

Search & verification

Modes and tiers

🕒 4 min read · Updated 26 October 2026 · Verified against [7a343c6](#) on 2026-09-25

The six search modes a caller can ask for and the twelve tiers auto mode runs through, each with its entry condition, the index it queries and the mode string it reports.

<https://docs.gurbanisoul.com/search/modes-and-tiers/>

`/api/search?q=&mode=&limit=&offset=` takes a query of at most 300 characters and a **mode**. The response always has `mode` (which tier answered), `results` (verbatim lines with their `Ang`, the transliteration and the English layer attached) and `related_themes`; a theme answer adds `concept`. `limit` is clamped to 0–200 and `offset` to 0–1,000,000, so a client can never dump the corpus by accident.

The modes

mode	What it does	Falls back to
auto (default)	detects the script and runs the waterfall	every tier in order
gurmukhi	the verbatim <code>text</code> column (FTS)	the <code>skeleton</code> column with vowel signs stripped (<code>gurmukhi-skeleton</code>)
roman	the <code>translit</code> column (FTS)	the phonetic fold over <code>translit_norm</code> (<code>roman-spelling-tolerant</code>)
first	first letters of each word, as a phrase: <code>fl_g</code> for Gurmukhi, <code>fl_r</code> for Roman	—
theme	the concept index: the concept's lines, with its description and total	the one concept whose name folds like the query (<code>waheguru</code> → <code>vahiguru</code>), then a <code>LIKE</code> over concept names and descriptions
english	the English translation layer, <code>fts_en</code>	—

The tiers of auto mode

Interactive on the web: a search, traced through each tier.

Tier	Entry condition	What it queries	mode
Mixed script	Gurmukhi and Latin in the query	<code>mixed_search</code> — one AND expression: Gurmukhi tokens on <code>text</code> , Latin tokens through variants, the lexicon and the fold; up to 10 tokens	<code>mixed-script</code>

Tier	Entry condition	What it queries	mode
Latin part alone	mixed script found nothing	the Latin tokens re-run through the waterfall	mixed-script (latin part: ...)
Gurmukhi first letters	every token is one letter, ≥ 2 tokens	fl_g phrase	first-letters
Gurmukhi text	Gurmukhi query	text (FTS, bm25)	gurmukhi
Gurmukhi skeleton	text found nothing	vowel signs stripped, skeleton LIKE	gurmukhi-skeleton
Concept name	the query is exactly a concept key (naam , haumai , ...), except vahiguru , which goes on to the lexicon	the concept index	theme
Roman first letters	≥ 2 tokens, all ≤ 3 letters	fl_r phrase, then translit	first-letters, roman
Transliteration	Roman query	translit (FTS, bm25)	roman
Seeker lexicon	nothing yet	a curated table: a word $\rightarrow$ a transliteration term or a theme	seeker-lexicon (term)
Variant index	nothing yet	per-token romanisation variants (≤ 3 by frequency $\times$ score), trailing-vowel retry, all-but-one AND	variant-match
English layer	nothing yet	fts_en — before the fold, so an English word reaches translations rather than colliding with a fold	english-translation
Honorifics dropped	nothing yet; ji , jee , jeo , jio , sahib , maharaj , shri , sri , baba , guru , dev , waale present; ≥ 2 tokens remain	the shorter query, through the waterfall	... (honorifics dropped)
Passage	nothing yet, ≥ 3 tokens	folded tokens AND in fts_shabad ; the composition's lines returned in order	passage-match (quote spans lines)
Phonetic fold	nothing yet	roman_norm(query) tokens of ≥ 2 characters against translit_norm ; headings excluded	roman-spelling-tolerant
Skeleton blob	nothing yet, ≤ 14 tokens, blob ≥ 8 characters	the query collapsed to one fold-skeleton blob; lines prefiltered by its head, fuzzy-ranked	skeleton-blob
Single-token theme	nothing yet, one token	the concept index	theme
Top-up	any tier, < 3 results, honorifics present	the honorific-dropped results appended	... + honorific-dropped

Ranking

Every FTS tier ranks with SQLite's bm25 and column weights **10 · 5 · 4 · 3 · 3 · 1** for text, translit, translit_norm, fl_g, fl_r, skeleton: the verbatim text outranks a transliteration hit, which outranks the fold, which outranks first letters and the skeleton. The passage tier ranks by the tightest span of matched lines, then bm25; the blob tier by fuzzy ratio.

What a client should rely on

- `results[].gurmukhi` is verbatim and `results[].ang` is its Ang; cite them as *Sri Guru Granth Sahib Ji · Ang N*. `translit` is a reading aid; `en` is a labelled translation.
- `mode` is informational — it names the tier and may carry a suffix; do not parse it for logic.
- The search route is no-store (each query is answered afresh); reader routes are cacheable. See [Versioning and caching](#).
- Search is a **contract**: `contract/golden_search.ndjson` records real queries and their results, replayed by every deploy — see [Harnesses and golden vectors](#).

Sri Guru Granth Sahib Ji — KnowledgeBase wiki. Generated from commit c3d9360. Updated 26 October 2026. Verified against 7a343c6 on 2026-09-25. This copy may be out of date; the page is at <https://docs.gurbanisoul.com/search/modes-and-tiers/>