Image generated by ChatGPT

AI Translation on the West Highland Peninsulas

Kendra Turnbull

Centre for Recreation and Tourism Research

February 2025

UHI | NORTH, WEST AND HEBRIDES
A TUATH, AN IAR IS INNSE GALL

Contents

Introduction

The UK is experiencing a growing number of foreign language speaking visitors with a 23% increase in international visitors from 2022-2023. On the Ardnamurchan peninsula there is a lack of provision to welcome foreign language tourists and assist them with their overall visitor experience. The importance of foreign language provision for Scottish tourism is highlighted by [Historic Environment Scotland](#) and featured in an article in the [Scotsman](#). Ardnamurchan is one of the least densely populated areas in Scotland and therefore has a lack of capacity to engage with time consuming translation. The ever-emerging Artificial Intelligence (AI) translation technology provides an opportunity to increase capacity and give the rural community a USP for its tourism industry. An online presence for tourism businesses has always been important, but the increasing emergence of AI itinerary generators, which 'learn' from online information, has meant that it is becoming even more important to be 'seen' online to attract visitors.

This report investigates the different types of AI translation technology and considers which tourism services across the Ardnamurchan peninsula could benefit from their application. The AI options outlined have been included in consideration of the fact that the potential adopters are rural tourism businesses and therefore focus on everyday usability and, where possible, free options.

The translation technologies investigated have been done so under the themes of:

- Images of text turned into text documents
- Audio recordings, or speech, transcribed to written text
- English written text translated to an alternate language text
- Foreign language written text turned into speech
- Real time translation or transcriptions

For each theme we consider the use of the technologies from the point of view of the tourism providers and, where appropriate, the visitors. Following the themed overview we then present how the technologies may be applied by the local Destination Management Organisation (DMO), West Highland Peninsulas, and the local businesses.

AI translation technologies

Images-to-text

Tourism context

Enables editable text to be created from a photograph or image

- **Transcribes interpretation, display boards, signs, menus or other text items which have no digital file.**
- **Text documents can be translated into different languages.**
- **Automatically translate text into a different language for immediate understanding.**

Technology has the ability to recognise characters within an image and then transcribe that into text. In its primary form this involves utilising Optical Character Recognition (OCR) to identify and extract text from images, but newer AI systems combine this with advanced computer vision techniques to increase accuracy and identify text from more complex images. Additionally Natural Language Processing, a field of AI, is also combined with some translation models. This enables computers to understand, interpret, and generate human language in a more nuanced manner than simple word for word translation.

Provider perspective-

ChatGPT: On a computer or smartphone attach a jpeg image and ask it to extract the text from the image.

Claude: On a computer or smartphone attach a jpeg image and ask it to extract the text from the image. Additional function of being able to take a photograph directly when using a smartphone, or a screenshot on a computer.

Google Translate: On a computer drag and drop an image and select copy text.

When this was tested, all three accurately transcribed the text from a photo (right) of a leaflet. Claude also accurately transcribed text from a photo of a handwritten note (left) but ChatGPT failed and Google translate contained errors. The
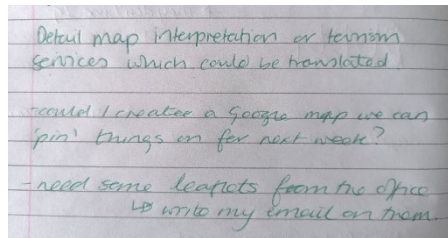
text created would then need to be copy and pasted into a separate document to preserve and/or edit it.
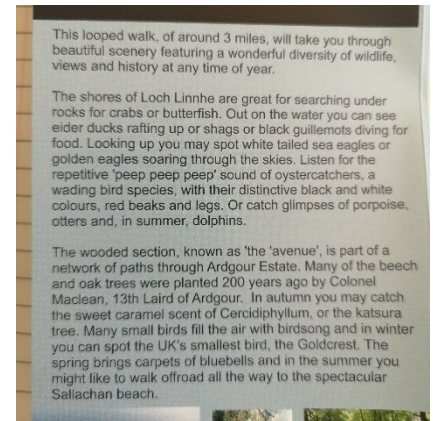
DeepL: In the App (which can be used on a computer or a smartphone) you can capture text on screen to create a transcription, on the desktop version you can only upload a pdf (Adobe), docx (Word), or ppt (PowerPoint) file. There are limited uploads for the free version.

ChatGPT and Claude require account registration for use of their free services. Google Translate and DeepL do not require any details.

Adobe Acrobat: An image can be opened in Adobe Acrobat Pro using the 'Create PDF' tool. The text is extracted from the image, using OCR, when the 'Edit PDF' tool is selected. If the document has a high-quality resolution and the original text is clear this option works well for editing text. This would be the case if original digital copies of an interpretation board exist or a high-quality scan of a leaflet is used. However, when tested using a photograph the text was as individual words and hard to edit. Adobe Acrobat Pro requires a subscription service.

Tourist perspective-

Google Lens: The Google Lens app can instantly translate text from an image. You can hover to see an instant change or you can take a snapshot photo (see image below). The image then enables the option of listening to the text (either as a whole or selecting the area to listen to). The voice has an electronic tone and is stilted.
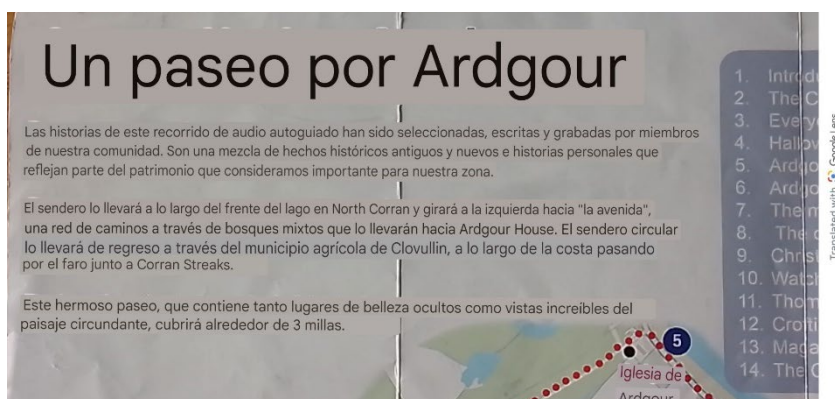
iTranslate: This app advertises itself as an OCR app which can instantly translate text from images using a smartphone. The basic version of the app are free or a more advanced version with more offline capabilities can be purchased as a monthly subscription. However current reviews for the app are poor so it was not tested.

Speech-to-text

<table>
<tr><td align="right">**Tourism context**</td></tr>
<tr><td>**Enables editable text to be created from an audio file or from a live recording**

- **Transcribes interpretation, display boards, signs, menus or other text items which have no digital file where the text is unclear for image transcribing.**
- **Transcribes audio recordings, tours, social history into text documents.**
- **Text documents can then be translated into other languages.**
- **Automatically translates speech into text for communication and understanding.**</td></tr>
</table>

Software transcription technology has utilised the AI subset of machine learning (computers analysing and learning from massive amounts of data). This has meant the accuracy of computer-generated speech to text transcriptions has risen significantly. All the transcription software requires careful checking as it is not 100% accurate and generally does not understand place names particularly well. To test the technology we use an audio recording of a story read by a female with a Scottish accent. The story was read at a gentle pace with a clear tone. It included Gaelic place names and heavily accented English words.

Provider perspective-

Notta.ai: Upload an audio file to get a transcription. The free version allows 50 files to be transcribed, or 120 minutes per month, but the text cannot be downloaded so would have to be copied and pasted into a separate document. The entry level paid version does allow for nuanced vocabulary to be included, files to be download and more files (100) and minutes (1800) for £7.20 ($9) per month. Of the options tested, Notta.ai most accurately transcribed an audio story of a west coast Scottish accent.

Otter.ai: Simply upload a voice file to get a transcription. The free version only allows 3 files per user and up to a maximum of 30 minutes per file. However, it will transcribe 300 minutes of conversation per month so an audio could be read 'live' to be transcribed.

Rev: Similar set up to Otter.ai

Microsoft Clipchamp: It is possible to upload an audio file to a free video editor like Clipchamp and then ask it to create captions. These captions can be copied and pasted into a word document. This is a clunky method of transcription but could be useful for short clips or if trying to avoid subscription payments for other software.

Adobe Premiere Pro: This software will transcribe the audio and enables a text file export. The file will require some editing as it is time stamps the text.

Visitor perspective-

See Real time transcription and translation section below.

Text-to-text

Tourism context
Enables text to be translated into editable foreign language text documents • **Foreign language text documents can be used for new online or offline interpretation, display boards, signs, or menus.** • **Enables the creation of location-specific marketing material** • **Enables translation of websites.** • **Provides a document from which foreign language audio files can be created.** • **Enables communication of important information such as waivers, safety information or contracts e.g, for tours or experiences.**

Machine Translation (MT) is a fast-evolving landscape as multilingual Large Language Models (LLMs), such as ChatGPT, have emerged and rapidly developed over the last few years. An independent report 'The State of Machine Translation 2024', by Intento, evaluated an additional 19 LLMs from 2023 to 2024 and found that of all the top-performing models LLMs accounted for 55% (up from 17% in 2023).

According to the Intento report, DeepL, GPT-4o, and Google are the three models that demonstrated the most translations with no or minor issues, but that GPT-4o and DeepL generally outperformed other models. Whilst the models performed well there is still a slight error rate. The below image shows DeepL's superior performance when translating from English to Italian, with a 90% no error or minor error rate and a 10% major error rate. Minor errors are classed as having only a slight impact on meaning.
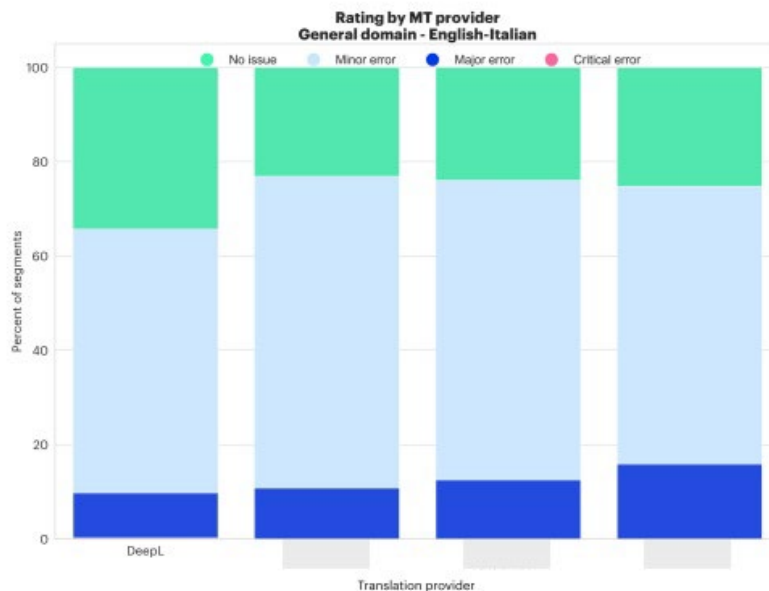
DeepL and Google Translate have frequently been compared against each other. DeepL is regarded as having more accurate and nuanced translations which can consider context and idioms, but has a more limited, European dominated, language selection. Google Translate has more languages and offers more within its free service, however it doesn't cope with nuance as well and is better for single word or casual translation. Google's more advanced model, Google Cloud Translation, is aimed at developers or business. This model has greater capabilities of customisation and integration.

Provider perspective-

DeepL and **Google Cloud** both have free monthly allowances. Google Cloud charges by usage and DeepL has a starting monthly subscription of £6-7. The Intento report placed Google Cloud as slightly cheaper. LLMs like **ChatGPT** have free to use services, the GPT-4o model is available on the free platform but with limited availability. The £16 monthly plan also has limited access, while full access costs £162 per month.

MTs and LLMs both improve if they are 'trained' using translation memories or through customisation through tools such as glossaries or vocabulary.

Visitor perspective-

Potential scenarios where a visitor might need to translate a large document from English into their own language could include when then are signing a legal agreement or a waiver for an experience. Google Translate, DeepL and LLMs like ChatGPT all have the ability to upload files for translation. For quick translation of a word or phrase Google Translate is the simplest free option.

Text-to-speech

Tourism context
Using AI generated voices to create audio files from text   • **Turning the text files of the transcribed interpretation/ information into audio files.**   • **Audio files can be listened to through the creation of an audio tour or linked via QR codes.**   • **Creates customer-facing services and enhanced visitor experiences.**   • **Provides opportunities for location-specific marketing.**

Text-to-Speech (TTS) is a type of AI technology that converts written text into spoken words. It uses natural language processing and speech creation techniques to generate human-like speech. Text can be converted into speech by analysing the text to recognise structures, punctuation, numbers and such like. It analyses the linguistics to determine pronunciation, grammar, intonation, and stress patterns and then turns this into natural sounding speech patterns. Technology has evolved to be able to clone a voice or there are a number of ready-made accents or languages available.

Provider perspective-

Elevenlabs creates realistic sounding AI audios from text. There are multiple different language voices to choose between and a variety of Scottish accents, if you have a paid for version, you can clone a voice of your choosing. The free version allows for around 10 minutes of audio per month, for £4/mo you get 30 minutes and a basic clone, for £18/mo you can create 2 hours of audio and access professional voice cloning.

Other text to speech software was investigated but not recommended:

- Blipcut- AI voices were too electronic

- Natural Readers- AI voices were good but to download an audio to use on a commercial basis required a paid account at $588/year or $99/mo

Visitor perspective-

See Real time transcription and translation section below.

Real-time transcription and translation

Tourism context
Immediate conversation and understanding • **Enables users to hear an immediate translation of what is being said.** • **Creates possibilities for foreign language tours/guides.** • **Enables in-location translation of interpretation or written text.** • **Creates opportunities for new customer-facing services and enhanced visitor experiences.**

Translating language in real time combines multiple AI technologies. Low latency technology enables large volumes of data to be processed with minimal delays. Speech is transcribed into text and then instantly translated into another language. Text-to-speech models then 'read' that text for the user.

Provider and visitor perspective-

Timekettle is a company which specialises in real time translation technology. They have a number of different products aimed at different scenarios.

1. WT2 Edge £350.00
   Using earbuds and a smartphone, speech is translated in real time through the earbuds and as text on a smartphone. There are different modes:
   - 'Simul mode' where speech is translated almost instantly between earbud wearers (up to 6 earbuds can be supported between two selected languages).
   - 'Touch mode' where people take it in turns to have control of the translation by touching the earbud.
   - 'Speaker mode' where the listener has both earbuds and a smartphone is used as a microphone for the speaker.
   - 'Group Chat' where a multilingual text conversation can happen.

Generally an internet connection is required but there are some languages which can be translated offline.

2. M3 £150.00
   The M3 is a less advanced version of the WT2 Edge with less modes. The user is required to tap the earbud to 'take control' of the conversation before they speak (touch mode). Listen mode provides one way translation to a user wearing both earbuds and speaker mode uses a smartphone to enable the conversation between two languages.

3. X1 £700.00
   The X1 Interpreter Hub has many similar features to the WT2 Edge but is a standalone translator and transcriber (no smartphone required) and can coordinate translation between 20 people in 5 languages. Each language user requires their own hub and earbuds. One hub can act as speaker/microphone if multiple people want to listen and communicate between two languages. Internet connection is required or there are some offline languages available.

4. T1 mini £150.00 / T1 £300
   This device is a simplified translator in that you talk into and it automatically repeats what you have said in the language of choice. It also has the option of image to text translation built into it. Voices need to speak in turn by pressing a button to activate.

The **Google Translate app** has a conversation mode where it listens to speech and translates it into text in another selected language. There is an option to play the text as a realistic sounding audio. Each time someone speaks a new text box is opened but the previous line of the conversation is left displayed. If the same language speaks twice the text is just added to the same section which may make it a little confusing. Voices need to speak in turn for the app to work. While it is not perfect it does cope with different accents if they speak at a reasonable pace.
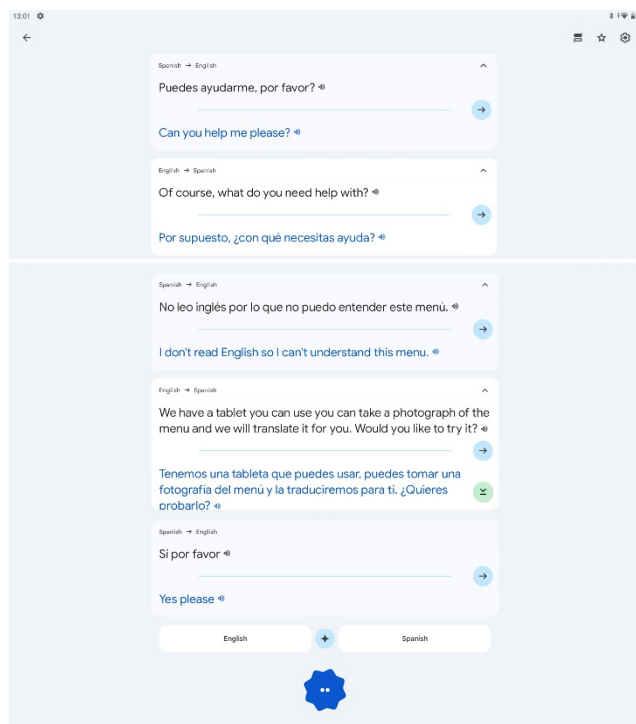
Enhancing Tourism Through AI Translation

Opportunities for the Destination Management Organisation

Multilingual digital experiences.

- Translate the existing peninsula **interpretation boards**
- Create new digital versions or translate the text into audio experiences.
- Embed the digital interpretation on Google Maps for ease of access and to experience in situ.
- Share the map to integrate the content with member websites.
- Use the audio content to develop geolocated audio tours which can be accessed on or offline.
- Use translated audios for location-specific marketing and promotional reels.

Advanced Website Translation

Implement API-driven translation for a fully optimised, **multilingual website** that go beyond basic Google Translate capabilities. This approach ensures:

- Higher accuracy in translations
- Improved Search Engine Optimisation (SEO) to attract foreign-language visitors
- Improved user experience with native language navigation.

Multilingual Visitor Resources

Create a series of **foreign language resources**. These could include:

- Targeted or themed blogs/travel guides (potentially using content from translated audios)
- Local phrases or food item translations
- Single track road driving guides
- Guides to using AI technology to improve the visitor experience

Translation Equipment Library

Establish a **shared resource centre** offering:

- AI translation devices like Timekettle products
- Tablets with pre-installed apps like Google Translate, Google Lens or GeoTourist

The resources could be promoted directly to visitors via the blogs/guides, available to hire for non-members and provided free of charge to members as incentive to increase membership.

Business implementation options

Real time translation services

Businesses could enhance guided tours and customer interactions with **real time translation devices**. Businesses could consider collaborative purchasing to share costs and resources.

Customer facing translation tools

Installing user friendly translation tools and solutions such as:

- Having a **tablet available** with pre-installed apps like Google Translate or Google Lens for translating displays or information boards on a large screen (eg. distilleries/ museums/ information centres/ restaurant menus)
- **QR code linked foreign language digital information packs** (to menus, facility information, DMO guides/maps)
- **Audio welcome** message or welcome pack
- **Sharing of the DMO created resources** or equipment with preloaded links, Google maps.

Website Translation for Global Reach

Use API driven translation to **make business websites multilingual**. Providing a user-friendly experience for global visitors which optimises search engine capabilities and sends a welcome message to foreign visitors.

Implementing the ideas

Partnership with UHI North, West and Hebrides

Taking advantage of knowledge exchange funding businesses can tap into opportunities to develop AI translation initiatives.

Covered by the funding are academic staff costs for advice, expertise and project development. Businesses contribute in-kind time and to direct implementation costs (e.g., QR code signage, equipment purchases). All costs would be agreed on prior to project application. Funding can be accessed as an individual business or combined for collaborative working.

Sample projects

1. Collaborative interpretation translation
   o Audio translation of all interpretation boards (including human accuracy checking)
   o Creation of a digital audio tour
   o Integration with existing tourism infrastructure
   o Promotion of translations
2. Business-specific AI solutions
   o Customised translation ideas and solutions
   o Translation and accuracy checking
   o Technical implementation support
   o Project management
3. AI implementation
   o Training and support for implementation of company purchased AI technology (e.g., Timekettle)
   o Seamless integration

Ready to enhance your international visitor experience?

Embrace AI translation technology and open your destination to a world of visitors. By removing language barriers, you enhance accessibility, improve visitor satisfaction, and drive global engagement.

Contact us to discuss how we can help implement these solutions for your organisation. Kendra.Turnbull@uhi.ac.uk

Let's work together to make our region more accessible to visitors from around the world.