer before inclusion in the final assessment review report submitted to the Assessment Office.

14. Development and Attribution

FIELD	DETAIL

Application name	AssessReview AI (ARAI)
Version	3.1 — October 2026
Institution	College of Pharmacy - Gulf Medical University
Assessment Lead	Dr. Mohammed Abdullatif · dr.abdullatif@gmu.ac.ae
Platform	Universal HTML5 / JavaScript — Client-Side · Gemini AI-Powered
Question formats	MCQ, SAQ and mixed ExamSoft examinations
Data storage	Browser localStorage only — no server-side storage