Voice-to-Text Language Services Outside the Courtroom

WORKSTREAM REPORT

OCTOBER 9, 2020

JUDICIAL COUNCIL OF CALIFORNIA
INFORMATION TECHNOLOGY ADVISORY COMMITTEE

Judicial Council of California

455 Golden Gate Avenue

San Francisco, California 94102-3688

www.courts.ca.gov

© 2020 by Judicial Council of California.

All rights reserved.

Except as permitted under the Copyright Act of 1976 and as otherwise expressly provided herein, no part of this publication may be reproduced in any form or by any means, electronic, online, or mechanical, including the use of information storage and retrieval systems, without permission in writing from the copyright holder. Permission is hereby granted to nonprofit institutions to reproduce and distribute this publication for educational purposes if the copies credit the copyright holder.

Document Revisions

Version	Date	Name	Change Description	Sections
1.0	10/2/2020	ITAC	Report for ITAC review	All
2.0	10/9/2020	Technology Committee	Report for Technology Committee review	No changes
3.0	1/22/21	Judicial Council	Report for Judicial Council review	No changes
4.0	2/1/21	Judicial Council	Final posted to website	No changes

Document Location

This document is maintained in the following location: https://www.courts.ca.gov/itac.htm

CONTENTS

1. EXECUTIVE SUMMARY 1
1.1 Recommendations 1
1.2 Report Structure 2
2. INTRODUCTION AND BACKGROUND 4
2.1 Directive from the Chief Justice and Formation of Workstream 5
2.2 Workstream Structure 6
2.3 Key Objectives and Concepts 6
2.4 Project Approach and Summary of Activities 7
3. METRICS AND EVALUATION TRACK 9
3.1 Workstream Track Considerations 9
3.2 Track Recommendation 11
4. TECHNOLOGY TRACK 12
4.1 Workstream Track Considerations 12
4.2 Track Recommendations 13
APPENDIX A: FUTURES COMMISSION DISCUSSION OF VOICE-TO-TEXT LANGUAGE SERVICES 15
APPENDIX B: WORKSTREAM MEMBERSHIP 18
APPENDIX C: EVALUATION SITE OVERVIEW 19
APPENDIX D: TEST RESULTS 22
APPENDIX E: SCRIPTS 23

1. EXECUTIVE SUMMARY

The Voice-to-Text Language Services Outside the Courtroom Workstream (workstream) was tasked by the Judicial Council's Information Technology Advisory Committee (ITAC) to explore available technologies to assist limited-English-proficient (LEP) customers at service counters and in self-help centers. The workstream's efforts were informed by recommendations of the Commission on the Future of California's Court System (Futures Commission) that the judicial branch pilot technology solutions allow "two individuals who speak different languages to converse without the assistance of an interpreter." (Commission on the Future of California's Court System, *Report to the Chief Justice* (Apr. 2017), p. 232 (hereafter *Futures Commission Report*).)

Chief Justice Tani G. Cantil-Sakauye directed ITAC, and by extension the workstream, "to explore available technologies and make recommendations to the Judicial Council on the potential for a pilot project using voice-to-text language interpretation services at court filing and service counters and in self-help centers." (Chief Justice Tani G. Cantil-Sakauye, mem. to Justice Douglas P. Miller, et al., "Addressing the recommendations of the Commission on the Future of California's Court System," May 17, 2017.)

This report provides the workstream's analysis of the current state of technology to enable real-time transcription and translation services and its potential for use in California courts and makes recommendations on the feasibility of a pilot project to test the technology in one or more courts.

1.1 Recommendations

The workstream approached its work and ultimately its recommendations with the following key concepts in mind:

- **Quality and Accuracy:** Any potential solution must be responsive, accurate, and accessible.
- **Security and Privacy:** Conversation confidentiality must be ensured.
- **Value:** The potential solution must be cost-effective.

The workstream recognized that existing consumer-grade solutions were available and widely used in web browsers, in stand-alone devices, and on mobile devices with minimal to no cost. However, these service providers leveraged the data gathered from free solutions for marketing and product development purposes. Given the confidential and sensitive topics discussed within the court, the workstream chose to focus its efforts on a customized solution that would allow for greater control of the data.

As a part of its evaluation, the workstream developed a proof-of-concept evaluation site in which to test existing offerings from vendors. Using scripts developed by workstream

members, the solutions were tested and scored for accuracy and responsiveness in both voice-to-text recognition and transcription as well as text-to-text translation.

In response to the Chief Justice's directive, the workstream determined that this technology could provide a significant benefit to the branch and developed its recommendations with the overall intent of identifying the path forward. Specifically, the workstream makes the following recommendations:

Recommendation #1: The Judicial Council should sponsor a project to deploy a pilot solution with the highest-scoring vendor from the proof-of-concept evaluation.

Given the importance of accurate translation and transcription within the court environment, the workstream recommends piloting with only the overall highest-scoring vendor from the evaluation phase.

Recommendation #2: Courts pursuing voice-to-text language services should consider enterprise solutions that offer a proven high level of accuracy and responsiveness while ensuring data privacy and confidentiality.

Enterprise-grade solutions allow the court to have full control over how—or if—conversation data is stored. This feature would offer the maximum level of court control of data to ensure privacy, security, and confidentiality. The solution should also take into account use at a socially distanced filing counter or self-help center.

Given the importance of communications between the public and court staff, the courts should pursue technology that offers the highest level of accuracy and is most feasible for use. The workstream recommends that courts use technology that features dictionaries that can continuously be expanded to include new and frequently repeated words and phrasing.

Recommendation #3: ITAC should collaborate with other appropriate advisory bodies to monitor the advances in voice-to-text language technology and advise how to expand its use in the branch.

The workstream believes it is important to continue to monitor the external environment for improvements and enhancements in voice-to-text technologies that would allow for additional or enhanced usage of the technology. Collaboration should continue between appropriate advisory bodies to develop guidelines for courts on the appropriate use of machine translation for web, computers, or tablets to assist LEP court users with accurate information in their language.

1.2 Report Structure

This final report provides the result from the workstream's work. Section 2 provides background information, the workstream structure, and key concepts that guided the workstream. Sections 3 and 4 provide an overview of the efforts made by the workstream's

two breakout teams (a Metrics and Evaluation track and a Technology track). The appendixes contain the original concept outlined by the Futures Commission, the workstream's membership, the evaluation site overview, the cumulative testing done and average scores by each vendor, and the English and Spanish scripts used to test the technology.

2. INTRODUCTION AND BACKGROUND

The Information Technology Advisory Committee formed the Voice-to-Text Language Services Outside the Courtroom Workstream to evaluate the feasibility of using automated voice-to-text transcription and translation services in non-courtroom settings. The workstream's evaluation was guided by the work of the Joint Working Group for California's Language Access Plan and the Futures Commission.

The language access plan (LAP), published in 2015, examined access to justice for LEP court users. "California is the most diverse state in the country, with approximately 7 million LEP residents and potential court users, dispersed over a vast geographic area, who speak more than 200 languages." (Judicial Council of Cal., *Strategic Plan for Language Access in the California Courts: Executive Summary* (Jan. 22, 2015), p. 1 (hereafter *Executive Summary*).) The LAP recommended that "[l]anguage access services must be available as an LEP court user enters the courthouse and at all points of contact within the courthouse, such as self-help centers, alternative dispute resolution services, and clerks' counters." (*Id.* at p. 2.)

Providing language access services, even in the courtroom, is a formidable challenge for courts of any size.

> Alpine County has 2 judges and 1 courthouse location, with no staff interpreters, and a total population of about 1,200. Los Angeles County, by contrast, has 477 authorized judges, 91 commissioners, and 26 referees. The Los Angeles court employs over 300 staff interpreters spread among its 600 courtrooms in 38 courthouses; they serve 10 million residents, spread across 4,800 square miles.

(Judicial Council of Cal., *Strategic Plan for Language Access in the California Courts* (2015), p. 12 [footnotes omitted].)

Budgetary and human resource constraints for courts of any size limit services for LEP court users in both small and large courts.

However, the LAP Joint Working Group saw that technology could help ease that burden. "Technologies such as video remote interpreting, telephonic interpretation, web-based access, and multilingual audiovisual tools have an important role to play in the statewide provision of language access." (*Executive Summary*, *supra,* at p. 3.) The Futures Commission agreed. "[A]dvances in technology, communications, and information processing all present opportunities for the judicial branch to give Californians greater, more efficient, and more responsive access to justice." (*Futures Commission Report, supra*, at p. 1.)

The Futures Commission sought practical ways to use technology to enhance the public's access to courts. Guided by the LAP, one of the Futures Commission's key recommendations was to develop a pilot project that would use voice-to-text language interpretation services at court filing and service counters and in self-help centers.

> Recent advances in voice-to-text language technology have been substantial and will continue to improve. Although these services are not yet accurate enough for hearings or trials, use of the technology within the courts for noncourtroom activities would greatly improve access for LEP court users. The technology can be customized, incorporating court-specific terms into the software. The voice-to-text language technology could be accessed by court staff on a tablet or other device to assist communication between court staff and LEP court users at clerk's counters, business offices, self-help centers, and other locations. Further, these translation services can be combined with intelligent chat technology to further enhance access for LEP court users. Use of this technology may replace other contracted services and their associated costs.

(*Id.* at p. 233; see Appendix A.)

Both the LAP Joint Working Group and the Futures Commission saw the value of emerging voice-to-text transcription and translation services to increase access to justice for LEP court users. The workstream sought to evaluate these technologies, consistent with one caveat from the LAP: "courts must exercise care to ensure that the use of technology is appropriate for the setting involved, that safeguards are in place for ensuring due process rights, and that high quality is maintained." (*Executive Summary*, *supra*, at p. 3.) Thus, the workstream evaluated the state of current voice-to-text and translation technologies with the key goals of speed, accuracy, availability, privacy, and security.

2.1 Directive from the Chief Justice and Formation of Workstream

After receiving the Futures Commission report, Chief Justice Cantil-Sakauye directed ITAC "to explore available technologies and make recommendations to the Judicial Council on the potential for a pilot project using voice-to-text language interpretation services at court filing and service counters and in self-help centers." (Chief Justice Tani G. Cantil-Sakauye, mem. to Justice Douglas P. Miller, et al., "Addressing the recommendations of the Commission on the Future of California's Court System," May 17, 2017.)

ITAC added this project to its 2018 annual agenda and launched the workstream in June 2018. The following tasks were included in the project:

- Define the standard of success and how to measure it as well as define the difference between translation and interpretation.
- Determine how, or if, the work for this initiative aligns with existing work of the Language Access Plan Implementation Task Force (LAPITF) and the work of The Legal Design Lab at the Stanford University Law School.
- Set up a technical lab environment at the Judicial Council or a local court to test the technical recommendations of the Futures Commission for this initiative.
- Test various voice-to-text language services in a lab environment, which will allow for exposure to more technologies and shorter learning cycles than if a specific technology is deployed at a court for piloting.

- Capture learnings and draft a white paper report on the lessons learned, findings, use cases, usage guidelines, and recommendations for next steps.
- At the completion of these directives, present findings and recommendations to, and seek approval from, ITAC, JCTC and, if appropriate, the Judicial Council. Formally sunset the workstream.

(Judicial Council of Cal., Advisory Com. mins., *Information Technology Advisory Committee (ITAC): Annual Agenda* (Jan. 2018).)

2.2 Workstream Structure

The roster of workstream members is included as Appendix B. The workstream was chaired by Judge James M. Mize, Superior Court of Sacramento County, and included participants from both appellate and trial courts, including judicial officers, technologists, operations staff, interpreters, and Judicial Council staff. During the exploration of the workstream, two tracks were formed: a Metrics and Evaluation track and a Technology track. The tracks met multiple times to develop initial recommendations on topics including technical requirements, minimum standards, evaluation scripts, accuracy, and confidentiality.

2.3 Key Objectives and Concepts

The objectives and concepts discussed below formed the foundation for the workstream's exploration into voice-to-text language services.

Quality and Accuracy

Given the recent advances in voice-to-text language technology, the workstream opted to set high standards for potential solutions: any solution considered should offer a proven demonstrated level of accuracy. In addition, because conversations at service counters and self-help centers include legal terms and usage that might be outside the vocabularies of consumer-grade products, the workstream decided that an acceptable voice-to-text solution must adapt to or learn domain-specific legal terms. However, as discussed in *Security and Privacy* below, any adaptation or machine learning must not allow third-party use of the data from stored conversations.

Security and Privacy

Several platforms offer voice-to-text language services, but many do not meet branch security and privacy needs. Even though conversations at court service counters and self-help centers are not legally confidential, these conversations often involve private matters that should be protected from unnecessary exposure. Consequently, any solution used by the courts should ensure the security and privacy of those conversations and vest control over the transmission and storage of those conversations with the court. In short, whether the transcription, translation, or storage of these conversations occurs in the courthouse or in the cloud, any solution must give the court control over who can access that data.

Value

Some counties struggle to provide sufficient language services because of limited resources; others do not have bilingual staff available. The ideal solution would provide a service to the branch at a more affordable cost than that of existing services, such as dial-in phone interpretation services.

2.4 Project Approach and Summary of Activities

The workstream undertook its work through three primary steps: market research, education sessions, and proof-of-concept development.

Market Research and Feasibility Evaluation

The workstream performed an initial assessment of current offerings in this area. Voice-to-text interpretation services, in a multitude of languages, were widely available on both mobile devices and web browsers. Many of these solutions used application program interfaces (APIs) to perform the tasks. In other words, the solution provider would create an interface for the user, but the actual transcription and translation would use a third-party service like those from Google or Amazon. These third-party APIs covered three phases of the process outside the user interface: first, converting spoken words to readable text; next, translating text from one language into text in another language; and finally, converting the translated text back to speech.

The workstream found that even though some solutions used their own translation engines, many products leveraged APIs offered by a handful of companies: Amazon Web Services Translate, Google Cloud Translate, IBM Watson Language Translator, and Microsoft Azure Translator. Each of these solutions met the workstream's privacy and security standards. They also had machine learning capabilities that could train the system to recognize legal terms in different languages. The workstream invited these vendors to hold education sessions about their platforms.

Education Sessions

Education sessions were held in June 2019 with the four vendors. Three vendors provided demonstrations in Sacramento, and one presented virtually. The sessions included an overview of the APIs, their functionality, and potential use cases. Workstream members were able to share potential use cases with the vendors, ask questions, and explore possible methods to deploy the technology.

Evaluation Site Development and Script Testing

Following the informational sessions, the workstream engaged with the Judicial Council's Information Technology department to develop a proof-of-concept evaluation site to further test the accuracy of the APIs. (See Appendix C.) A simple 0 to 5 scoring methodology was created to help evaluate the APIs. The functionality was limited to voice-to-text transcription, and text-to-text translation. For the purpose of the evaluation, the workstream opted to focus on English to Spanish. The Superior Court of Orange County provided scripts that

represented common interactions between court users and staff, translated from English to Spanish. The workstream recorded the scripts into the evaluation site and observed the test results of the four vendor solutions, scoring their accuracy. The transcription and translation accuracy results were compiled and used as a basis for the recommendation to conduct a pilot.

3. METRICS AND EVALUATION TRACK

To best evaluate available technology, the workstream determined early on that a quantifiable evaluation would best support any recommendations presented to the Judicial Council. Potential scoring and evaluation approaches and methodologies were discussed. Several members provided valuable insight from their experiences in language access services and input on existing standards and requirements for court interpreters. Because the directive from the Chief Justice specifically referenced creating solutions for service counters and self-help centers outside the courtroom (which have no established standard of translation), for the purpose of evaluating the proof-of-concept solutions, the workstream opted not to align criteria to courtroom standards.

3.1 Workstream Track Considerations

Segregating Transcription and Translation

In discussions with vendors offering these kinds of technology solutions the vast majority offered separate products or APIs for providing voice-to-text transcription, text-to-text translation, and text-to-speech output. Of the three functions, the workstream recognized that text-to-speech output enabling systems to "speak" is an extremely mature technology and thus did not evaluate any offerings by vendors. Instead, the workstream focused its evaluation on the other two components of an end-to-end solution.

The evaluation site developed by Judicial Council Information Technology staff allowed for the use of one vendor for voice-to-text transcription and a separate vendor for text-to-text translation. Hence, the workstream determined to evaluate these two steps separately.

Scoring Approach

Based on the recordings uploaded to the proof-of-concept evaluation site, the capacity of each solution to both (1) "hear" and transcribe the source language and (2) translate it into either English or Spanish was evaluated according to a five-point scale. The purpose of the five-point scale was to enable testers—who may not have a background in or experience with the evaluation of translation or interpretation—to evaluate the quality of transcriptions and judge the ability of the various products to convey meaning in another language against a standardized ranking procedure.

Transcription

Using the scripts provided by members from the Superior Court of Orange County, workstream testers uploaded a variety of recordings in English, ranging from simple one-sentence commands, to more complicated explanations and instructions consisting of several sentences. Each vendor's product was assigned one of the following rankings for transcription:

0 - Something went wrong

1 - Majority or all words recognized inaccurately

2 - Some words recognized accurately, but meaning not conveyed

3 - Majority of words recognized accurately; some follow-up questions could be asked to clarify the meaning

4 - All words were recognized accurately, but imperfect contextual meaning communicated

5 - Perfect word recognition and contextual meaning communicated

For the purpose of evaluating the transcription, rankings of 0 and 1 were reserved for situations in which there was some technical difficulty; rankings of 2 and 3 meant that the product was somewhat successful with an accurate transcription, but not enough to provide an intelligible message. Rankings of 4 and 5 recognize accurate 1:1 transcription of the discrete words and, in some cases, reflect accurate punctuation into coherent sentences.

Translation

Only perfect transcriptions were forwarded to the translation stage, to avoid confusing the evaluation of two separate functions. The translations were evaluated on a similar five-point scale.

0 - Something went wrong

1 - Majority or all words translated inaccurately

2 - Some words translated accurately, but meaning not conveyed

3 - Majority of words recognized accurately; some follow-up questions could be asked to clarify the meaning.

4 - All words translated accurately, but imperfect contextual meaning communicated

5 - Perfect translation and contextual meaning communicated

For translation, the purpose of the five-point scale was to acknowledge the possibility of variable performance in both 1:1 translation of discrete words that may be terms of art in the legal field, and the communication of meaning, which may require adjusting for the differences in syntactical structures between English and Spanish. Rankings of 0 and 1 were reserved for situations in which there was some technical difficulty or a complete inability of the product to discern the words; rankings of 2 and 3 reflected some number of accurately translated words (not all) without a coherent conveyance of meaning. There may also be in these two rankings the presence of false cognates, where a word that looks like a word in the other language is given a literal translation that does not take into account the legal or conversational context. Rankings of 4 and 5 were assigned to those products that displayed high levels of accuracy in both 1:1 translation and overall meaning.

Results of PoC Evaluation and Pilot Feasibility

It was quickly apparent to the workstream that a pilot program featuring voice-to-text language solutions assisting court customers with services outside the courtroom is feasible given current technology and should be pursued. Although continuously evolving, solutions

from several vendors demonstrate a sufficient level of both accuracy and responsiveness to provide a benefit to court customers and the branch. Low scores or challenges experienced while using the proof-of-concept evaluation site could potentially be attributed to technical difficulties by the user and not necessarily the lack of solution maturity.

Test results of the four vendors evaluated, for both transcription and translation, are available in Appendix D.

3.2 Track Recommendation

After concluding the evaluations, the Metrics and Evaluation track produced the following recommendations:

Recommendation #1: The Judicial Council should sponsor a project to deploy a pilot solution with the highest-scoring vendor from the proof-of-concept evaluation.

Based on the scoring results of the four vendors evaluated through the proof-of-concept site, the Judicial Council should move forward with deploying a pilot to several courts of varying sizes to further test the technology. Although all the solutions tested offer a large number of languages, the track recommends that the pilot focus on English-to-Spanish translation because Spanish is the most prevalent non-English language spoken in California.

The workstream recommends that a custom solution be developed using the technology of the highest-scoring vendor (cumulative of transcription and translation), using its APIs to facilitate dialog between court staff and the public. This pilot should include quantifiable methods to evaluate the interactions, with scripts of the dialog saved for further study of any trends in misspellings, omissions, structural mistakes, or other errors. These trends would assist the vendor in maturing the solution to best serve the public. Courts participating in a pilot would need to work cooperatively with any affected unions, consistent with any existing labor agreements.

4. TECHNOLOGY TRACK

The workstream began its investigation into the technology by exploring existing consumer-level services. This research resulted in the identification of a variety of potential solutions, including web-based applications, assistive technology devices for deaf and hard of hearing persons, smartphone applications, and purpose-built custom solutions. Additionally, the workstream experimented with home assistant smart speaker devices, collaborating with bilingual members and holding a mock discussion between a customer and court staff.

The workstream engaged with several vendors to hold education sessions to further explore available technology. Following these sessions, a custom proof-of-concept evaluation site was developed to test the APIs offered by four vendors: Amazon, Google, IBM, and Microsoft. Using a rubric developed by the Metrics and Evaluation track, the APIs were scored by workstream members. The scripts used for testing can be found in Appendix E (Scripts), with detailed results provided in Appendix D (Test Results).

4.1 Workstream Track Considerations

Available Technology

A fundamental finding of the workstream is that the technology exists today to facilitate a conversation between an LEP customer and court staff. The investigatory process found several widely available services, including offerings from Google, Microsoft, Amazon, and IBM. Additionally, these large vendors offered the services for custom-built solutions catered to meet unique needs. For example, the workstream met with SpeakSee, a vendor offering an assistive technology device that transcribes text to speech in real time. This vendor leveraged the API from Google to perform these services. Other products—such as the Pocketalk, a small two-way translation device—use their own custom software for transcription and translation services.

Product and Vendor Maturity

The workstream opted to engage four vendors—Amazon, Google, IBM, and Microsoft—in education sessions. Through these sessions, the workstream concluded that any potential solution should be in partnership with an established, mature vendor offering flexible enterprise solutions to meet the unique needs of the branch. This position was reinforced during the latter part of the workstream's efforts, when it was discovered that the partner company that attended the education session with Microsoft ceased operations altogether, with the caveat that the APIs used by this vendor are still available through Microsoft. Thus, the recommendation from the workstream is for courts to leverage proven solutions from mature vendors in this space.

Data Confidentiality

Although many free web-based solutions offer real-time voice-to-text language services, they do so at the cost of data confidentiality. Conversations captured by these solutions are harvested and used for analytical and third-party marketing purposes. Given the sensitive

nature of conversations at service counters and in self-help centers, the workstream recognized that these free solutions are not a viable option for use by the branch. Should a court opt to use voice-to-text language services, it should select a solution that gives the court control over the data captured from the interactions (whether locally or in the cloud).

Hardware Agnostic

The workstream recognized the diverse and unique needs of the courts throughout the state and identified the need for a flexible solution when considering service counter and self-help center layouts, staffing needs, privacy, and confidentiality. Any solution should be adaptable to these unique needs and not reliant on specific hardware. This finding is even more relevant given the COVID-19 health emergency and the need for physical distancing. The workstream opted to focus on cloud-based solutions, using applications that can be launched within a web browser. This setup would allow courts to deploy a solution—including the screen that displays the conversation and speakers that read out the translated dialog—in a way that best meets their needs.

Constantly Evolving Landscape

The workstream recognized the rapidly evolving technological capabilities in this market. In the education sessions with vendors, several shared a road map of additional languages to be added to their services in the coming months. Given these rapid advances in both accuracy and language offerings, the workstream recognized the need to monitor this technology segment to identify additional use cases for the branch.

4.2 Track Recommendations

Following the conclusion of its efforts, the Technology track produced two additional recommendations.

Recommendation #2: Courts pursuing voice-to-text language services should consider enterprise solutions that offer a proven high level of accuracy and responsiveness while ensuring data privacy and confidentiality.

As previously discussed, any voice-to-text solution procured should be a mature product sourced from a stable vendor. By leveraging enterprise solutions, courts can ensure that the conversations facilitated by the solution are confidential and the data produced from the interaction is managed by the court themselves. Additional functionality could be realized through enterprise solutions, including the use of machine learning and predictive analysis for greater accuracy and responsiveness, as well as the ability to expand and apply domain-specific contexts (such as in a legal setting). Furthermore, data gathered from the use of the enterprise solution could be used to identify trends—including frequency and type of questions, common challenges by the public, or other data points—that can help to improve services offered to the public. If the conversations are stored, a public record may be created. Accordingly, the Judicial Council or the courts should consider developing policies on not just storage and retention but also notice to users.

Recommendation #3: ITAC should collaborate with other appropriate advisory bodies to monitor the advances in voice-to-text language technology and advise how to expand its use in the branch.

In addition to recommending a Judicial Council–sponsored pilot, the workstream believes that the benefit of this technology can be expanded as it continues to mature. Furthermore, additional or overlapping considerations and efforts could be pursued and explored by other advisory bodies. The workstream therefore recommends that ITAC, with support from Judicial Council staff, collaborate with any related branch efforts in the area of voice-to-text language services.

Voice-to-text language technology is improving at a rapid pace, and other organizations in both the public and private sector are looking to voice-to-text technology to enhance their services. The workstream believes it is important to continue to monitor the external environment for improvements and enhancements in voice-to-text technologies that would allow for additional or enhanced usage of the technology.

This workstream supports ongoing collaboration between ITAC and the Language Access Subcommittee of the Advisory Committee on Providing Access and Fairness. Additionally, the voice-to-text pilot will provide helpful information for the branch on the development of guidelines and build on the findings of a previous pilot project conducted in 2018–2019 by the Stanford Law School Legal Design Lab, which worked with the Superior Court of Santa Clara County Self-Help Center and the Judicial Council's Language Access Implementation. (Stanford University Legal Design Lab, *Design Report: Language Access Innovations in Court: How Can Courts Use Technology & Design to Support People in Court When They're Not Proficient in English?* (Feb. 2019).)

APPENDIX A: Futures Commission Discussion of Voice-to-Text Language Services

(An excerpt on the Commission on the Future of California's Court System, *Report to the Chief Justice* (Apr. 2017), pp. 248–251, citations omitted.)

Rationale for Recommendation #5: Voice-to-Text Language Services Outside the Courtroom

California residents are among the most diverse in the country, with approximately 7 million speaking more than 200 languages. Without proper language assistance, LEP court users may be excluded from meaningful participation. Many courts have bilingual staff to assist some non-English-speaking users; however, they are usually limited to the most frequently used languages in that community. No court has staff fluent in the multitude of languages spoken by all court users. Court interpreters are also used when possible, but courts prioritize their services for in-court proceedings. Because court users can appear any time, scheduling interpreters on short notice is virtually impossible. Another limitation is the availability of interpreters for emerging languages spoken by newly arrived immigrants. Typically, these court users come to the public filing counters, self-help centers, and information desks. Court staff often find themselves assisting LEP individuals without an interpreter present.

In the absence of an interpreter, many court users rely on the help of a family member or friend. Often these individuals do not, themselves, understand legal terminology or court procedures. Friends and family members may also experience LEP, limiting their own availability to assist.

Some courts use telephonic interpreter services provided by a third party. The services are provided on demand in such settings as customer service counters, self-help centers, and other areas. These services can be provided in multiple languages. The cost for a certified telephonic language interpreter ranges from $1.49 to $1.99 per minute and $0.99 per minute for a noncertified interpreter. The vendor provides a single, toll-free number. From March 2016 to February 2017, the services under this master agreement were used by 17 courts.

Current technology can combine speech recognition technology and translation software. Speech recognition turns spoken language into text by a computer or other device. Speech recognition technology is used successfully by business organizations in various applications, including voice dialing for smart phones, data entry by phone in customer service calls, word processing by dictation, and language learning. More complex applications include military use of voice commands for fighter aircraft.

This technology integrated with translation software now allows two individuals who speak different languages to converse without the assistance of an interpreter. The process works as follows:

- When an individual speaks, his or her words are heard by the other participant. The text of the spoken words is displayed on screen in the speaker's language and immediately translated into the listener's language. When an individual finishes speaking, the software also provides an audio interpretation in the listener's language.
- At the end of the conversation, a transcript of the conversation is available, which includes a record of the conversation in each speaker's language.

Recent advances in voice-to-text language technology have been substantial and will continue to improve. Although these services are not yet accurate enough for hearings or trials, use of the

technology within the courts for non-courtroom activities would greatly improve access for LEP court users. The technology can be customized, incorporating court-specific terms into the software. The voice-to-text language technology could be accessed by court staff on a tablet or other device to assist communication between court staff and LEP court users at clerk's counters, business offices, self-help centers, and other locations. Further, these translation services can be combined with intelligent chat technology to further enhance access for LEP court users. Use of this technology may replace other contracted services and their associated costs.

The Futures Commission recommends developing a pilot project for the use of voice-to-text language interpretation services to serve court users at court filing and service counters and in self-help centers.

Successful application of this technology would enhance access in multiple languages conveniently, without court users having to wait for an interpreter, family member, or friend to translate for them. Use of this technology also allows court staff to print out the conversation for later reference by the court user, and to serve as a record of the information given. This technology can also enhance information available at self-help centers.

Pilot Project
The pilot project should include several courts, preferably of different sizes. The courthouses participating in the project should serve a large number of LEP court users, at the clerk's counter and in self-help centers.

Authorization needed to implement
No existing statutes or rules of court preclude the use of voice-to-text language services outside the courtroom. However, to implement the pilot project, participating courts would need to work cooperatively with any affected unions. Voice-to-text translation services must be used in a manner consistent with:

- The court's obligations under their respective regional interpreter memoranda of understanding;
- All applicable sections of the Trial Court Interpreter Employment and Labor Relations Act, as well as the Trial Court Employment Protection and Governance Act;
- The payment policies for contract court interpreters; and
- The Government Code sections, California Rules of Court, and Judicial Council forms applicable to the use of noncertified and nonregistered interpreters during court proceedings.

The pilot project will require funding to implement and to evaluate. Evaluation factors include:

- Frequency of use by location, case and proceeding type, and the duration of each session;
- Actual cost of devices and software for the court and a comparison to previous expenses for telephonic interpreter services from LanguageLine Solutions, if applicable; and
- Satisfaction of court staff and court users with the effectiveness of the interpretation in the various locations of use and proceeding types.

Evaluation of the pilot project will allow the judicial branch to assess the technology's usefulness and define best practices for using voice-to-text language services. If the pilot project is successful, minimum standards for its use should be developed and implemented branchwide to achieve the goals of Access 3D.

Costs to implement
Costs for the pilot project will vary based on size of court, number of courthouses, and number of clerk counters, as well as the device the court uses for this technology. The estimated cost of a laptop is $500, or $400 per tablet. Currently, voice-to-text language software is available on most devices at no charge.

Some courts currently use LanguageLine when the need for interpretation arises. Use of voice-to-text translation technology would replace use of LanguageLine and the associated costs.

Public comment
Public comment on this proposal was generally supportive. Some comments highlighted the need for funding assistance for some courts. No comments were received in opposition.

Feasibility of branchwide implementation or pilot project
The Futures Commission recognizes that with certain new processes, implementing a pilot project is more feasible and prudent than implementing a branchwide program. A pilot project provides the opportunity to gauge the impact on court and user interaction and to fine-tune a branchwide program. As such, a pilot project to provide voice-to-text translation services would be more feasible than branchwide implementation. The pilot project would provide information vital to future expansion.

The pilot project could include a few courts or a single court. If a single court is chosen, a medium-sized court with a known LEP court user population would be optimal. A participating court should have the flexibility to select the specific hardware to be used to access the voice-to-text translation service software.

This recommendation supports Goal 3 of California's language access plan, which states: "By 2020, courts will provide language access services at all points of contact in the California courts. Courts will provide notice to the public of available language services." The use of this technology will further assist LEP court users when prepared information, either electronic or printed in their language, may not address their particular questions.

If the pilot project is successful, extending its use, in conjunction with intelligent chat technology, would also support ITAC's SRL E-Services initiative, included in the *Tactical Plan for Technology (2017–2018).*

APPENDIX B: Workstream Membership

Hon. James Mize, Executive Sponsor
Judge
Superior Court of Sacramento County

Mr. Rick Walery, Business Lead
Chief Information Officer
Superior Court of San Mateo County

Mr. Richard Blalock, Project Manager
Senior Business Systems Analyst
Information Technology, Judicial Council

Hon. Jackson Lucky
Judge
Superior Court of Riverside County

Mr. Ryan Burkhart
Information Technology Manager
Superior Court of Sonoma County

Mr. Brian Cotta
Clerk/Executive Officer
Court of Appeal, Fifth Appellate District

Ms. Cynthia Gonzalez
Manager, Family Court Services
Superior Court of Sacramento County

Ms. Ana Parrack
Language Access Services Manager
Superior Court of Orange County

Mr. David Schlothauer
Chief Information Officer
Superior Court of Nevada County

Ms. Heather Pettit
Chief Information Officer
Information Technology, Judicial Council

Ms. Diana Glick
Attorney II
Center for Families, Children & and the Courts, Judicial Council

Ms. Camilla Kieliger
Senior Business Systems Analyst
Information Technology, Judicial Council

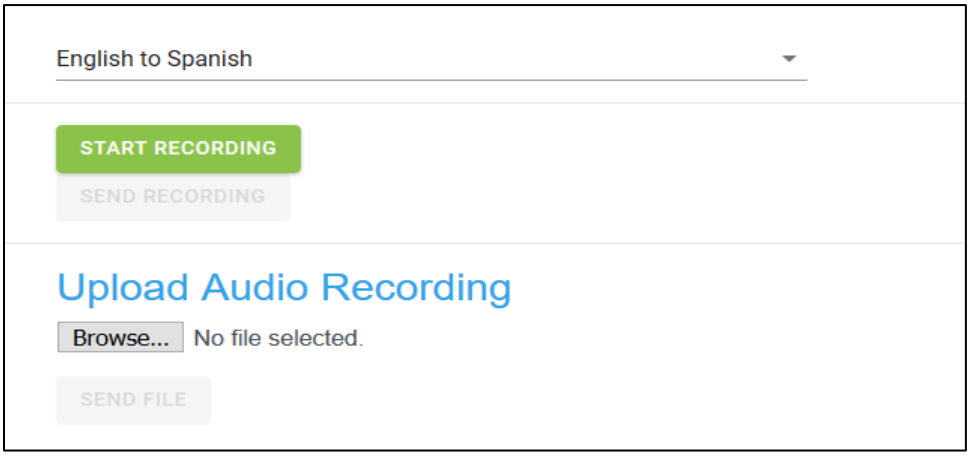
Ms. Claudia Ortega
Supervising Analyst
Court Operations Services, Judicial Council

Mr. Juan Palomares
Administrative Coordinator
Center for Families, Children & and the Courts, Judicial Council

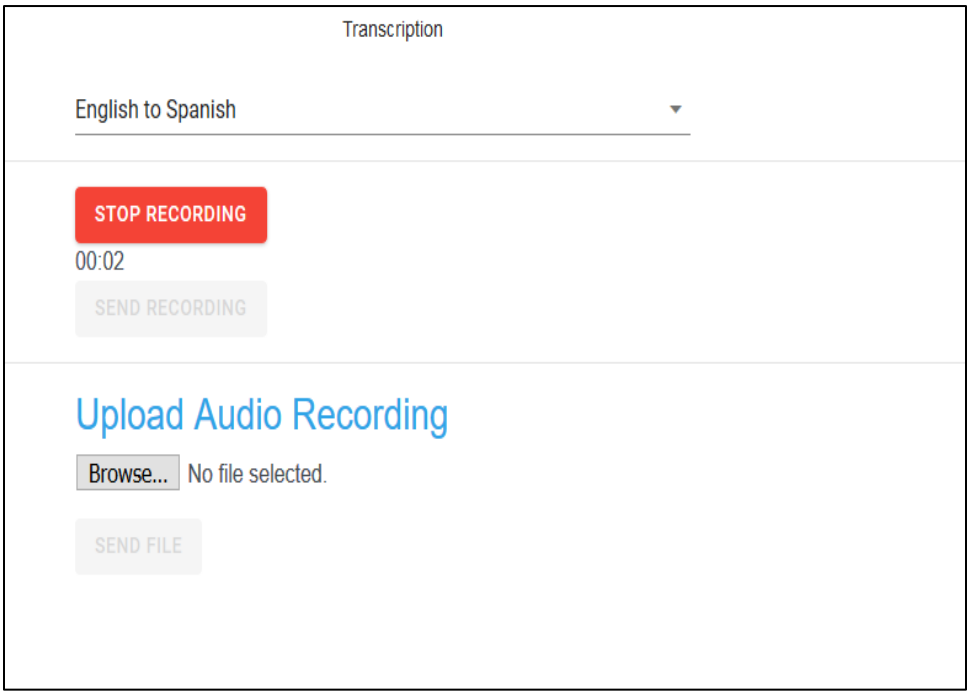
Mr. Glen Souza
Enterprise Architect
Information Technology, Judicial Council

APPENDIX C: Evaluation Site Overview

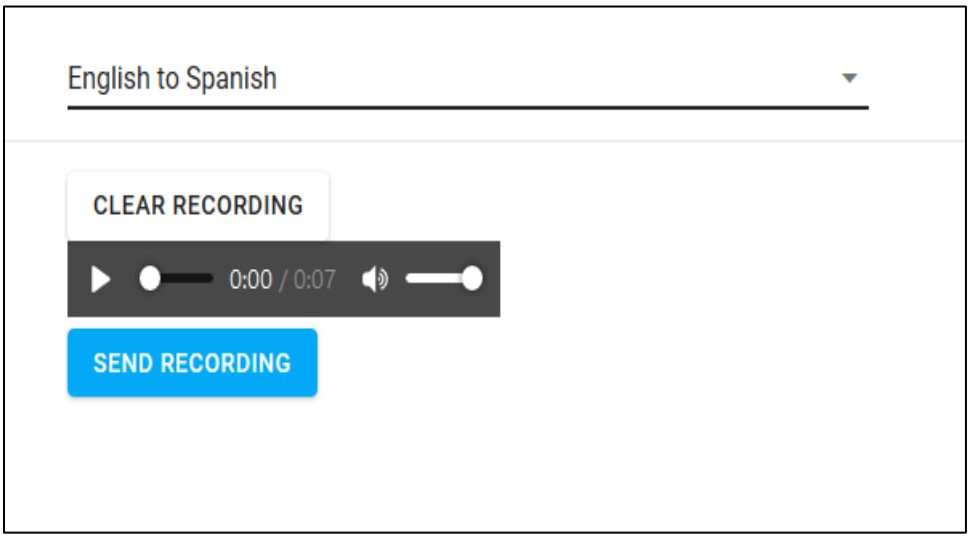
User begins recording by pressing "Start Recording." Alternatively, the user can upload an audio file:



When finished speaking, press "stop recording":

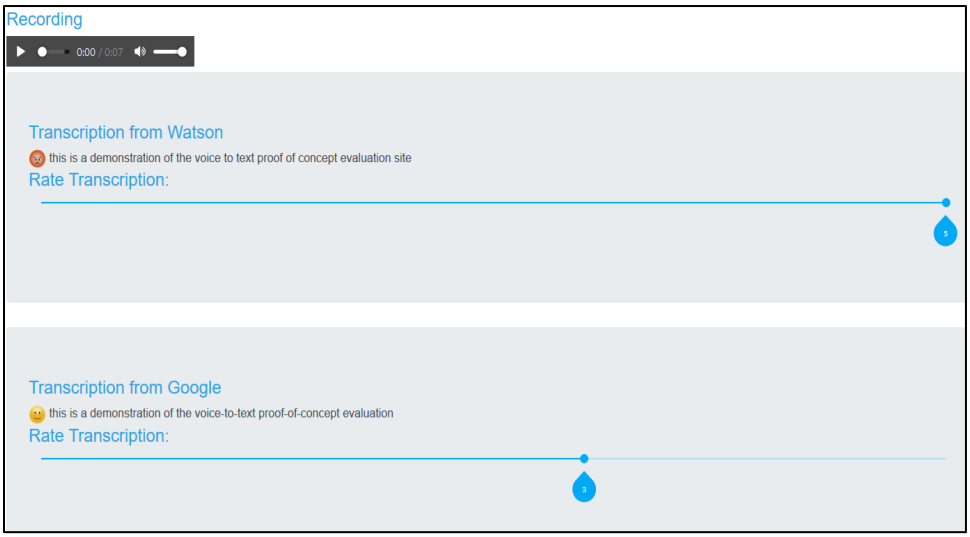


The user can then send the recording to be transcribed:



The user is taken to a new screen that shows the transcription of the recording by the four vendors in random order:

The transcription can be scored on a 0–5 scale. Following the rating, the vendor is shown:



The highest-scoring transcription result is then forwarded to the next page for translation by the four vendor solutions to be scored:

APPENDIX D: Test Results

Vendor	# of Transcriptions	Average Score
Amazon (AWS)	93	3.8
Google (GCS)	114	3.8
IBM (Watson)	112	3.6
Microsoft (Azure)	104	3.5
Total Transcriptions	**423**	

Vendor	# of Translations	Average Score
Amazon (AWS)	37	3.6
Google (GCS)	37	4.0
IBM (Watson)	34	0.5
Microsoft (Azure)	43	3.5
Total Translations	**151**	

APPENDIX E: Scripts

S = Staff

P = Public

S: Hello, how can I help you.
Hola, ¿cómo puedo ayudarte?

P: I need to see the status of my case.
Necesito ver el estado de mi caso.

S: Please give me your name and case number.
Por favor dáme tu nombre y número de caso.

P: My name is Nicole Robles, and my case number is 14d035999.
Mi nombre es Nicole Robles y mi número de caso es el 14d035999.

S: Ok, give me a second so I can look up your case. This is an open case. Custody. Correct?
Ok, dáme un segundo para buscar tu caso. Este es un caso abierto. Custodia. ¿correcto?

P: My ex-husband will not leave me alone.
Mi esposo no me deja en paz

S: Do you want to file a restrating order?
¿Quieres presentar una orden de restricción?

P: I don't think I will get a restraining order. I just need him to stop.
No pienso conseguir una orden de restricción. Solo necesito que (él) pare.

S: I'll give you the form that you can take to the restraining order department.
Te voy a dar el formulario que peudes llevar al departememto de órdenes de restricción.

P: I just need him to stop.
Sólo necesito que (él) pare.

S: Nicole, right?
Nicole, ¿correcto?

P: Yes.
Sí.

S: You live in Brea?
¿Vives en Brea?

P: Yeah. I tried to go to the police. I just need him to stop.
Ey. Trate de ir con la policia. Solo necesito que (él) pare.

S: Do you need a fee waiver?
¿Necesitas una exencion de cuotas?

P: No.
No.

S: Do you need a referral to our paralegals? Today we close at 3 p.m.
¿Necesita una referencia a nuestros paralegales? Cerramos hoy a las 3 p.m.

P: Yes, thank you.
Sí, gracias.

S: Here you go, we have appointments at 12:30, 1:30, and 2:30.
Ahí te va, tenemos citas a las 12:30, 1:30, y 2:30.

P: 1:30 please.
1:30 por favor.

S: Here is your referral. Please bring this with you and wait in that line at 1:30 pm. A paralegal will come get at that time.
Aqui esta tu referencia. Por favor traigala contigo y espere en la cola a la 1:30 pm. Un paralegal vendrá a esa hora.

P: Thank you.
Gracias.

S: You're welcome.
De nada.

S = Staff

P = Public

S: Hello, how can I help you.

Hola, ¿cómo puedo ayudarte?

P: I need a fee waiver.

Necesito una exención de cuotas.

S: Is this for your divorce?

¿Esto es para tu divorcio?

P: Yes.

Sí.

S: Did you already file your divorce?

Ya presentó (los documentos/papeles) para tu divorcio.

P: No.

No.

S: Do you live in Orange County?

¿Vives en el condado de Orange?

P: Yes.

Sí.

P: Do have the forms you need to file your divorce?

¿Tiene los formularios que necesita para presentar su divorcio?

P: No.

No.

S: I'm going to give you the forms you need to start your case and the fee waiver form. If you need help filling out the forms, we have workshop on Wednesdays. Here is the information on the workshop.

Te voy a dar los formularios que necesitas para iniciar tu caso y el formulario de exención de cuotas. Si necesitas ayuda para llenar los formularios, tenemos talleres en los Miercoles. Aquí esta la información sobre el taller.

P: Thank you.

Gracias.

S = Staff

P = Public

S: Hello, how can I help you.

Hola, ¿cómo puedo ayudarte?

P: Do you speak Spanish?

¿Hablas español?

S: Yes. How can I help you?

Sí ¿cómo puedo ayudarte?

P: I need to file for custody of my son without the father knowing.

Necesito presentar (pedir) custodia de mi hijo sin que el padre lo sepa.

S: Someone still has to serve the father to give him notice. Here are the forms you need to start your case. Do you need a fee waiver?

Alguien todavia tiene que entregar los documentos (papeles) al padre para darle aviso. Aquí están los formularios que necesitas para iniciar el caso. ¿necesitás una exención de cuotas?

P: Yes. Where do I submit these forms?

Sí. ¿dónde presento los formularios?

S: You can get assistance of a paralegal to explain the process and you can attend the family law workshop.

Puedes conseguir ayuda de un paralegal para explicar el proceso y puedes asistir al taller de derecho de familia.

P: I also need to ask for child support.

También necesito pedir manutención de hijos.

S: The forms are included here. And here is the information on the family law workshop.

Los formularios están incluidos aquí. Y aquí esta la información sobre el taller de derecho de familia.

P: Thank you.

Gracias.

S = Staff

P = Public

P: ¿Habla español?

Do you speak Spanish?

S: Sí, como le puedo ayudar?

Yes, how can I help you?

P: No sé qué hacer. Tengo dos niñas. El mánager de donde vivo viene y me estresa. Me dice que me tengo ya que salir. Me fui al hospital por tan mal que esto me está haciendo. Tengo dos niñas. Yo les afecto. No se cuáles son mis derechos en este caso. Ya me llamaron de la escuela. Y mis hijas no se pueden concentrar por todo este estrés.

I don't know what to do. I have two daughters. The manager where I live comes and stresses me out. He tells me that I have to leave. I went to the hospital for how ill this is making me. I have two daughters. I affect them. I don't know what my rights are in this situation. They've already called me from school. And my daughters cannot concentrate because of all this stress.

S: ¿La mandaron a qui para buscar una orden de restrinjo?

Did they send you here to seek a restraining [sic] order?

P: Me desalojo la semana pasada. El jueves me dio una carta diciéndome que tenía que darle mantenimiento a la yarda. Me empezó a decir muchas cosas.

I was evicted last week. On Thursday I was given a letter telling me that I have to pay him for yard maintenance (gardening). He started saying a lot of things.

S: ¿Le dio documentos?

Did he give you any documents?

P: Me dio tres días para limpiar. Ya no se que hacer. Tengo miedo.

He gave me three days to clean up. I don't know what to do. I'm scared.

S: No podemos dar le aviso legal. Pero puede pedir una orden para que no las molesten. La corte le da una opción para poner una restricción. Aquí a unos números de personas y servicios que le pueden ayudar. Esto son los formularios para una restricción. Esta se lleva a la corte en Fullerton allí ay otro centro. Alguien allí le puede ayudar también.

We cannot give you (him) legal notice. But you (he) can ask for an order not to bother them (daughters). The court gives you (him) an option to place a restriction (restraint). Here are some numbers of people and services who can help you (him). These are the forms for a restriction (restraint). This can be taken to the Fullerton court there is another center there. Someone there can also help you.

P: Sí, es que a el no le importa. El me amenaza, y tengo mucho miedo que el me vaya a dar un mal golpe. Tiene a mis niñas con mucho miedo. Les está afectando demasiado ya en la escuela. ¿Tengo derecho a llamarle a la policía?

Yes, well he doesn't care. He threatens me, and I am very afraid that he is going to hit me. He makes my daughters very scared. It is really affecting them at school. Do I have the right to call the police?

S: Uno siempre tiene derecho a llamar a las autoridades.

One always has the right to call the authorities.