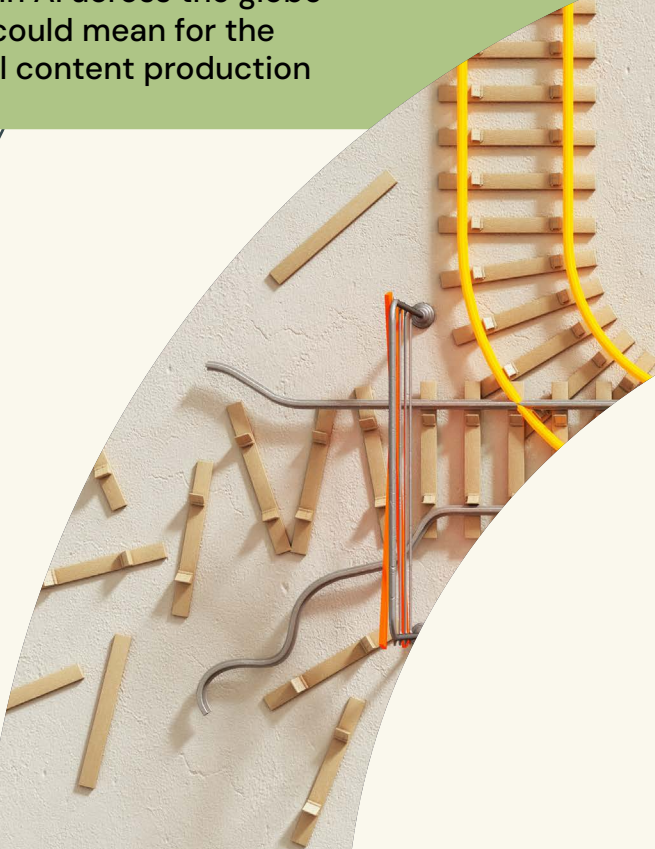


# CHATGPT AND LOCALISATION - FRIEND OR FOE?

A look at trust in AI across the globe  
– and what AI could mean for the  
future of global content production



# CONTENTS

<b>Foreword</b>	3
<b>The Research</b>	5
AI vs human: Performance test	6
AI content and consumer trust	9
The potential of AI-generated content	11
The limitations of AI-generated content	13
<b>A view from Croud's Head of Language Services</b>	18
Current opportunities for AI global content production	22
<b>Conclusion: An exciting future</b>	23



# FOREWORD

The rapidly developing world of AI opens up many opportunities for brands when it comes to content creation and localisation. With concerns around accuracy and trust, it is essential that the use of AI tools is based on substantive research and clear frameworks. This research into the effectiveness and trustworthiness of AI-generated content aims to shed light on the present opportunities and future potential of AI tools.

Our research into the use of AI for global content production involved two strands – an extensive consumer survey and an in-depth editorial test.

Our editorial test focused on the accuracy of AI translations. ChatGPT and human translators localised a complex piece of written content into three different languages: French, Spanish and Simplified Chinese. We used a comprehensive scoring

system on the translated copy to identify multiple types of error.

The second element of our research was a 2,500 person consumer survey. Involving people across the UK, US, France, Spain and China, the survey evaluated consumer trust in content produced by AI.

The findings of this research showed a high degree of receptiveness from consumers towards AI-generated content, with some concerns around accuracy. These concerns are matched by the editorial results, which highlighted the existing limitations of AI translation in comparison to human content.

Whilst there are clearly still limitations around AI's applications to content production and localisation, and human input remains vital, ChatGPT and other AI tools open up exciting possibilities for brands and agencies alike. Those companies that embrace these opportunities will be the ones that get ahead.



**Emil Bielski**

Managing Director, Croud UK



## Section 01

# THE RESEARCH

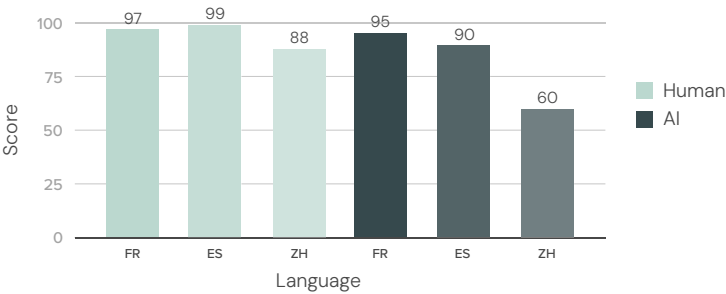
# AI vs human: Performance test

The editorial test involved a challenging source text that included emotive language and local references. The human translators and ChatGPT worked from the same source, using the same prompt to translate into three different languages. To evaluate the results, our scoring system incorporated:

- Accuracy
- Language
- Terminology
- Style
- Country/regional specifics
- Consistency

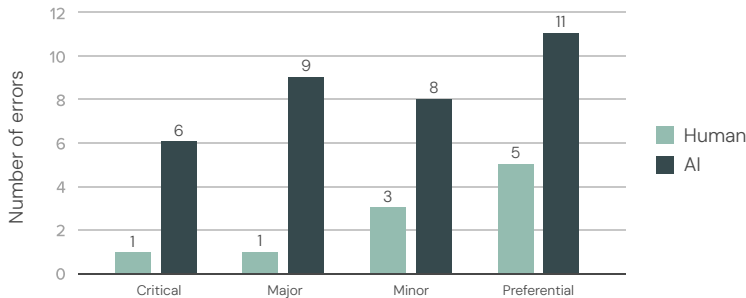
Overall, scores showed that human translators performed better than ChatGPT in all three languages. The difference was least apparent in French, with just a 2% difference between human and AI. Chinese showed the largest gap, with humans outperforming AI by 28%. This likely reflects the high level of complexity in Chinese translation, with the human score for Chinese 9% below the human score for French.

# Overall translation scores



In terms of errors, ChatGPT made double the number of human translation errors in Spanish and Chinese. Across all three languages, six critical errors were made by AI, compared to just one from the human translators.

# Types of error



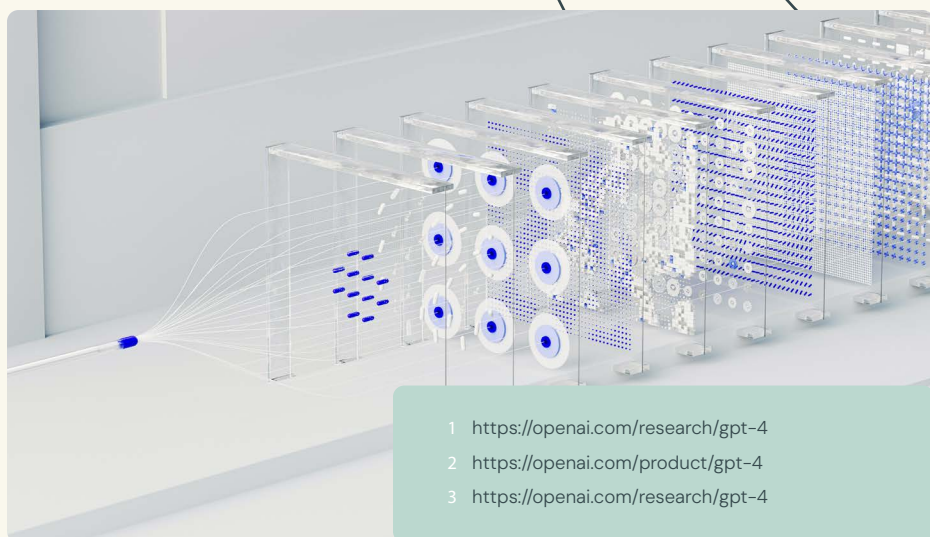
In Spanish translations, ChatGPT was able to produce readable, grammatically correct text. However, some of the nuances of local phrases were lost and the output felt 'plain' and lacking emotion.

In French, the AI translations were readable but the arrangement of the words stuck too closely to the original English source, leading to rigid text that didn't flow well.

ChatGPT had the most problems with Chinese translation, including mistranslations and broken phrases, along with a lack of local audience considerations.

Our editorial testing clearly shows that ChatGPT is not as accurate as human translators. However, it is essential to consider the rate of development of ChatGPT specifically and AI in general.

ChatGPT-3 was launched in November 2022. Just four months later, ChatGPT-4 emerged, outperforming version 3 by up to 19% on standard machine learning benchmarks<sup>1</sup>, improving accuracy of responses by 40%<sup>2</sup> and displaying specific improvements in multilingual capabilities<sup>3</sup>. With this rate of improvement, it is reasonable to expect the performance in localisation to close the gap on human translators in the near future.



1 <https://openai.com/research/gpt-4>

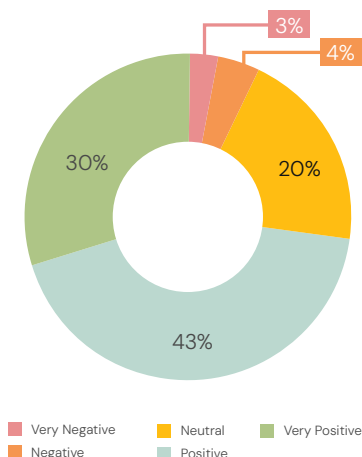
2 <https://openai.com/product/gpt-4>

3 <https://openai.com/research/gpt-4>

# AI content and consumer trust

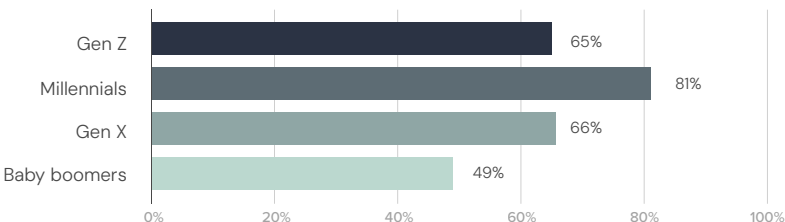
## How do you feel about AI tools generally?

Our consumer survey questioned 2,500 people, who all had some familiarity with AI. The respondents were split equally between the UK, US, France, Spain, and China. The survey showed that a large majority of people (73%) felt positively or very positively about AI tools in general.



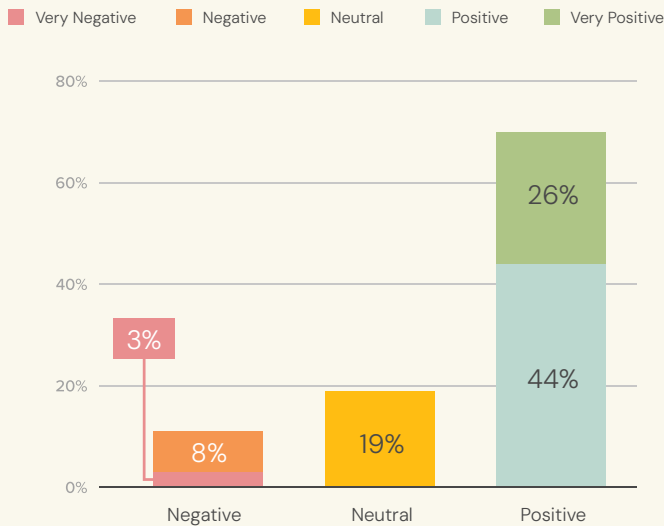
Interestingly, the youngest and oldest respondents of the consumer survey were the least positive in their feelings towards AI in general – with less than half of baby boomers saying they feel positive or very positive.

## Percentage of respondents who feel positive or very positive about AI in general, by age



This positivity extended to feelings about companies using AI to produce written materials for consumers.

**How do you feel about companies using AI tools to produce written materials or communications to customers like yourself?**



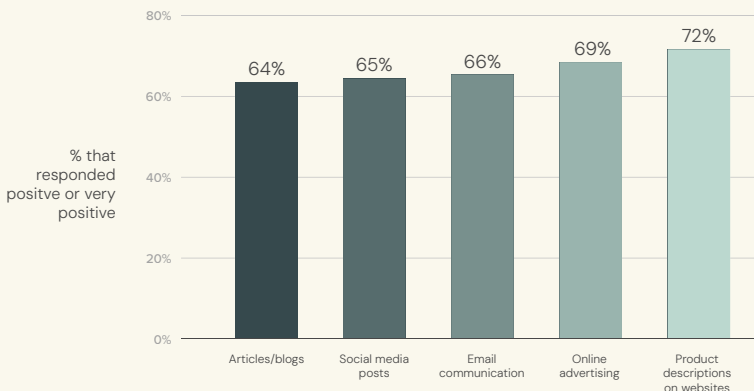
The positive feelings towards AI content were notably high in China (94%) and correlated strongly with people’s prior knowledge of AI. 81% of those who described themselves as knowing a lot about AI were positive or very positive; while those who knew less came in at 59%. Again, this suggests that, as time passes and awareness of AI tools increases, consumer trust in content produced by those tools may continue to rise.

# The potential of AI-generated content

Our research highlighted several key areas which demonstrate the promise of AI tools for content production and localisation.

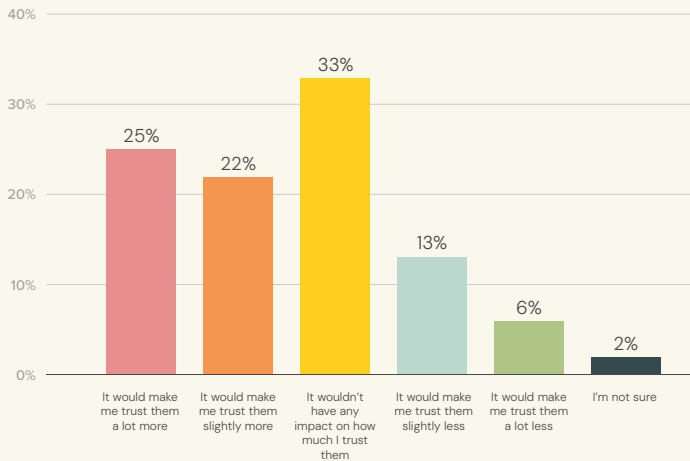
- Seven in ten people feel positive towards AI tools and content generated using AI. As described above, this correlates with the level of awareness of AI tools and therefore is likely to increase.
- Particular regions, such as China, are especially receptive towards AI-generated content, with 94% being likely to purchase from a brand that used AI-generated content.
- Respondents were particularly positive (72%) towards functional content, such as product descriptions, generated by AI. Whilst sentiment towards other types of content was also high overall, respondents were more hesitant about articles and blogs, or social media posts, being produced by AI – with only 55% of French and 56% of British respondents saying they feel positive or very positive about the possibility of AI-generated articles.

## How do you feel about companies using AI tools to write each of the following materials?



- By industry, the most well-received areas for AI-generated content were travel (67%) and retail (66%).
- 80% of people surveyed said that discovering a favourite brand had been using AI-generated content would not affect their trust in that brand, or would even positively impact (47%) their trust in that brand.

## How do you feel about companies using AI tools to write each of the following materials?



- AI-translation was also well trusted overall, with 55% of respondents trusting the information more if it had been translated by AI.
- When it came to our editorial testing, ChatGPT produced mostly grammatically correct text that was understandable to the reader. The AI-generated translations were faithful to the original text for the most part, following basic grammatical rules, with ChatGPT generating translations that are similar to other machine translation software.

# The limitations of AI-generated content

Despite the many promising findings related to AI content, our research highlighted a number of significant limitations.

- Although large numbers were receptive to AI content, a significant proportion of people in the UK (14%), the US (14%), and France (15%) describe themselves as unlikely or very unlikely to purchase from brands using AI-generated content or communication.
- Transparency is very important to consumers, with 80% of respondents stating that brands should disclose their use of AI.



14%



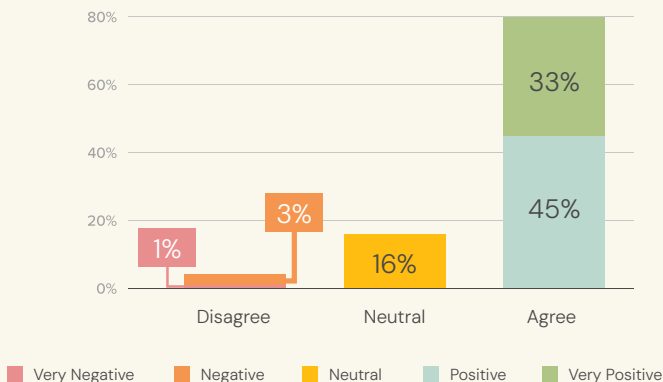
14%

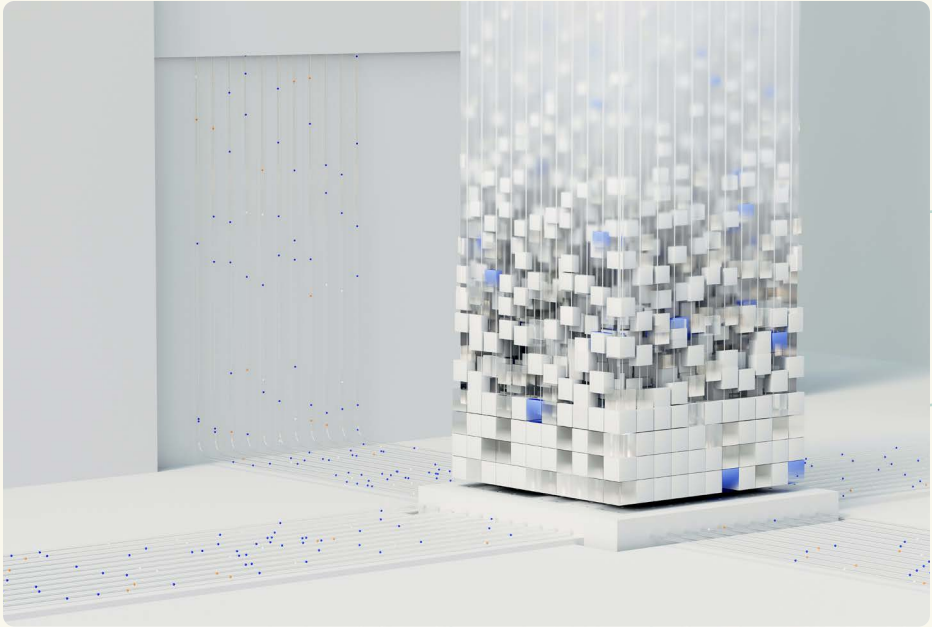


15%

describe themselves as unlikely or very unlikely to purchase from brands using AI-generated content or communication.

To what extent do you agree that a brand should disclose if it's using AI in its written materials and communication?





- Unsurprisingly, consumers are less enthusiastic about AI-generated content produced within the finance and medical industries, which is most likely due to concerns about accuracy. This aligns with Google's YMYL (your money or your life) content – topics that could significantly impact readers' health, financial stability, safety, or well-being. Common YMYL topics include finance, medicine, and current events.
- 35% would have a negative perception of a brand if they noticed mistakes in AI-translated content in their language, with 35% of people in the UK and Spain being unlikely to purchase from a brand if they noticed mistakes in AI-generated content.
- Social media posts and blog articles are the least popular potential use of AI with consumers, who tend to feel more positive towards the use of AI in conjunction with a human editor.

In editorial testing, AI-generated translations scored worse than human translations across all three languages – French, Spanish and Simplified Chinese – with a few key areas in which ChatGPT faltered.

## Factual and grammatical errors

More critical translation errors were found in the AI-translated copy compared with the human-translated copy, particularly in Chinese. Whilst most vocabulary selected by ChatGPT was technically correct, it often didn't use the most natural terminology for the target language, and awkward grammar also often interrupted the flow.

In the Spanish AI translation for instance, there were calques from the English language, punctuation errors, lack of italicisation of some English terms, and a verb that lacked a preposition and a complement to complete its meaning.

Similarly, the AI-generated translation for French stuck too closely to the English original in terms of punctuation. For example, quotation marks were used incorrectly in the French translation, with "ton" being left as in the original text rather than correctly updated to « ton ».

Furthermore, because ChatGPT is trained on outdated data, this can also contribute to factual inaccuracies and inconsistencies, alongside the real risk of plagiarism.

## Stilted language

In all three translations, AI translated the text rigidly following the English way of speaking, which made it appear 'broken' and in some cases incomprehensible. With English and Chinese languages, in particular, having drastically different sentence structures, this approach meant that the AI versions do not read well to native speakers.

In addition, AI often failed to recognise the nuances between different translations, not choosing the most appropriate word for the target audience. For instance, in the Chinese translation, 'motorcycles' are 摩托 instead of 摩托车, as the former is more commonly used among young people.

## Lack of emotive response

Editors felt that the AI copy was obviously machine-translated. This led to a flat feeling with a lack of emotional response. Across all three AI translations, the writing style was said to be 'plain' and, because it stuck too closely to the original English source, it struggled to truly engage readers. Failing to use an engaging tone and convey the excitement of the original text, ChatGPT's efforts were felt to be largely 'robotic'.

Notably, even though both the human and machine translations would inevitably require a further layer of human QA to truly localise the copy, reviewers commented that the human-translated text would be much quicker to edit than the AI-generated version – calling into question potential time savings through using ChatGPT.

## What would Google think?

Google has stated that use of AI-generated content does not necessarily harm SEO performance. The focus is on quality rather than human or AI source. Nevertheless, there are still several SEO concerns for AI-generated content, particularly when considering the E-E-A-T guidelines:

- Experience: AI-generated content can harvest a lot of prior information but lacks original thought.
- Expertise: Factual inaccuracies are still common in AI-generated content.
- Authoritativeness: As AI models 'learn' from existing content, content generated using these tools does not have the same authority as thought-leaders in a particular space.
- Trustworthiness: As our research shows, consumers are generally receptive to the idea of AI-generated content but retain some scepticism, particularly regarding accuracy and transparency.

When it comes to creating and localising content, the main concern is around ranking in the top positions. Those that rank in the top three places have huge levels of interesting, thought-provoking original content. This generates backlinks, which in turn increases the keyword rankings. For the best quality content and best results, AI should be used with care, as it will likely dilute original thought by paraphrasing what's already out there.



## Section 02

# A VIEW FROM CLOUD'S HEAD OF LANGUAGE SERVICES

While ChatGPT and similar technologies can provide reasonable accuracy for content with straightforward language and limited cultural or contextual nuance, there is still room for improvement when it comes to localising for a brand's tone of voice.

In particular, we find that ChatGPT can still fall short in the following areas when it comes to localisation.

## Getting the locale right

ChatGPT lacks the ability to ask for clarification, which can lead to automatic assumptions about the desired locale if not explicitly specified. For example, it defaults to American English when asked to translate into English, and to Simplified Chinese when asked to translate into Chinese. Such assumptions can lead to confusion among local audiences due to differences in language variations. To avoid this problem, translators should always provide precise prompts that specify the intended locale.

**Example of  
ChatGPT  
automatically  
translating into  
American English:**

Translate this sentence into English: 生日聚会着装要求: 粉色裤子

Birthday party dress code: pink pants.

In British English please

Birthday party dress code: pink trousers.

## Idiomatic expressions

Idiomatic expressions can have multiple interpretations and require context to decide which is the most appropriate. Usually, idiomatic expressions have linguistic elements that involve puns or metaphors that rely on the peculiarities of the source language, making them more difficult to translate correctly.

## Cultural references

Cultural references often have meanings that are specific to a particular culture and may not have a direct equivalent in the target language. Different cultures have different values and norms, which may not correspond to the world view of the local audience. Failure to account for these differences can lead to misinterpretation or offence.

## Localisation of currencies and addresses

When translating, ChatGPT often retains the original currency from the source language rather than converting it to the local target currency. In addition, when localising addresses, it may produce a mixture of the source and target languages, which can be impractical for users trying to enter the translation into services such as Google Maps.

### Example of how ChatGPT translates addresses:

Translate this into Chinese: 9th & 10th floors, The Bard Building, 20 Curtain Rd, London EC2A 3NG

Here's the translation of the address into Simplified Chinese:

9th & 10th 楼, 巴尔德大厦, 20 Curtain Rd, 伦敦 EC2A 3NG



## Ethical oversight

As with all machine-translated content, ChatGPT raises concerns about misinformation, bias and ethical implications. For example, it could inadvertently mistranslate a term or phrase for a religious text or a text related to a particular cultural belief, which could lead to a misrepresentation of the original meaning and potentially offend readers. In which case translators can play a crucial role in providing ethical oversight by carefully reviewing the content to ensure that it is consistent with ethical norms, cultural awareness and social responsibility.

As ChatGPT and similar AI language models develop, the roles of editors and translators may evolve to embrace these emerging technologies. We expect the role of translators to be reshaped to emphasise their expertise in creative content and quality control, as well as ethical oversight. For example, as AI handles more repetitive tasks, human translators may focus on more creative and complex projects that require a deeper understanding of a brand's tone of voice, glossary, context and culture. This could lead to greater demand for specialised language skills in the translation and editing of creative content, such as marketing content, video game dialogues, literary works, poetry, and other artistic expressions. For now, human translators remain crucial in capturing a brand's unique tone of voice and ensuring that translations are culturally appropriate.

# Current opportunities for AI global content production

In a rapidly changing environment, the limitations described above are likely to improve with each iteration of AI tools. Nevertheless, opportunities to leverage these tools already exist for global content production.

- **Keyword research:** ChatGPT does not access real-time keyword data but is able to compile historic keyword volume as a starting point for keyword research. It can effectively categorise keywords and is useful for finding long-tail and question keywords.
- **Ideation:** Generating ideas for article titles based on keywords enhances the efficiency of the ideation process, particularly through the careful use of prompts to encourage ChatGPT to represent your brand tone of voice.
- **Research:** Rapid generation of research for content is another effective use of AI in the writing process, which customers receive well. Human editing is essential while concerns around accuracy remain.
- **Structure:** Based on keywords and titles, using AI to generate an outline of content can be a valuable way of providing detailed writing briefs.
- **Localisation:** As our research shows, localising content will still require human input to identify errors, but the effectiveness of AI can dramatically speed up the localisation process in specific languages and contexts.
- **Proofreading:** ChatGPT can make good quality suggestions around grammar and sentence structure, although is less effective with matching brand tone of voice.



# CONCLUSION

# An exciting future

Based on this research, it is clear that AI tools offer an exciting future for content production and localisation. Most encouragingly, consumers are open to the use of AI-generated content by brands and are likely to become more open to it as public awareness of AI tools continues to grow.

Currently, many practical limitations exist, with human input still essential to maintain accuracy and consumer trust. Transparency is also extremely important to consumers and should be a key consideration for brands utilising AI-generated content.

Brands can maximise the use of AI tools by thinking through frameworks for their use and by identifying specific types of content and tasks to develop an effective process. Human input is important in 'training' the AI tool to produce the desired output, and in editing content to ensure accuracy and emotional connection.

The fast-moving nature of AI tools mean that governance is often playing catch-up. The consumer desire for transparency regarding AI-generated content means that brands should carefully consider their use of AI, putting frameworks and guidelines in place to maintain consumer trust.

It seems fitting to let ChatGPT have the final word on how brands can benefit from AI content:



ChatGPT can be used by brands to enhance their content production and localisation efforts, improve customer engagement, and streamline communication channels. Brands can leverage ChatGPT's natural language processing capabilities to create high-quality content, localise content for different languages and regions, engage with customers on various platforms, develop chatbots for automated customer support, and provide personalised responses to customer inquiries.

# About Croud

Croud is a global, full-service digital agency that helps businesses drive sustainable growth in the new world of marketing. With a rich heritage in performance, we apply that mindset to everything we do; brand planning, strategy, integrated media, social, creative, and data.

Founded in 2011 with the mission to reinvent the agency model, we combine 500+ in-house digital experts with a global, on-demand network of marketing specialists. This agile, scalable model gives us more time to really understand our clients' businesses, and deliver truly incremental growth.

Part of the Croud Group, which also includes luxury specialists VERB Brands and social-first agency Born Social, Croud is proud to partner with leading global brands, including Audible, Vans, IWG, and AllSaints.



[www.croud.com](https://www.croud.com)

020 8017 7723

croud