PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Reported — Title, p. 1: identifies the article as “A Systematic Review.”
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Reported — Abstract, p. 1: review design, PRISMA 2020, retrospective OSF registration 7fazq, sources, selection counts, qualitative synthesis, main findings, and implications.
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Reported — Sections 1 and 2: rationale, conceptual boundaries, prior-review positioning, and the focused Tiny Transformer evidence gap.
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Reported — Section 1 and Methodology, “Review objective and research questions”: explicit objective and RQ1–RQ6.
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Reported — Methodology, “Eligibility criteria”: inclusion and exclusion criteria are specified, including the explicit exclusion of methodological reporting standards and registration records from the qualitative evidence corpus and study-level appraisal. Included reports are grouped by research question, deployment strategy, hardware class, trustworthiness dimension, benchmark maturity, evidence-strength category, application domain, and evidence role.
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Methodology, “Information sources and search timing,” source-architecture table, and Appendix A identify all sources/interfaces. The registered review period is 2 March–1 June 2026 and the evidence set was frozen on 1 June 2026; source-specific execution dates are retained in the review audit trail rather than individually printed in the article.
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Reported — Methodology, “Search strategy and validation,” search-block table, and Appendix A with source-specific strategies, field syntax, filters, and sensitivity searches.
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation	Reported — Methodology, “Record management, screening, and

Section and Topic	Item #	Checklist item	Location where item is reported
		tools used in the process.	retrieval": two independent screeners, third adjudicator, pilot calibration, manual/identifier-based deduplication, no blinding, and Cohen's kappa used to assess agreement.
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Reported — Methodology, "Data collection process and data items": piloted extraction matrix; principal extractor; second-reviewer verification of high-impact fields; third-reviewer adjudication; ambiguous values checked against full text and coded as not reported when unresolved.
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Reported — "Data collection process and data items": accuracy/F1/AUROC/RMSE, latency, throughput, Flash/working memory, energy/power, trustworthiness and reproducibility outcomes; priority rules are stated for multiple configurations.
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Reported — Same subsection: bibliographic, model, hardware, runtime/toolchain, task, dataset, comparator, application, validation, and limitation variables; missing data coded as "not reported"; unsupported inference prohibited.
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Reported — Methodology, evidence-strength/deployment-readiness appraisal: seven engineering-relevant domains scored 0–2 (total 0–14), with verification/adjudication. The manuscript explains why a clinical risk-of-bias instrument is not appropriate for this heterogeneous engineering corpus.
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Not applicable for pooled effects — No meta-analysis or common pooled effect measure was used. Study-reported technical metrics were normalized where appropriate and synthesized qualitatively.
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Reported — "Eligibility criteria" and "Synthesis strategy and inference criteria": evidence roles/tier, research-question grouping, deployment thresholds, and direct-versus-adjacent evidence rules.
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Reported — "Data preparation and evidence coding": memory, latency and energy normalization; original context retained; no missing data imputed.

Section and Topic	Item #	Checklist item	Location where item is reported
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Reported — Structured evidence tables, source/study-selection tables, PRISMA flow diagram, appraisal tables, and RQ-level synthesis.
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Reported — Structured qualitative synthesis used because tasks, hardware, model families, precision regimes, and outcomes were heterogeneous; no meta-analysis performed.
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Reported — Heterogeneity interpreted by hardware class, precision/runtime, deployment strategy, evidence tier, application domain, benchmark maturity, and trustworthiness mechanism.
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Reported — "Reporting-bias assessment, robustness checks, and confidence in evidence": four pre-specified qualitative robustness checks (peer-reviewed only; real-device only; removal of adjacent edge-AI/SLM evidence; implemented/measured trustworthiness and sustainability only).
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Reported — Qualitative assessment of positive-results bias, hardware-reporting bias, benchmark-selection bias, incomplete energy reporting, and missing security-overhead measurements; effect-size publication-bias methods were not applicable.
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Reported — Narrative confidence based on evidence-strength results, hardware directness, measurement transparency, reproducibility, consistency, and evidence tier; no formal GRADE/CERQual was applied.
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Reported — Results, "Study selection"; PRISMA flow diagram and selection-summary table: 355 identified, 142 duplicates removed, 213 screened, 97 sought, 87 assessed, 6 excluded, 81 included.
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Full-text exclusion reasons and counts are reported in the "Study selection" subsection and exclusion table; the six excluded reports are not individually cited in the manuscript.

Section and Topic	Item #	Checklist item	Location where item is reported
Study characteristics	17	Cite each included study and present its characteristics.	Reported — All 81 reports included in the qualitative synthesis are individually identifiable and cited in the manuscript. Table A12 provides one study-level appraisal row for each of the 81 included reports. References [44] (PRISMA 2020) and [45] (OSF methodological registration) are methodological references only and are not part of the qualitative evidence corpus or Table A12. Detailed characteristics of the direct and representative evidence are additionally presented in the main-text evidence tables and RQ-level syntheses.
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Reported using the review's engineering evidence-appraisal framework — Study-level appraisal scores and evidence-strength classifications are presented in the appraisal appendix/table.
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Study-level technical outcomes are reported in structured tables and RQ synthesis. Conventional effect estimates/precision intervals are not uniformly applicable because the corpus comprises heterogeneous systems, benchmarking, and deployment studies rather than a common intervention-outcome design.
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Reported — RQ1–RQ6 synthesis, quality-appraisal results, evidence roles/tiers, and narrative-confidence judgments.
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Not applicable — No meta-analysis or pooled statistical synthesis was performed.
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Reported qualitatively — Cross-study differences are interpreted by task, hardware, precision, runtime/toolchain, deployment strategy, evidence tier, and measurement practice.
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Reported — Results, "Methodological robustness and confidence in the synthesis": conclusions remained stable under the four registered qualitative robustness checks, with more conservative trustworthiness interpretations under stricter evidence restrictions.
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Reported qualitatively — Reporting-bias patterns and their implications

Section and Topic	Item #	Checklist item	Location where item is reported
			are incorporated in the methodological robustness/confidence discussion; no funnel-plot or effect-size method was appropriate.
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Reported — Narrative-confidence table for RQ1–RQ6: moderate for model design/hardware feasibility, limited for security/privacy, uncertainty/reliability and application maturity, and moderate–limited for benchmarking/sustainability.
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Reported — Cross-Cutting Discussion and Conclusions.
	23b	Discuss any limitations of the evidence included in the review.	Reported — Limitations: evidence-base novelty, heterogeneity, benchmark comparability, sparse trustworthiness implementation, and limited field validation.
	23c	Discuss any limitations of the review processes used.	Reported — Limitations: retrospective registration, terminology/search sensitivity, heterogeneous evidence, and limits of qualitative synthesis.
	23d	Discuss implications of the results for practice, policy, and future research.	Reported — Open Challenges and Future Research Agenda plus Conclusions.
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Reported — Methodology, “Protocol, registration, and review period”: Open Science Framework, identifier 7fazq, retrospective transparency record.
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Reported — Public OSF record: https://osf.io/7fazq.
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	The manuscript identifies the OSF registration as retrospective and explains its role as a protocol-transparency record. No separate post-registration amendment history is reported.
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Reported — Funding statement: no external research funding; article processing charge funded by Escuela Politécnica Nacional.
Competing interests	26	Declare any competing interests of review authors.	Reported — Conflicts of Interest statement: authors declare no conflicts of interest.
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Reported in the revised Data Availability Statement — public OSF protocol/methodological record at https://osf.io/7fazq; evidence base is the cited published literature; search

Section and Topic	Item #	Checklist item	Location where item is reported
			strategies, source-level counts, PRISMA trail, exclusion categories, extraction framework, evidence-appraisal rubric, study-level scores, evidence-strength classifications, and synthesis procedures are in the article and appendices.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. This work is licensed under CC BY 4.0. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/