FUNDAMENTALS

The 10 Biggest Issues Facing Natural Language Processing

With the help of complex algorithms and intelligent analysis, Natural Language Processing (NLP) is a technology that is starting to shape the way we engage with the world. NLP has paved the way for digital assistants, chatbots, voice search, and a host of applications we've yet to imagine.

Of course, NLP isn't on an unlimited upward trajectory just yet. There are still many issues faced by NLP developers... but we're already starting to find ways to resolve them.

What Is Natural Language Processing?

Essentially, NLP systems attempt to analyze, and in many cases, "understand" human language. These systems are often responsible for recognizing human speech (in other words, being able to "hear" what someone is saying), understanding it (figuring out the context of what's been said), and generating a natural language response (i.e. talking back).

If you've had a conversation with a digital assistant or chatbot recently, you've likely seen firsthand just how far we've come in terms of technological sophistication. However, we still have a long way to go.

NLP, these are some of the most important issues we need to resolve:

1. Language differences

In the United States, most people speak English, but if you're thinking of reaching an international and/or multicultural audience, you'll need to provide support for multiple languages.

Different languages have not only vastly different sets of vocabulary, but also different types of phrasing, different modes of inflection, and different cultural expectations. You can resolve this issue with the help of "universal" models that can transfer at least some learning to other languages. However, you'll still need to spend time retraining your NLP system for each language.

2. Training data

At its core, NLP is all about analyzing language to better understand it. A human being must be immersed in a language constantly for a period of years to become fluent in it; even the best AI must also spend a significant amount of time reading, listening to, and utilizing a language. The abilities of an NLP system depend on the training data provided to it. If you feed the system bad or questionable data, it's going to learn the wrong things, or learn in an inefficient way.

3. Development time

Along similar lines, you also need to think about the development time for an NLP system.

To be sufficiently trained, an AI must typically review millions of data points. Processing all those data can take lifetimes if you're using an insufficiently powered PC. However, with a distributed deep learning model and multiple GPUs working in coordination, you can trim down that training time to just a few hours. Of course, you'll also need to factor in time to develop the product from scratch—unless you're using NLP tools that already exist.

4. Phrasing ambiguities

Sometimes it's hard even for another human being to parse out what someone means when they say something ambiguous. There may not be a clear concise meaning to be found in a strict analysis of their words. In order to resolve this, an NLP system must be able to seek context to help it understand the phrasing. It may also need to ask the user for clarity.

5. Misspellings

Misspellings are a simple problem for human beings. We can easily associate a misspelled word with its properly spelled counterpart, and seamlessly understand the rest of the sentence in which it's used. But for a machine, misspellings can be harder to identify. You'll need to use an NLP tool with capabilities to recognize common misspellings of words, and move beyond them.

6. Innate biases

7. Words with multiple meanings

No language is perfect, and most languages have words that have multiple meanings. For example, a user who asks, "how are you" has a totally different goal than a user who asks something like "how do I add a new credit card?" Good NLP tools should be able to differentiate between these phrases with the help of context.

8. Phrases with multiple intentions

Some phrases and questions actually have multiple intentions, so your NLP system can't oversimplify the situation by interpreting only one of those intentions. For example, a user may prompt your chatbot with something like, "I need to cancel my previous order and update my card on file." Your AI needs to be able to distinguish these intentions separately.

9. False positives and uncertainty

A false positive occurs when an NLP notices a phrase that should be understandable and/or addressable, but cannot be sufficiently answered. The solution here is to develop an NLP system that can recognize its own limitations, and use questions or prompts to clear up the ambiguity.

10. Keeping a conversation moving

Many modern NLP applications are built on dialogue between a human and a machine. Accordingly, your NLP AI needs to be able to keep the conversation moving, providing additional questions to collect more information and always pointing toward a solution.

Find out more

If you have any Natural Language Processing questions for us or want to discover how NLP is supported in our products please get in touch.
Contact Us

Share this post

Fundamentals   NLP

Back to Fundamentals page

Previous   Next

i2 GROUP

Contact:

hello@i2group.com

Company

About i2

Careers

Partners

Contact

PR and Media Enquiries

Case Studies

Solutions

Analyst's Notebook

TextChart

iBase

Analysis Hub

Analysis Studio

Solution Packages

Subscription Bundle

Sectors

Defense

Government

Banking and Finance

Law Enforcement

National Security

Support

Raise a Support Ticket

KnowledgeBase

Product Documentation

Supported Software

Latest Release

Tips and Tricks

Newsletter Sign Up

Sign Up

CYBER ESSENTIALS   bsi ISO/IEC 27001 Information Security Management CERTIFIED   AICPA SOC

Privacy Policy   License terms