Oceanic Bhasha Renaissance – Methods Appendix (One Page)

Scope: Curate small, citable units—poem/inscription/itinerary snippets—
linked to concrete places in the Indian Ocean World (IOW), 2nd millennium BCE–1st millennium CE,
plus living ritual memories (e.g., Bali Jatra).

Selection:
• Prioritize primary sources: Vedic hymns, Sangam corpus, Greco-Roman itineraries,
  epigraphy (Kutai, Takuapa), archaeology (Arikamedu, Mantai), and festival songs.

Annotation:
• Each record = {work, snippet_en (paraphrase), original_lang, theme, place, lat/lon, period, source_type, citation_hint, tags}.
• Snippets are paraphrases to avoid copyright and to minimize misquotation risk;
  add precise line/verse IDs at editorial pass with source-in-hand.

Geocoding:
• Use well-known coordinates for excavated sites; for textual ports with uncertain locations (e.g., Muziris),
  pin the most widely accepted locus (e.g., near Pattanam) and mark as 'approximate'.

Quality Flags:
• Add columns later: confidence_site (0–1), confidence_date (0–1), citation_url (if open-access),
  and peer_reviewed (boolean).

Licensing & Provenance:
• Dataset released CC BY 4.0; paraphrases are original. Coordinates from public domain sources.

Pipeline Suggestions:
1) Enrich with multilingual text (Tamil/Odia/Sanskrit/Brahmi transliterations).
2) Link to facsimiles (images) where permissible.
3) Expose as GeoJSON and CSV for the Maps & Data page; power a Leaflet/MapLibre layer.

Caveat:
• All historical claims should be verified against editions/translations held in your archive or trusted libraries.