
SLAT PRELIMINARY REVIEW

SL2T 1.0

Google DeepMind

A criterion-level review of the public evidence record for Google DeepMind's sign-language-to-text model, conducted under the Sign Language Access Trust (SLAT) Index four days after public announcement.

Author	Heather M. Grizzle
Issued	16 August 2026
Subject announced	12 August 2026
Method anchor	SLAT Index Evaluation Standard v2.0
DOI	10.5281/zenodo.21537271
Publisher	Novara Consulting Group, Elk River, Minnesota

STANDING OF THIS DOCUMENT

This is a preliminary review, not a SLAT Snapshot and not a SLAT Trust Assessment. It issues no trust rating. It reports the state of the public evidence record as of the date of issue. A full Snapshot of SL2T 1.0 to the SLAT Snapshot specification is in preparation.

01 Purpose and standing

On 12 August 2026, Google DeepMind announced SL2T, a multilingual sign-language-to-text model, and began shipping it in consumer products through Gboard and Live Transcribe on the Pixel 11. The release is a significant event in sign language AI, and it arrived with an unusually substantial public governance record, including a joint impact report co-authored with an advisory committee of Deaf organizations. This document reviews that record under the Sign Language Access Trust (SLAT) Index, a criterion-based evaluation framework for sign language AI systems published by Novara Consulting Group.

This review is deliberately limited in standing. The SLAT method reserves trust ratings for a full assessment conducted against written evidence conditions, and reserves the Snapshot designation for documents built to the SLAT Snapshot specification, including a complete sweep of mandatory search locations and indicator-level scoring. Neither has yet been performed for this subject. What this document offers is the framework's first pass: a demonstration of what a structured governance instrument surfaces from the public record of a major release within days of launch, stated at the level of the Index's ten criteria across its five domains of Linguistic Trust, Transparency Trust, Governance Trust, Accessibility Trust, and Deployment Trust.

Two principles govern everything that follows. First, this review reports evidence availability, not system quality: a finding that the public record is silent on a question is a finding about the record, and the absence of published evidence is not evidence that a practice is absent. Second, favorable findings are stated with the same directness as unfavorable ones. On several criteria the SL2T record is the strongest this programme has examined.

02 Independence and interest

Novara Consulting Group has no commercial relationship with Google, Google DeepMind, or any affiliate, sought and received no payment or consideration for this review, and holds no financial interest in any sign language AI vendor. This review rests entirely on the public record cited in section 07. Google DeepMind and the AI Sign Language Advisory Committee are welcome to correct any error of fact or of location, and corrections will be recorded in any subsequent version.

03 Subject profile

SL2T 1.0 is a sign-language-to-text translation model announced by Google DeepMind on 12 August 2026 and deployed as an input method in Gboard and as a conversation feature in Live Transcribe on the Pixel 11, with additional devices and sign languages stated as planned. The initial release translates American Sign Language into English. The announcement describes training on more than 100,000 hours of data across more than 50 sign languages, with roughly a quarter of the data in ASL. Model input is not raw video but 130 two-dimensional landmark coordinates covering the face, body, and hands, extracted per frame using

MediaPipe Holistic. Google DeepMind and its AI Sign Language Advisory Committee (AISLAC) published a joint impact report alongside the release, describing the intended scope of SL2T 1.0 as a low-stakes assistive input tool and enumerating settings in which it should not be used.

The subject differs from most systems previously examined under this programme in two respects that shape the review. It translates from sign language rather than into it, so the criteria addressing rendered signing output read differently here. And it is distributed directly to consumers rather than procured by institutions, so the Index's procurement questions have, at present, no procuring counterparty to answer them.

04 Method and limitations

This review was conducted against three public sources: the Google DeepMind announcement of 12 August 2026, the AISLAC Joint Impact Report for SL2T 1.0, and contemporaneous press coverage. It was conducted with reference to the published SLAT Index Evaluation Standard v2.0. Findings are stated at the criterion level, in qualitative terms, rather than as indicator-level scores; indicator scoring is reserved for the Snapshot in preparation, where each of the instrument's forty indicators is assessed against a written sufficiency condition. Experience in this programme has shown that scoring without written conditions in hand drifts systematically toward leniency, and that discipline is precisely what the instrument exists to enforce.

Three limitations are material. The consumer legal surface of the deployment, including Google's terms of service, the Gboard and Live Transcribe privacy disclosures, and application store data safety labels, has not yet been examined, and for a consumer deployment those documents carry real weight. Sources have not yet been captured to archive. And this review reflects the public record as of 16 August 2026, four days after announcement; the record should be expected to develop.

05 Findings by domain

LINGUISTIC TRUST

LT-1 Linguistic Integrity

The record here is unusually candid and structurally compromised in the same breath. The impact report discloses that the model's input representation does not track the tongue, which the report itself identifies as a critical articulator in ASL, and that non-manual markers, including facial expressions, eyebrow position, and head movement conveying negation, questioning, and grammatical affect, can be missed. These are not peripheral features; they are load-bearing grammar. Variety coverage within ASL is not disaggregated, and the report defers evaluation across sociolects, naming Black ASL specifically, to a long-term research agenda. Comprehension-adjacent testing exists, including internal testing with Deaf employees, an external multi-day diary study with Deaf participants, and hands-on access by North American AISLAC members, but it is described at the level of process rather than results. The disclosure is strong; the demonstrated linguistic integrity is not yet.

LT-2 Evidence and Validation

This is the most complete benchmark disclosure this programme has examined: named evaluation corpora (FLEURS-ASL including a one-handed variant, PRESTO-ASL, and FSboard for fingerspelling), stated scores including a zero-shot BLEURT of 70 on FLEURS-ASL, and a stated construction method for the principal benchmark involving Certified Deaf Interpreters recorded in studio conditions. One point of precision is worth stating in the subject's favor: because SL2T generates English text rather than signed output, text-similarity metrics of the BLEURT class are an appropriate instrument here, and the familiar objection to such scores being offered as proof of signing quality does not apply. The limitation is of a different kind. Every evaluation in the record was conducted by the developer or within the project's own governance structure. Disaggregation across gender, age, ethnicity, region, and sociolect is explicitly deferred. Signers under 18 were neither trained on nor evaluated, and evaluation across signers with motor disabilities or atypical movement patterns is acknowledged as limited. No independent, third-party validation of any reported result was located.

TRANSPARENCY TRUST

TT-1 Transparency and Disclosure

Failure modes are named with mechanisms rather than gestured at: hallucinated text the report calls ghost text, predictive commitment before an utterance is complete, environmental sensitivity, single-subject tracking, short context windows, and difficulty with grammatical complexity. The disclosure is version-bound to SL2T 1.0 and the input architecture is described. The open question is placement. This disclosure lives in a report linked from a research announcement, and whether any material part of it reaches the surface where a Gboard user actually encounters the feature is exactly what the unexamined consumer legal surface would establish. Disclosure that does not appear where a user would ordinarily read it is a recurring weakness across this field, and it is the first thing the full Snapshot will test.

TT-2 Procurement Readiness

The subject's corporate standing, support apparatus, and versioning practices are documented at a scale without precedent in this field, and are presumably governed by Google's consumer terms, which this review has not yet examined. The finding of substance is structural: there is no procurement posture at all, because this is direct-to-consumer distribution. No solicitation, no contract, no acquiring authority. The instrument's procurement questions do not fail here; they land on nobody, which becomes significant at institutional accountability below.

GOVERNANCE TRUST

GT-1 Deaf and Disabled Governance

The record exceeds the field's norm, where claims of Deaf involvement typically rest on assertion alone. AISLAC's membership is named: the National Association of the Deaf, Rochester Institute of Technology's National Technical Institute for the Deaf, the World Federation of the Deaf, the Deaf Professional Arts Network, and a named independent subject-matter expert. The joint impact report is co-authored. But the

report describes AISLAC as an advisory body convened by Google, its external engagements are managed through two facilitators from Google's accessibility organization, and final decision authority rests with Google DeepMind. Member selection criteria and compensation are not disclosed. Advisory feedback is stated to directly inform findings, mitigations, and the roadmap, but no specific design change attributable to the committee's intervention is identified in the record reviewed. The consultation is well structured. Authority does not transfer.

GT-2 Data Governance and Privacy

This is the sharpest silence in the record. The announcement states that SL2T was trained on more than 100,000 hours of data across more than 50 sign languages. The impact report, which is the governance document of this release, does not state where any of that data came from, under what license it was obtained, what consent the signers captured in it gave to training use, or whether any contributor was compensated. Signed language data is not abstract: it is video of identifiable people, in a language where the face itself carries grammar, drawn from a community with long experience of being studied without consent. For a release this carefully governed in other respects, the training corpus is the question the record declines to answer. The inference-time record is considerably better: the landmark input representation is presented as a privacy safeguard replacing photorealistic detail with abstract coordinates, and the report states that no logs of user inputs or outputs are retained unless the user explicitly authorizes it.

ACCESSIBILITY TRUST

AT-1 Accessibility Impact

The intended population and settings are stated with precision, and the service-displacement question, which most of this field avoids, is answered head-on. The report states that the model must not be used by third parties to evade statutory accessibility obligations, and that Google DeepMind and AISLAC do not believe SL2T 1.0 satisfies obligations under the Americans with Disabilities Act, Section 504 or 508, or international frameworks that are met by qualified human interpreters. That is the strongest anti-substitution statement this programme has located in a vendor's own governance record. Fallback is inherent to the deployment surface, since the feature sits inside a keyboard whose ordinary function remains available throughout.

AT-2 Human Oversight

The oversight design relocates the human in the loop from the institution to the user. Generated text is presented for review and correction before it is used, and in Live Transcribe nothing is shown or voiced to the conversation partner until the Deaf user elects to release it. For live output, elective release is a real compensating control, and the report then does something unusual: it names the safeguard's own failure condition, stating that the design presumes functional English literacy and device proficiency, and that where those are absent, translation errors may pass unnoticed. That prerequisite will not hold uniformly across the population the tool serves, and for the users beyond it the oversight mechanism does not exist. No indication of model uncertainty is surfaced to the user, and no measurement of intervention or correction rates is reported.

DT-1 Deployment Context

The out-of-scope statement is the most detailed this programme has examined: healthcare including triage, emergency medical services, and informed consent; legal settings including interrogations, courtroom proceedings, and depositions; academic evaluation and classroom instruction including IEP meetings; employment decisions including interviews and disciplinary hearings; and government benefits determinations. The finding is the gap between that statement and the deployment. Under the SLAT stakes method, a deployment is classified by its highest foreseeable use, not its intended one, and a translation tool resident on a phone forecloses nothing. The emergency room is not an edge case of this deployment; it is a foreseeable use of it. The only control standing between the intended classification and the foreseeable one is a prohibition stated in a document most users will never encounter, with no technical enforcement described.

DT-2 Institutional Accountability

The weakest criterion in the record after data governance. The accountable contact in the impact report is a team email address in the document header. No user-facing route for reporting errors or harms is described, no named accountable role, no monitoring specifics, no public reporting cadence. Commitments to external evaluation are aspirational rather than binding: programmatic API access and open-weights release are described as having potential, and questions of model access governance are committed to further consultation. Independent researchers currently have no stated path to evaluate the system, and affected users have no stated path to be heard when it fails.

06 What this record demonstrates

Three observations reach beyond this subject.

First, SL2T presents the strongest day-one public evidence record this programme has examined, and the two deficits that recur across the entire field survive it essentially unimproved: training-data provenance and independent validation. Every system this programme has reviewed, without exception, lacks independent third-party validation of its output, and that now includes the best-resourced release in the field's history. Google's own scope restrictions concede the practical consequence: a system without independent validation is not deployable unassisted where the stakes are high. On that conclusion, the SLAT Index and Google DeepMind are in agreement.

Second, this is the first subject in which the privacy architecture and the linguistic deficit are the same engineering decision, disclosed as such. Reducing signers to abstract landmark coordinates is offered, credibly, as a privacy safeguard, and it is simultaneously the reason the tongue and much non-manual grammar are invisible to the model. Capture scope and linguistic integrity are usually separate questions in evaluation. This subject demonstrates that they can be a single fact, and that a design choice made for a protective reason can carry a linguistic cost that falls on the language itself.

Third, the oversight model deserves the field's attention. Placing review and release authority in the hands of the Deaf user is a genuine control, and it is more respectful of user agency than oversight architectures that interpose an institution. But its stated prerequisite, functional English literacy, is unevenly distributed across the population served, for reasons inseparable from the educational inequities this technology is meant to help address. An oversight mechanism whose availability tracks the very disadvantage the product addresses is a design question the field has not yet confronted, and no benchmark will surface it.

07 Questions institutions should ask now

SL2T is a consumer feature today. It will arrive in institutional settings regardless, carried through the door on the phones of the people those institutions serve, and the pressure to treat it as an accommodation will follow. An institution considering any reliance on this class of technology should be able to obtain answers to the following before doing so.

- 01 What evidence exists of the system's accuracy for the specific population the institution serves, including signers of Black ASL and regional varieties, older signers, children, and signers with motor disabilities, as distinct from aggregate benchmark performance?
- 02 Which error types has the developer measured in deployed use, at what rates, and which of those errors change meaning rather than wording?
- 03 What does the user see when the model is uncertain, and what evidence shows that users detect and correct errors in practice rather than in principle?
- 04 Where did the training data come from, under what consent from the signers captured in it, and with what compensation?
- 05 What route exists for an affected Deaf user to report a harmful failure, and to whom, with what obligation to respond?
- 06 What is the developer's stated position on use of the system in the institution's setting, and does the proposed use fall inside or outside it?
- 07 If the answer to the previous question is outside, what is the institution's documented justification for proceeding, and who is accountable for that decision?
- 08 What arrangement replaces the qualified interpreter this deployment would otherwise displace, and does the arrangement survive the developer's own statement that the system does not satisfy statutory accommodation obligations?

08 Sources

Google DeepMind, "Putting sign language AI into users' hands," 12 August 2026.
deepmind.google/blog/putting-sign-language-ai-into-users-hands/

Google DeepMind and AISLAC, "AISLAC Joint Impact Report for SL2T 1.0," 2026.

storage.googleapis.com/deepmind-media/DeepMind.com/Blog/putting-sign-language-ai-into-users-hands/aislac-joint-impact-report-for-sl2t-1-0.pdf

Novara Consulting Group, "SLAT Index Evaluation Standard v2.0," July 2026.

doi.org/10.5281/zenodo.21537271

Contemporaneous press coverage of the SL2T announcement, 12 to 14 August 2026, was consulted for context and no finding in this review rests on it.

NOVARA CONSULTING GROUP · GOVERNING TRUST BEFORE DEPLOYMENT

The Sign Language Access Trust (SLAT) Index is a criterion-based governance evaluation framework for sign language AI systems, developed under Deaf-centered governance principles and published by Novara Consulting Group. NCG advises procuring institutions, accessibility offices, disability organizations, and vendors on the evaluation of sign language AI before deployment. Consultation requests and the SLAT Executive Brief are available at novaracg.com.

© 2026 Novara Consulting Group LLC. SLAT and the SLAT Index are trademarks of Novara Consulting Group LLC (registration pending). This document is a preliminary review based on public sources, issues no trust rating, and does not constitute legal advice. Findings describe the state of the public record as of 16 August 2026 and may reflect limits of search rather than limits of practice; corrections are welcome and will be recorded.