



## DEVELOPING A MULTILINGUAL GRAMMAR AND SPELL CHECKING SYSTEM

As part of an ongoing, multi-year project with one of the world's largest technology companies, Lionbridge AI provides linguistic rule development, quality assurance and maintenance for the development of textual proofing components, from spell checking and grammar to more advanced features.

### THE CHALLENGE

Proofing tools enable users to write with precision by automatically detecting grammar, spelling, punctuation, word choice, and style mistakes in text. Like other natural language machine learning applications, a great deal of linguistic and engineering resources is required to develop reliable proofing tools in any language.

Our client licenses its word processor to over one billion users worldwide. As a multinational corporation, they need to ensure that their multilingual proofing tools are accurate and up to date, reflecting not only language trends but also cultural and social changes. To do this, they required a partner with extensive linguistic expertise, who also had access to a specialized team with the ability to support software development in 16 languages.

### THE SOLUTION

The client selected Lionbridge from a range of vendors because of our experience in providing multilingual training data. In addition to recruiting and training linguistic resources, Lionbridge provides the client with end-to-end software development, quality assurance and engineering support.

#### Linguistic Rule Development

- 1. Grammar Creation:** Project managers and computational linguists worked closely with the client to write and deploy clear syntactic rules for each language.
- 2. Ontology Creation:** Lionbridge regularly analyzes language trends and new terms before adding words to each language-specific dictionary.



- 3. Part of Speech (POS) Tagging:** Language specialists break down text data, tagging individual words with their morphological characteristics to train the client's machine learning based models.
- 4. Scaling:** Lionbridge collaborates with the client in the development of advanced new writing assistance components, building out their tested ideas and implementing them across multiple languages.
- 5. Data Annotation:** Each month, the Lionbridge team curates tens of thousands of words and/or sentences based on a wide range of requirements.

#### Quality Assurance

Lionbridge's project management team coordinates with dozens of full-time quality assurance engineers to check for errors and suggest corrections.

#### Component Maintenance

Lionbridge provides continued engineering and technical support following the client's software releases to ensure that all requirements are met.

## THE RESULTS

By partnering with Lionbridge, the client was able to **dramatically improve their proofing tool's accuracy rate in over 16 languages.**

Since the project demonstrated Lionbridge's ability to complete large-scale, linguistically complex projects, the client ramped up our involvement in other projects across the entire scope of their services. Lionbridge now processes millions of data points on a monthly basis for multiple machine learning teams.



**16+** Languages



**100+** Full-time Data Engineers



**150+** Grammar Rules per Language



**300,000+** Hours of Work Completed

We sourced data for the following languages:

- English
- Italian
- Russian
- German
- Brazilian Portuguese
- Danish
- Finnish
- Korean
- Chinese (Traditional)
- Chinese (Simplified)
- Hebrew
- Hungarian
- ... And more



## ABOUT THE CLIENT

Our client is a multinational technology company that develops, manufactures, licenses, and sells computer software, consumer electronics, and related services. As part of their product offering, they license productivity tools to over one billion customers worldwide.

## LIONBRIDGE AT A GLANCE

- **300+** languages and dialects
- **500,000** global contributors
- **20+** years of experience providing human-annotated training data

→ Learn more at [LIONBRIDGE.AI](https://LIONBRIDGE.AI)

