## AI ??????????????

# AI ??????????????

???Ilia Shifrin



Successful businesses such as Spotify have shown how artificial intelligence (AI) can build brand loyalty through hyper-personalization. Spotify famously uses AI to help its subscribers explore new music by suggesting songs and artists that users might be curious to learn more about. AI is crucial to keeping users engaged and delighted by opening up their worlds to new content. In fact, AI shapes users' tastes and makes people aware of new artists, thus making Spotify an even more influential brand. A streaming service platform that can effectively target its audience by offering highly relevant content selection based on users' individual preferences naturally attracts the world's largest recording labels and professional musicians. Most if not all media streaming platforms are already offering or about to roll out personalized content recommendation engines. The next frontier for AI is creating hyper-personalized experiences at a local level. And there is still a lot of work to be done here.

**Customizing Content for Different Countries**

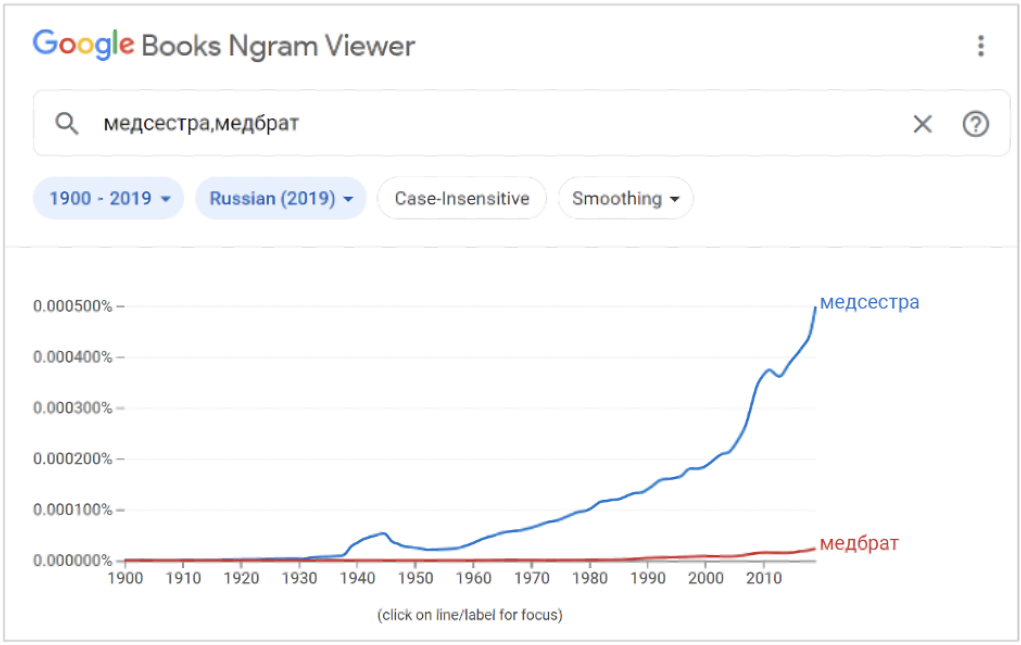
Localization means adapting and customizing content for different locales – a complex process that requires an understanding of idioms and cultural norms, among many other matters. If you have ever used Google to search for content in different countries, you've probably seen how the world's

leading search engine attempts to offer results relevant to the region you are searching from and the language your search query is written in. If you type “car” in the United States, ideally Google should serve images of automobiles that reflect your own tastes as an American consumer based on Google’s algorithm that understands, among other things, what kinds of car brands are popular in your area. Ideally, Google would serve up a completely different set of images for “car” in areas of Russia, where smaller and more fuel-efficient automobiles such as the locally manufactured Lada Granta and Lada Vesta are especially popular, with the vehicles produced under the Lada brand reportedly constituting for almost 30% of all registered vehicles in the country as of the article publication date.

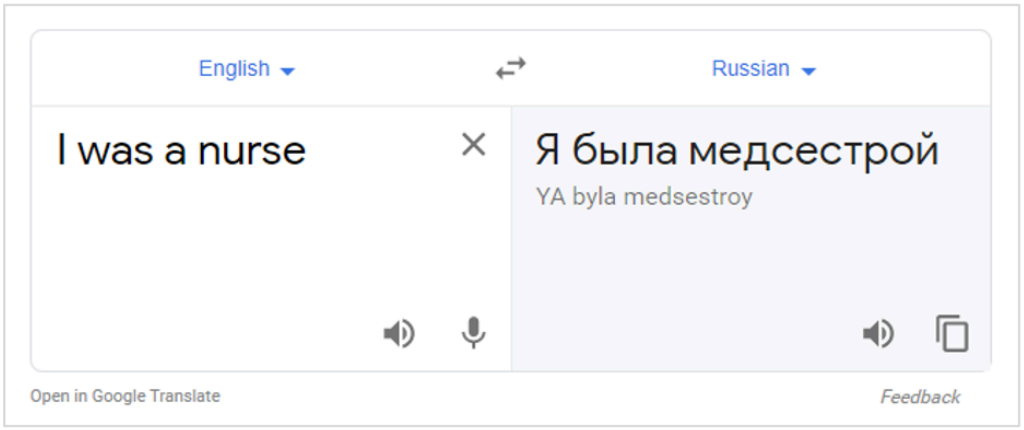
To cite another example, a major home improvement retailer is currently working with Centific to adjust its merchandise catalogues and localize its AI recommendation engine for multiple regions the customer is purchasing from. That way, the retailer provides a lot more relevant and personalized experience beyond the rather generic “also frequently bought together” type of recommendation that many retailers already use.

**Strong Potential with Proper Training**

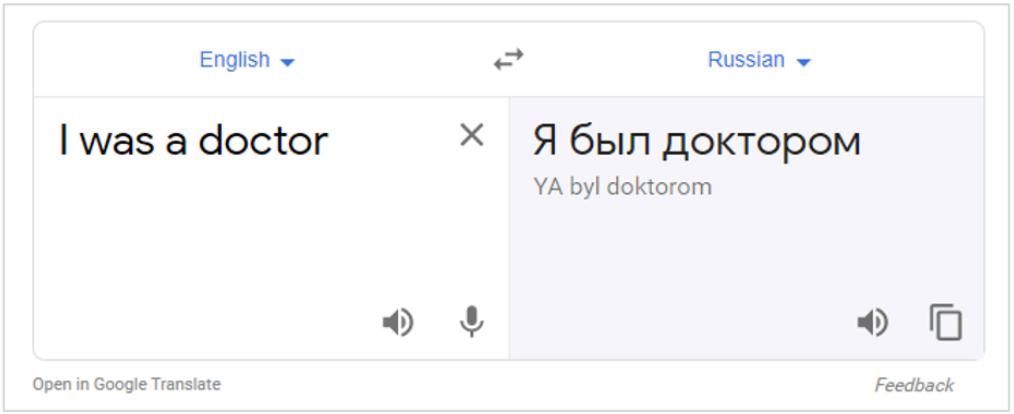
The potential for AI to deliver hyperlocal, hyper-personalized experiences is vast. These kinds of experiences can help businesses such as Google with its highly targeted ad services and adaptive multilingual digital assistant or Netflix with its newly announced shuffle play functionality build emotional trust because users feel like the content is a more relevant and natural reflection of their own cultures. But AI must be trained properly to deliver these kinds of experiences. Without a proper human-in-the-loop machine learning methodology, AI will deliver highly flawed experiences. For example, we have become painfully aware of how AI-generated content reflects race and gender bias, and those shortcomings become magnified at a global level. It does not take a data scientist to spot these problems. Google still translates the sentence “He was a nurse” into Russian as “?? ??? ??????????,” for instance. To those readers here who do not speak Russian, the word “?????????” literally translates as “a medical sister.” Since historically most of nurses were females, this has apparently significantly affected Google’s training datasets for the Russian language. The word “???????” (“medical brother”) came into use only much later towards the end of the 20th century to refer to a male medical assistant. Even now because of female workers dominating this medical specialization, the word “?????????” is used roughly 21 times more frequently in Russian publications, thus solidifying gender bias in training data.



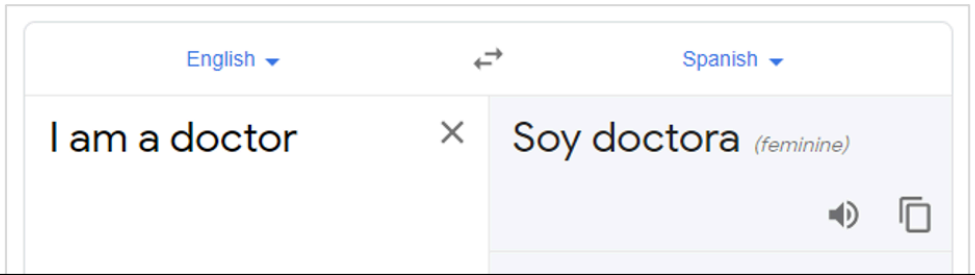

The word ‘?????????’ (‘medical sister’) referring to a nurse for the occupation historically predominated by women has been used over 21 times more frequently in the past century than ‘???????’ (‘medical brother’).




Translation into Russian of the gender-neutral phrase ‘I was a nurse’ gets feminine.




Translation of the gender-neutral phrase ‘I was a doctor’ gets masculine, inheriting, solidifying, and further promoting the existing societal imbalance. Currently two thirds of all doctors in the USA are


English
Spanish
I am a doctor
Soy doctora (feminine)

Comparison of the frequency of use of phrases ‘he was a doctor’ vs. ‘she was a doctor’ in Russian and English texts.

Political bias not infrequently finds its way into machine translation algorithms too through flawed training corpora used in natural language processing solutions. In late 2020, it was reported that Google started translating the phrase “Thank you, Mr. President” from English to Russian as "Thank you, Vladimir Vladimirovich.” Also interesting to note that some time earlier Google was translating the phrase “Russian Federation” as “Mordor”, the fictional evil land in *The Lord of the Rings*; and “Russians” to “occupiers.” Google blamed the glitches on a translation bug and corrected them, but their existence in the first place underscored the magnitude of the challenge of preventing bias in machine translation algorithms – complex interconnected hierarchical dependencies that are deduced from language data that is in its nature a product of human culture, subjective and prejudiced by definition.

**The Crucial Role of AI Localization**

The solution is for businesses to incorporate proper AI localization services. Localization requires high-quality training data that reflects the way that people speak, sound, and express themselves – and lots of it. AI localization as defined by Centific fully encompasses language, image, and voice AI data services, including multilingual corpora creation for natural language processing solutions, e.g. machine translation and speech-to-text engines, content curation for market-specific product and merchandise search relevance, data collection for locale- and language-aware multimedia recommendation engines, culturalized image search support, human-centered chatbots development and AI-powered locale-relevant eLearning solutions.

When all those many moving parts are managed well, businesses achieve the benefits of AI trained on hyperlocal content and in-market user experiences-generated data. For example, Swisscom successfully developed a multilingual platform to serve trilingual customers with highly tailored digital experiences. Through measuring international customers’ behavior data to support internal decisions, such as which colors to avoid on banners and where and in which language calls to action should be placed the reported average digital marketing efficiency increased by 40 percent.

AI localization can make a big difference in the service and product personalization journey as brands are planning to enhance their customer experiences in the coming years. To get started, we recommend that companies:

- Build a strong foundation of professionally localized and responsibly curated data to help train AI models relevant to your target markets audiences – whether you are working on a custom neural machine translation engine getting it to use the correct account-specific terminology and style, a personal digital assistant speaking 30+ languages naturally and effortlessly understanding its users requests regardless of their regional accents, a smart home security solution alerting of the pet’s agitation or misbehavior, or an intelligent chatbot capable of tracking the person’s emotional state by their response and adjusting communication strategy and content on the fly.
- Inquire with your data vendor on their AI localization competencies, global scale, and reach to understand whether their production throughput is sufficient to continuously keep your custom models and localized AI solutions up to date.
- Designate an AI center of excellence to drive R&D.
- Partner with AI-first localization and data agencies.

**Contact Centific**

Centific's AI Data Localization services help businesses build and adapt AI-powered systems to new markets and languages using the most reliable thoughtfully collected and curated data. With tens of thousands of resources across our in-house and outsourced localization and data collection network, we connect the most qualified in-market individuals with the task at hand, based on their availability and validated skills. Our translation and data curation teams do not work alone. The framework of battle-tested processes and intelligent systems streamline the workflow, from resource profiling, onboarding, training, and testing to delivery and QA. We know that data quality is mission-critical in training AI. This makes Centific's robust quality assurance process all the more essential – detecting and removing bias in data of any kind, as well as fixing slips or errors that may easily lead the model astray.

Our globe-spanning network of in-house and external translation resources covers 200+ languages and accents relevant in countless regional markets. With skills honed for over 20 years, our teams have developed a vast body of data localization knowledge. From cloud technology, to machine learning and AI integration, Centific is constantly advancing.

**Contact us** to succeed with hyper-personalized content strategy and market-relevant localized AI solutions.

- 
- 
- 
- 
-