IMPLEMENTATION OF AI : NETFLIX

AI-POWERED MULTILINGUAL CONTENT LOCALIZATION

@Team 3

2026

Pain Points

Fully reliant on human translation and LSPs

INCONSISTENT QUALITY

Each LSP recruits its own translators, resulting in varying quality standards and limited quality control by Netflix.

COMPLEX WORKFLOW

Translators and subtitle editors are different people, requiring frequent coordination and increasing overhead

LIMITED QUALITY METRICS

Metrics focus on vendor-level KPIs (e.g., rejection and on-time rates) and cannot accurately measure translation quality at the title or translator level.

LONG TURNAROUND TIME

Each language requires separate vendor management, scheduling, translation, and review, delaying global releases

Why This Matters ?

(01) **TRANSLATION NEEDS KEEP UP WITH GLOBAL EXPANSION**

(02) **SLOW LOCALIZATION KILLED THE GLOBAL LAUNCH STRATEGY**

(03) **POOR SUBTITLE QUALITY DIRECTLY HURTS RETENTION**

(04) **MANAGING LSP VENDORS AT SCALE BECAME UNMANAGEABLE**

AI READINESS ASSESSMENT

WAS NETFLIX READY?

ORGANIZATIONAL AUDIT ACROSS CULTURE, DATA, LEADERSHIP & TALENT

READINESS CONDITIONS THAT ENABLED DEPLOYMENT

CULTURE & DATA INFRASTRUCTURE

✦ CULTURAL READINESS

Netflix's engineering culture runs hundreds of A/B tests concurrently. When Hermes — their 2017–18 human translator certification system — failed, leadership shut it down rather than defend sunk costs. That willingness to kill a failing program and pivot to tech is precisely the cultural condition AI adoption requires. The internal narrative framed AI as freeing humans from repetitive work, not replacing them — reducing internal resistance.

- **FAILURE-TOLERANT**
- **EXPERIMENTATION CULTURE**
- **AUGMENTATION FRAMING**

✦ DATA INFRASTRUCTURE

Netflix's data moat was not volume alone — it was domain specificity. When the team tested Wikipedia data (WikiLarge), translation quality degraded across all 7 tested languages. Only Netflix's own subtitle corpus worked. Their content infrastructure also ensured data was well-labelled and accessible from day one.

- **9.5M LABELLED PAIRS**
- **12,301 TITLES**
- **IN-DOMAIN ONLY**
- **BLACK-BOX MT DEPENDENCY**

READINESS CONDITIONS THAT ENABLED DEPLOYMENT

LEADERSHIP & TALENT

✦ LEADERSHIP ALIGNMENT

AI localization had clear executive mandate — not a grassroots initiative. Netflix's 190-country expansion made fast multilingual content a competitive necessity. The Q4 2025 shareholder letter explicitly cited GenAI-powered promotional localization as a strategic priority. Operations teams were realigned from "translation owners" to "post-editors," a structural change requiring top-down authority.

- **BOARD-LEVEL VISIBILITY**
- **SHAREHOLDER LETTER CITED**
- **STRUCTURAL TEAM REALIGNMENT**

✦ TALENT & CAPABILITIES

Netflix used a hybrid build model — partnering with Virginia Tech (researcher Sneha Mehta) while leveraging internal NLP, MLOps, and NVIDIA V100 GPU infrastructure. The Transformer seq2seq architecture is mainstream, making talent acquisition feasible. Key gaps: no internal HITL workflow design capability, and LSP vendors needed reskilling from full translation to MTPE post-editing.

- **NLP + MLOPS TEAM**
- **GPU COMPUTE READY**
- **VIRGINIA TECH PARTNERSHIP**
- **MTPE RESKILLING GAP**
- **NO HITL DESIGN CAPABILITY**

READINESS SCORECARD

6 OF 8 CRITERIA MET

KEY GAPS THAT SHAPED THE HYBRID (HUMAN-IN-THE-LOOP) DESIGN

1. Error tolerance is partial — subtitle errors directly impact viewer retention, so AI drafts remain subject to human MTPE review. Full automation is not yet viable.
2. Personalization is market-level only — the agent adapts at the language/region level, not the individual viewer level. Viewer-level personalization remains a future state.
3. LSP vendor reskilling — vendors pivoting from full translation to post-editing required process redesign and change management that Netflix had to manage externally.

A strong fit for an AI agent

This use case checks 6 of 8 agent-readiness criteria . The error-tolerance split explains why humans stay in the loop.

FIT

SPECIALIZATION BOTTLENECK

Scarce, Costly translation.

FIT

24/7 AVAILABILITY

Pipeline runs nonstop

FIT

REPETITIVE/TEDIOUS

High-volume rote work

FIT

WELL-DEFINED PROCESS

Transcribe → translate → QC

FIT

MEASURABLE

Turnaround, quality/title

FIT

ERROR TOLERANCE

In draft, not final

PARTIAL

COMPLEX DECISION

Nuance left to humans

PARTIAL

PERSONALIZATION

Market-level only

Efficiency at Global Scale

This is an efficiency play. AI scales localization without scaling headcount, sustaining Netflix's global strategy.

EFFICIENCY **INNOVATION**

PRIMARY **ENABLER**

PERSONALIZATION **DISRUPTION**

Efficiency: Less waiting, less vendor management, less spend per language.

Innovation: The pipeline does the heavy lifting: listen, simplify, translate

Personalization: The adaptation happens at the market level, not the viewer level.

Disruption: It sustains Netflix's growth but it dismantles the traditional vendor model.

Augmentation, not automation

SIMPLIFY THEN TRANSLATE

PROOF OF CONCEPT

AN AI PROOF-OF-CONCEPT THAT IMPROVES TRANSLATION OF ENGLISH SLANG AND IDIOMS. IT SIMPLIFIES SOURCE SENTENCES BEFORE SENDING THEM THROUGH STANDARD MACHINE TRANSLATION, PRODUCING CLEANER LOCALIZED SUBTITLES.

The Implimentation Framework

(01) DEPLOYMENT GOAL

SCALE SUBTITLES WITHOUT ADDING HEADCOUNT, WITH HUMAN REVIEW

(02) TARGET

BREAKDOWN TO SIMPLIFIED ENGLISH FOR BETTER TRANSLATION

(03) PERFORMANCE

- MODEL LEVEL: QUALITY SCORES
- BUSINESS LEVEL: COST
- USER LEVEL: VARIETY OF TRANSLATION

(04) DATA

- EXTENSIVE
- PERENNIAL
- IMMUTABLE

(05) MODELLING

- TRANSFORMER SEQ2SEQ SENTENCE-SIMPLIFICATION
- BLACK-BOX NMT, PLUS ASR/TTS FOR DUBBING

(06) LAUNCH

BUILT ON THE GROUNDWORKS OF HERMES TO THRIVE

THANK YOU

REFERENCE

Alarcon, N. (2020, July 15). Netflix builds proof-of-concept AI model to simplify subtitles for translation. NVIDIA Developer Blog. https://developer.nvidia.com/blog/how-netflix-uses-ai-to-simplify-subtitles-for-translation-updated/

DigitalDefynd. (n.d.). 10 ways Netflix uses AI. https://digitaldefynd.com/IQ/ways-netflix-uses-ai/

ILC Worldwide. (2019, January 30). HERMES was a flash in the pan for Netflix's growing translation needs. Now what?https://ilcworldwide.com/2019/01/30/hermes-was-a-flash-in-the-pan-for-netflixs-growing-translation-needs-now-what/

Localization Institute. (n.d.). Case study: Netflix's AI-powered multilingual content localization. https://www.localizationinstitute.com/case-study-netflixs-ai-powered-multilingual-content-localization/

Mehta, S., Azarnoush, B., Chen, B., Saluja, A., Misra, V., Bihani, B., & Kumar, R. (2020). Simplify-then-translate: Automatic preprocessing for black-box machine translation (arXiv:2005.11197). arXiv. https://doi.org/10.48550/arXiv.2005.11197

Netflix. (2026, January 20). Q4 2025 shareholder letter. Netflix Investor Relations. https://ir.netflix.net/

Siegel, E. (2024). The AI playbook: Mastering the rare art of machine learning deployment. MIT Press.

Slator. (2018, October 16). Why Netflix shut down its translation portal HERMES. https://slator.com/why-netflix-shut-down-its-translation-portal-hermes/

University College London. (2022). Improving conditions for subtitlers and increasing Netflix's subtitled content. https://www.ucl.ac.uk/impact/case-studies/2022/apr/improving-conditions-subtitlers-and-increasing-netflixs-subtitled-content