



The State Bar *of California*

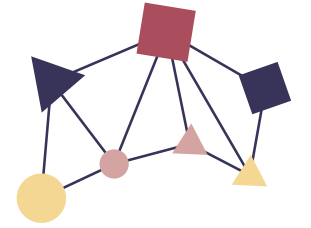
AI Translation Pilot Workshop

Tech & AI in Legal Aid Workshop Series

Yeonwoo Lee, Intern, Office of the Executive Director

Fredi Castillo, Senior Translator – Interpreter, Office of Chief Trial Counsel

May 28, 2025



- The Office of Access and Inclusion is excited to launch its **statewide Legal Technology Initiative** to help grantees understand and safely adopt emerging technologies to expand to access to justice at scale. As part of the initiative's launch, staff are forming a working group.
- Grantees interested in serving on the working group should email Jennifer Zelnick, Lead Program Analyst, at jennifer.zelnick@calbar.ca.gov, and Sylvia Nam, Senior Program Analyst, at sylvia.nam@calbar.ca.gov by June 4, 2025.
- The initiative includes a monthly workshop series in Tech & AI in Legal Aid. Stay tuned for more information about future workshops, which will be led by grantees.



EXTERNAL RECOMMENDATIONS

- 1** OA&I Structure and Governance
 - AI Models & Frameworks
 - Key Supplier Network
- 2** Resource & Network Sharing
- 3** People Resources & Training
 - Future Innovation Systems
 - Tech-Enabled Pro Bono Engagement

AI TRANSLATION PILOT WORKSHOP AGENDA



INTRODUCTIONS



**USE CASE AND
AI OPPORTUNITY**



**PILOT APPROACH AND
RESULTS**




**LESSONS LEARNED
AND TAKEAWAY FOR
GRANTEES**



INTRODUCTIONS

 Yeonwoo Lee, Intern, Office of the Executive Director

 Fredi Castillo, Senior Translator – Interpreter, Office of Chief Trial Counsel

Use Case

- The State Bar receives a large volume of complaints from non-English speaking witnesses
- In turn, the State Bar must translate those complaints and then communicate with the witness in the language the complaint was submitted

Considerations

- This information is statutorily confidential
- Many translated documents become court record, making accuracy critical
- The complaint process, including the translation step, has strict deadlines
- The State Bar has limited translators on staff and contracting these services is expensive

Most Common Languages

- Spanish
- Simplified Chinese
- Vietnamese
- Korean
- Russian

PILOT APPROACH AND RESULTS



**Step 1: Select
generative AI
models to test**



**Step 2:
Create testing
and evaluation
criteria**



**Step 3: Conduct
testing for each
AI model**

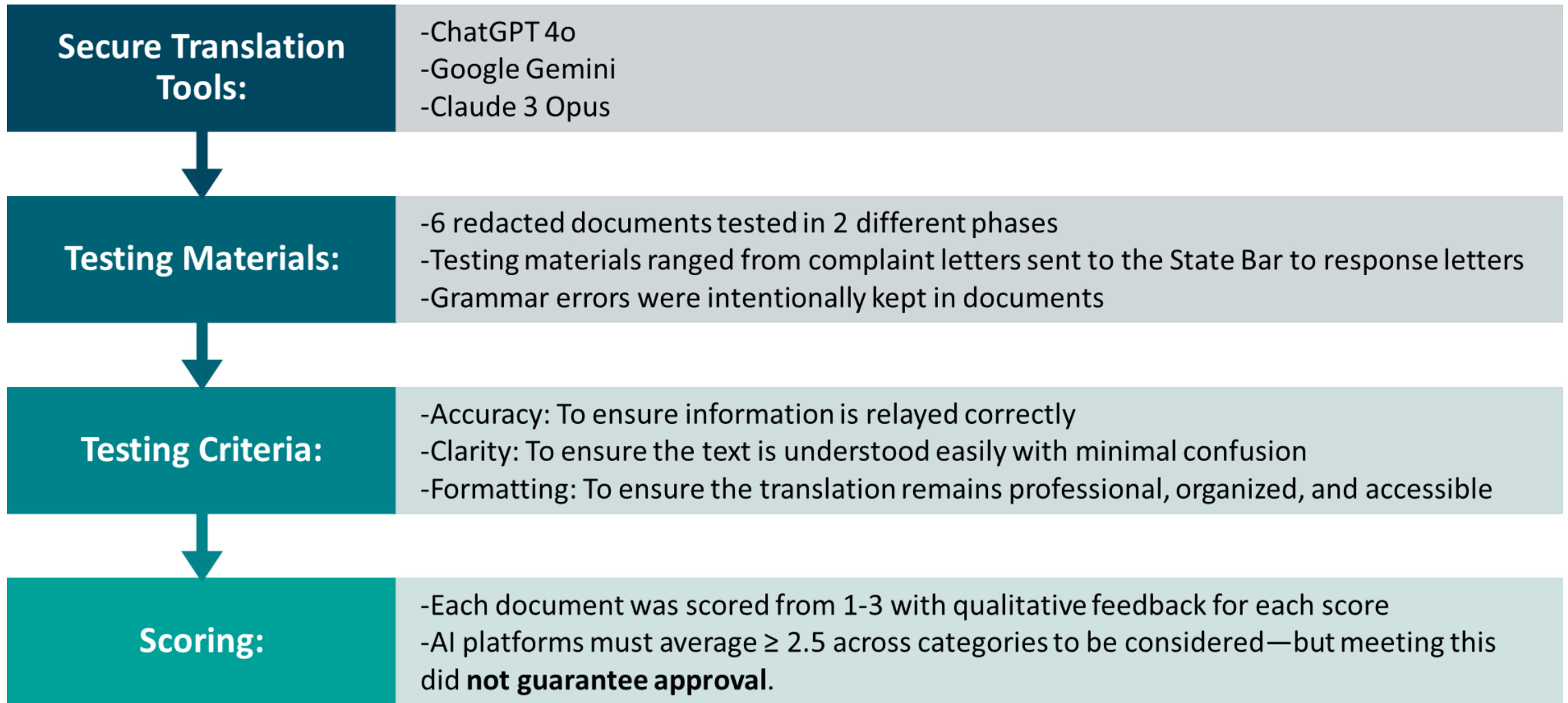


**Step 4: Design
and conduct
pilot**



**Step 5: Evaluate
and
operationalize**

STEP 1: SELECT AI MODEL & STEP 2: EVALUATION CRITERIA



STEP 3: THE TESTING PROCESS

Divide document in 300-word sections

Use same prompt for all AI models
“You are a native _____ speaker.
Translate so a native _____ speaker who
doesn’t speak English can understand.”

Translate each section with all models

Submit to third-party service for scoring

Analyze feedback and make formal
recommendations

STEP 3: TEST RESULTS

Language	Top Performer	Average Accuracy Score for Claude	Average Clarity Score for Claude	Average Format Score for Claude
Spanish	Claude 3 Opus	3	3	2.8
Simplified Chinese	Claude 3 Opus	3	2.6	2.5
Korean	Claude 3 Opus	2.7	2.5	2.6
Russian	None	2.2	2.4	2.6
Vietnamese	None	1.8	1.4	2.2

Each document was scored from 1-3 with qualitative feedback for each score

STEP 4: DESIGN PILOT, OVERVIEW

- Identified Spanish as the pilot language and engaged internal State Bar participant
- Claude procurement process
 - Sign off with Office of General Counsel
- Developed supporting documentation
 - Policy Document
 - Definition of major vs. minor errors
 - Error Documentation Tracker
 - Requirements for moving past human review
 - How-to-Guide
- Provided training

STEP 4: DESIGN PILOT, DEFINING ERRORS

Major Errors

Major Errors in translations are:

(1) Mistranslations that could alter the legal, technical, or contextual meaning of the original document

Example: Omitting a sentence that outlines important terms of a policy or content of the communication

(2) Translations that contain ambiguous or awkward phrasing that might confuse the reader regarding the meaning of the original document

Example: Changing the flexibility of a clause (e.g., translating "may terminate" as "must terminate")

Minor Errors

Minor Errors in translation are stylistic or formatting issues that do not impact the reader's understanding of the meaning of the original document. Minor Errors may involve small deviations from perfection but do not alter the meaning, confuse the reader, or compromise the professional quality of the document.

Example: Using generic instead of specific terms (e.g., "vehicle" instead of "car") in a scenario where the distinction is not essential.

STEP 4: DESIGN PILOT, POLICY DOCUMENT

Objective: Create a standardized policy to ensure high quality translations.

- Framework for stages of human review for AI translations

Level of Human Review	Criteria to advance to the next stage
Initial Stage: All documents human reviewed	10 consecutive translations of different documents all of which are free from any Major Errors
Progressive Sampling Stage: Periodic human review Daily --> Weekly --> Biweekly --> Monthly	5 consecutive random samples with no Major Errors, the sampling frequency will decrease to the next level in the progression.

- Privacy and security for AI translations
- AI use disclaimers for each level of AI usage
 - Fully AI generated with human review
 - Partially translated with human review
 - Fully AI translated with no human review

STEP 4: DESIGN PILOT, POLICY DOCUMENT

Objective: Develop tools to evaluate the State Bar Authorized AI Translation Tool.

Error Tracking

- Major v. Minor Errors
- Errors are documented to collect data on the progress of the State Bar Authorized AI Translation Tool

Date of completion	Case Number	Approximate time taken to complete the translation	Translation Case Number	How many minor Errors?	How many major errors?	Number of pages of the document that needs translation
7-Mar				0	13	2-Jan
3/10/2025				3	11	3
3/10/2025				0	3	3
3/10/2025		1.5 hours (othr wrk necsry)	-----	0	0	1
3/11/2025		< 1 hour		0	2	1
3/11/2025		2 hours (othr wrk necsry)				
3/11/2025		2 hours (othr wrk necsry)				
3/11/2025		< 3 hours				
3/13/2025		4 hours				
3/13/2025		4 hours				
3/14/2025		> 4 hours				
3/14/2025		3 hours				

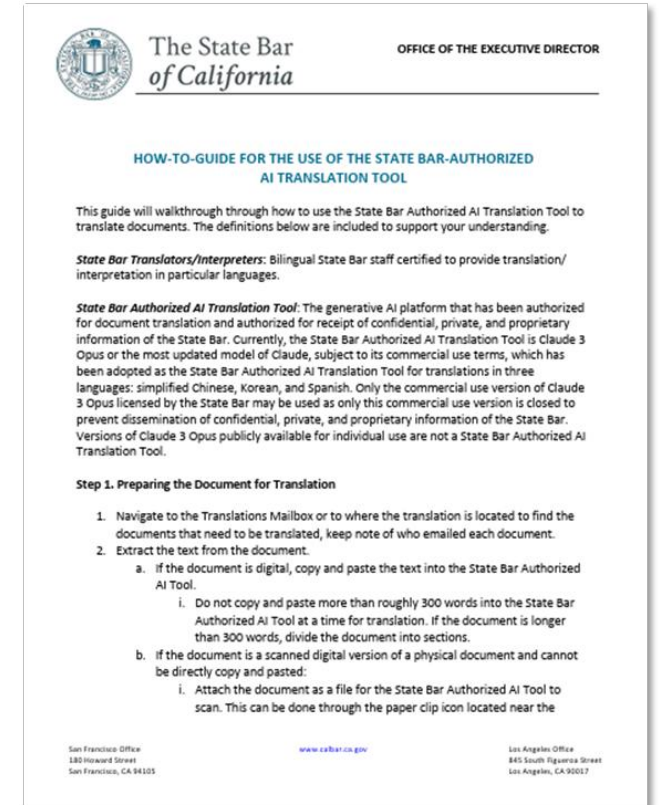
Description of the Error(s) Present in the Translation:	How to Correct the Error	Chatbot Model, Date of Error, and Language of Translation (Spanish to English, English to Korean, etc):	Did the chatbot correct the error after receiving feedback? If not, please describe its response or actions instead:
C omits the name "Juan;" eliminates witness	Added it, myself.	Claude. 3/7/25. Spanish to English.	Easier to add it, myself, than writing a command to Claude.
C omits "which were 2 processes"	Added it, myself.	Claude. 3/7/25. Spanish to English.	Easier to add it, myself, than writing a command to Claude.
C incorrectly placed "and he helped me;" thus altering statement.	Corrected it, myself.	Claude. 3/7/25. Spanish to English.	Easier to change it, myself, than writing a command to Claude.
C omitted "he was also paid in full;" important part of witness testimony.	Added it, myself.	Claude. 3/7/25. Spanish to English.	Easier to add it, myself, than writing a command to Claude.
C omitted "answered a single call;" important part of witness testimony.	Added it, myself.	Claude. 3/7/25. Spanish to English.	Easier to add it, myself, than writing a command to Claude.
C wrote "sign" instead of "pay," significantly altering statement.	Corrected it, myself.	Claude. 3/7/25. Spanish to English.	Easier to change it, myself, than writing a command to Claude.
C omits "I also don't know if he;" important part of witness testimony.	Added it, myself.	Claude. 3/7/25. Spanish to English.	Easier to add it, myself, than writing a command to Claude.
C writes "he has not done anything" instead of "I have not heard from them." Alters stmt.	Corrected it, myself.	Claude. 3/7/25. Spanish to English.	Easier to change it, myself, than writing a command to Claude.
C writes "keeps" instead of "takes." Serious.			Easier to change it myself, than

Error Documentation Tracker

STEP 4: DESIGN PILOT, HOW-TO-GUIDE

How-to-Guide

- Outlines how to use the State Bar Authorized AI Translation Tool to translate documents.
 - Step 1: Prepare the Document for Translation
 - Step 2: Using the State Bar Authorized AI Tool to Translate the Document
 - Guidelines on how to prompt the AI for the most accurate results
 - Tips on how to avoid errors during the translation process such as:
 - Unintended additions the model may add
 - Translating addresses
 - Step 3: Sending out Translations



How-to-Guide

STEP 4: CONDUCT PILOT

Pilot Overview

- Began in early March and will run through May
- To date, approximately 50 documents ranging from 1 page to 30 pages
- Utilized Error Documentation Table on an ongoing basis
- Regular check-ins between project team to answer questions

Learn and Adapt

- Errors were common and human review is crucial
- Translations of handwritten documents was more helpful than initially anticipated

STEP 5: EVALUATE AND OPERATIONALIZE

Results

- **Increased productivity by approximately 40%**
- Allows for more work to be done in less time
- Human review is still required
- Consistently experiencing major and minor errors

Common issues to be aware of when using a generative AI translation tool

- Doesn't accurately identify jargon and nuances
- Can make unnecessary corrections
- Is not always precise
- Makes assumptions or creates hallucinations

Operationalize

- After pilot, the State Bar will update policies based on results
- Continued periodic quality reporting
- Stay up to date on new products on the market as generative AI is rapidly evolving

Lessons learned for how to conduct generative AI pilots

- Establish framework for handling errors prior to start
 - Policies and tools for error management
- Customize for the organization and have pilot goals established early
- Generative AI can be trained based on an organization's priorities and use cases
- Need to be able to adapt

Lessons learned for using AI Translation tools

- Human review is required but can still enhance overall efficiency
 - Many generative AI tools will claim to provide translation that is “better than humans” but for legal documents, a certified translator is needed
- Can improve the process for handwritten translations

A pair of metal scales of justice is shown in a dark teal, semi-transparent overlay on the left side of the slide. The scales are centered vertically and horizontally within the teal area. The text "THANK YOU!" is overlaid on the scales in white, bold, sans-serif font.

THANK YOU!

Any questions?

