A PILOT STUDY FOR BAIKAL-GVD

GigaChat as a scientific assistant

Author corrections in astronomical reports

Can an AI turn a chain of scientific messages into a precise correction that a researcher can verify?

256 circulars **20 event chains** **12 recommendations**

Where the AI enters the scientific workflow

Baikal-GVD searches for possible links between neutrino candidates and events observed by other instruments

Astronomers publish

GCN Circulars

Text reports of measurements, follow-up observations and later author corrections.

→

GigaChat

Matches a correction to the right observation and proposes an exact record update.

→

Researcher reviews

The proposed change

Checks the cited source. Applies the correction if the old information was used.

↓

Reconstructed neutrino candidates

Time, direction, energy and uncertainties

→

Scientific association analysis

Temporal and angular consistency, then the probability of a chance coincidence.

Pilot boundary: public report text. GigaChat had no access to the researcher's working analysis.

Manuscript §1. GCN Circulars supplement the machine-readable Notices already used in follow-up workflows.

What GigaChat receives and returns

Example: 19 English circulars about GRB 221028A, with a Russian instruction to propose justified changes

Input: the full event chain

GCN 32893

Swift/BAT reports fluence and peak photon flux for 15–150 keV.

GCN 32895

The authors correct the band to 50–150 keV for both quantities.

Output: an actionable recommendation

Update the energy band in both records

15–150 keV → **50–150 keV**

Keep the numerical values and their timing.
Retain the duration and spectral information.
Attach the author's supporting paragraphs.

Quantity	Numerical value retained
Fluence	$(1.51 \pm 0.14) \times 10^{-6}$ erg cm^{-2}
1-s peak photon flux	0.8 ± 0.2 ph cm^{-2} s^{-1}

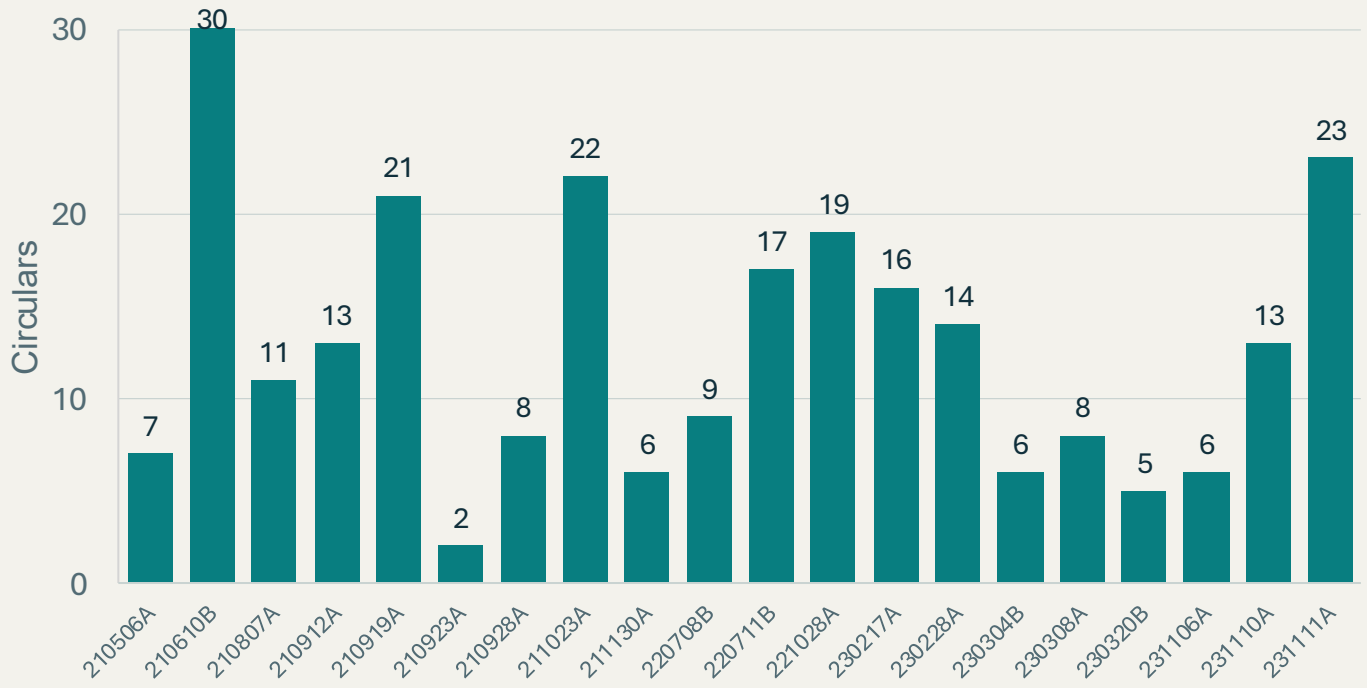
The AI must identify the scope of the correction and preserve valid information around it.

GCN 32893, 32895. Manuscript §§1, 3.2. On-slide description paraphrases the archived recommendation.

The pilot study

256 full-text circulars from 20 event chains, fixed before the model calls

How much text the model had to reconcile



A targeted test of corrections

15 chains selected for correction signals
5 ordinary observation chains
Development events excluded

Material and model

Sources: May 2021–November 2023
Up to 30 messages in the first 30 days
Absent linked documents not fetched

GigaChat-3-Ultra, temperature 0.2
API calls on 16 October 2026
English sources, Russian instructions

A targeted sample supports case analysis. General model accuracy remains unmeasured.

Manuscript §2.1. Per-chain counts from the retained corpus total 256. All chart event identifiers have prefix GRB.

How the AI recommendations were produced

GigaChat proposes and revises the actions. Another AI agent checks the sources and helps select the final list.

Each generation and checking request receives the full source chain

A Propose record updates

B Find corrections and retractions

Independent first readings
Neither sees the other answer

→

GigaChat reconciles and self-checks

Up to two refinements after its self-check

→

Another AI checks each recommendation

Observation, source support and scientific qualifications
Up to two further revisions

→

Freeze the final list

Separate evaluation

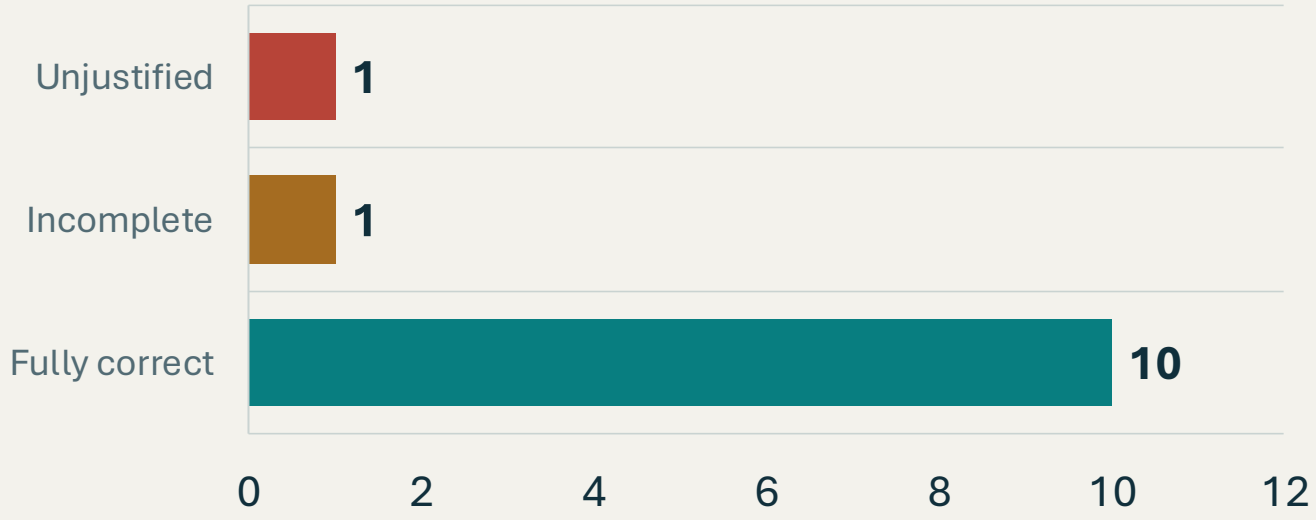
Two other AI agents prepare reference labels before the calls.
Compare with the final list only after freezing it. No evaluation feedback.

The model selects paragraph IDs. Code attaches the original excerpts verbatim for verification.

Manuscript §2.2. The reported result belongs to the combined GigaChat and AI-review workflow.

The result: 10 fully correct recommendations out of 12

Assessment covers each complete recommendation, including all affected fields and scientific qualifications



What the correct recommendations cover

3 event-name corrections
2 retractions or classification changes
5 numerical values or measurement conditions

What “fully correct” means

The right observation
The exact replacement, with units
All material qualifications preserved

Why two recommendations fail

One drops “approximate” and “very preliminary” from a flux result.

One rejects a tentative redshift without an author retraction.

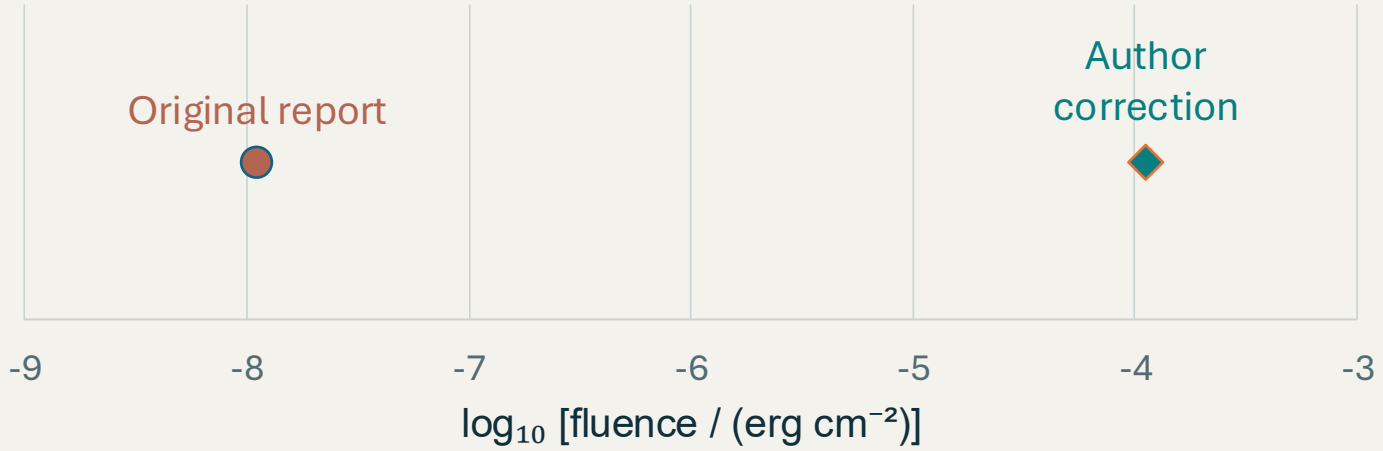
The reference labels are AI-generated and still await expert astrophysical review.

Manuscript §§2.2, 3.1 and Table 1. Completeness, researcher time saved and real-analysis performance were not measured.

A 10,000-fold correction to the fluence

GRB 211023A, Fermi/GBM. The authors correct GCN 30965 in GCN 30998.

Fluence: received energy per unit area



(1.114 ± 0.007) × 10⁻⁸ erg cm⁻²

(1.114 ± 0.007) × 10⁻⁴ erg cm⁻²

What the AI proposes

Link the two reports.
Replace the fluence only.
Retain the spectral parameters, duration and peak photon flux.

Why the correction matters

$$E' / E = 10^4$$

For an estimate proportional to fluence, with distance, spectral corrections and other assumptions unchanged.

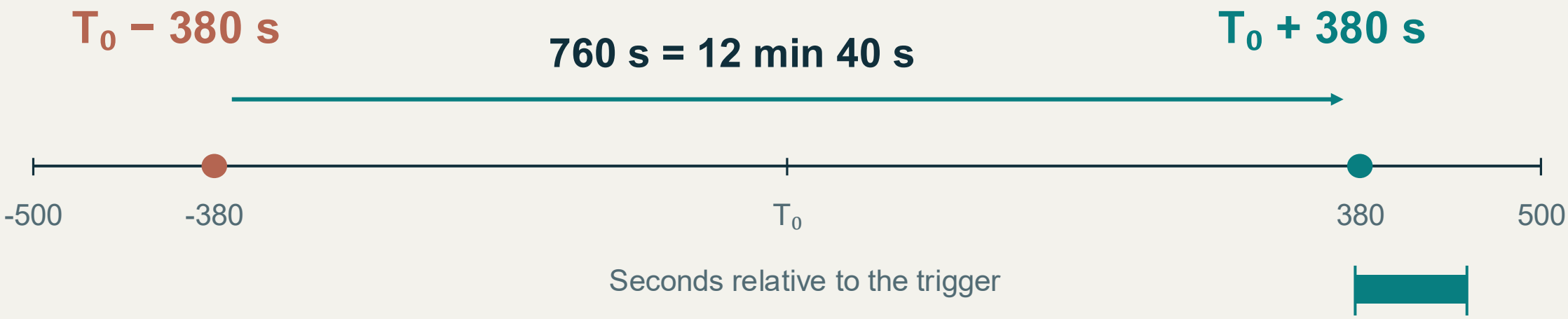
A missed author correction could propagate into an energy estimate by a factor of 10,000.

Manuscript §3.2. Same 10–1000 keV measurement. The ratio is an arithmetic illustration calculated by ordinary code.

A timing correction moves the search reference

GRB 220708B, Konus-Wind. GCN 32359 corrects the second-episode start reported in GCN 32358.

The author changes the sign of the episode time



AI action: correct the episode marker

Review a coincidence window anchored to the old time, if it was used.

AI action: preserve the spectral interval

$T_0 + 377.088$ s to $T_0 + 450.816$ s

Keep the preliminary status of the report.

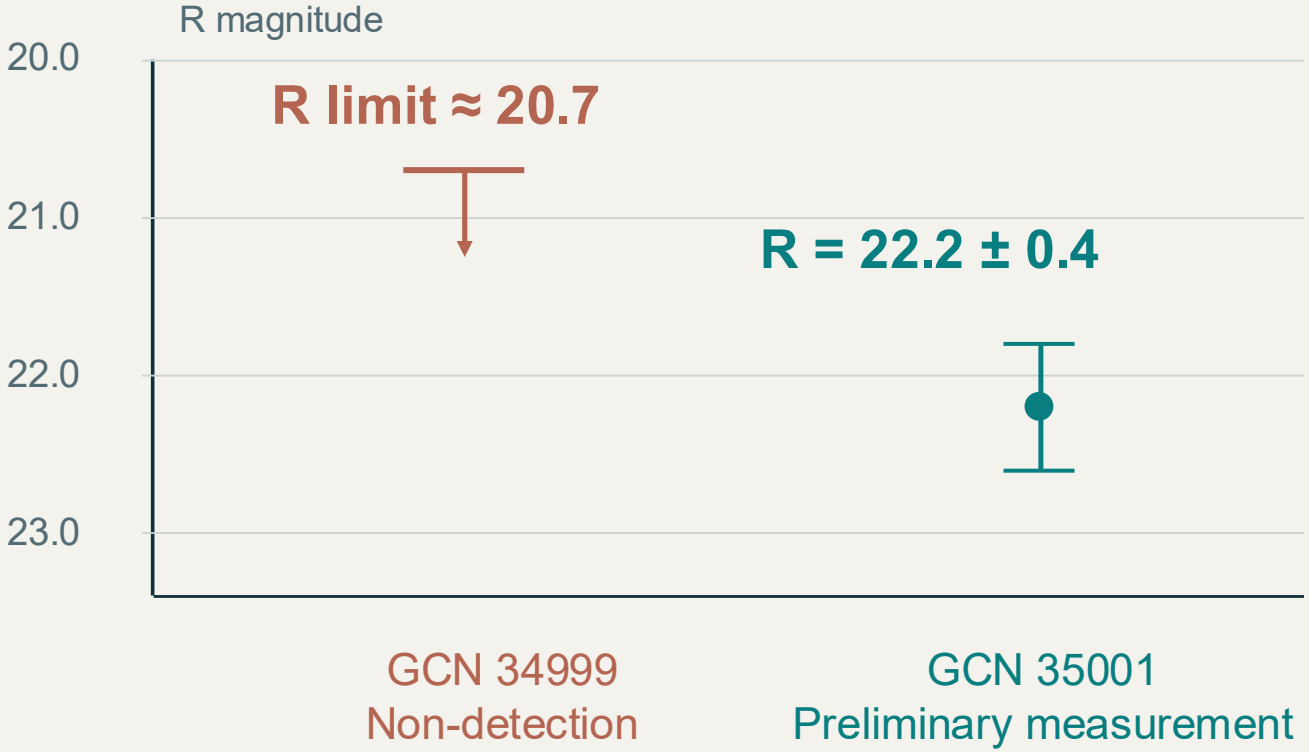
The AI changes the episode time while preserving the separate spectral interval.

Manuscript §3.2. The 760 s displacement is arithmetic. The pilot did not rerun a neutrino coincidence search.

One observation, two versions of its report

GRB 231111A, GMG. GCN 35001 replaces a non-detection in GCN 34999 with a preliminary measurement.

The AI must recognise the same observing epoch



Match the observation

Same instrument and R filter
Same start: 12 Nov 2023, 11:29:03 UT
About 21.2 hours after the trigger

Propose the right replacement

Replace the limit with the measured point at that epoch. Keep its error and preliminary qualification.

This avoids treating a revised report as evidence of fading over time.

A revision replaces a record. A new observing epoch adds a point to the light curve.

Manuscript §3.3. Horizontal positions show report versions, not time. The 1.5-mag difference is not a fading measurement.

The failures reveal the need for scientific review

The combined AI workflow can edit the right number while losing a qualification or making an unsupported choice

Incomplete: a lost qualification

GRB 231110A, GECAM-B
GCN 34992

The AI corrects the flux units and the T_0 to T_0 + 4 s interval.

It drops “approximate” and “very preliminary”.

The record would overstate the certainty of a result derived from realtime data.

Unjustified: selecting a redshift

Independent report	Value and status
TNG, GCN 35007	$z \approx 1.39$, tentative
GTC, GCN 35009	$z \geq 1.179$, lower limit

The AI proposes rejecting the TNG result.

GTC reports no lines at z = 1.39 in its own spectrum. The sources contain no TNG author retraction.

Correct values, preserved qualifications and justified scope all require review.

Manuscript §3.4. GTC also suggests 1.179 is probably the source redshift. Expert review must resolve the two reports.

What the researcher gets from the AI workflow

A specific proposal with its source, ready for checking against the inputs actually used in the analysis

Example: the proposed update for GRB 220708B

What the workflow supplies	What the researcher can verify
Target observation GRB 220708B, Konus-Wind, second-episode start	Does this identify a field that I used in my own analysis?
Exact replacement Change T_0 − 380 s to T_0 + 380 s	What must change in my record and in calculations that depend on it?
Preserved context Spectral interval and preliminary status	Which parts of the record remain valid after I apply the correction?
Source evidence GCN 32359 linked to the earlier GCN 32358	Can I check the author's statement before accepting the recommendation?

Intended benefit: reduce the risk of carrying obsolete information into the next calculation.

The pilot demonstrates reviewable proposals. Researcher time saved still needs measurement.

Manuscript §§1, 3.2, 4. Apply the recommendation only if the superseded information was actually used.

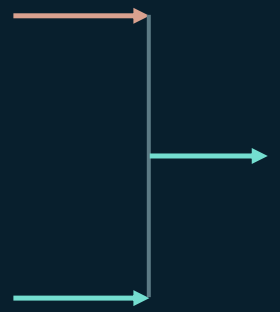
Next test: preparing a verified analysis request

Proposed validation with an existing Baikal-GVD association program

A researcher first verifies and accepts the correction

Original input
T_0 − 380 s

Corrected input
T_0 + 380 s



GigaChat prepares the program arguments

Time and window definition
Sample and configuration IDs

The analysis program executes the request

Compare its returned output with a controlled reference run

Hold fixed

Neutrino sample, window definition, code and configuration.
Change only the intended input. Expert validation and access are still pending.

Pilot result: reviewable update proposals. Next: expert validation and a controlled rerun.

Manuscript §§4–5. The proposed program test and any change in physics results remain future work.