

AIRefAssist

User Manual | Engine version refassist 0.9 | Style target: IEEE Editorial Style Manual

Application: <https://ictect-refassist.azurewebsites.net>

Revision date: 21 August 2026

Introduction

AIRefAssist verifies bibliographic references against fourteen scholarly databases, corrects the fields that disagree with the published record, flags retracted and fabricated works, and formats the result to IEEE style. Every change it makes is reported with the source that justified it.

This manual covers everyday use of the web application, how to interpret the reports it produces, the guarantees that govern how your data is treated, the IEEE conventions applied to the output, and the HTTP API for automated workflows.

WHO THIS IS FOR Authors preparing IEEE-style manuscripts, editors and reviewers screening submissions, and anyone who needs to trust a bibliography — including one drafted with AI assistance, where invented citations are a known failure mode.

Contents

Introduction.....	1
1 What AIRefAssist does	3
Is it real?	3
Is it right?	3
Is it IEEE?	3
2 Quick start	3
A worked example	3
3 The workspace.....	4
Input panel	4
Results panel	4
4 Checking references	4
Input rules	4
How batches run	5
5 Uploading documents	5

6 Reading your results	5
Status badges.....	5
Confidence	6
7 The full report	6
Sample report (abridged)	8
8 Corrections and guarantees	8
Reading a correction	9
9 Integrity checks	9
Retracted works	9
Fabricated references	9
Other checks.....	10
10 IEEE output conventions.....	10
Patterns by reference type	11
11 Exports	12
12 Developer API	12
Single reference	13
Batch job lifecycle	13
Error codes	13
Appendix A Limits at a glance.....	14
Appendix B Troubleshooting and FAQ	15
Troubleshooting	15
Frequently asked questions	15
Appendix C Sources consulted.....	17

1 What AIRefAssist does

AIRefAssist takes the reference list you were about to submit and answers three questions about every entry.

Is it real?

Each reference is searched across fourteen scholarly databases — Crossref, IEEE Xplore, OpenAlex, PubMed, arXiv, DBLP and others; the full list is in Appendix C. References that cannot be found anywhere, cite a venue that does not exist, or carry a dead identifier are flagged as possibly fabricated.

Is it right?

Fields that disagree with the published record — pages, years, author spellings, volumes, DOIs — are corrected, and every correction names the source that justified it. Retracted works are flagged prominently, with the reasons on record at Retraction Watch.

Is it IEEE?

The verified reference is rendered in IEEE style by a deterministic formatter: correct author initials, *et al.* rules, quoted titles, italic venues, volume and page notation, IEEE month forms, and doi: notation.

The output is a ready-to-paste IEEE reference list plus a per-reference report showing what was verified, what was changed, what could not be confirmed, and what needs your attention before you submit.

2 Quick start

1. Open the application at <https://ictect-refassist.azurewebsites.net>. The first request after a quiet period can take up to a minute while the service wakes; later requests are fast.
2. Paste one or more references into the input box, one reference per line. Numbering such as [1] is fine and is stripped automatically.
3. Press **Check**, or use Ctrl+Enter.
4. Read the result card for each reference: a status badge, the formatted IEEE entry, and the corrections that were applied. Use **Copy Output** for the finished list, or **Download full report** for the complete audit as a Word document.

A worked example

Paste this reference, which contains a deliberate page error:

Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 430-444, May 2015.

AIRefAssist returns the status **corrected** with this entry:

Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

The page range was corrected against the publisher's record, the missing DOI was added, and the corrections list names the source for each change.

3 The workspace

The screen is split into an input panel on the left and results on the right.

Input panel

- **Paste References** — the default tab: a text box for one reference per line, a live count of detected references, and the Check button. **Try example** fills the box with sample references; **Clear** (Ctrl+K) empties it.
- **Upload Files** — drop manuscripts or bibliography files and press Process Files. See chapter 5 for formats and limits.
- While a batch is running, a progress bar appears with a Cancel button. Cancelling keeps the results of references that already finished.

Results panel

- **Result cards** — one per reference, numbered in input order: status badge, confidence badge, the formatted entry, and per-reference details. A RETRACTED badge or a “duplicate of [n]” badge appears where relevant.
- **Copy Output** copies the formatted reference list; **Copy BibTeX** copies the same entries as BibTeX records.
- **Full Report Preview** and **Download full report** give the complete per-reference audit, on screen and as a Word document.
- The moon/sun button in the header switches between dark and light mode.

4 Checking references

Input rules

- One reference per line. Leading numbering, quotes from CSV exports, and doubled quote marks are cleaned up automatically.
- Any citation style is accepted as input — IEEE, APA, Chicago, or an unstyled scrawl. Output is always IEEE.

- Partial references are fine: a DOI alone is enough to reconstruct the full entry when the record exists.
- Up to **200 references** are accepted per request.

How batches run

Batches of up to 10 references are processed while you wait. Larger batches switch automatically to a background job: the progress bar tracks completion, result cards appear as each reference finishes, and the batch can be cancelled at any point. A typical reference takes 4 to 20 seconds, and several are processed in parallel.

CAUTION Background jobs are held in server memory. Download your results and report promptly — a job can expire and cannot be re-fetched after that.

5 Uploading documents

The Upload Files tab extracts references directly from files, so you can check a whole manuscript's bibliography without copying it out by hand.

Accepted formats	Per-file limit	Files per request
.docx, .pdf, .tex, .bbl, .txt	15 MB	10

5. Drag files onto the upload area, or click it to browse.
6. Press **Process Files**. Reference sections are located and split into individual entries; the combined batch then runs exactly like pasted input, including the automatic switch to a background job for more than 10 references.

NOTE Extraction quality follows the file's own structure. A clearly delimited reference section — a numbered list under a References heading, or a .bbl file — extracts best; references buried in running prose may need to be pasted manually.

6 Reading your results

Status badges

Badge	Meaning	What you should do
verified	Matched an authoritative record; no corrections were needed.	Use as shown.

Badge	Meaning	What you should do
corrected	Matched an authoritative record; the listed corrections were applied.	Skim the corrections list, then use.
unverified	No trustworthy online match. Your original text is preserved — nothing was invented to fill the gaps.	Verify against the original document yourself; common for theses, technical reports, and niche or very new works.
likely fabricated	The fabrication check found strong signals that this work does not exist (see chapter 9).	Do not cite unless you can produce the original document.
not a reference	The input line is not a bibliographic reference — a sentence, a heading, a note.	Remove or rewrite the line.

Confidence

Level	What it means
high	Identity is anchored by DOI-level agreement across sources and nearly all fields verified. Wrongful edits at high confidence are treated as the worst possible failure, and the engine is calibrated accordingly.
medium	Good match with some fields unconfirmed, or corrections that contradicted several of the values you supplied.
low	Weak or conflicting evidence — including the label “Verified with conflicts — review before use” when the matched record disagrees with identifiers you supplied. Review before using.

READING TIP Status answers “what happened to this reference?”; confidence answers “how sure is the engine that it picked the right work?”. An unverified result at low confidence is honest ignorance, not an error — the reference may simply not be indexed anywhere.

7 The full report

Every reference gets a complete audit. The sections appear in this order.

Section	Contents
Action required	What you must do before submitting, most severe first: retraction warnings, fabrication verdicts, fields to verify manually. “None” means the reference can be used as shown.
Input	Your reference exactly as submitted.
Final reference (IEEE)	The formatted entry, plus a “Produced by” line naming the formatter that generated it.
Corrections applied	Every change as field: old to new, with the source. An empty list means the reference matched the record as provided.
Fields	Every field with its value, its source, and a verification mark: confirmed by sources, checked but not confirmed, or informational only.
Sources consulted	Each database that was queried and what it returned, including “no matching records” and timeouts. Nothing is hidden.
Source disagreements	Where databases disagreed on a field: every value seen, who said it, and which value was used.
Evidence	Links (DOI, publisher, database pages) supporting the selected record, so you can verify the match yourself.
Warnings	Anything suspicious that did not block processing: odd years, single-page ranges, failing ISBN checksums, multiple published versions.
Processing log	Engine and model used, retraction-database status, lookup and correction rounds, processing time.

Sample report (abridged)

```
=====
                        REFASSIST REFERENCE REPORT
=====
Status      : Verified against online sources - corrections applied
Confidence: High
Type        : Journal Article
DOI         : 10.1002/j.1538-7305.1948.tb01338.x

ACTION REQUIRED: none
  This reference can be used as shown below.

FINAL REFERENCE (IEEE)
  C. E. Shannon, "A Mathematical Theory of Communication," Bell System
  Technical Journal, vol. 27, no. 3, pp. 379-423, July 1948,
  doi: 10.1002/j.1538-7305.1948.tb01338.x.
  Produced by: Rule-based IEEE formatter (deterministic)

CORRECTIONS APPLIED (1)
  - doi: (missing) -> 10.1002/j.1538-7305.1948.tb01338.x    [Crossref]

FIELDS (v = confirmed, ! = not confirmed, . = informational)
  v title      : A Mathematical Theory of Communication    (Crossref)
  v volume     : 27                                       (Crossref)
  v issue      : 3                                       (Crossref)
  v pages      : 379-423                                 (Crossref)
```

Shannon's 1948 paper was published in two parts with the same title, authors, volume and year. The engine used the issue and pages supplied in the citation to select Part 1 — see the identifier-anchoring guarantee in chapter 8.

8 Corrections and guarantees

Corrections are the product's core value, and its core risk. A tool that “fixes” a correct reference into the wrong work does more damage than no tool at all. AIRefAssist therefore operates under explicit guarantees.

GUARANTEE 1 — YOUR IDENTIFIERS ANCHOR THE MATCH The volume, issue, pages, ISBN and DOI you supply are treated as evidence of which work you meant. When two published works share a title and authors, your coordinates decide between them. Coordinates are a tiebreaker, not a straitjacket — a reference whose pages are simply wrong still gets corrected when only one plausible work exists.

GUARANTEE 2 — CO-AUTHORS ARE NEVER DELETED ON A GUESS A matched record with fewer authors than you cited (an excerpt, a solo essay, a different edition) is never allowed to delete your co-authors unless an exact DOI or ISBN match proves the citation really is that smaller-author record.

GUARANTEE 3 — VERIFIED FIELDS NEVER REGRESS Once a field has been confirmed against sources during processing, a later lookup round cannot rewrite it to something else.

GUARANTEE 4 — NOTHING IS INVENTED When a reference cannot be verified, your original text is preserved untouched. The formatter also refuses to emit an entry that would keep fewer than two fields; degraded output is replaced by your original text. Journal abbreviations are only ever taken from the NLM Catalog or the matched record, never invented.

GUARANTEE 5 — CONFLICTS CAP CONFIDENCE A match that contradicts identifiers you supplied is labelled “Verified with conflicts — review before use” at low confidence. It will never present itself as a clean high-confidence verification.

Reading a correction

pages: 430-444 -> 436-444 [Crossref]

Each correction names the field, the before and after values, and the source that justified the change. Sources include the individual databases, *DOI agreement* (multiple databases agreeing on the same DOI), *Consensus* (a weighted vote across matching records), and *Provided input* (your own data, kept).

9 Integrity checks

Retracted works

Every verified reference is checked against Crossref, OpenAlex, PubMed and the Retraction Watch database. A retracted work gets an unmissable banner with the retraction date, the reasons on record, and the DOI of the retraction notice. Expressions of concern and corrections are reported as advisories.

Fabricated references

A reference that cannot be verified is additionally screened for the signature pattern of an invented citation. Five independent signals feed the verdict.

Signal	What it checks
No title match	No record with a similar title exists in any of the fourteen databases.

Signal	What it checks
Unknown venue	The cited journal or conference matches nothing in the Crossref registry — the venue itself may not exist.
Dead DOI	The citation carries a DOI that no registry resolves.
DOI collision	The DOI resolves — to a completely different work.
Disjoint authors	The title matches a real work, but the cited authors are entirely different people: the classic AI-hallucination pattern.

Two or more strong signals produce the **likely fabricated** verdict: the report leads with the evidence, and the entry is **not** polished into IEEE form. Dressing up a fake makes it more credible, which is the opposite of this product's job. One signal produces a “possible fabricated reference” advisory.

Other checks

- **Author mismatch** — when the title matches a published work but your cited authors do not, the report shows the real authors so you can decide whether the citation is misattributed or fabricated.
- **Multiple published versions** — a work published both as a conference paper and a journal article, or as preprint and published version, triggers a warning to confirm you are citing the version you actually used.
- **Duplicates** — identical references in one batch are marked as duplicates of the first occurrence.

10 IEEE output conventions

Formatting is performed by a deterministic rule-based engine: the same input always produces the same output. The conventions follow the IEEE Editorial Style Manual.

- **Authors** — initials before surname. Two authors are joined by “and” with no comma; three to six use the serial comma; seven or more become the first author followed by *et al.* in roman, never italic.
- **Titles** — article and paper titles in quotation marks with the comma inside the closing quote; book, standard and venue titles in italics.
- **Numbers** — vol. 27, no. 3, pp. 379–423, with an en dash in page ranges; article numbers as Art. no. where journals use e-locations.
- **Dates** — IEEE month forms: Jan., Feb., Mar., Apr., May, June, July, Aug., Sept., Oct., Nov., Dec.
- **DOIs** — rendered as doi: 10.1109/5.771073 at the end of the entry, not as a URL.

Patterns by reference type

Journal article

S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

Conference paper

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in Proc. NAACL-HLT, Minneapolis, MN, USA, 2019, pp. 4171-4186.

Book

T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, Introduction to Algorithms, 3rd ed., Cambridge, MA, USA: MIT Press, 2009, ISBN: 978-0262033848.

Book chapter

C. M. Bishop, "Graphical Models," in Pattern Recognition and Machine Learning, New York, NY, USA: Springer, 2006, pp. 359-422.

Standard

IEEE Standard for Floating-Point Arithmetic, IEEE Std 754-2019, 2019, doi: 10.1109/ieeestd.2019.8766229.

Thesis

I. Sutskever, "Training Recurrent Neural Networks," Ph.D. dissertation, Univ. of Toronto, Toronto, ON, Canada, 2013.

Preprint

J. Kaplan et al., "Scaling Laws for Neural Language Models," arXiv:2001.08361, 2020.

Technical report

A. Krizhevsky, "Learning Multiple Layers of Features from Tiny Images," Univ. of Toronto, Toronto, ON, Canada, Tech. Rep., 2009.

Supported types: journal article, conference paper, book, book chapter, thesis, technical report, dataset, standard, software and preprint. The type is detected automatically from the citation and the matched records.

11 Exports

Export	Where	Contents
Reference list	Copy Output button	The formatted IEEE entries, ready to paste into a manuscript. Retracted entries are suffixed [RETRACTED].
BibTeX	Copy BibTeX button	The same entries as BibTeX records for LaTeX workflows.
Word report	Download full report	The complete audit — status, corrections, fields, warnings — one section per reference, as .docx.
Job archive	Background jobs	Large batches deliver a ZIP containing the formatted list, the Word report and the per-reference details.
CSL-JSON and BibTeX data	API responses	Structured records for reference managers such as Zotero and Mendeley, and for pipelines. See chapter 12.

12 Developer API

Everything the interface does is available over HTTP at the same host. Responses are JSON unless noted.

Endpoint	Purpose
POST /v1/resolve	Verify one reference. JSON body {"reference": "..."}; returns the formatted entry, full report_data, CSL-JSON and BibTeX.
POST /api/process	Verify a batch. Form field references (text, one per line) and/or files. Up to 10 references run synchronously; larger batches return 202 with a job.
POST /api/jobs	Queue a batch as a background job unconditionally. Same inputs as /api/process.
GET /api/jobs/{id}	Job progress: status (running, completed, cancelled, failed), done and total counts, and progressive results as they finish.
DELETE /api/jobs/{id}	Cancel a running job; completed references are kept.
GET /api/jobs/{id}/report	ZIP archive of a completed job; returns 409 while the job is still running.
POST /api/extract	Extract reference text from uploaded files without processing them.

Endpoint	Purpose
GET /healthz	Liveness probe.

Single reference

```
curl -s https://ictect-refassist.azurewebsites.net/v1/resolve \  
-H "Content-Type: application/json" \  
-d '{"reference": "C. E. Shannon, \"A Mathematical Theory of  
Communication,\" Bell System Technical Journal, vol. 27,  
no. 3, pp. 379-423, July 1948.\"}'
```

The response includes formatted, retracted, the plain-text report, and report_data — the structured audit with status, confidence, corrections, fields, fabrication, warnings and evidence.

Batch job lifecycle

```
# 1. Queue (form-encoded; one reference per line)  
curl -s -X POST https://ictect-refassist.azurewebsites.net/api/jobs \  
--data-urlencode "references@refs.txt"  
# -> {"job_id": "...", "status_url": "/api/jobs/...", "report_url": "..."}  
  
# 2. Poll until completed  
curl -s https://ictect-refassist.azurewebsites.net/api/jobs/JOB_ID  
# -> {"status": "running", "done": 12, "total": 40, "results": [...]}  
  
# 3. Fetch the archive  
curl -s -o refassist_report.zip \  
https://ictect-refassist.azurewebsites.net/api/jobs/JOB_ID/report
```

Error codes

Code	Meaning
400	No references detected, or more than 200 in one request.
404	Unknown or expired job.
409	Job report requested before the job completed.
413	Request body exceeds the upload size limit.
429	Rate limit reached (default 30 processing requests per minute per client), or the job queue is full. Honour the Retry-After header.

Appendix A Limits at a glance

Limit	Default
References per request	200
Synchronous batch size (larger batches auto-queue)	10
Files per upload	10
Size per file	15 MB
Processing requests per minute per client	30
References processed in parallel	6
Typical time per reference	4 to 20 seconds

Server operators can adjust these through environment variables; the values above are the deployed defaults.

Appendix B Troubleshooting and FAQ

Troubleshooting

Symptom	Cause and fix
First request hangs or fails, then works	The service was asleep (cold start). Wait up to a minute and retry.
429 Too many requests	Rate limit. Pause for the Retry-After interval, then continue. For large workloads, use one background job instead of many small requests.
Job link stopped working	Jobs live in server memory and expire. Re-submit the batch and download the report as soon as it completes.
A correct reference came back unverified	The work is not indexed in the consulted databases — common for theses, standards bodies outside the DOI system, and regional venues. Your text is preserved; verify manually and use it.
File uploaded but few references found	The reference section was not recognisable. Paste the references directly into the Paste tab.

Frequently asked questions

Does it support styles other than IEEE?

Input, yes — any style is parsed. Output is IEEE only in this release; the BibTeX and CSL-JSON exports let a reference manager restyle the verified data.

Will it “fix” a reference into the wrong paper?

This failure mode is what the engine is specifically calibrated against: identifiers you supply anchor the match, conflicting matches are demoted to low confidence, and unverifiable references are preserved rather than guessed at. See the guarantees in chapter 8.

Where do retraction flags come from?

Crossref, OpenAlex and PubMed retraction markers, cross-checked against the Retraction Watch database, with the retraction notice DOI included where available.

Can it tell if a bibliography was AI-generated?

It detects the two signatures invented citations leave: works that exist nowhere (fabrication signals, chapter 9) and real titles with invented authors (author-mismatch detection).



125 East State St
Burlington, WI 53105
Ph: 262-898-7277
Web: www.ictect.com

Is my data stored?

References are processed in memory; batch results are held only until downloaded or expired. No account is required.

Appendix C Sources consulted

Every reference is searched in parallel across the following sources.

Source	Strength
Crossref	DOI registry; the backbone for journals, conferences and books.
IEEE Xplore	IEEE journals, conferences and standards.
OpenAlex	Broad scholarly index with retraction markers.
Semantic Scholar	Computer science and cross-disciplinary coverage.
PubMed and Europe PMC	Biomedical literature and retraction types.
arXiv	Preprints in physics, mathematics and computer science.
DBLP	Computer science bibliography of record.
DataCite	Datasets and software DOIs.
Unpaywall, DOAJ, bioRxiv/medRxiv	Open-access locations and preprint servers.
Google Books and Open Library	Books and ISBNs.
NLM Catalog	Authoritative journal-title abbreviations.
Retraction Watch	Retraction reasons, dates and notices.

The report's "Sources consulted" section shows exactly which of these were queried for your reference and what each returned.

AIRefAssist — refassist engine v0.9. Verification powered by Azure OpenAI and the sources listed above. Formatted output follows the IEEE Editorial Style Manual. Always review action items before submission: the tool assists judgement, it does not replace it.