Wearable Edge-AI Assistant for Multilingual Safety Communication on Construction Sites: A Heat Stress Prevention Application

Yuting Chen[1], Gongfan Chen[2], Lingguang Song[3], and Don Chen[4]

[1,2,3,4] Albert School of Construction, The University of North Carolina at Charlotte at Charlotte, NC, USA 28223

ychen106@charlotte.edu, gchen12@charlotte.edu, lsong2@charlotte.edu, dchen9@charlotte.edu

Abstract –

Language discordance remains a persistent barrier to effective safety communication on U.S. construction sites, particularly for time-critical hazards such as heat stress. Existing digital safety tools and translation applications rely on scripted content, text-based interfaces, or cloud connectivity, limiting their effectiveness in dynamic, multilingual job-site environments. This paper presents a wearable, fully offline edge-AI assistant for real-time bilingual safety communication, with a specific focus on heat-stress prevention. The proposed system integrates offline speech recognition and lightweight LLM-based bilingual translation on a low-power NVIDIA Jetson Nano platform. To ensure fidelity to safety-critical terminology, an 80-sentence English–Spanish heat-stress dataset was developed from OSHA-aligned safety guidance. System performance was evaluated through a real-time feasibility demonstration and offline benchmarking. This work represents the first wearable, edge-only bilingual safety assistant demonstrated for feasibility in heat-stress prevention scenarios and designed for construction jobsite use.

Keywords –

Wearable Edge AI; Multilingual Safety Communication; Domain-adapted LLM; Language Barriers; Construction Workers

1 Introduction

Construction remains one of the most hazardous industries in the United States, accounting for the highest number of occupational fatalities across all sectors. In 2023 alone, 1,075 construction workers lost their lives due to work-related injuries [1]. Despite extensive regulatory frameworks and mandatory safety training requirements, fatal injury rates remain persistently high, indicating that safety failures are often rooted not in the absence of rules but in how safety information is communicated, understood, and acted upon in dynamic jobsite environments.

One persistent challenge undermining construction safety communication is language discordance, defined as the lack of a shared language sufficient to effectively transmit safety-critical information. U.S. construction sites are increasingly multilingual, with English-speaking supervisors frequently working alongside crews whose primary language is Spanish. In 2023, fatalities in the private construction sector accounted for 37.5% (315) of the 839 deaths among foreign-born Hispanic or Latino workers [1]. At the same time, Hispanic workers represented 30.31% of the U.S. construction workforce—approximately 3.7 million workers—highlighting the structural nature of language-related safety risks in the industry [2].

Prior research has demonstrated that language barriers directly degrade the effectiveness of safety training and on-site communication. Trajkovski and Loosemore [3], in a study of 400 non-English-speaking background migrant construction workers, reported that 85.7% primarily used a language other than English on site, 48.7% misunderstood work or safety instructions due to limited English proficiency, and 66.7% attributed work-related mistakes to miscommunication. English-only safety induction programs were found to be largely ineffective, with more than two-thirds of participants struggling to understand training materials and assessments.

In response, recent studies have explored the use of digital tools to support Spanish–English safety communication on construction sites. Cappella et al. [4] evaluated construction-specific applications such as Need2Say, JHA Safety App, and Toolbox Talks Safety App, noting their advantages over general-purpose translators using industry terminology and structured safety content. However, existing solutions remain fragmented and insufficient for real-world jobsite conditions. Most tools rely on scripted phrases, text-based interfaces, or cloud-dependent translation services, limiting their usability for low-literacy workers, offline environments, and fast-paced, safety-critical interactions. Moreover, mistranslation of specialized safety terminology, such as lockout/tagout, fall arrest, or heat exhaustion, can introduce immediate and severe safety risks. Collectively, the literature reveals a critical gap: the absence of domain-specific, offline-capable, real-time bilingual safety communication systems designed for construction operations.

This gap is particularly consequential for heat-stress prevention, which has emerged as a growing occupational safety concern. Since 2011, at least 436 work-related deaths in the United States have been attributed to environmental heat exposure [5]. Construction workers face elevated risk due to prolonged outdoor labor, high metabolic workloads, and the use of personal protective equipment (PPE). While authoritative resources, such as OSHA guidance and Susan Harwood Training Grant materials [6], provide evidence-based recommendations, these materials remain largely static and text-based, limiting their effectiveness during real-time field operations where timely, clear, and linguistically accessible communication is essential.

Recent advances in edge computing and lightweight large language models (LLMs) offer new opportunities to address these limitations. Edge computing chips, such as the NVIDIA Jetson Nano, can now support various computing tasks entirely offline, such as automatic speech recognition, bilingual language understanding, and text-to-speech synthesis. However, existing construction safety research has predominantly focused on computer vision applications, such as PPE detection and hazard recognition, rather than multilingual, worker-centered safety communication. To date, no system has demonstrated a wearable, real-time, fully offline bilingual safety assistant suitable for frontline construction workers.

To address this unmet need, this paper presents the first wearable, edge-AI–based assistant for multilingual safety communication on construction sites, with a specific focus on heat-stress prevention. The proposed system integrates offline automatic speech recognition and lightweight LLM-based bilingual language understanding on a low-power Jetson Nano platform equipped with a microphone, speaker, and display. To ensure fidelity to safety-critical terminology, an 80-sentence bilingual heat-stress communication dataset was developed, directly derived from OSHA-aligned safety guidance. The dataset consists of aligned English–Spanish instructions validated by a native Spanish speaker and serves as a domain-specific reference for evaluating translation accuracy and consistency.

This paper describes the system architecture, dataset development process, and a two-phase evaluation in which two Spanish-speaking undergraduate participants conducted real-time interaction tests and recorded all 80 Spanish instructions for offline benchmarking. System performance is evaluated using Bilingual Evaluation Understudy (BLEU) scores and terminology fidelity relative to OSHA-aligned references. This work represents the first hands-free, edge-only bilingual safety communication system for heat-stress prevention, explicitly designed for the operational realities of construction work. This study is framed as a proof-of-concept feasibility demonstration, focusing on validating system functionality rather than conducting a large-scale human-subject evaluation.

2 Methodology

2.1 System Overview

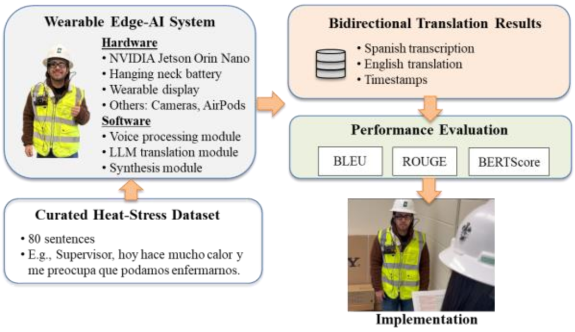

Figure 1. System Overview

Figure 1 illustrates the end-to-end workflow of the proposed wearable edge-AI system. The system integrates on-device computing, power, audio, and display components to enable fully offline speech translation of heat-stress safety concerns from Spanish to English. Translated outputs are logged locally and analyzed offline using standard machine translation evaluation metrics, supporting reproducible performance assessment without reliance on cloud connectivity or automated notification mechanisms.

2.2 Dataset Construction

An 80-sentence bilingual dataset was developed to support real-time Spanish-to-English safety communication for heat-stress prevention on construction sites. Unlike traditional bilingual datasets, this dataset targets spoken, worker-initiated communication, transforming OSHA-endorsed safety content into realistic, first-person Spanish expressions suitable for wearable speech-based systems.

Heat-stress messages were extracted from three OSHA-endorsed sources: Susan Harwood Training Grant heat-stress slide decks (English and Spanish, SH29650-SH6 [6]), Work Safely in Heat factsheets (English 3422-07R [7]; Spanish 3423-07R [8]), and Prevent Heat Illness at Work posters (English 3431-05R [9]; Spanish 3432-05R [10]). All relevant text, including directives, warnings, and infographic labels, was segmented into concise, standalone statements, forming an initial set of standardized Spanish sentences. Then each Spanish sentence was rewritten into a conversational, first-person Spanish worker expression (e.g., "Supervisor, me estoy mareando…"), preserving the original safety intent while reflecting realistic jobsite communication. These rewritten sentences serve as spoken inputs to the wearable edge-AI assistant and were optimized for Automatic Speech Recognition (ASR) robustness and practical deployment. For evaluation purposes, each rewritten Spanish expression was paired with a reference conversational English translation, representing the expected output displayed to supervisors. These reference translations are not used during system operation; instead, they provide ground-truth targets for evaluating translation accuracy, terminology fidelity, and semantic consistency.

The final dataset consists of three aligned layers: (1) original Spanish text, (2) rewritten conversational Spanish worker speech (system input), and (3) reference English translation (evaluation benchmark). All 80 sentences were validated for fidelity, clarity, and ASR compatibility and assigned unique identifiers (1–80). This structured dataset provides a standardized foundation for evaluating offline speech recognition and bilingual safety communication in heat-stress applications.

2.3 Wearable Edge-AI Hardware Configuration

Prior work has demonstrated the feasibility of deploying edge-computing frameworks for real-time construction applications under resource-constrained jobsite conditions [11]. Table 1 summarizes the hardware components used to develop the wearable edge-AI platform. The system integrates a Jetson-based processing unit, a wearable display (Vufine VUF-110) for visual output, an audio interface (Apple AirPods), a camera module, onboard storage, and a portable power supply (ZyberVR Neck Battery Pack). This configuration enables fully offline speech recognition and translation while maintaining mobility and usability in construction environments. The total system cost of $613.53 demonstrates the feasibility of a low-cost and scalable edge-AI solution for on-site applications.

Table 1 Hardware used in this study

Product	Function	Price (USD)
NVIDIA Jetson Orin Nano Super Developer Kit	Single-board computer	249.00
Vufine VUF-110 Wearable Display	Visual interface	159.00
ZyberVR Neck Battery Pack	Wearable power supply	32.99
IMX219-160 8-MP Camera Module	Camera	23.59
Apple AirPods (2nd Gen)	Audio module	129.00
SanDisk 128GB Extreme microSDXC	Storage	19.95
	Total	613.53

Figure 2. Wearable system configuration and physical implementation

Figure 2 illustrates the configuration of the proposed wearable edge-AI system and its appearance when deployed by construction personnel. The visualization demonstrates a one-way communication workflow in which Spanish-speaking workers communicate verbally, while construction managers or site supervisors who do not speak Spanish receive real-time English translations on a wearable display. In this setup, the worker requires only an audio interface (e.g., AirPods) for voice capture, whereas construction managers or site supervisors wear the NVIDIA Jetson–based computing module, a neck-mounted power supply, and a heads-up display to receive and view translated information. The current one-directional pipeline supports translation from any input language into English; however, the architecture can be readily extended to bidirectional communication by duplicating the hardware–software stack and specifying the appropriate output language (e.g., Spanish) on the second device.

2.4 Voice Processing Module

The voice processing module encompasses audio detection and speech transcription. Audio input is processed using the Whisper model [12], which performs multilingual speech recognition in near real-time. The model incorporates built-in voice activity detection (VAD) to distinguish meaningful speech from silence or background noise and can automatically identify the spoken language without requiring manual specification. The resulting transcription is generated in the detected source language for subsequent translation.

Whisper provides multiple model sizes, each with different computational requirements. When deployed on edge hardware such as the Jetson Orin Nano, only lightweight models (e.g., tiny and base) were found to operate reliably under memory and resource constraints, while also supporting low-latency inference.

2.5 Large Language Model Inference Module

Transcriptions produced in the previous step may contain errors or incomplete segments. However, modern LLMs can still interpret and translate imperfect text by leveraging contextual cues rather than relying solely on exact character-level accuracy. Prior work has shown that LLMs can effectively interpret unstructured and imperfect construction discourse to support real-time issue identification and mitigation, even when inputs are noisy or incomplete [13]. In this module, a local lightweight LLM continuously monitors the rough transcription file and processes each newly appended line for deterministic translation (e.g., multilingual-to-English). The module "tails'' the transcription file so that every new entry is immediately captured and wrapped in a predefined translation prompt. To ensure reproducible outputs, the model is configured with temperature = 0 and a fixed random seed. The lightweight LLM used in this study is only 815 MB (Gemma3: 1b), making it highly cost-effective. The following prompt is used to guide the translation behavior:

"Your task is to translate the detected text into English. Do not do anything else; output ONLY the transcribed English; keep it simple. Below is the detected text:"

A software module was developed to integrate the voice-processing components with the lightweight LLM inference system. This module provides a simple voice-controlled interface for managing the two-stage processing pipeline. Users can say "start" to initiate the multilingual translation workflow and "stop," "quit," or "exit" to end it. When "start" is detected, the system activates the voice-processing module and begins real-time, bidirectional multilingual communication, with translated text displayed on a small screen attached to the eyeglasses. When a "stop" command is spoken, the controller safely terminates the running tasks. This hands-free setup allows users to operate the entire multimodal pipeline through voice commands, while the controller handles process monitoring, error management, and cleanup in the background.

2.6 Experimental Design

The experiment was designed as a controlled feasibility demonstration to validate end-to-end system functionality, rather than to perform statistical evaluation across a large participant pool. To demonstrate the functionality of the bilingual heat-safety communication system, a controlled demonstration involving two volunteer participants was conducted. The experiment simulated a realistic on-site interaction between a Spanish-speaking worker and an English-speaking supervisor. One participant acted as the Spanish-speaking worker, reading aloud ten selected sentences from the rewritten dataset (Table 2). The other participant served as the supervisor, wearing the edge-AI wearable assistant equipped with an integrated microphone, speaker, and augmented-reality glasses. To approximate real field conditions and enhance situational realism, both participants wore PPE, including high-visibility safety vests and hard hats (See Figure 2). However, this setup does not fully integrate with all PPE components or dynamic field environments.

Table 2 Spanish heat-stress communication sentences used in the demonstration

ID	Rewritten Spanish
1	Supervisor, hoy hace mucho calor y me preocupa que podamos enfermarnos.
7	Supervisor, estoy pendiente de mis compañeros por si alguien se siente mal.
10	Supervisor, ¿podemos repasar qué hacer si alguien se desmaya por el calor?
13	Me duele la cabeza; creo que puede ser por el calor.
16	Supervisor, mi cuerpo está muy caliente; algo no está bien.
19	Supervisor, mi compañero está actuando extraño; creo que es una emergencia.
20	Supervisor, voy a llamar al 911; alguien tiene señales de golpe de calor.
33	Supervisor, alguien tiene dificultad para hablar; necesitamos actuar ya.
46	Mi corazón está latiendo más rápido de lo normal.
75	Supervisor, como soy nuevo, solo debería hacer alrededor del veinte por ciento del trabajo hoy.

During the experiment, the worker read each Spanish sentence naturally, and the system processed the speech locally to perform real-time Spanish-to-English translation. The translated English text was then displayed instantly on the supervisor's glasses screen. The ten sentences used in the demonstration were selected to represent a diverse range of heat-stress communication scenarios, including personal symptoms, observations of coworkers, and emergency situations (Table 2).

This setup demonstrated the system's ability to perform fully offline, real-time, bilingual communication, in which the Spanish-speaking worker's voice input was accurately recognized, translated, and displayed as English text on the supervisor's heads-up display within seconds. The test confirmed both the system's translation reliability and its potential for field deployment in multilingual construction environments.

2.7 Evaluation Metrics

Three complementary evaluation metrics were employed to assess translation accuracy. First, the BLEU score [14] assesses the quality of machine-generated translations by measuring their similarity to human-produced reference texts via n-gram overlap. BLEU computes precision over 1-gram to 4-gram matches between the candidate and reference translations and applies a brevity penalty to discourage overly short outputs, with higher scores indicating closer alignment with professional human translations. Second, the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) [15] score measures lexical overlap between generated and reference texts by quantifying shared words and n-gram sequences, reflecting how well the core content of the reference text is preserved. Unlike BLEU, ROUGE emphasizes recall (i.e., how much of the reference content is captured) while also considering precision, making it widely used for evaluating summarization and machine translation quality. Both BLEU and ROUGE primarily assess word-level accuracy. Finally, BERTScore [16] evaluates semantic similarity between the generated translation and the reference text by leveraging contextualized embeddings from a pretrained BERT model. Instead of relying on exact word matches, BERTScore compares token-level semantic representations, enabling a more robust assessment of whether the generated translation preserves the underlying meaning of the source text.

3 Results

3.1 Transcription and Translation Accuracy

To enable a controlled and reproducible evaluation, the 80 sentences were synthesized into Spanish audio clips using text-to-speech generation with varied speaker characteristics and speech rates. This approach allowed systematic assessment of transcription and translation performance across consistent inputs, while isolating system-level behavior from environmental variability. The generated audio files were then processed through the proposed framework to produce Spanish transcriptions and corresponding English translations. Both outputs were quantitatively compared against researcher-validated ground-truth transcripts and translations. During transcription, three audio files contained no detectable speech and were excluded from subsequent analysis.

Evaluation metrics were computed using BLEU, ROUGE, and BERTScore. In this study, BLEU is reported at the corpus level, assessing overall translation quality by measuring n-gram overlap between the complete set of generated translations and the reference corpus. In contrast, ROUGE scores and BERTScore are reported as averages across individual test samples. Specifically, ROUGE-1 (R-1) measures unigram (word-level) overlap, ROUGE-

2 (R-2) measures bigram (two-word sequence) overlap, and ROUGE-L (R-L) measures the longest common subsequence, capturing sentence-level structural similarity. Table 3 summarizes the resulting transcription and translation accuracy metrics.

Table 3 Transcription and translation accuracy

Language	BLEU	R-1	R-2	R-L	BERT
Spanish	0.669	0.837	0.796	0.836	0.928
English	0.287	0.577	0.372	0.559	0.938

The results indicate that the proposed framework performs reliably across both speech transcription and cross-lingual translation tasks. For Spanish transcription, the high BLEU and ROUGE scores demonstrate strong word-level accuracy, phrase continuity, and sentence-level structural alignment with ground-truth references, suggesting reliable capture of spoken safety instructions despite variations in speech rate and speaker voice. The consistently high BERTScore further confirms that both lexical content and semantic intent are well preserved during transcription.

For English translation, BLEU and ROUGE scores are comparatively lower, reflecting expected linguistic divergence introduced by translation, including paraphrasing, word reordering, and syntactic differences between Spanish and English. However, the high BERTScore indicates that the translated outputs maintain strong semantic similarity to the reference texts. This contrast highlights the limitations of n-gram-based metrics for fully capturing translation quality in safety-critical communication, where meaning preservation is more important than exact word matching.

Collectively, these findings suggest that the proposed framework can support multilingual heat-stress safety communication by delivering accurate transcription and semantically consistent translation under controlled conditions.

3.2 Real-Time Implementation

Figure 3 illustrates the real-time testing scenario, in which translated output is displayed directly on the wearable device. In this demonstration, Spanish-speaking workers verbalize heat-stress–related concerns (e.g., "OK, supervisor, mi cuerpo está muy caliente"). The system automatically detects the speech, transcribes it, and generates an English translation. The wearable display displays the translated message (e.g., "My body is very hot"), enabling supervisors to immediately understand safety-critical information despite language barriers. All conversational records are stored locally for subsequent analysis. This implementation demonstrates the feasibility of a wearable, edge-based AI assistant for near real-time multilingual safety communication on construction sites.

Figure 3. Real-time testing with results displayed on the screen.

One participant noted: *"It translated most of it. A couple of words were mistranslated. I think for this to work well, you have to speak slowly and pause between sentences so it can catch up. But it does translate most scenarios."* This feedback is reasonable, given that the system requires processing time to execute Whisper's voice activity detection and to invoke a lightweight LLM for semantic inference, even with limited computing resources. Nonetheless, the framework is highly scalable and can support more streamlined, real-time performance as more capable hardware and more powerful models are integrated. In addition, the end-to-end latency was observed to be on the order of a few seconds during testing, supporting near real-time interaction.

4 Discussions

This study demonstrates the feasibility of a fully offline, wearable edge-AI system for multilingual safety communication on construction sites, with a focus on heat-stress prevention. By integrating on-device speech recognition and deterministic lightweight LLM-based translation, the proposed framework enables near real-time communication without reliance on cloud connectivity. This capability directly addresses a well-documented safety challenge in construction, where heat stress remains a leading cause of work-related illness and fatalities, particularly among outdoor workers engaged in physically demanding tasks [5]. The results from the controlled evaluation suggest that the system has the potential to support the communication of safety-critical information across language barriers. This capability may assist supervisors in interpreting worker inputs in dynamic job-site environments, although further validation under real-world conditions is needed.

The real-time demonstration highlights the practical value of this approach for construction safety operations. Translated messages were displayed on a wearable heads-up display within seconds, allowing supervisors to immediately understand workers' heat-stress concerns without interrupting ongoing tasks. This real-time, hands-free interaction addresses limitations identified in existing digital safety communication tools, which often rely on scripted content or text-based interfaces that are poorly suited to fast-paced field conditions [4].

The proposed system is designed as an augmentation tool to support multilingual safety communication, rather than a substitute for existing safety practices. In construction environments, effective safety communication relies on multiple complementary strategies, such as formal training and standardized procedures. Within this context, the wearable edge-AI assistant aims to enhance communication in situations where language barriers may hinder understanding, particularly during time-sensitive interactions, while preserving the central role of human judgment and supervisory decision-making.

Non-verbal communication also plays a critical role in safety-critical construction environments, where gestures, visual cues, and situational awareness complement verbal interaction. The proposed system is therefore intended to support verbal communication within this broader communication framework rather than replace established practices. However, presenting translated text on a wearable display may introduce additional attentional demands, potentially affecting how users allocate visual attention during tasks. Its impact on attention, distraction, and overall communication effectiveness requires further investigation. Future work will examine how wearable translation systems can be integrated with natural communication behaviors to minimize disruption and support both verbal and non-verbal interaction in dynamic construction environments.

As stated previously, this study is intended as a proof-of-concept demonstration of system feasibility rather than a comprehensive human-subject evaluation. The real-time experiment involved a limited number of participants in a controlled setting, which constrains the generalizability of user-level findings. In addition, the offline evaluation relied on synthetic speech data generated through text-to-speech methods. While this approach enabled controlled and reproducible testing, it does not fully capture the acoustic complexity of real construction environments, such as background noise, overlapping speech, and irregular speaking patterns.

Another limitation is the focus on one-directional Spanish-to-English communication. Although the system architecture is readily extensible to bidirectional interaction and additional safety domains, these capabilities were not evaluated in this work. Furthermore, the practical compatibility of the proposed system with real construction environments remains to be fully assessed. While the feasibility demonstration incorporated basic PPE elements (e.g., hard hats and safety vests), integration with standard protective equipment (e.g., safety glasses and hearing protection) and its effects on cognitive load and situational awareness require further investigation.

This work focuses on translation accuracy and system feasibility rather than comprehensive system benchmarking. As such, direct comparison with cloud-based translation tools is outside the scope of this study, given the system's fully offline edge-based design. Additionally, latency performance was not systematically benchmarked, although the observed response time was on the order of a few seconds during testing. Future work will therefore focus on large-scale field deployment, real-world noise testing, bidirectional communication, and rigorous evaluation of system performance, usability, and human factors in active construction environments.

5 Conclusions

This study presented a wearable, fully offline edge-AI assistant designed to support multilingual safety communication on construction sites, with a focus on heat-stress prevention. By integrating on-device speech recognition, deterministic Spanish-to-English LLM translation, and wearable display and logging, the proposed system enables real-time communication without reliance on cloud connectivity. Experimental results demonstrate initial performance of speech transcription and semantically faithful translation, as reflected in strong BLEU, ROUGE, and BERTScore scores, while real-time testing confirms the feasibility of deploying the system in practical job-site scenarios. The proposed framework addresses a critical gap in construction safety communication by providing a hands-free, human-in-the-loop solution that augments supervisors' situational awareness without automating decision-making or enforcement. As edge computing and lightweight language models continue to mature, wearable edge-AI systems such as the one presented in this study offer a promising pathway toward more inclusive, responsive, and resilient safety practices in the construction industry.

References

[1] U.S. Bureau of Labor Statistics. National Census of Fatal Occupational Injuries in 2023 (USDL-24-2564). U.S. Department of Labor, 2024.

[2] National Hispanic Construction Alliance. 2025 State of Hispanics in Construction Report. National Hispanic Construction Alliance, 2025.

[3] S. Trajkovski and M. Loosemore. Safety implications of low-English proficiency among migrant construction site operatives. International Journal of Project Management, 24(5):446–452, 2006.

[4] A. Cappella, S. Clarke, and J. Lucas. Digital solutions for bridging language barriers: A review of apps for construction Spanish safety communication on construction jobsites. In Proceedings of the 61st Annual International Conference of the Associated Schools of Construction (ASC 2025), 2025.

[5] U.S. Bureau of Labor Statistics. 36 work-related deaths due to environmental heat exposure in 2021. U.S. Department of Labor, 2023. Online: https://www.bls.gov/opub/ted/2023/36-work-related-deaths-due-to-environmental-heat-exposure-in-2021.htm, Accessed: 2026.

[6] Occupational Safety and Health Administration. Susan Harwood Training Grant Heat-Stress Materials (SH-29650-SH6). Online: https://www.osha.gov/harwoodgrants/grantmaterials/fy2016/sh-29650-sh6, Accessed: 2026.

[7] Occupational Safety and Health Administration. Work Safely in Heat. Online: https://www.osha.gov/sites/default/files/publications/3422_factsheet_en.pdf, Accessed: 2026.

[8] Occupational Safety and Health Administration. Trabaje de forma segura en el calor. Online: https://www.osha.gov/sites/default/files/publications/3423_factsheet_sp.pdf, Accessed: 2026.

[9] Occupational Safety and Health Administration. Prevent Heat Illness at Work Poster. Online: https://www.osha.gov/sites/default/files/publications/3431_WKSITEPOSTER_EN.pdf, Accessed: 2026.

[10] Occupational Safety and Health Administration. Prevenga las enfermedades por calor en el trabajo. Online: https://www.osha.gov/sites/default/files/publications/3432_WKSITEPOSTER_SP.pdf, Accessed: 2026.

[11] G. Chen, A. Alsharef, M. Pedrazza, E. Albert, and E. Jaselskis. Construction jobsite image classification using an edge computing framework. Sensors, 24(20):6803, 2024.

[12] OpenAI. Introducing Whisper. Online: https://openai.com/index/whisper/, Accessed: 2026.

[13] G. Chen, A. Alsharef, A. Fadi, E. Albert, and E. Jaselskis. Meet2Mitigate: An LLM-powered framework for real-time issue identification and mitigation from construction meeting discourse. Advanced Engineering Informatics, 64:103068, 2025.

[14] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu. Bleu: A method for automatic evaluation of machine translation. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL '02), pages 311–318, Philadelphia, Pennsylvania, USA, 2002.

[15] C.-Y. Lin. ROUGE: A package for automatic evaluation of summaries. In Text Summarization Branches Out, pages 74–81, Barcelona, Spain, 2004.

[16] T. Zhang, V. Kishore, F. Wu, et al. BERTScore: Evaluating text generation with BERT. In Proceedings of the International Conference on Learning Representations (ICLR 2020), pages 1–43, 2020.