AI Interpreting Solutions Evaluation Toolkit

Ensuring the Integrity of Communication

PART A: Organization, Implementation and Management

Developed by:

Stakeholders Advocating for Fair and Ethical AI in Interpreting Task Force (SAFE AI)

Coalition for Sign Language Equity in Technology (CoSET)

Publication Date: August 2025

Document Classification: For Public Distribution

Recommended Citation

AI Interpreting Solutions Evaluation Toolkit: PART A: Organization, Implementation and Management. Developed by Stakeholders Advocating for Fair and Ethical AI in Interpreting Task Force (SAFE AI) & Coalition for Sign Language Equity in Technology (CoSET). August 2025.

Contact Information

For questions about this Toolkit or implementation guidance contact:
www.safeaitf.org or www.coset.org

Table of Contents

PART A: Introduction

About This Toolkit

This Toolkit helps AI adopters—organizations that are considering automated interpreting by artificial intelligence (AI)—understand the potential benefits while keeping safety, fairness, and ethical concerns as top priorities for the principal communicators who will use these technologies for multilingual communication.

The Toolkit has three parts that cover evaluating and purchasing AI interpreting tools, technical specifications and capabilities of these tools, and impacts on limited English proficient (LEP) and Deaf and Hard of Hearing communities. These impacts include English speakers in every part of society who need multilingual assistance to work with and understand people who use the many different spoken and signed languages in our communities and who need meaningful and equitable access.

How to Use Part A of This Toolkit

Part A: Organization, Implementation and Management of the *AI Solutions Interpreting Evaluation Toolkit* provides a comprehensive, risk-informed approach for evaluating AI and hybrid AI-human language access solutions. It is designed to be a practical resource for both strategic planning and solution implementation. The checklists included in this PART A provide useful evaluation questions relevant for all languages and language combinations.

Other publications in the *AI Interpreting Solutions Evaluation Toolkit* series include:

> **PART B: Technical Specifications** – A detailed description of User Experience / User Interface (UX/UI) controls and language model baselines, including metrics related to AI and hybrid language solutions.
>
> **PART C: Legalities and Practical Considerations** – A detailed consideration of the contemporary legal landscape and practical considerations for early adoption of AI and hybrid language solutions.

PART A of this toolkit series has *two sections*: a Toolkit overview and the Toolkit checklists.

Comprehensive Overview (The "Why" of a Risk-Informed Approach)

The Overview for *Part A* explains the foundational approach and framework for this resource. It provides the "why" behind both the structure of the evaluation checklists and the emphasis on risk management, which are grounded in the following core principles:

- **SAFE:** Ensure that solutions will not harm primary communicators and use robust safeguards for all encounters.
- **ACCOUNTABLE:** Establish clear and transparent lines of responsibility for the technology's performance and impact. Patterns of user feedback and performance monitoring are publicly disclosed, including changes stemming from these metrics, on a regular, frequent basis.
- **FAIR:** Promote autonomy for all primary communicators and actively work to reduce algorithmic bias to achieve the integrity of communication.
- **ETHICAL:** Guarantee transparency about safety, accountability, fairness, and user consent (opt-out / prefer human controls) when integrating AI interpreting.

This Toolkit draws attention to known limitations in current AI interpreting technology which:

- Functions optimally only under controlled conditions.
- Requires adequate audio and visual quality for all participants.
- Cannot replicate human interpreters' cultural competency, contextual understanding, and problem-solving abilities.
- Performs inconsistently across different language pairs (the combination of languages used in the interpreting interaction, such as ASL and English, or English and Vietnamese).
- May create access barriers for languages with minimal training data, which are called "low resource languages."

Toolkit Checklists: Implementation (The "How")

When you are ready to begin the evaluation process, proceed to the practical tools located in the Checklists section, which are organized in a three-step sequence:

Step 1: Assess Your Organization's Readiness

Use **CHECKLIST 1: Organizational Readiness Evaluation** to determine if your organization is prepared to implement and manage this technology responsibly.

Use **CHECKLIST 2: Setting-Specific Considerations for AI Interpreting Implementation** as appropriate to your setting and evaluation needs.

Step 2: Evaluate Risk Factors to Determine Risk Levels

Use **CHECKLIST 3: Risk Factor Assessment Framework for AI Interpreting Solutions** to categorize your specific use cases and determine which encounters are broadly appropriate for AI versus a human interpreter, prioritizing the ability for primary communicators to opt out of AI interpreting and switch to a human solution, if preferred.

Step 3: Select and Evaluate Vendors

Use **CHECKLIST 4: Vendor Assessment Checklist** to systematically evaluate risk levels posed by potential AI solutions against your specific requirements and the SAFE principles.

Use **CHECKLIST 5: Guidance for Request for Proposals (RFPs) that Include AI Interpreting** if you are writing RFPs for interpreting services that include AI interpreting tools.

Match the AI Interpreting Solution to the Task

Before selecting an AI interpreting solution, it is essential to understand the dynamics of using it in the real world. A successful implementation balances the benefits of technology with the needs of the primary communicators. To do this, consider the following:

Align Technology with the Environment.

AI interpreting solutions perform best in controlled settings. They are most suited for simple, transactional tasks in highly constrained, predictable contexts.

Reserve Less Predictable and Nuanced Tasks for Human Interpreters.

Situations involving measurable risk require cultural competency, ethical judgment, empathy and facilitation from human interpreters.

Validate Performance for Your Specific Communities.

Always pilot an AI interpreting solution with your primary communicators, especially in non-English languages, before implementing to ensure it can be reasonably expected to produce understanding among principal communicators

Prioritize Effective Understanding among Principal Communicators.

Ensure that pathways to a human interpreter remain visible and easy to enact, so that technology does not become a barrier for any community member.

PART A: Overview

This Toolkit provides organizations with a systematic framework to evaluate AI interpreting platforms and applications. It is meant to help decision-makers identify potential problems, assess risks and liability, and determine appropriate implementation strategies. It outlines the requirements for both spoken language interpreting and sign language interpreting to ensure accessible communication for all.

The *AI Interpreting Solutions Evaluation Toolkit* is based on the following publications:

- *Interpreting SAFE AI Task Force Guidance on AI and Interpreting Services* published by the Stakeholders Advocating for Fair and Ethical AI in Interpreting Task Force (SAFE AI) and the Coalition for Sign Language Equity in Technology (CoSET).
- *Deaf-Safe AI: A Legal Foundation for Ubiquitous Automatic Interpreting* by CoSET.
- *Perceptions on Automated Interpreting and Automated Speech-to-Speech Interpreting, produced* by CSA Research.

Three-Pillar Assessment Framework

PART A: Organization, Implementation and Management of the *AI Interpreting Solutions Evaluation Toolkit* addresses the convergence of changes to federal language access policy, rapidly evolving AI interpreting technologies, and AI vendors entering the market at an accelerating pace. These factors are fundamentally reshaping the landscape for organizations that have traditionally structured and funded their language access services based on prior federal guidance and regulatory frameworks.

Three-Pillar Assessment Framework

This Toolkit is designed for decision makers who need to systematically evaluate AI platforms and applications used for language services, especially in regulated industries such as healthcare, education and the justice system.

The evaluation framework is built on three decision pillars:

1. **Organizational Readiness**: Is your organization prepared to adopt and manage this technology?
2. **Technical Fitness**: Does the AI interpreting solution effectively support the needs of your organization and primary communicators across all encounters for each use case?
3. **Total Cost of Implementation**: How can you determine the real (total) costs for early adoption of an AI interpreting solution?

Risk Factors, Risk Levels, and Organizational Readiness

Determining risk levels for AI interpreting requires understanding three dimensions: the scenario's typical conditions, individual characteristics of primary communicators (such as accent, dialect, or communication style), and the technical capacity of the specific language pairing. By definition, early adoption of new technology carries higher risk than using established methods for providing interpreting services. Integrating AI solutions incorrectly can create significant risk due to both technical limitations and the challenges of adapting machine-learning translation technologies to live interpreting contexts, including how to recognize and repair misunderstandings when they occur. Early adoption increases the likelihood of overlooking these critical evaluation factors. *Part A's* framework and practical checklists are designed to help organizations identify and manage these risks effectively.

This Toolkit does not formally define thresholds for "low risk," "moderate risk," or "high risk" scenarios, nor does it prescriptively try to categorize "low complexity," "moderate complexity," or "high complexity" communications. Every interaction has potentially complicating variables depending upon the language pairing and individual characteristics of the primary communicators.

Instead, this Toolkit is designed to enable your organization to assess the risks of early adoption based on your institutional context. Presently less than five percent of all live interpreting, globally, is performed with AI assistance.[1] AI interpreting solutions represent an emerging technology: the challenges and

[1] Nimdzi Insights, LLC. "The 2023 Nimdzi Interpreting Index: Ranking of the Top Interpreting Companies." Accessed August 24, 2025. https://www.nimdzi.com/interpreting-index-top-interpreting-companies/

consequences of implementation are not yet fully clear. *Part A*'s framework recognizes that conversations may begin with low-risk, routine topics that can escalate quickly and unpredictably to higher-risk scenarios. Vendors and organizations must prioritize immediate escalation protocols and human oversight mechanisms, including complete, quick rollover to a qualified human interpreter initiated by a primary communicator, given the potential jeopardy to individuals relying on AI interpreting solutions. Contracting organizations are accountable for running pilots to support the thorough evaluation of their specific communication use cases, collecting preliminary data, identifying the presence of risk factors, determining acceptable, risk-informed risk levels, and establishing appropriate safeguards based on pilot study results.

Practical Assessment and Implementation Tools

The previous sections outlined a comprehensive framework to help your organization assess its preparedness to implement AI interpreting by examining opportunities versus risks, regulatory contexts, and governance structures. The checklists represent the functional heart of this document. They take practical considerations and turn them into actionable steps, listing specific capabilities that alert you to everything needed for safe, accountable, fair and ethical implementation of an AI interpreting solution:

CHECKLIST 1: Organizational Readiness Evaluation - A self-assessment tool that helps organizations find readiness gaps before implementing an AI interpreting solution across 8 key categories: strategic alignment; use case analysis; technical infrastructure; data privacy; human resources; quality assurance; financial readiness; and implementation planning.

CHECKLIST 2: Setting-Specific Considerations for AI Interpreting Implementation - Targeted guidance for healthcare, legal, education, and business environments, including sector-specific compliance requirements, implementation protocols, and specialized considerations for both spoken and signed language interpreting needs.

CHECKLIST 3: Risk Assessment Framework for AI Interpreting Solutions - A tool for categorizing interpreting use cases by risk factors to determine risk levels. Includes examples, risk mitigation recommendations for each level, and clear documentation requirements for accountability.

CHECKLIST 4: Vendor Assessment Checklist - A 10-category evaluation framework for assessing AI interpreting vendors across usability, technical

capabilities, security, ethics, customization, support, escalation protocols, compliance, cost structure, and vendor stability.

CHECKLIST 5: Guidance for Request for Proposals (RFPs) that Include AI Interpreting - Step-by-step guidance for procurement professionals writing RFPs that include AI interpreting solutions, with template language, evaluation criteria, compliance requirements, and pilot testing protocols for both spoken and signed language services.

Foundational Considerations for AI Implementation

Before diving into the evaluation checklists, it's important to understand the key considerations that shape the approach of this Toolkit. The following seven areas represent critical factors that organizations consistently encounter when evaluating AI interpreting solutions—from understanding what the technology can and cannot do, to navigating legal requirements, to calculating true costs. These foundational considerations provide the context you need to use the checklists effectively and help explain why certain evaluation criteria are essential for making sound decisions about AI interpreting solutions.

Opportunities and The Risk Picture

AI language solutions can provide interpreting services that are cheaper, easier to scale, and more immediately available on-demand. This can make routine interactions more accessible to more people. However, these benefits come with risk factors that need careful consideration. When organizations implement AI interpreting solutions poorly, they may face serious operational, legal, and ethical problems that can actually hurt meaningful access for principal communicators, leading to adverse financial consequences.

The main risk factors include mistakes and misunderstandings caused by poor audio or video quality or flawed translation. AI systems have algorithmic bias from their training data, which typically disadvantages people speaking indigenous or less-resourced languages, those with accents, or of certain genders. AI interpreting solutions are not designed to identify cultural or contextual information that professional human interpreters naturally recognize. Some AI solutions can also create data security risks by storing sensitive user information.

A critical operational risk occurs when multiple risk factors combine, transforming a typically simple, low-risk conversation into a situation where achieving mutual

understanding becomes difficult. Whatever is not being understood by principal communicators can range from trivial matters to high-stakes information with life-altering consequences. AI struggles with these combined risk factors – which is why the software interface should include user-initiated switching to human interpreters in as timely a manner as possible.

Legal and Regulatory Context

In a time of changing federal priorities and budget pressures, your organization needs to carefully evaluate language access strategies while understanding legal obligations. Core civil rights and accessibility laws—including Title VI of the Civil Rights Act, the Americans with Disabilities Act (ADA), Section 1557 of the Affordable Care Act, and the Individuals with Disabilities Education Act (IDEA), among others—remain unchanged and are legally binding. These foundational protections differ from executive orders and accompanying federal guidance memos, which can shift with each presidential election. While federal enforcement approaches and funding priorities may change, the underlying legal mandates for language access continue to apply at federal, state, and local levels.

In this shifting policy and budgetary landscape, your organization currently faces strategic decisions about maintaining, reducing, or restructuring language access services. When considering AI interpreting as part of your language access continuum, you need to consider both current budget pressures and the potential costs of rebuilding services when priorities shift again. Your organization remains accountable for understanding how policy changes affect your specific compliance obligations and should stay informed through relevant professional associations and language access organizations that provide updated guidance during transitional periods. (Legal requirements and compliance strategies are explored in depth in *Part C* of the *AI Interpreting Solutions Evaluation Toolkit*.)

Governance and Collaborative Assessment

Successfully adopting AI interpreting technology requires teamwork across your organization. Don't let IT or procurement departments make these decisions alone. Instead, create a team that includes people from procurement, legal and compliance, IT and security, language access program managers, and frontline staff from the departments that will actually use the technology.

This team approach is essential for several reasons. It helps you define what you want the AI system to accomplish, ensures you properly weigh the risks against the benefits, and creates a clear record of why you made specific decisions. Working together this way also helps establish clear policies for using the technology and makes sure everyone understands their responsibilities for managing it effectively.

AI Quality and Performance Considerations

The quality of AI interpreting solutions can vary significantly depending on the situation and languages involved. One key limitation is that AI systems work much better when translating into English than from English to other languages. This happens because most AI training data uses English as the base language. This performance gap affects both spoken and signed languages.[2] [3] [4]

As explained above, low resource spoken languages are not translated in a balanced, bidirectional way as accurately as higher resourced spoken languages. For sign language interpreting, AI can be accurate for simple, controlled tasks like recognizing individual letters spelled out by hand. However, it struggles with real-world conversations and cannot reliably interpret essential parts of sign language communication like facial expressions and the use of space to show relationships between ideas.

Because AI performance is so inconsistent, your organization cannot rely on vendor marketing claims alone. You must pilot any AI interpreting solution with your own communities and languages before making decisions. (Evaluating technical specifications is explored in depth in *Part B* of the *AI Interpreting Solutions Evaluation Toolkit*.)

---

[2] Bragg, Danielle, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreault, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler, and Meredith Ringel Morris. "Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective." Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19), October 24, 2019. https://dl.acm.org/doi/10.1145/3308561.3353774

[3] Desai, Aashaka, Maartje De Meulder, Julie A. Hochgesang, Annemarie Kocab, and Alex X. Lu. "Systemic Biases in Sign Language AI Research: A Deaf-Led Call to Reevaluate Research Agendas." arXiv preprint arXiv:2403.02563v1 cs.CV, March 5, 2024. https://aclanthology.org/2024.signlang-1.6/

[4] Moghe, N., Fazla, A., Amrhein, C., Kocmi, T., Steedman, M., Birch, A., Sennrich, R., & Guillou, L. (2024). Machine Translation Meta Evaluation through Translation Accuracy Challenge Sets. *Computational Linguistics*, 51(1). https://doi.org/10.1162/coli_a_00537 [1]

Identifying and Testing Solutions

Choosing the right AI interpreting solution requires evaluating how well it works, how accurate it is, how secure it is, and whether it follows ethical design principles. Your organization should go beyond what vendors promise in their marketing materials. Instead, evaluate vendors on practical factors like whether they actually support the specific languages and language variants your community uses, whether the system fits your industry's needs, whether they can show you real performance data, and whether they have proper security certifications.

The most important step is running your own pilot test before making any decisions. This means identifying the real situations where you would use AI interpreting, finding people from your community who represent the languages and language variants you serve, and testing how well the technology actually works in real-world conditions—not in perfectly controlled laboratory settings. During the pilot, evaluate whether the system functions properly, provides accurate interpretations, reduces bias against certain groups, is easy for people to use, and keeps information secure.

Total Cost of Implementation

Calculating the real cost of AI interpreting requires looking beyond vendor subscription or usage fees to understand the full organizational impact. The actual additional costs your organization faces will depend on factors like the specific AI solution you choose, how broadly you plan to use it, and your current IT capabilities and infrastructure.

This assessment helps you identify how implementing AI interpreting solutions affect different departments across your organization and what investments may be needed beyond the technology itself. You'll need to evaluate what resources you currently have, where those resources will be needed for successful implementation, and what gaps exist that require additional investment. These considerations might include staff training, pilot testing, quality monitoring systems, human oversight processes, and maintaining human interpreting services. Understanding these cross-departmental impacts and resource requirements upfront helps you make informed decisions about whether and how to proceed with AI interpreting solutions.

Implementation Guardrails and Ongoing Management

To use AI interpreting responsibly, your organization needs clear guidelines and ongoing management processes across multiple departments. This multi-layered approach ensures the technology works effectively while protecting the people who use it.

Senior leadership must provide strategic oversight and make sure adequate resources are available for human review of the AI interpreting solution when needed. Staff who use the AI system directly need thorough training on what the technology can and cannot do, plus clear understanding of why principal communicators may decide to switch between AI and human interpreting during the course of an encounter. The person or department responsible for Language Services plays an essential role in monitoring the quality of interpretation and managing how AI solutions and human interpreters work together. The IT department must ensure the technical systems are reliable, secure, and meet all data privacy requirements.

Each department has specific responsibilities, but they must work together to ensure AI interpreting solutions support and strengthen language access rather than impeding the integrity of effective communication for your community.

Summary

Together, these foundational considerations and practical checklists form a comprehensive framework for evaluating and managing AI interpreting solutions. This Toolkit is designed not as a one-time checklist for purchasing decisions, but as a guide for the ongoing process of evaluation, implementation, and continuous improvement that responsible AI adoption requires.

By using this risk-informed approach, your organization can better navigate the complexities of this rapidly evolving technology. The three-pillar framework helps you understand both the opportunities and limitations of AI interpreting, assess your organizational readiness, evaluate vendors effectively, and implement solutions that truly serve your community's communication needs. Most importantly, it ensures you can harness the potential benefits of AI interpreting while maintaining your commitment to safe, accountable, fair, and ethical communication for everyone you serve.

CHECKLISTS

Checklist 1: Organizational Readiness Evaluation

Checklist 2: Setting-Specific Considerations for AI Interpreting Implementation

Checklist 3: Risk Factor Assessment Framework for AI Interpreting Solutions

Checklist 4: Vendor Assessment Checklist

Checklist 5: Guidance for Request for Proposals (RFPs) that Include AI Interpreting

Checklist 1: Organizational Readiness Evaluation

A self-assessment tool that helps organizations find readiness gaps before implementing an AI interpreting solution across 8 key categories:

1. Strategic alignment and governance
2. Use case analysis and language needs
3. Technical infrastructure and audio environments
4. Data privacy and security controls
5. Human resources and training requirements
6. Quality assurance and monitoring processes
7. Financial resources and sustainability
8. Implementation and contingency planning

Organizations can use the straightforward scoring system to calculate readiness in each area and determine if they are fully ready for implementation, need minor adjustments, require significant preparation, or are not yet ready to proceed with AI interpreting technology.

As you complete the checklist, note that some questions assess concrete readiness factors while others identify areas requiring organizational consideration and discussion. In both cases, mark "*Ready*" or "*Not Ready*" in the appropriate columns—marking "Ready" when you have either achieved readiness or thoroughly considered the discussion point for your context.

At the bottom of the table, assign one point for each "Ready" checkmark to compare to the potential total. The sub-totals can be added after the last checklist for an overall total.

1.1 Strategic Alignment and Governance

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Strategic Alignment	Do expected benefits align with organizational goals?	☐	☐	
	Is there a clear vision for how AI fits into your language access plan?	☐	☐	
	Has leadership formally endorsed AI interpreting exploration	☐	☐	
	Does your language access strategy address both spoken and sign language needs?	☐	☐	
Decision Authority	Have you identified who has final decision authority for AI adoption?	☐	☐	
	Is there a cross-functional team with designated representatives?	☐	☐	
	Are all relevant departments represented (e.g., IT, Compliance, Language Services), and Primary Communicators)?	☐	☐	
	Does the decision-making team include accessibility expertise?	☐	☐	
Governance Framework	Is there a documented governance framework for AI technologies?	☐	☐	
	Are there clear policies for AI usage, limitations, and escalations?	☐	☐	
	Has legal counsel reviewed and approved the governance structure?	☐	☐	
	Does governance address disability rights compliance?	☐	☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Change Management	Is there a change management plan for introducing AI interpreting? Have key immigrant stakeholders been engaged and informed? Have you engaged Deaf and disability community stakeholders? Are there designated champions in each affected department?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Equity and Disability	Is there executive-level commitment to disability and equity inclusion?	☐	☐	
Inclusion Governance	Are there designated leaders responsible for Title VI/ADA/508 compliance? Do you have a Title VI / disability advisory group or committee? Is there engagement with external LEP and disability advocacy organizations?	☐ ☐ ☐	☐ ☐ ☐	
READINESS SCORE TOTAL (1.1)	**(add 1 for each checkmark)**	**__ / 20** **Total Ready**	**__ / 20** **Total Not Ready**	**READINESS SCORE TOTAL (1.1)**

1.2 Use Case Analysis and Volume Assessment

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Use Case Inventory	Have you created an inventory of potential use cases for AI interpreting?	☐	☐	
	Have you classified use cases by risk level (low, moderate, high)? (*See Checklist #3*)	☐	☐	
	Have you documented the use case requirements and constraints?	☐	☐	
	Have you identified use cases requiring sign language interpretation?	☐	☐	
Volume Assessment	Do you have sufficient volume to justify AI implementation?	☐	☐	
	Have you quantified the number of potential interactions by language?	☐	☐	
	Is there a clear understanding of peak vs. average demand?	☐	☐	
	Have you assessed volume for sign language vs. spoken language needs?	☐	☐	
Language Needs	Have you mapped all languages required for your organization?	☐	☐	
	Have you identified which languages are most critical?	☐	☐	
	Have you identified language pairs that AI cannot manage due to training data limitations?	☐	☐	
	Have you assessed dialectal and regional accent requirements?	☐	☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
	Have you documented specific sign language variation needs (ASL, BSL, etc.)?	☐	☐	
User Demographics	Have you analyzed your user demographics (age, comfort using technology, etc.)? Have you considered accessibility needs for primary communicators with disabilities? Have you gathered input from those who will use the technology?	☐ ☐ ☐	☐ ☐ ☐	
READINESS SCORE TOTAL (1.2)	**(add 1 for each checkmark)**	**__ / 16 Total Ready**	**__ / 16 Total Not Ready**	**READINESS SCORE TOTAL (1.2)**

1.3 Technical Infrastructure

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Network Capability	Does your network infrastructure support AI interpreting requirements? Is your internet bandwidth sufficient for video/audio streaming?	☐ ☐	☐ ☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
	Have you tested network performance in all implementation locations?	☐	☐	
Audio Environment Assessment	Have you evaluated the acoustics of spaces where interpretation will occur?	☐	☐	
	Have you identified potential noise interference sources in your facilities?	☐	☐	
	Do you have a plan for mitigating background noise in interpretation areas?	☐	☐	
Hardware Readiness	Do you have appropriate microphones designed for capturing multiple speakers?	☐	☐	
	Are your audio input devices capable of filtering background noise?	☐	☐	
	Have you planned for device management and security?	☐	☐	
	Do you have appropriate video equipment for sign language communicators?	☐	☐	
System Integration	Can AI interpreting solutions integrate with your existing systems?	☐	☐	
	Have you identified integration requirements for key platforms?	☐	☐	
	Are there APIs available for necessary customizations?	☐	☐	
	Does the solution integrate with assistive technologies?	☐	☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
IT Support	Does your IT team have capacity to support AI implementation? Is there a documented support process for technical issues? Have you planned for ongoing maintenance requirements? Is IT staff trained on accessibility requirements?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Maintenance Planning	Have you planned for ongoing maintenance requirements? Is there a protocol for regular testing of audio and video equipment? Have you established quality thresholds for acceptable audio and video performance?	☐ ☐ ☐	☐ ☐ ☐	
Accessibility Infrastructure	Does your physical environment support accessible interpreting? Have you evaluated lighting, background, and camera positioning for sign language? Do you have sufficient display screens of appropriate size? Have you tested compatibility with hearing aids/cochlear implants?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
READINESS SCORE TOTAL (1.3)	**(add 1 for each checkmark)**	**__ / 25** **Total Ready**	**__ / 25** **Total Not Ready**	**READINESS SCORE TOTAL (1.3)**

1.4 Data Privacy and Security

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Privacy Framework	Do you have a privacy framework that can accommodate AI interpreting? Have you updated privacy policies to reflect AI use? Are there clear guidelines for data handling and retention?	☐ ☐ ☐	☐ ☐ ☐	
Security Controls	Do you have appropriate security controls for AI implementation? Have you conducted a security assessment for new technology? Are there protocols for security incident response?	☐ ☐ ☐	☐ ☐ ☐	
Compliance Verification	Have you verified compliance with relevant regulations (e.g., HIPAA, Title VI, ADA)? Is there a process for ongoing compliance monitoring? Has legal counsel reviewed compliance considerations? Have you assessed Section 504/508 compliance for federally funded programs?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Risk Assessment	Have you conducted a formal risk assessment for AI interpreting? (*See Checklist 3*) Are there mitigation strategies for identified risks? Is there a process for ongoing risk monitoring?	☐ ☐ ☐	☐ ☐ ☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
	Have you evaluated specific risks related to serving Deaf/hard of hearing populations?	☐	☐	
READINESS SCORE TOTAL (1.4)	**(add 1 for each checkmark)**	**__ / 14 Total Ready**	**__ / 14 Total Not Ready**	**READINESS SCORE TOTAL (1.4)**

1.5 Human Resources and Training

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Staff Capacity	Do you have staff designated to oversee AI interpreting implementation? Are there resources available for ongoing management? Is there clear ownership of AI interpreting outcomes?	☐ ☐ ☐	☐ ☐ ☐	
Training Plan	Are there a comprehensive training plan and budget for all users? Have you developed role-specific training materials? Is there a plan for refresher training and updates? Does training include disability etiquette and accessibility awareness?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Staff Acceptance	Have you assessed staff willingness to adopt AI interpreting? Have you addressed concerns about AI replacing human interpreters? Is there a plan to manage resistance to change?	☐ ☐ ☐	☐ ☐ ☐	
Human Interpreter Integration	Have you planned how AI and human interpreters will coexist? Is there a clear rollover path from AI to human interpreters? Have human interpreters been involved in planning? Are qualified sign language and indigenous language interpreters readily available when needed?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
READINESS SCORE TOTAL (1.5)	**(add 1 for each checkmark)**	**__ / 14** **Total Ready**	**__ / 14** **Total Not Ready**	**READINESS SCORE TOTAL (1.5)**

1.6 Quality Assurance and Monitoring

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Quality Framework	Do you have quality scores from your existing human interpreting processes?	☐	☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
	Do you have a quality assurance framework for AI interpreting? Have you established performance metrics and standards? Is there a process for quality improvement? Do your standards consider accessibility quality metrics?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Monitoring Plan	Is there a plan for ongoing monitoring of AI interpreting quality? Have you identified who will be responsible for monitoring? Are there automated and manual monitoring components? Does your monitoring plan include disability-specific metrics?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Feedback Mechanisms	Have you established mechanisms for user feedback? Is there a process for addressing negative feedback? Will feedback be incorporated into system improvements? Are feedback mechanisms accessible to primary communicators with disabilities, or lack of access to technology?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Error Resolution	Is there a documented process for addressing AI translation errors? Have you established error severity classifications? Is there a clear protocol for critical error remediation?	☐ ☐ ☐	☐ ☐ ☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
	Is the party liable for errors in the AI output clearly identified?	☐	☐	
Disability Specific Quality Control	Have you established quality standards specific to sign language interpretation? Are there quality metrics for various disability accommodations? Have you involved Deaf/disabled consumers in quality evaluation? Is there a process for resolving disability-related complaints?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
READINESS SCORE TOTAL (1.6)	**(add 1 for each checkmark)**	**__ / 21 Total Ready**	**__ / 21 Total Not Ready**	**READINESS SCORE TOTAL (1.6)**

1.7 Financial Readiness

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Budget Allocation	Has the budget been allocated for all aspects of AI interpreting implementation, maintenance, staff training, improvement, and remediation? Is there funding for ongoing maintenance and updates? Is there a budget specifically for accessibility-related costs?	☐ ☐ ☐	☐ ☐ ☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Return on Investment (ROI) Analysis	Have you conducted a return on investment analysis? Does ROI account for tangible and intangible benefits? Is there a timeline for expected return on investment? Does the ROI consider benefits from improved accessibility?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Cost Modeling	Have you developed a total cost of ownership model? Does your model include direct and indirect costs? Have you compared costs with current interpreting solutions? Have you modeled costs for disability-specific accommodations?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Funding Sustainability	Is there sustainable funding for AI interpreting over time? Have you planned for potential cost increases? Is there contingency funding for unexpected expenses? Is there funding for human interpreters when AI is inappropriate or declined by primary communicators?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Accommodation Funding	Is there dedicated funding for reasonable accommodations? Have you budgeted for specialized sign language interpreting? Is there funding for qualified Deaf Interpreters when needed?	☐ ☐ ☐	☐ ☐ ☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
	Have you allocated funds for specialized equipment/technology?	☐	☐	
READINESS SCORE TOTAL (1.7)	**(add 1 for each checkmark)**	**__ / 16 Total Ready**	**__ / 16 Total Not Ready**	**READINESS SCORE TOTAL (1.7)**

1.8 Implementation Planning

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Rollout Strategy	Have you developed a phased rollout strategy?	☐	☐	
	Is there a timeline with defined milestones and specific people accountable for them?	☐	☐	
	Have you identified pilot groups and AI adopters?	☐	☐	
	Does your rollout plan address accessibility-specific use cases?	☐	☐	
Success Metrics	Have you established clear success metrics?	☐	☐	
	Are there qualitative and quantitative measures?	☐	☐	
	Is there a baseline for comparison?	☐	☐	
	Have you defined accessibility-specific success criteria?	☐	☐	

Readiness Factor	Assessment Questions	Ready	Not Ready	Notes
Documentation	Have you developed implementation documentation? Are there clear procedures for each implementation phase? Is documentation accessible to all relevant stakeholders? In their language? Is documentation available in accessible formats?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Contingency Planning	Is there a contingency plan for implementation challenges? Have you identified potential technical and procedural failure points? Is there a rollback strategy if needed? Are there backup plans for accessibility-related failures?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
READINESS SCORE TOTAL (1.8)	**(add 1 for each checkmark)**	**__ / 16** **Total Ready**	**__ / 16** **Total Not Ready**	**READINESS SCORE TOTAL (1.8)**

Total Scores for Checklist 1

Subtotals	Ready	Not Ready
1.1	_____ / 20	_____ / 20

1.2	_____ / 16	_____ / 16
1.3	_____ / 25	_____ / 25
1.4	_____ / 14	_____ / 14
1.5	_____ / 14	_____ / 14
1.6	_____ / 21	_____ / 21
1.7	_____ / 16	_____ / 16
1.8	_____ / 16	_____ / 16
Grand Total:	_____ **/ 142**	_____ **/ 142**

1.9 Organizational Readiness Assessment

Based on the evaluation above:

[ ] Fully Ready to Implement

[ ] Not Ready for Implementation

Priority Action Items

1. _____

2. _____
3. _____
4. _____
5. _____

Evaluator(s): _____

Date: _____

Checklist 2: Setting-Specific Considerations for AI Interpreting Implementation

The risks and requirements for AI interpreting vary significantly across different sectors. This checklist provides targeted guidance for implementing solutions in four key settings.

- For **healthcare**, it focuses on Health Insurance Portability and Accountability Act (HIPAA) compliance and patient safety. Checklists 2.1–2.3.
- For **legal settings**, it emphasizes due process protection and attorney-client privilege. Checklists 2.1, 2.4–2.5.
- For **education**, it addresses requirements under Individuals with Disabilities Education Act (IDEA) and Family Educational Rights and Privacy Act (FERPA). Checklists 2.1, 2.6–2.7.
- For **business**, it centers on customer satisfaction and compliance with Americans with Disabilities Act (ADA) Title III. Checklists 2.1, 2.8.

Each section uses tables to organize key considerations, compliance requirements, and implementation steps with specific attention to both spoken and sign language requirements.

As you complete the checklist, note that some questions assess concrete readiness factors while others identify areas requiring organizational consideration and discussion. In both cases, mark "Ready" or "Not Ready" in the appropriate columns—marking "Ready" when you have either achieved readiness or thoroughly considered the discussion point for your context.

At the bottom of the table, assign one point for each "Ready" checkmark to compare to the potential total. The sub-totals can be added after the last checklist for an overall total.

2.1 Universal Implementation Checklist

Regardless of setting, all AI interpreting implementations should include:

Phase	Implementation Steps	Ready	Not Ready	Notes
Pre-Implementation	Conduct setting-specific risk assessment using Checklist 3	☐	☐	
	Map language pair needs across all user groups, including sign languages, spoken language variants, and mixed languages	☐	☐	
	Identify domain-specific terminology for AI training	☐	☐	
	Review relevant regulations and compliance requirements	☐	☐	
	Establish clear initial policies on when AI vs. human interpreters will be used that will be updated post-pilot	☐	☐	
Implementation	Provide setting-appropriate training for all staff using the system	☐	☐	
	Develop clear user instructions in multiple formats and languages	☐	☐	
	Implement streamlined rollover or escalation processes to human interpreters	☐	☐	
	Establish quality monitoring protocols specific to the setting	☐	☐	
	Create feedback mechanisms for primary communicators of all language backgrounds	☐	☐	

Phase	Implementation Steps	Ready	Not Ready	Notes
Post-Implementation	Conduct regular compliance audits specific to the setting's regulations	☐	☐	
	Gather and analyze user feedback by language group and use case	☐	☐	
	Measure actual performance against established quality thresholds	☐	☐	
	Compare method of interpreting delivery to outcomes for language groups and language pairings	☐	☐	
	Document all incidents and resolution processes	☐	☐	
	Schedule regular reviews and updates to policies and procedures	☐	☐	
READINESS SCORE TOTAL (2.1)	**(add 1 for each checkmark)**	**__ / 16 Total Ready**	**__ / 16 Total Not Ready**	**READINESS SCORE TOTAL (2.1)**

The following sections provide specific considerations for different settings, building upon this universal implementation framework.

2.2 Healthcare Settings

Key Considerations and Implementation Steps

Category	Healthcare-Specific Considerations	Ready	Not Ready	Notes
Critical Considerations	**Risk Stratification**: Clearly delineate which healthcare interactions are appropriate for AI versus those requiring qualified human interpreters	☐	☐	
	Integration with Electronic Health Records (EHR): Evaluate and assess how AI interpreting solutions can integrate with EHR systems (and track patient interpretation preferences) while maintaining HIPAA compliance	☐	☐	
	EHR Language Tracking Ability: Ensure that the EHR correctly assigns the primary communicator's preferred language.	☐	☐	
	Patient-Centered Care: Maintain focus on patient experience and outcomes when implementing technology solutions	☐	☐	
	Provider Acceptance: Develop strategies to address healthcare provider concerns about AI accuracy and patient safety	☐	☐	
	Language Pair Limitations: AI performs inconsistently depending on a) each language pair and b) the direction of translation into or out of English	☐	☐	
Implementation Steps	**Medical Terminology Training**: Ensure AI systems are trained on specialized medical vocabulary	☐	☐	

Category	Healthcare-Specific Considerations	Ready	Not Ready	Notes
	Emergency Protocol: Establish clear protocols for immediate rollover to human interpreters in emergent situations	☐	☐	
	Oversight Model: Implement a supervision model where qualified medical interpreters monitor AI performance in moderate-risk scenarios	☐	☐	
	Transparent Documentation: Develop clear documentation processes that record whether AI or human interpreters were used in patient encounters	☐	☐	
Signed Language Considerations	**Medical Sign Vocabulary**: Verify AI signed language capabilities include specialized medical signs and classifiers.	☐	☐	
	Visual Privacy: Ensure proper positioning of video equipment for patient sightlines and to maintain patient privacy	☐	☐	
	DI Access: Establish protocols for quickly engaging qualified Deaf Interpreters for complex medical situations	☐	☐	
	Visual Clarity: Implement proper lighting, backgrounds, and technical specifications for clear visual communication	☐	☐	

Category	Healthcare-Specific Considerations	Ready	Not Ready	Notes
READINESS SCORE TOTAL (2.2)	**(add 1 for each checkmark)**	**__ / 14** **Total Ready**	**__ / 14** **Total Not Ready**	**READINESS SCORE TOTAL (2.2)**

2.3 Healthcare Regulatory Compliance Requirements

Regulation	Compliance Considerations	Ready	Not Ready	Notes
HIPAA	Ensure end-to-end encryptions for all interpreted sessions	☐	☐	
	Implement zero-retention policies for Protected Health Information (PHI) when possible	☐	☐	
	Verify vendor Business Associate Agreements (BAA) address all interpreting scenarios	☐	☐	

Regulation	Compliance Considerations	Ready	Not Ready	Notes
Section 1557 of ACA	Ensure meaningful access regardless of language or disability Provide notices of language services in top 15 languages Document all language access decisions and accommodations	☐ ☐ ☐	☐ ☐ ☐	
Centers for Medicare and Medicaid Services (CMS) Requirements	Address interpreter qualifications in policies and procedures Comply with Conditions of Participation for language access Document language preferences in permanent medical records	☐ ☐ ☐	☐ ☐ ☐	
Joint Commission	Align with effective communication standards Implement competency assessment for interpreter services Integrate language services into quality improvement activities	☐ ☐ ☐	☐ ☐ ☐	
State Requirements	Ensure compliance with state medical interpreter qualification requirements Follow state-specific patient rights regulations Address any state-specific consent requirements	☐ ☐ ☐	☐ ☐ ☐	

Regulation	Compliance Considerations	Ready	Not Ready	Notes
READINESS SCORE TOTAL (2.3)	**(add 1 for each checkmark)**	**__ / 15 Total Ready**	**__ / 15 Total Not Ready**	**READINESS SCORE TOTAL (2.3)**

2.4 Legal Settings

Key Considerations and Implementation Steps

Category	Legal-Specific Considerations	Ready	Not Ready	Notes
Critical Considerations	**Due Process Protection**: Ensure technology implementation doesn't compromise constitutional or statutory rights	☐	☐	
	Evidentiary Value: Consider the evidentiary weight of AI-interpreted communications	☐	☐	
	Specialized Terminology: Legal vocabulary presents unique challenges for AI interpretation	☐	☐	
	Confidentiality Preservation: Maintain attorney-client privilege and other legal confidentiality requirements	☐	☐	

Category	Legal-Specific Considerations	Ready	Not Ready	Notes
Implementation Steps	**Limited Scope Deployment**: Restrict AI interpreting to low-risk administrative processes	☐	☐	
	Judicial Approval: Obtain proper judicial authorization before implementing AI in court-adjacent processes	☐	☐	
	Legal Oversight: Establish attorney review of AI-interpreted communications before they become part of the official record	☐	☐	
	Clear Disclaimers: Implement explicit notices about the use of AI technology in legal contexts	☐	☐	
Sign Language Considerations	**Court Certified Interpreters**: Maintain roster of certified legal sign language interpreters	☐	☐	
	Adversarial Proceedings: Ensure proper positioning and protocols for sign language interpreters in courtrooms	☐	☐	
	Legal Terminology in Sign: Verify capacity to accurately convey specialized legal concepts	☐	☐	
	Recording Considerations: Address challenges of simultaneously recording signed and spoken testimony	☐	☐	
READINESS SCORE TOTAL (2.4)	**(add 1 for each checkmark)**	**__ / 12 Total Ready**	**__ / 12 Total Not Ready**	**READINESS SCORE TOTAL (2.4)**

2.5 Legal Regulatory Compliance Requirements

Regulation	Compliance Considerations	Ready	Not Ready	Notes
Court Interpreter Acts	Identify applicable federal and state court interpreter requirements Verify compliance with court interpreter certification requirements Document all interpreter qualifications according to judicial standards	☐ ☐ ☐	☐ ☐ ☐	
ADA Title II	Ensure effective communication in all court proceedings Provide appropriate auxiliary aids and services Maintain documentation of all accommodations and modifications	☐ ☐ ☐	☐ ☐ ☐	
Rules of Evidence	Verify compliance with hearsay and evidence authentication rules Address chain of custody issues for interpreted communications Ensure interpreter qualifications meet evidentiary standards	☐ ☐ ☐	☐ ☐ ☐	
Ethics Rules	Comply with attorney ethics obligations for client communication Address confidentiality requirements for legal proceedings Ensure conflicts of interest are properly managed with interpreting services	☐ ☐ ☐	☐ ☐ ☐	

Regulation	Compliance Considerations	Ready	Not Ready	Notes
Confidentiality	Implement additional safeguards for privileged communications Ensure interpreter confidentiality agreements are in place Train all parties on legal confidentiality requirements	☐ ☐ ☐	☐ ☐ ☐	
AI Use Policies	Establish clear policies regarding AI tool usage by interpreters. Define permitted uses of AI for case file preparation and summarization Set guidelines for automated terminology extraction from case materials Specify rules for Computer-Assisted Interpreting (CAI) during proceedings Address AI use for post-session notes synthesis and documentation Ensure all AI applications comply with confidentiality and security requirements	☐ ☐ ☐ ☐ ☐ ☐	☐ ☐ ☐ ☐ ☐ ☐	
READINESS SCORE TOTAL (2.5)	**(add 1 for each checkmark)**	**__ / 21** **Total Ready**	**__ / 21** **Total Not Ready**	**READINESS SCORE TOTAL (2.5)**

2.6 Education Settings

Key Considerations and Implementation Steps

Category	Education-Specific Considerations	Ready	Not Ready	Notes
Critical Considerations	**Developmental Appropriateness**: Ensure AI interpreting solutions are appropriate for different educational levels and consider limitations of certain language pairs	☐	☐	
	Educational Equity: Implement technology in ways that reduce rather than amplify disparities	☐	☐	
	Usage Policy Development: Establish clear and transparent policies for AI interpreting use, including specific timeframes and contexts (e.g., social interaction periods vs. instructional time), with clear parental consent and communication protocols	☐	☐	
	Pedagogical Integrity: Maintain educational quality and learning objectives when incorporating technology	☐	☐	
	Language Pair Limitations: Human interpreting is automatically provided when the language pairing is not balanced in both translation directions	☐	☐	
Implementation Steps	**Curriculum Alignment**: Train AI systems on subject-specific vocabulary and academic terminology	☐	☐	
	Tiered Support Model: Implement a tiered model where routine communications may use AI while complex content uses qualified interpreters	☐	☐	

Category	Education-Specific Considerations	Ready	Not Ready	Notes
	Feedback Loop: Create mechanisms for families and students to provide feedback on interpreting quality	☐	☐	
	Accessibility Integration: Ensure AI interpreting tools integrate with existing educational accessibility systems	☐	☐	
Signed Language Considerations	**Educational Interpreting Standards**: Adhere to established standards for educational interpreters	☐	☐	
	Academic Sign Vocabulary: Ensure capacity for subject-specific sign vocabulary within each discipline	☐	☐	
	Classroom Positioning: Consider optimal positioning of technology for visual access by primary communicators	☐	☐	
	Multi-party Communication: Address challenges of interpreting classroom discussions with multiple participants	☐	☐	
IDEA	Ensure interpreting services align with IEP requirements	☐	☐	
	Document how technology meets Free Appropriate Public Education (FAPE) obligations	☐	☐	
	Address dispute resolution processes for interpreting concerns	☐	☐	
Section 504	Align technology implementation with 504 accommodation plans	☐	☐	
	Document reasonable modifications for effective communication	☐	☐	

Category	Education-Specific Considerations	Ready	Not Ready	Notes
	Ensure equal access to educational programs and activities	☐	☐	
FERPA	Implement privacy protections for student educational records	☐	☐	
	Address recording and transcript retention policies	☐	☐	
	Ensure parental consent for technology use when required	☐	☐	
State Educational Requirements	Comply with state educational interpreter qualification standards	☐	☐	
	Address state-specific educational accessibility requirements	☐	☐	
	Follow state department of education technology guidelines	☐	☐	
Educational Technology Standards	Align with educational technology accessibility guidelines	☐	☐	
	Ensure compatibility with existing assistive technologies	☐	☐	
	Document technology selection criteria and assessment	☐	☐	
READINESS SCORE TOTAL (2.6)	**(add 1 for each checkmark)**	**__ / 28 Total Ready**	**__ / 28 Total Not Ready**	**READINESS SCORE TOTAL (2.6)**

2.7 Business Settings

Key Considerations and Implementation Steps

Category	Business-Specific Considerations	Ready	Not Ready	Notes
Critical Considerations	**Customer Satisfaction**: Balance efficiency gains with customer experience metrics	☐	☐	
	Brand Representation: Consider how AI interpreting quality reflects on organizational brand	☐	☐	
	Internal vs. External Use: Differentiate standards for employee-facing versus customer-facing implementations	☐	☐	
	Scalability Requirements: Develop frameworks that can scale across multiple locations	☐	☐	
	Language Pair Limitations: Human interpreting is automatically provided when the language pairing is not balanced in both translation directions	☐	☐	
	Data Security and AI Solution Selection: Distinguish between built-in AI software (automated captioning/subtitling in videoconferencing platforms) and dedicated AI interpreting solutions that don't use conversation data for engine training. Establish policies prohibiting use of built-in software unless data security protocols are specifically approved.	☐	☐	
	Automated Captioning Limitations: Recognize that automated captioning has significant error rates and performs poorly with	☐	☐	

Category	Business-Specific Considerations	Ready	Not Ready	Notes
	accents, requiring careful evaluation before implementation in customer-facing scenarios.			
Implementation Steps	**Use Case Prioritization**: Identify high-volume, low-risk customer interactions as initial targets	☐	☐	
	Customer Consent: Develop transparent consent processes for AI interpreting us	☐	☐	
	Feedback Integration: Implement systems to capture customer feedback about interpretation quality	☐	☐	
	ROI Measurement: Establish clear metrics to measure return on investment	☐	☐	
	AI Solution Policy Development: Create clear guidelines differentiating between readily available built-in captioning software and secure, trained AI interpreting solutions, with specific approval processes for each category.	☐	☐	

Category	Business-Specific Considerations	Ready	Not Ready	Notes
Signed Language Considerations	**Visual Customer Service**: Ensure virtual customer service platforms support high-quality video	☐	☐	
	Employee Accessibility: Address needs of Deaf/Hard of Hearing employees in workplace communication	☐	☐	
	Marketing Accessibility: Ensure marketing materials are accessible to signed language communicators	☐	☐	
	Multi-channel Support: Provide multiple communication modalities for Deaf/Hard of Hearing customers	☐	☐	
	Turn-taking protocols: Enforce turn-taking in meetings, as overlapping speech reduces understanding.	☐	☐	
READINESS SCORE TOTAL (2.7)	**(add 1 for each checkmark)**	**__ / 17 Total Ready**	**__ / 17 Total Not Ready**	**READINESS SCORE TOTAL (2.7)**

2.8 Business Regulatory Compliance Requirements

Regulation	Compliance Considerations	Ready	Not Ready	Notes
ADA Title III	Ensure effective communication in places of public accommodation Provide appropriate auxiliary aids and services Document accommodation policy for customers with disabilities	☐ ☐ ☐	☐ ☐ ☐	
Section 255	Ensure telecommunications services are accessible Address compatibility with TTY[5] and other specialized equipment Document accessibility features of communication technology	☐ ☐ ☐	☐ ☐ ☐	
Employment Laws	Address reasonable accommodations for employees (ADA Title I) Ensure interpreting technology supports equal employment opportunity Document interactive process for employee accommodation requests	☐ ☐ ☐	☐ ☐ ☐	

[5] Modern day equivalents of TTY include equipment such as RTT, VRS, internet relay, etc.

Regulation	Compliance Considerations	Ready	Not Ready	Notes
Consumer Protection	Ensure clear disclosure of AI technology use Address accuracy claims in marketing materials Document complaint resolution processes for service failures	☐ ☐ ☐	☐ ☐ ☐	
Industry-Specific Requirements	Identify and comply with sector-specific communication requirements Address industry standards for customer service Document conformance with industry best practices	☐ ☐ ☐	☐ ☐ ☐	
READINESS SCORE TOTAL (2.8)	**(add 1 for each checkmark)**	**__ / 15** **Total Ready**	**__ / 15** **Total Not Ready**	**READINESS SCORE TOTAL (2.8)**

Total Scores for Checklist 2

Subtotals	Ready	Not Ready
2.1	_____ / 16	_____ / 16
2.2	_____ / 14	_____ / 14

2.3	_____ / 15	_____ / 15
2.4	_____ / 12	_____ / 12
2.5	_____ / 21	_____ / 21
2.6	_____ / 28	_____ / 28
2.7	_____ / 17	_____ / 17
2.8	_____ / 15	_____ / 15
Grand Total:	_____ **/ 138**	_____ **/ 138**

2.9 Organizational Readiness Assessment

Based on the evaluation above:

[ ] Fully Ready to Implement

[ ] Not Ready for Implementation

Priority Action Items

1. _____
2. _____
3. _____
4. _____

5. _____

Evaluator(s): _____

Date: _____

Checklist 3: Risk Factor Assessment Framework for AI Interpreting Solutions

Not every situation is appropriate for AI interpreting. This risk assessment framework provides a structured methodology for organizations to categorize interpreting scenarios by risk level. It helps organizations evaluate specific use cases and assess the potential for harm from errors or uncorrected misunderstandings. Low-risk scenarios, such as simple transactional tasks with primary communicators without communication-related disabilities, may be suitable for AI if the language pairing is balanced. However, higher-risk situations involving (for instance) medical consent, legal proceedings, or trauma-informed care, require qualified human interpreters due to the complexity, nuance, and severe consequences of potential errors. The framework also accounts for the critical challenge of conversations that begin as low risk but can unpredictably escalate, requiring robust protocols to switch to a human interpreter.

The Risk Factor Assessment Framework Assessment includes:

- **Three key areas** of attention for determining risk levels
- A **risk factor evaluation matrix** for scoring use cases: no, low, moderate, significant, and high risk
- A **use case analysis worksheet template** for documenting setting-specific assessments
- **Four example use case assessments** demonstrating application across different interpreting settings (healthcare, legal, education and business)
- **Tailored risk mitigation strategies** for each risk level
- **Documentation and accountability requirements**

By applying this framework, organizations can make evidence-based decisions about how important it is for organizations and primary communicators to be able to manage rollover from AI to human or hybrid solutions (and vice-versa). Specifically, these conditions can inform the vendor when the tool can be initiated in a specific encounter with full AI, a hybrid approach with human backup, or human interpreters.

3.1 Three Areas of Attention for Determining Risk Levels

When determining the overall risk score, there are three areas requiring attention, which are the main dimensions of interpreted communication. For each risk factor within these areas, we recommend consulting with subject matter experts who have relevant expertise.

3.1.1 Attention Area: Basic Specs of the Technology

a. **Language Pair Accuracy:** Investigate the AI solution's ability to interpret accurately within language pairs. Accuracy is significantly reduced when working between language pairs that include Indigenous and other languages that are underrepresented in AI training data. Working with languages beyond the most common ones (often referred to as low-resource languages) increases the risk of mistranslation in both directions—whether translating into English or from English into these languages.

b. **Privacy/Confidentiality**: Assess sensitivity of information being communicated

c. **Regulatory Requirements**: Consider applicable legal and compliance requirements. (see Checklist 2 above, and *Part C* of this toolkit series).

d. **Accessibility Requirements**: Does the software satisfy Web Content Accessibility Guidelines (WCAG) requirements, which ensure digital content is accessible to people with disabilities? Evaluate complexity of disability accommodation needs.

e. **Disability-Related Legal Compliance**: Consider applicable disability rights laws and mandates (see Checklist 2 above, and *Part C* of this toolkit series).

3.1.2 Attention Area: The Principal / Primary Communicators

a. **Potential Harm from Uncorrected Misunderstandings and Errors**: Assess severity of consequences if errors occur in translation and when primary communicators do not understand the full meaning and implications of automated translations.

b. **Communication Skills of the English Speaker**(s): Will they recognize misunderstandings? Do they know how to follow up for clarification and repair when responses don't make sense or are insufficient?

c. **Cultural/Contextual Mediation**: Assess need for cultural mediation and contextual understanding, for instance, based on continuous residence and/or the amount of time physically present in the country.

d. **Emotional/Psychological Context**: Evaluate emotional sensitivity and psychological impact.

e. **User Vulnerability**: Consider vulnerability of individuals involved (children, elderly, immigrants, etc.).

3.1.3 Attention Area: Environmental Context / Scenario

a. **Communication Complexity**: Consider the number of speakers (primary communicators), potential for cross talk, need for nuance or specialized vocabulary, and the probability for simple conversation to escalate to high-stakes scenarios.

b. **Environmental Settings**: Consider the room acoustics, potential for echo or background noise, if there is adequate lighting for signed language communication, the amount of visual noise especially for low vision communicators, etc.

c. **Equipment Availability**: Check if there is technical equipment needed such as iPads, computers, availability and bandwidth of working network connections and strength of Wi-Fi.

3.2 Risk Factor Evaluation Matrix

The matrix below is designed to identify in a simple way, a use case's potential risk level and recommendation of AI, Hybrid, or Qualified Human solution. Use it to help you complete the *Use Case Analysis Worksheet (3.3)* that follows:

Risk Factors by Areas of Attention (see 3.1)	Probable Risk Level	Base Recommendation
3.1.1 Basic Tech, e.g., balanced language pair accuracy **3.1.2 Primary Communicators**, e.g., trained English speakers and no vulnerability for other primary communicators **3.1.3 Typical Scenario**, e.g., very simple, predictable scenarios. No sensitive information. No significant consequences from errors.	No Risk	AI interpreting is appropriate.
3.1.1 Basic Tech, e.g., satisfies all required disability accommodations **3.1.2 Primary Communicators**, e.g., English-speaking providers (principal communicators) able to recognize and address conversational misunderstandings, service recipients (principal communicators) known to be familiar with the institutional setting, and a low chance of emotional/psychological harm **3.1.3 Typical Scenario**, e.g., involves minimal sensitive information, and limited consequences from errors for any principal communicators	Low Risk	AI interpreting appropriate with rollover mechanism to a qualified human that can be initiated by any primary communicator.

Risk Factors by Areas of Attention (see 3.1)	Probable Risk Level	Base Recommendation
3.1.1 Basic Tech, e.g., satisfies all required disability accommodations **3.1.2 Primary Communicators**, e.g., may not be familiar with the institutional setting requiring contextual or cultural mediation, yet not a member of an otherwise vulnerable population **3.1.3 Typical Scenario**, e.g., somewhat complex scenarios, leading to follow-up / additional sessions. Consequences from errors and uncorrected misunderstandings can be managed without adverse effects on well-being or potential of lawsuits	Moderate Risk	AI interpreting with qualified human backup; rollover can be initiated by any primary communicator
3.1.1 Basic Tech, e.g., balanced language pair accuracy **3.1.2 Primary Communicators**, e.g., increased chance of harm from errors and uncorrected misunderstandings, including due to vulnerability and or emotional/psychological factors **3.1.3 Typical Scenario**: Complex, involving sensitive information, with significant consequences from errors or uncorrected misunderstandings.	Significant Risk	Use qualified human interpreters only. AI or hybrid solutions may only be considered for primary communicators who opt in, and only when all of the following conditions are met: (1) all primary communicators demonstrate native fluency and literacy/education levels, (2) the AI translation algorithms are fully balanced and bidirectional, and (3) real-world evidence from formal research supports their effectiveness in similar scenarios.

Risk Factors by Areas of Attention (see 3.1)	Probable Risk Level	Base Recommendation
3.1.1 Basic Tech, e.g., may have balanced language pair accuracy **3.1.2 Primary Communicators**, consequences from errors and uncorrected misunderstandings are very likely to cause harm. **3.1.3 Typical Scenario**: Highly complex, sensitive, with severe consequences from errors and uncorrected misunderstandings, very likely to cause adverse effects on well-being and potential lawsuits	High Risk	Qualified human interpreters only.

3.3 Risk Mitigation Strategies by Risk Level

Before you begin conducting individual use case assessments, it's important to understand what each risk level means in practical terms—specifically, what safeguards and interventions will be required for different risk categories. This understanding will help you make more informed decisions when categorizing your specific scenarios using the assessment tools that follow.

The mitigation strategies below correspond directly to the risk levels identified in the evaluation matrix (section 3.2) and provide the operational framework for managing AI interpreting solutions across different risk scenarios.

Risk Level	Mitigation Strategies
Low Risk Level (AI Appropriate)	**Obtain explicit acceptance (for use of the solution) and consent (for use of data) from all parties for AI** Implement clear labeling of AI-interpreted communications Provide simple feedback mechanisms for primary communicators Conduct regular quality sampling Prominent display of the mechanism to request human rollover initiated by any primary communicator Train staff on system limitations Use pre-programmed templates and scripts where possible Prominently display confidence levels for low resource languages Implement system monitoring for unexpected responses that, when triggered, immediately initiate verification and possible intervention by a human interpreter Ensure basic accessibility features are available (screen reader compatibility, adjustable text sizing, adjustable background color) Establish protocols for identifying when disability accommodations may be needed Provide multiple communication modalities (text, audio, video)

Risk Level	Mitigation Strategies
Moderate Risk Level (Hybrid Approach Recommended)	**Obtain explicit acceptance (for use of the solution) and consent (for use of data) from all parties for AI** Require human oversight and real-time monitoring Implement automatic confidence scoring with threshold alerts Establish seamless rollover pathways initiated by primary communicators both ways: to AI from human interpreters and to human interpreters from AI Provide specialized training in context-specific terminology Conduct regular audits of AI performance in these scenarios Develop explicit systems that automatically transition from AI to human (and vice-versa) upon request Obtain explicit consent from all parties for AI interpretation with opt-out and automatic rollover to human interpreting services Ensure recording and documentation of all sessions only when primary communicators consent Implement post-session quality reviews by qualified interpreters (only when primary communicators consent to recording), with notable errors and evident misunderstandings shared with primary communicators for their remediation Confirm processes for providing disability accommodations are enacted when requested by primary communicators Ensure availability of matching/appropriate sign language interpreters who can be engaged quickly when needed Train staff on recognizing when accessibility needs require human interpretation Develop protocols for handling requested reasonable accommodations

Risk Level	Mitigation Strategies
High Risk Level (Human Interpreters Required)	**Obtain explicit acceptance (for use of the solution) and consent (for use of data) from all parties for AI** Use only certified/qualified human interpreters Use AI as a support tool for the interpreter (computer-aided interpreting) Ensure interpreters have specialized training for the context Implement cultural competency requirements Provide trauma-informed interpreting where applicable Establish clear documentation protocols Maintain quality assurance processes Implement interpretation team approaches for complex scenarios (for instance, relay interpreting and interpreter team open process and/or backchanneling) Conduct regular professional development for interpreters Never use AI interpreting systems for these scenarios Provide qualified sign language interpreters for Deaf individuals Make qualified Deaf Interpreters or Certified Deaf Interpreters (CDIs) available for complex situations or upon request (an example of relay interpreting) Provide specialized interpreters for DeafBlind individuals when needed Establish protocols for accommodating requests Ensure compliance with all applicable disability laws and regulations Provide multiple-modality communication options based on individual preferences

3.4 Documentation and Accountability

Effective risk management requires systematic documentation and accountability measures. Before you begin your use case assessments, establish these documentation protocols to ensure you can track, evaluate, and improve your AI interpreting solution implementation over time.

These requirements apply regardless of risk level and should be built into your organizational processes from the outset:

1. Document the risk assessment process and scores
2. Record justification for risk categorization
3. Identify who performed the assessment and when
4. Schedule regular re-evaluations (at least annually)
5. Document any incidents or errors that occur
6. Communicate immediately to primary communicators (all of them) with specific details about errors/misunderstandings identified during audits, reviews, and other mechanisms
7. Track changes in risk categorization over time
8. Maintain records of mitigation strategies implemented
9. Document user feedback specific to each use case
10. Create clear audit trail for compliance purposes
11. Include disability advocates in the review process
12. Establish review process with stakeholder input
13. Document accessibility accommodations requested and provided

14. Maintain records of disability-related compliance measures
15. Include disability advocates in the review process
16. Track escalations related to accessibility needs
17. Document training provided on disability accommodation protocols

3.5 Use Case Analysis Worksheet Template

Now that you understand the risk levels, corresponding mitigation strategies, and documentation requirements, you're ready to assess your specific use cases. This worksheet template provides a structured approach to categorize interpreting scenarios and determine appropriate implementation strategies.

Use this template in conjunction with the risk evaluation matrix (section 3.2) and the mitigation strategies outlined above to systematically evaluate each interpreting scenario in your organization.

Use Case:	**Describe the interpreting scenario**	
Scenario Description	Detailed description of the interpreting scenario, including: • Purpose of the interaction • Typical duration • Number of participants/primary communicators • Profile of the primary communicators (native or main language fluency, literacy, education, time in the US, disability status) • Physical or virtual setting • Type of session moderation	Interpreting Scenario:

Use Case:	Describe the interpreting scenario	
	• Required languages/modalities • Typical content • Specific accessibility requirements • Other, based on your context	
Overall Risk	□ No Risk □ Low Risk □ Moderate Risk □ Significant Risk □ High Risk	Notes:
Key Risk Factors	List the primary risk factors that contributed to the score:	
Recommended Approach	• **AI Interpreting Only** (No Risk) • **AI with Basic Safeguards** including primary communicator or AI-initiated rollover (Low Risk, perhaps Moderate Risk)	

Use Case:	Describe the interpreting scenario
	• **AI with Human Backup** (primary communicator or AI-initiated rollover (Moderate Risk, Significant Risk) • **Human Interpreter Only** (Significant Risk, High Risk)
Required Safeguards	If using AI, list necessary safeguards:
Opt out / Rollover Triggers	Specific conditions that would trigger an automatic rollover by the AI or the intervention of human interpreter who may be providing real-time monitoring/oversight

Use Case:	Describe the interpreting scenario
Compliance Requirements	Relevant regulatory and legal requirements, including disability accommodations (See Checklist 2 and *Part C*):
Accessibility Considerations	Specific accessibility needs that must be addressed:
Notes	Additional considerations:

3.6 Example Use Case Risk Assessments

The following examples demonstrate how to apply the risk assessment framework using the worksheet template. Each example shows how different scenarios map to risk levels and corresponding mitigation strategies, illustrating the practical application of the assessment process.

These examples span different settings and risk levels to help you understand how to categorize your own use cases. The four examples cover: customer service (low risk), healthcare/pharmacy (moderate risk), healthcare/medical follow-up (significant risk), and legal/court proceedings (high risk):

3.6.1 Example 1-LOW RISK: Customer Service Appointment Scheduling

Use Case:	**Customer Service Appointment Scheduling**
Scenario Description	Scheduling a routine appointment over the phone with a customer service representative. Conversation typically lasts 3-5 minutes, involves dates/times, basic personal information, and simple clarification questions.
Overall Risk	□ No Risk ✓Low Risk □ Moderate Risk □ Significant Risk □ High Risk

Use Case:	Customer Service Appointment Scheduling
Key Risk Factors	Potential consequences from errors (quickly rescheduled appointments) Simple, predictable dialogue with limited vocabulary - depends on native language fluency of primary communicators and their personal familiarity with the service Basic personal information only Limited regulatory requirements Minimal accessibility complexity
Recommended Approach	□ AI Interpreting Only ✓ AI with Basic Safeguards □ AI with Human Backup □ Human Interpreter Only
Required Safeguards	Clear labeling as AI-interpreted Bank of pre-translated common questions and answers in every language pair Quality monitoring for common scheduling terms Ability for customer to request clarification Basic accessibility features available Ability to switch to a human interpreter at any point or the lack of availability of a human interpreter is clearly communicated

Use Case:	Customer Service Appointment Scheduling
Escalation Triggers	User explicitly requests human interpreter Conversation veers to non-scheduling topics Multiple failed understanding attempts Complex accessibility needs identified
Compliance Requirements	ADA/Title VI (if applicable) General customer service standards Basic Section 504 requirements if federally funded
Accessibility Considerations	Ensure interface works with screen readers Provide text alternatives to audio - are the captions accurate and functional on all devices, e.g., iPads? Support TTY/telecommunications relay services, if applicable
Notes	Well-suited for AI interpreting given the routine, predictable nature with low stakes

3.6.2 Example 2 - MODERATE RISK: Pharmacy Visit for Medication Refill

Use Case:	Pharmacy Visit for Medication Refill
Scenario Description	Patient visiting a pharmacy to refill an existing prescription. Conversation typically involves verifying patient identity, confirming medication details, discussing basic usage instructions, and addressing any simple questions about side effects. Usually 5-10 minutes at the pharmacy counter.

Use Case:	Pharmacy Visit for Medication Refill
Overall Risk	□ No Risk □ Low Risk ✓ Moderate Risk □ Significant Risk □ High Risk
Key Risk Factors	• Moderate potential harm if any medical instructions have changed • Protected health information (HIPAA) • Some pharmaceutical terminology • Structured and predictable conversation flow • Limited emotional content • Moderate accessibility considerations
Recommended Approach	□ AI Interpreting Only □ AI with Basic Safeguards ✓Human with AI Backup □ Human Interpreter Only

Use Case:	Pharmacy Visit for Medication Refill
Required Safeguards	HIPAA-compliant AI solution Pharmaceutical terminology optimization Bank of pre-translated common questions and answers Quality monitoring for medication names and dosages Immediate pharmacist or human interpreter availability Verification protocols Basic accessibility features Ability to switch to a human interpreter at any point
Intervention Triggers	Questions about new medications or changes in dosage Complex medication interactions Side effect reports or adverse reactions Accuracy confidence of the translation algorithm falls below threshold for medication names or allergies in either language Patient with hearing loss or deafness requiring additional accommodations, including sign language
Compliance Requirements	HIPAA- Pharmacy Board regulations ADA requirements for effective communication Section 504 if federally funded pharmacy State pharmacy practice requirements

Use Case:	Pharmacy Visit for Medication Refill
Accessibility Considerations	Text alternatives to verbal instructions Clear visual information for medication dosage Multiple communication modalities Support for hearing aid compatibility
Notes	Well-suited for AI with human backup as conversation is fairly structured but involves medication information and relevant health history that requires accuracy

3.6.3 Example 3 - SIGNIFICANT RISK: Routine Medical Follow-up Visit

Use Case:	Routine Medical Follow-up Visit
Scenario Description	Patient follow-up for chronic condition management with physician. Discussion of medication adherence, symptom changes, and basic care plan adjustments including behavior change. Typically 15 minutes, in-person at the medical office.
Overall Risk	□ No Risk □ Low Risk □ Moderate Risk ✓ Significant Risk □ High Risk

Use Case:	Routine Medical Follow-up Visit
Key Risk Factors	Significant potential harm from misunderstanding medical instructions Protected health information (HIPAA) Complex medical terminology and specialized vocabulary Detailed discussion of symptoms and treatment plans Potential for emotional content related to health conditions Significant regulatory compliance requirements Potentially vulnerable patients (elderly, chronically ill) Various accessibility considerations
Recommended Approach	□ AI Interpreting Only □ AI with Basic Safeguards □ AI with Human Backup ✓ Human Interpreter Only
Required Safeguards	HIPAA-compliant AI solution Comprehensive medical terminology training in the language pair Real-time quality monitoring by medical interpreter Human interpreter with medical interpretation experience Provider and patient informed consent Comprehensive accessibility features

Use Case:	Routine Medical Follow-up Visit
	Clear protocol for interpreter intervention
Escalation Triggers	Not applicable
Compliance Requirements	HIPAA Section 1557 of ACA ADA requirements for effective communication Section 504 if federally funded provider State medical interpreting requirements Medical ethics guidelines
Accessibility Considerations	Qualified medical sign language interpreters for Deaf patients Interface compatible with hearing aids/cochlear implants Visual alternatives/accompaniments to audio information Multiple communication modalities available Appropriate font sizes and contrast for visual content Specialized accommodations based on patient needs
Notes	Human interpreters should lead these interactions. Human supervised AI support for basic information may be appropriate. Medical follow-up visits involve critical healthcare information and nuanced discussions that can significantly impact patient care.

3.6.4 Example 4 - HIGH RISK: Legal Proceeding - Court Testimony

Use Case:	Legal Proceeding - Court Testimony
Scenario Description	Witness testimony during court proceedings, under oath. May include detailed accounts of events, cross-examination, and legal terminology. Duration varies from 30 minutes to several hours. High-stakes legal environment.
Overall Risk	□ No Risk □ Low Risk □ Moderate Risk □ Significant Risk ✓ High Risk
Key Risk Factors	Severe consequences from errors (affecting legal outcomes) Complex legal terminology and concepts High emotional content possible Critical need for precision and nuance Significant regulatory and due process requirements Strict ADA and disability law requirements Potential for complex disability accommodation needs
Recommended Approach	□ AI Interpreting Only □ AI with Basic Safeguards

Use Case:	Legal Proceeding - Court Testimony
	□ AI with Human Backup ✓ Human Interpreter Only
Required Safeguards	Certified court interpreters only Legal terminology expertise Ability to capture cultural nuances Process for requesting clarification Specialized deaf/disability accommodations if needed
Escalation Triggers	N/A (human interpreter required)
Compliance Requirements	Court interpreting certification requirements Due process protections ADA Title II requirements for courts Section 504 compliance if federally funded Court rules of procedure State/federal language access laws
Accessibility Considerations	Qualified sign language interpreters for Deaf individuals Certified Deaf Interpreters when appropriate CART services (real-time captioning) when needed

Use Case:	Legal Proceeding - Court Testimony
	Visual/tactile accommodations for DeafBlind individuals Strategic positioning of interpreters in courtroom
Notes	Not appropriate for AI interpreting due to high stakes, complexity, legal requirements, and disability accommodation needs

---

Evaluator(s): _____

Date: _____

Next Review

Date: _____

Checklist 4: Vendor Assessment Checklist

Selecting a high-quality AI solution requires evaluating its functionality, accuracy, security, and ethical design. This comprehensive checklist provides ten categories to systematically assess vendors across these areas.

The ten assessment categories are organized into individual checklist tables (below), each containing detailed evaluation criteria and guiding questions to support thorough vendor evaluation.

The assessment is designed to help organizations:

1. **Evaluate Technical Performance** - Assess language coverage, speech and sign recognition accuracy, translation quality, and system adaptability across various environments.
2. **Ensure Security and Compliance** - Verify proper data handling, privacy protections, and regulatory compliance with standards such as HIPAA and ISO27001.
3. **Verify Ethical Standards** - Examine transparency in AI usage, bias mitigation strategies, primary communicator controls, and disclosure practices.
4. **Assess Customization Capabilities** - Evaluate terminology management, domain specialization, and system learning capabilities.
5. **Determine Operational Readiness** - Review usability, accessibility compliance, technical support, and escalation protocols.

Organizations can use the scoring system to objectively compare vendors against their specific requirements and priorities. The checklist provides a clear recommendation pathway that guides decision-makers toward pilot testing, leading to implementation or further evaluation, or rejection based on comprehensive assessment results.

This checklist supports evidence-based procurement decisions for AI interpreting technologies by prioritizing quality, security, ethics, and practical implementation considerations.

- Note: You may choose to create a weighted system based on the setting and use cases. For example, in healthcare settings "Backup and Escalation" and "Compliance & Legal" would have higher weight than in some other settings.

4.1 User Experience and Accessibility

Criterion	Assessment Questions	Ready	Not Ready	Notes
Opting in and out	Does the system provide primary communicators with clear options/toggles to opt in (accept) and out of (decline) using AI for interpreting throughout the entire usage? Separately, does the system provide clear and accessible informed consent of data captured during use? How does the system behave or prioritize if there are multiple primary communicators involved with differing preferences (i.e. one opts in, but the other opts out), ensuring the system does not pressure specific choices?	☐ ☐ ☐	☐ ☐ ☐	
User Interface (UI)	How intuitive is the user interface? Is it clear, efficient, consistent, and easy to use? Has the UI been tested with target users? Are there reports on user satisfaction metrics? What training is required for effective use by primary communicators?	☐ ☐ ☐	☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
	How long is the average waiting time when switching from AI to human interpreting or vice versa?	☐	☐	
	Is there a way for users to submit feedback, suggestions, and/or complaints to the system? To neutral third parties?	☐	☐	
	While the system is processing (e.g., listening, interpreting, switching between AI and human), is there clear visibility of system status?	☐	☐	
	Does the UI display real-time confidence levels in the automated translation? Does it show if there are errors or breakdowns in a user-friendly way?	☐	☐	
Accessibility	Is the system interoperable with other telecommunications devices?	☐	☐	
	Does the system meet WCAG accessibility standards?	☐	☐	
	Are user interfaces and user controls customizable to meet user preferences and accessibility requirements?	☐	☐	
	How are primary communicators with disabilities accommodated?	☐	☐	
	Has accessibility been independently verified?	☐	☐	
	What are the credentials of the third-party evaluator?	☐	☐	
	Can UI be localized to the user's language?	☐	☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Mobile Compatibility	Is there a mobile application or mobile-friendly interface? What are the minimum device requirements? How is battery consumption managed?	☐ ☐ ☐	☐ ☐ ☐	
Hardware Requirements	What hardware is required for optimal performance? Are there specific microphone/camera requirements? What connectivity is needed? Is hardware equipped with theft-prevention tags (e.g., RFID)?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Offline Capabilities	Does the system function offline? What features are limited in offline mode? How does it sync when connectivity returns?	☐ ☐ ☐	☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
READINESS SCORE TOTAL (4.1)	**(add 1 for each checkmark)**	**__ / 27 Total Ready**	**__ / 27 Total Not Ready**	**READINESS SCORE TOTAL (4.1)**

4.2 Technical Capabilities

Criterion	Assessment Questions	Ready	Not Ready	Notes
Language Coverage	Does the vendor support all languages required by your organization? Can the system handle multiple language pairs in the same session? Are mixed languages (like Spanglish or Hinglish) properly supported? If all the language pairs needed are not provided, what is the strategy to meet the need? For each language pair, does the vendor support the pair bidirectionally?	☐ ☐ ☐ ☐ ☐ ☐ ☐	☐ ☐ ☐ ☐ ☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
	Is there data to show accuracy by reliability across languages and per language pair? How well does the solution handle low-resource languages? Can the system handle multiple languages in the same session?	☐	☐	
Speech/Sign Recognition	Is the system error rate minimal across all language pairs, including with fast articulation? How does the system perform with utterances that are accented, expressed by children, dysarthric, disfluent, or by individuals with varying disabilities (i.e. speech or physical impairments)? How does the system handle various qualities of input data (e.g., background audio or visual noise, lighting, angles and distances of speakers?	☐ ☐ ☐	☐ ☐ ☐	
Translation Quality	What are the evaluation metrics, (e.g., machine translation metrics, primary communicator satisfaction metrics, dataset sizes, language pair accuracy), for each language pair? Can the vendor provide sample transcripts of translations? Does the system backtrack and make corrections with more context? How are these communicated to primary communicators? How does the system handle idioms, slang, and cultural references?	☐ ☐ ☐ ☐ ☐	☐ ☐ ☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
	How does translation quality compare to available qualified human interpreters?	☐	☐	
	How does the system process unrecognized terms (new words or proper nouns)?	☐	☐	
	Does the system degrade gracefully when there are communication breakdowns (rather than hallucinate)?	☐	☐	
Speech/Sign Synthesis	How natural is the generated output data? What are the evaluation metrics for each language pair?	☐	☐	
	How does the attention span and data comprehension change with AI output compared to human output?	☐	☐	
	Does the system maintain appropriate intonation and emphasis, and match the intentions of the primary/principal communicators involved?	☐	☐	
	Are there options for characteristic (e.g., voice, gender) customization?	☐	☐	
	Can the speed of voice-to-signing be regulated, customized by primary communicators?	☐	☐	
Performance Metrics	What is the average latency (delay) between inputs, processes, and outputs?	☐	☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
	What is the system's uptime guarantee?	☐	☐	
	How does performance degrade under heavy or unstable network loads?	☐	☐	
	Does the UI alert primary communicators when confidence drops below a threshold?	☐	☐	
	Does the UI provide customization options (e.g., customizing or selecting from automatically generated outputs)?	☐	☐	
Environmental Adaptability	What background noise mitigation strategies does the tool include?	☐	☐	
	Can it handle crosstalk or multiple speakers?	☐	☐	
	Does it work effectively both onsite and remotely or for hybrid solutions?	☐	☐	
Integration Capabilities	Can the system integrate with your existing platforms?	☐	☐	
	If not, what equipment and infrastructure are needed?	☐	☐	
	What APIs are available for custom integration?	☐	☐	
	Is there compatibility with telehealth or video conferencing systems?	☐	☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Predictive Translation	Does the system support predictive translation suggestions to interpreters?	☐	☐	
Mobile Compatibility	Is there a mobile application or mobile-friendly interface?	☐	☐	
	What are the minimum device requirements?	☐	☐	
	How is battery consumption managed?	☐	☐	
Hardware Requirements	What hardware is required for optimal performance?	☐	☐	
	Are there specific microphone/camera requirements?	☐	☐	
	What connectivity is needed?	☐	☐	
	Is hardware equipped with theft-prevention tags (e.g., RFID)?	☐	☐	
Offline Capabilities	Does the system function offline?	☐	☐	
	What features are limited in offline mode?	☐	☐	
	How does it sync when connectivity returns?	☐	☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
READINESS SCORE TOTAL (4.2)	**(add 1 for each checkmark)**	**__ / 47** **Total Ready**	**__ / 47** **Total Not Ready**	**READINESS SCORE TOTAL (4.2)**

4.3 Data Security and Privacy

Criterion	Assessment Questions	Ready	Not Ready	Notes
Data Handling Policies	Does the vendor store recordings or transcripts according to the principal communicator's opt out/opt in decisions? How long is data retained? Who retains ownership of the data? What is the vendor's data re-use policy? Are primary/principal communicators informed of data handling policies? Can they change their minds after initiating system use?	☐ ☐ ☐ ☐ ☐	☐ ☐ ☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Encryption	What encryption standards are used in transit and at rest? Is end-to-end encryption available? How are encryption keys managed?	☐ ☐ ☐	☐ ☐ ☐	
Certifications	Does the vendor have relevant certifications that apply to your sector (e.g., HIPAA, ISO27001, SOC 2, FERPA) or other relevant certifications? Can they provide documentation of compliance? When were certifications last renewed?	☐ ☐ ☐	☐ ☐ ☐	
Access Controls	Who at the vendor has access to your data? How is staff access managed and audited? How can user access get disabled? What authentication methods are used?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Data Breach Protocols	What is the vendor's data breach notification process? What mitigation strategies are in place? Has the vendor experienced previous breaches?	☐ ☐ ☐	☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Third-Party Sharing	Is data shared with third parties for any purpose? How is data anonymized before sharing? Can data sharing be opted out of while retaining permission for data storage to be exclusively used internally?	☐ ☐ ☐	☐ ☐ ☐	
READINESS SCORE TOTAL (4.3)	**(add 1 for each checkmark)**	**__ / 21 Total Ready**	**__ / 21 Total Not Ready**	**READINESS SCORE TOTAL (4.3)**

4.4 Transparency and Ethics

Criterion	Assessment Questions	Ready	Not Ready	Notes
AI Labeling	How is AI-generated output translation clearly labeled? Is a welcome message played/stated that AI interpreting is in use? Are recipients continuously provided an actionable an opt-out option, i.e., switching to a human interpreter? Does the system meet ASTM F2575-23e2 standards?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Error Disclosure	How are system errors disclosed to primary communicators? Is there a clear indication when confidence is low? Are there warnings for potentially inaccurate translations?	☐ ☐ ☐	☐ ☐ ☐	
Bias Mitigation	Has the system been audited for bias? What measures address gender, racial, or cultural bias? What measures address bias stemming from inadequate language pair accuracy? Describe various diversity parameters applied to the training datasets.	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Training Data Transparency	What data was used to train the system? How was training data selected and vetted? Is the training process documented? Does the interface disclose what data is collected and why?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Human Oversight and Review	What human oversight exists for the AI system? What credentials do human reviewers hold? What credentials do human interpreters hold? How is human interpreter quality assured? When does human oversight occur? How are corrections provided?	☐ ☐ ☐ ☐ ☐ ☐	☐ ☐ ☐ ☐ ☐ ☐	
READINESS SCORE TOTAL (4.4)	**(add 1 for each checkmark)**	**__ / 21 Total Ready**	**__ / 21 Total Not Ready**	**READINESS SCORE TOTAL (4.4)**

4.5 Customization and Learning

Criterion	Assessment Questions	Ready	Not Ready	Notes
Terminology Management	Can industry-specific terminology be customized?	☐	☐	
	How are custom terms integrated into the system?	☐	☐	
	What are the limits regarding term customization?	☐	☐	
	Can you customize the pronunciation of proper nouns (e.g., names)?	☐	☐	
	Does training occur upfront in the pre-launch phase or by session?	☐	☐	
	Can terminology be imported from existing glossaries?	☐	☐	
	Can the organization support the cost implications of customization?	☐	☐	
Adaptive Learning	Does the system improve based on corrections?	☐	☐	
	How quickly are improvements implemented?	☐	☐	
	Is learning specific to your organization or shared?	☐	☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Domain Specialization	Can the system be optimized for specific domains? What domains are currently optimized? How is domain specialization implemented?	☐ ☐ ☐	☐ ☐ ☐	
Data Ownership	Who owns data used to train or improve the system? Can you export your custom improvements? Are your improvements shared with other clients or companies?	☐ ☐ ☐	☐ ☐ ☐	
Feedback Mechanisms	How can primary communicators provide feedback on translations? How is feedback incorporated into the system? Is there a formal feedback loop documented?	☐ ☐ ☐	☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
READINESS SCORE TOTAL (4.5)	**(add 1 for each checkmark)**	**__ / 19 Total Ready**	**__ / 19 Total Not Ready**	**READINESS SCORE TOTAL (4.5)**

4.6 Support and Service

Criterion	Assessment Questions	Ready	Not Ready	Notes
Technical Support	What support hours and response times are guaranteed?	☐	☐	
	What support channels are available?	☐	☐	
	Is support available in multiple languages?	☐	☐	
	Who has access to the support?	☐	☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Implementation Assistance	What assistance is provided during implementation? Is there a dedicated implementation manager? What is the initial setup time promised in the Service Level Agreement (SLA)? What training resources are provided?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	
Documentation	How comprehensive is system documentation? Are there user guides for different roles, covering behind-the-scenes support roles, management, and primary/principal communicators? How often is documentation updated?	☐ ☐ ☐	☐ ☐ ☐	
Service Level Agreements	What uptime percentage is guaranteed? What are the penalties for SLA violations? How is SLA compliance monitored and reported? What is the guaranteed response time for service tickets? Escalation process?	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Updates & Maintenance	How frequently are updates released? How are updates communicated and implemented? Is there a roadmap for future improvements?	☐ ☐ ☐	☐ ☐ ☐	
READINESS SCORE TOTAL (4.6)	**(add 1 for each checkmark)**	**__ / 18 Total Ready**	**__ / 18 Total Not Ready**	**READINESS SCORE TOTAL (4.6)**

4.7 Backup and Escalation

Criterion	Assessment Questions	Ready	Not Ready	Notes
Human Interpreter Access	How quickly can sessions escalate to human interpreters? Is there automatic escalation? Can the platform run hybrid mode (AI continues while a human interpreter joins)?	☐ ☐ ☐	☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
	Can our organization program parameters or keywords for triggering automatic escalation (rollover) to human interpreters?	☐	☐	
	How seamless is the transition process?	☐	☐	
Escalation Protocols	What is the process for user-initiated escalation (rollover)?	☐	☐	
	How are escalation events documented?	☐	☐	
	Is there a limit to escalation (rollover) frequency?	☐	☐	
Backup Systems	What redundancy exists in case of system failure? Request system uptime historical data and determine if it meets our organizational requirements.	☐	☐	
	How quickly do backup systems activate?	☐	☐	
	How are primary communicators notified of system issues?	☐	☐	
	What remedies are offered due to harms that arise because of system issues?	☐	☐	
READINESS SCORE TOTAL (4.7)	**(add 1 for each checkmark)**	**__ / 12 Total Ready**	**__ / 12 Total Not Ready**	**READINESS SCORE TOTAL (4.7)**

4.8 Compliance and Legal

Criterion	Assessment Questions	Ready	Not Ready	Notes
Regulatory Compliance	How does the system support HIPAA compliance? How does it support Title VI requirements? How does it address ADA obligations?	☐ ☐ ☐	☐ ☐ ☐	
Liability & Accountability	Who is liable for interpretation errors and uncorrected misunderstandings promoted by the AI's presentation of unquestionable confidence? How is liability documented in contracts? What indemnification is provided?	☐ ☐ ☐	☐ ☐ ☐	
Contract Terms	Are there minimum commitment periods? What are the termination conditions? How are disputes resolved?	☐ ☐ ☐	☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Documentation & Records	What records are kept of interpreting sessions? How long are records maintained? How can records be accessed if needed?	☐ ☐ ☐	☐ ☐ ☐	
READINESS SCORE TOTAL (4.8)	**(add 1 for each checkmark)**	**__ / 12** **Total Ready**	**__ / 12** **Total Not Ready**	**READINESS SCORE TOTAL (4.8)**

4.9 Cost Structure

Criterion	Assessment Questions	Ready	Not Ready	Notes
Pricing Model	Is pricing per minute, per session, or subscription-based? Are there minimum usage requirements? How are unused minutes/credits handled?	☐ ☐ ☐	☐ ☐ ☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Hidden Costs	Are there setup or onboarding fees? Are there charges for customization or training? Are there fees for human escalation (rollover, monitoring, intervention)?	☐ ☐ ☐	☐ ☐ ☐	
Volume Discounts	What volume discount tiers are available? How is volume calculated (monthly, annually)? Are there enterprise pricing options?	☐ ☐ ☐	☐ ☐ ☐	
Cost Comparison	How does cost compare to human interpreters? What is the total cost of ownership? What ROI can be expected?	☐ ☐ ☐	☐ ☐ ☐	
READINESS SCORE TOTAL (4.9)	**(add 1 for each checkmark)**	**__ / 12** **Total Ready**	**__ / 12** **Total Not Ready**	**READINESS SCORE TOTAL (4.9)**

4.10 Vendor Stability and Reputation

Criterion	Assessment Questions	Ready	Not Ready	Notes
Company History	How long has the vendor been in business?	☐	☐	
	What is their financial stability?	☐	☐	
	What is their market share/position?	☐	☐	
	What is their expertise in providing language services?	☐	☐	
	Are they primarily a tech vendor or language services company?	☐	☐	
Client References	Can they provide references from similar organizations?	☐	☐	
	What is their client retention rate?	☐	☐	
	Are case studies available?	☐	☐	
	Are third-party user-feedback data (e.g., public ratings) available?	☐	☐	
Industry Recognition	Have they received industry awards or recognition?	☐	☐	
	Are they mentioned in analyst reports?	☐	☐	
	Do they participate in standards development?	☐	☐	

Criterion	Assessment Questions	Ready	Not Ready	Notes
Future Outlook	What is their product roadmap? What is their funding situation? Are there acquisition risks?	☐ ☐ ☐	☐ ☐ ☐	
READINESS SCORE TOTAL (4.10)	**(add 1 for each checkmark)**	**__ / 15** **Total Ready**	**__ / 15** **Total Not Ready**	**READINESS SCORE TOTAL (4.10)**

Total Scores for Checklist 4

Subtotals	Ready	Not Ready
4.1	_____ / 27	_____ / 27
4.2	_____ / 47	_____ / 47
4.3	_____ / 21	_____ / 21
4.4	_____ / 21	_____ / 21

4.5	_____ / 19	_____ / 19
4.6	_____ / 18	_____ / 18
4.7	_____ / 12	_____ / 12
4.8	_____ / 12	_____ / 12
4.9	_____ / 12	_____ / 12
4.10	_____ / 15	_____ / 15
Grand Total:	_____ **/ 204**	_____ **/ 204**

4.11 Decision Recommendation

Based on the assessment above:

- [ ] Recommended for Implementation for the following use cases:
    - _____
    - _____
- [ ] Recommended for Pilot Testing for the following use cases:
    - _____

- o _____

- [ ] Requires Further Evaluation
- [ ] Not Recommended

Justification: _____

Evaluator(s): _____

Date: _____

Checklist 5: Guidance for Request for Proposals (RFPs) that Include AI Interpreting

This checklist provides structured guidance for procurement professionals writing Requests for Proposals (RFPs) that include AI interpreting tools or automated language services. It translates the Toolkit's assessment framework into actionable RFP language and requirements, ensuring vendors provide the information necessary for thorough evaluation while maintaining compliance with federal procurement regulations. The guidance includes template language for key RFP sections, evaluation criteria that align with the Toolkit's three-pillar framework, and specific requirements for AI technology procurement that address both spoken and sign language interpreting needs.

The guidance helps organizations define the scope of services, set clear technical specifications, establish mandatory compliance requirements, and structure a mandatory pilot testing phase to validate vendor claims in real-world scenarios. It also helps calculate the total cost of implementation, which includes not only subscription fees but also indirect costs like staff training, IT upgrades, and maintaining dual systems for human interpreters.

The guidance includes template language for key RFP sections, evaluation criteria that align with the Toolkit's three-pillar framework, and specific requirements for AI technology procurement that address both spoken and sign language interpreting needs.

5.1 Pre-RFP Planning Using the Toolkit

Before drafting your RFP, complete the following assessments using other appendices:

Step 1: Complete Organizational Readiness Assessment (Checklist 1)	• Identify gaps that must be addressed before implementation • Determine if your organization is ready to manage AI interpreting technology • Document specific training and infrastructure needs
Step 2: Conduct Risk Assessment (Checklist 3)	• Categorize your use cases by risk level (no, low, moderate, significant, high) • Identify which scenarios are appropriate for AI interpreting, hybrid AI-human interpreting or human-only interpreting • Establish clear escalation requirements
Step 3: Define Technical and Compliance Requirements	• Map language needs across all user groups • Identify regulatory compliance requirements (HIPAA, ADA, Title VI, etc.) • Determine integration needs with existing systems

5.2 Essential RFP Sections for AI Interpreting Services

5.2.1 Scope of Services and Requirements	**Template Language:** *[Agency Name] seeks qualified vendors to provide AI-powered interpreting services to complement human interpreting capabilities. Services must support both spoken language interpretation and sign language interpretation as specified in this RFP.*
	Core Service Requirements: • AI interpreting technology for [specify languages and modalities] • Seamless escalation/rollover to qualified human interpreters • Integration with existing [healthcare/legal/education/business] systems • Compliance with all applicable federal, state, and local regulations
	Use Case Categories: The vendor must demonstrate capability across the following risk levels: • Low-risk conditions: [list specific use cases from your assessment] • Moderate-risk conditions: [list specific use cases requiring human backup] • High-risk conditions: [clarify that human interpreters are required]
	Technology Requirements and Vendor Response: The client will be responsible for providing and maintaining all necessary IT infrastructure, including sound systems, visual equipment, and lighting adequate for interpretation services. However, applicants must

	demonstrate in their proposals how their AI interpreting solution will perform under varying technical conditions and address potential challenges including: ● System response to technical failures or equipment malfunctions ● Performance degradation due to poor audio quality, background noise, and/or limited bandwidth ● Functionality limitations when lighting conditions and/or bandwidth are insufficient for sign language interpretation ● Protocols for communicating technical requirements to client IT staff ● Recommended minimum technical specifications for optimal system performance ● Contingency plans when technical conditions fall below optimal thresholds Vendors must acknowledge that AI interpreting technology functions optimally only under controlled conditions and must specify how their solution adapts to or compensates for suboptimal technical environments that may occur despite the client's IT infrastructure efforts.

5.2.2 Technical Specifications	**Template Language:** *Vendors must demonstrate that their AI interpreting solutions meet the technical performance standards outlined below. All proposed systems must be capable of operating effectively in real-world conditions while maintaining accuracy, reliability, and accessibility standards. Vendors shall provide detailed documentation of system capabilities, performance metrics, and technical requirements necessary for successful deployment and ongoing operation.*
	Language Coverage Requirements: ● Spoken Languages: [specify required languages, dialects, and regional variations] ● Sign Languages: [specify ASL, BSL, and other signed language variants]

	• Mixed-language support for code-switching scenarios [such as Spanglish or Hinglish] • Minimum accuracy thresholds: [specify percentages by language pair]
	Performance Standards: • Provide method of testing, testing conditions and parameters, the volume tested, the date of the last evaluation, and the evaluation frequency • Maximum latency: [specify milliseconds] for real-time interpretation • Uptime guarantee: [specify percentage] with documented SLA penalties • Transcription accuracy of source • Word Error Rate (WER): Maximum [specify percentage] for each language pair • Confidence scoring with automatic alerts below [specify threshold]
	Technology Requirements and Vendor Response: The client will be responsible for providing and maintaining all necessary IT infrastructure, including sound systems, visual equipment, and lighting adequate for interpretation services. However, applicants must demonstrate in their proposals how their AI interpreting solution will perform under varying technical conditions and address potential challenges including: • System response to technical failures or equipment malfunctions • Performance degradation due to poor audio quality or background noise • Functionality limitations when lighting conditions are insufficient for sign language interpretation • Protocols for communicating technical requirements to client IT staff

	• Recommended minimum technical specifications for optimal system performance • Contingency plans when technical conditions fall below optimal thresholds Vendors must acknowledge that AI interpreting technology functions optimally only under controlled conditions and must specify how their solution adapts to or compensates for suboptimal technical environments that may occur despite the client's IT infrastructure efforts.
	Technical Requirements: • Hardware and software requirements • APIs for integration with [specify existing systems] • Compatibility with [specify telehealth, video conferencing, or other platforms] • Data export capabilities for quality monitoring and reporting (with details on what can be reported) • Accessibility compliance with WCAG 2.1 AA standards • IT support and training needed to deploy the solution
	Security and Privacy: • End-to-end encryption for all interpreted sessions • Dual authorization framework: explicit opt-in acceptance for AI use and separate consent for data use (with opt-out capability for data consent) • Zero data retention policies for sensitive information • HIPAA compliance with executed Business Associate Agreement • SOC 2 Type II certification or equivalent security standards

5.2.3 Vendor Qualifications and Experience	**Template Language:** *Vendors must demonstrate that their AI interpreting solutions meet the technical performance standards outlined below. All proposed systems must be capable of operating effectively in real-world conditions while maintaining accuracy, reliability, and accessibility standards. Vendors shall provide detailed documentation of system capabilities, performance metrics, and technical requirements necessary for successful deployment and ongoing operation.*
	Required Qualifications: ● Demonstrated ability to deliver AI interpreting services at expected volumes for [government/healthcare/legal/education] organizations, with documented performance metrics and capacity planning ● Demonstrated compliance with applicable regulations in your sector ● Technical support capabilities including 24/7 availability for critical scenarios ● Financial stability documentation (audited financial statements for past 2 years)``
	Past Performance Requirements: Vendors must provide detailed case studies demonstrating: ● Accuracy metrics and conditions used for measurement ● User satisfaction data ● Incident management and resolution procedures ● Escalation protocol effectiveness
	Reference Requirements: Provide contact information for three (3) current clients with similar requirements, including: ● Organization name and primary contact

	• Contract value and duration • Scope of services provided • Volume of AI interpreting delivered in hours • Performance metrics and satisfaction levels"

5.2.4 Evaluation Criteria Aligned with Toolkit Framework	**Template Language:** *Proposals will be evaluated using a comprehensive scoring framework that aligns with the three-pillar assessment approach outlined in this Toolkit. Each criterion includes specific metrics and documentation requirements. Vendors must address all evaluation categories to be considered for award. Scoring will be based on demonstrated capabilities, measurable outcomes, and alignment with organizational requirements.*
	Technical Fitness (35% of total score): • Language coverage and accuracy metrics (10%) • Performance benchmarks and reliability (10%) • Security and compliance capabilities (10%) • Integration and accessibility features (5%)
	Total Cost of Implementation (30% of total score): • Initial licensing and setup costs (10%) • Ongoing operational expenses (10%) • Training and support costs (5%)

	• Hidden costs and fee transparency (5%)
	Reference Requirements: Provide contact information for three (3) current clients with similar requirements, including: • Organization name and primary contact • Contract value and duration • Scope of services provided • Volume of AI interpreting delivered in hours • Performance metrics and satisfaction levels"
	Organizational Readiness Support (25% of total score): • Implementation planning and support (10%) • Training programs and materials (5%) • Quality monitoring and reporting tools (5%) • Escalation and backup human interpreter services (5%)
	Vendor Stability and Support (10% of total score): • Financial stability and market position (5%) • Technical support and maintenance capabilities (5%)

5.2.5 Mandatory Compliance Requirements	**Template Language:** *All vendors must meet mandatory compliance requirements as a condition of contract award. Failure to demonstrate full compliance with any requirement will result in proposal disqualification. Vendors must provide detailed documentation and certifications as evidence of compliance. These requirements are non-negotiable and reflect legal obligations that cannot be waived or modified during the procurement process.*
	Regulatory Compliance (Pass/Fail): All proposals must demonstrate compliance with: ● [Include applicable regulations based on your sector] ● Title VI of the Civil Rights Act (for federally funded programs) ● Americans with Disabilities Act (ADA) requirements ● Section 508 accessibility standards (for federal agencies) ● HIPAA requirements (for healthcare settings) ● State-specific interpreter qualification requirements
	AI Technology Disclosure: Vendors must provide: ● Baseline performance metrics for all supported language pairs, documented as current capabilities at the time of proposal submission ● Current training data sources and existing bias mitigation strategies, with acknowledgment that these will continue to evolve ● Algorithm transparency and explainability measures ● Quality assurance and error correction procedures, with prioritization for on-demand accessibility while documenting any current limitations that prevent immediate availability

	Human Interpreter Integration: Demonstrate capability to: • Provide mechanism for immediate escalation to qualified human interpreters and the triggers (e.g. manual by host, manual by any participant, triggered by keywords, triggered by number of participants, triggered by low quality flag, etc.] • Maintain roster of qualified interpreters for high-risk scenarios • Support hybrid delivery models (AI with human oversight) • Ensure seamless transition between AI and human services

5.2.6 Pilot Testing Requirements	**Template Language:** *All vendors selected for final consideration must successfully complete a mandatory pilot testing phase before contract award. The pilot serves as the final validation of vendor claims and system performance under real-world conditions. Pilot results will directly inform the final selection decision, and vendors must demonstrate measurable success across all defined criteria to proceed to contract award.* The pilot must include:
	Testing Scenarios: • Representative use cases from each risk category identified • Various environmental conditions (noise, lighting, etc.) • Multiple language pairs and communication modalities • Escalation scenarios requiring human interpreter backup

	Success Metrics: • Accuracy rates meeting or exceeding specified thresholds • User satisfaction scores of [specify minimum percentage] • Technical performance within specified parameters • Successful escalation protocol execution
	Pilot Evaluation Process: • Joint evaluation team including technical and end-user representatives • Structured feedback collection from all stakeholders • Documentation of any issues and resolution procedures • Final pilot report with go/no-go recommendation

5.2.7 Contract Terms and Risk Management	**Template Language:** *The contract will establish clear performance standards, risk allocation, and management procedures to ensure successful service delivery and protect both parties' interests. All terms are designed to maintain accountability, ensure quality outcomes, and provide mechanisms for addressing performance issues throughout the contract lifecycle.*
	Performance Standards and Penalties: • Service Level Agreement with clearly defined uptime requirements • Accuracy thresholds with financial penalties for non-compliance • Response time guarantees for technical support and escalation

	• Quality monitoring and reporting requirements
	Risk Allocation: • Vendor liability for interpretation errors in specified scenarios • Professional liability insurance requirements • Indemnification clauses for compliance failures • Data breach notification and mitigation procedures
	Contract Management: • Regular performance reviews and adjustment mechanisms • Technology update and improvement requirements • Termination clauses for performance failures • Transition assistance for contract end or termination

5.3 RFP Checklist Using Toolkit Components

Criterion	Assessment Questions	Notes
Pre-Release Overview	☐ Organizational readiness assessment completed (CHECKLIST 2) ☐ Risk assessment framework applied to use cases (CHECKLIST 3) ☐ Vendor assessment criteria developed (CHECKLIST 1) ☐ Setting-specific requirements included (CHECKLIST 4)	

Criterion	Assessment Questions	Notes
	▯ Legal and compliance review completed ▯ Stakeholder input gathered and incorporated	
Pre-Release Review - RFP Content Verification	▯ Three-pillar evaluation framework clearly articulated (Described in the Introduction of this document) ▯ Technical specifications aligned with organizational needs □ Compliance with requirements in your sector ▯ Pilot testing successful □ Escalation and human interpreter successful as planned ▯ Accessibility and disability accommodation needs addressed	
Post-Award Implementation	▯ Vendor assessment using Checklist 4 criteria ▯ Pilot testing following risk assessment protocols in Checklist 3 ▯ Quality monitoring procedures established ▯ Staff training programs planned and scheduled ▯ Regular performance review schedule implemented	
Sample RFP Timeline	**Week 1-2:** Complete Toolkit assessments and stakeholder consultations **Week 3-4:** Draft RFP using Toolkit guidance and template language **Week 5:** Internal review and legal compliance verification **Week 6:** RFP release with minimum 30-day response period	

Criterion	Assessment Questions	Notes
	Week 10-12: Proposal evaluation using Toolkit criteria **Week 13-14:** Vendor selection and pilot testing preparation **Week 15-26:** 90-day pilot testing and evaluation (12 weeks) **Week 27:** Final vendor selection and contract award	

Glossary of Terms

This glossary provides plain-language definitions for technical terms used throughout the AI Interpreting Solutions Evaluation Toolkit.

Term	Definition
AI (Artificial Intelligence)	Technology that mimics human abilities like understanding speech or generating / recognizing signs. In this Toolkit, AI is used for interpreting between languages—either spoken or signed.
AI Interpreting Tool	Software or a system that automatically converts language from one spoken or signed language to another using AI.
Accessibility	Making sure that everyone can use tools or services, including people who are Deaf, have low vision, mobility challenges, or use different languages.
ASL (American Sign Language)	A complete visual language used primarily by the Deaf community in the United States and parts of Canada, with its own grammar and sentence structure.
Backchanneling	Communication between interpreters during an interpretation session to clarify meaning, correct errors, or provide support. This can include visual cues, whispered corrections, or brief consultations to ensure accuracy.
Bias (AI Bias)	When an AI system works better for some groups of people than others due to who or what it was trained on. This can affect fairness and understanding in interpretation.
Certified Deaf Interpreter (CDI)	A Deaf or hard-of-hearing interpreter certified by RID, fluent in ASL and trained in culture-based and visual techniques (e.g. gesture, mime, props). CDIs enhance communication access when a Deaf person's language or communication mode is unique—often working with hearing ASL interpreters as a team but sometimes working independently.
Data Privacy	Protecting personal or sensitive information, especially health, legal, or identity-related data.

Term	Definition
Deaf+	A term for Deaf people who also have additional disabilities, such as being blind or having cognitive disabilities.
Escalation (Protocol)	A planned process to switch from AI to a human interpreter when the situation becomes too complicated, risky, or unclear for the AI.
Fingerspelling	Spelling out words by signing the letters of a manual alphabet. Used in ASL and many other signed languages.
High-Risk Scenario	A situation where mistakes in interpretation can cause serious problems (for example, in medical or legal settings, or for those who have experienced trauma or abuse).
Hybrid Model	A system where AI and human interpreters both play a role. For example, AI interprets first, and a human joins or corrects when needed.
Interpreter (Human)	A trained professional who helps people communicate by converting spoken or signed language from one language to another in real time.
Interpreter Team Open Process	A collaborative approach where multiple interpreters work together openly during a session, with clear communication between team members and transparency about when they are consulting with each other or making corrections.
Language Access	The ability for people to get information and services in a language they use and understand.
Language Pair	The combination of languages used in the interpreting interaction, such as ASL and English, or English and Mandarin
Language Pair Accuracy	Refers to the translation accuracy rates when going from Language A to Language B and Language B to Language A
LEP (Limited English Proficient)	People who do not speak English as their first language and may not be fluent enough to understand complex spoken or written English.
Low Resource Languages	Low resource languages are those that do not have a large amount of written text available on the internet.

Term	Definition
Machine Translation	When a computer program translates text or speech from one language to another using datasets and programmed rules.
Modality	The "type" of communication—spoken, written, or signed.
Multilingual Communication	An interaction where people use more than one language. It can include spoken and signed languages.
Non-manual signals	Facial expressions, head movement, or body posture used in sign languages for grammatical structure and to show meaning or emotion.
Opt-Out	The right to say "no" to using AI and choose a human interpreter instead.
Pilot Test (or Pilot)	A small test of a technology before fully using it. Done to make sure it works well in real-life situations.
Plain Language	Clear and simple writing that is easy to understand and reduces possible dual meanings.
Primary/Principal Communicator(s)	Each of the individuals involved in the interpreted interaction, who have the most to gain from full comprehension of each other and the most to lose by being misunderstood.
Proof of Concept	A demonstration of technology but not yet tested in real life or under rigorous conditions.
Ready/Not Ready	A basic way to evaluate if your organization has what it needs to safely use AI interpreting tools for a specific situation.
Relay Interpreting	An interpreting method where communication passes through multiple interpreters in sequence. In **consecutive relay interpreting**, a message is interpreted from one language to another, then from that language to a third language in turn-taking fashion. In **simultaneous relay interpreting**, multiple interpreters work at the same time to facilitate communication across languages. For example, a Deaf person communicates in ASL to a Certified Deaf

Term	Definition
	Interpreter (CDI), who then conveys the message to a hearing ASL interpreter, who then interprets into spoken English. This process reverses for the return communication.
Risk Factor	A condition, characteristic, or behavior that increases the likelihood of a negative event occurring. Some risk factors can be modified, and fall in categories such as environmental, behavioral, social, psychological, etc.
Risk Score (Risk Level)	A number or label that shows how risky it would be if an interpreter service fails or is wrong. High-risk means a serious mistake could happen.
Sign Language Interpreting	Facilitating communication between a spoken language and a sign language, or between two sign languages.
Technical Fitness	How well an AI interpreting tool performs—how accurate, fast, stable, and secure it is.
Total Cost of Implementation	The full cost of using AI tools—not just buying it, but also training staff, improving equipment, and making sure it works safely in your setting.
Translation Accuracy	How close the interpreted message is to the original meaning.
Turn-taking (in interpreting)	Recognizing when one person finishes speaking so the next person can respond smoothly. Important for spoken and sign language AI tools.
Use Case	A real-world example or situation where interpreting tools are used.
Vendor	A company that sells or provides the AI tool, service, or related support.

Bibliography

ADA Requirements: Effective Communication Regulations. Available at: ADA.gov.

ASTM F2089-14: Standard Practice for Language Interpreting Services.

ASTM F2575-23e2: Standard Practice for Language Translation Services.

Bragg, Danielle, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreault, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler, and Meredith Ringel Morris. "Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective." *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*, October 24, 2019. https://dl.acm.org/doi/10.1145/3308561.3353774.

Coalition for Sign Language Equity in Technology (CoSET). *Deaf-Safe AI: A Legal Foundation for Ubiquitous Automatic Interpreting.*

CSA Research. *Automated Speech-to-Speech Interpreting: Six Evaluation Dimensions for Professional Deployments.*

CSA Research. *Perceptions on Automated Interpreting: Results of a Large-Scale Study of End-Users, Requestors, and Providers of Interpreting Services and Technology.*

CSA Research. *What Language Access Teams Must Know about Automated Speech-to-Speech Interpreting.*

Desai, Aashaka, Maartje De Meulder, Julie A. Hochgesang, Annemarie Kocab, and Alex X. Lu. "Systemic Biases in Sign Language AI Research: A Deaf-Led Call to Reevaluate Research Agendas." *arXiv preprint* arXiv:2403.02563v1 cs.CV, March 5, 2024. https://aclanthology.org/2024.signlang-1.6/.

European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*, L 1689, 12 July 2024.

HIPAA Privacy & Security Guidance. U.S. Department of Health and Human Services. Available at: HHS.gov.

International Organization for Standardization. (2022). Information security, cybersecurity and privacy protection — Information security management systems — Requirements (ISO/IEC 27001:2022). Available at: https://www.iso.org/standard/27001

International Organization for Standardization. ISO 31000:2018 Risk Management Guidelines. Available at: https://www.iso.org/standard/65694.html.

Moghe, N., Fazla, A., Amrhein, C., Kocmi, T., Steedman, M., Birch, A., Sennrich, R., & Guillou, L. (2024). Machine Translation Meta Evaluation through Translation Accuracy Challenge Sets. *Computational Linguistics*, 51(1). https://doi.org/10.1162/coli_a_00537.

Murray, Malcolm. (April 9, 2025). "AI Risk Management Can Learn a Lot From Other Industries." *AI Frontiers*. Available at: https://www.ai-frontiers.org/articles/ai-risk-management-can-learn-a-lot-from-other-industries.

National Council on Interpreting in Health Care (NCIHC). *Guidance for Contracting AI-generated Interpreting.*

Nimdzi Insights, LLC. "The 2023 Nimdzi Interpreting Index: Ranking of the Top Interpreting Companies." Accessed August 24, 2025. https://www.nimdzi.com/interpreting-index-top-interpreting-companies/.

Patient Protection and Affordable Care Act, Section 1557. 45 CFR 92.

Stakeholders Advocating for Fair and Ethical AI in Interpreting Task Force (SAFE AI). *Interpreting SAFE AI Task Force Guidance on AI and Interpreting Services.*

Title VI Compliance Guidelines. U.S. Department of Justice Civil Rights Division. Available at: DOJ.gov.