



SCALING SOLUTIONS FOR SOFTWARE IN MULTIPLE LANGUAGES

17 ADDITIONAL
LANGUAGES

25 PROJECTS

COMPONENTS FROM
SCRATCH

Our client needed a partner that could help them scale their software implementation framework into multiple languages. We delivered—in 17 additional languages, in the context of 25 projects.

The Challenge

Our client has built a framework to allow the implementation of a software for multiple languages. This framework relies on both supervised and unsupervised machine learning methods as well as text analysis and pre- and post-processing procedures, including tokenization and Part of Speech tagging.

Our client developed and tested a set of tools and processes, using them to implement working components for several languages. The client sought a partner to document, develop, and maintain the framework, as well as to implement the components in additional languages.



About the Client

One of the largest enterprise software providers in the world, our client is a leader in natural language text technology. Among its suite of products, our client builds efficient algorithms that process text, which can then be used by other applications.

Key Terms

Named Entity Recognizer

In a given text, locates named entities and identifies which pre-defined categories they belong to (e.g., person names, place names).

e.g., Peter is an engineer and lives in Toronto
Peter = person, Toronto = place

Key Phrase Extractor

Extracts a list of key phrases or key words in relation to their relevance in the context of a given text.

e.g., Based on the sentence above, 1. Key Phrase Extractor
2. List of key phrases 3. Given text

The Results

Lionbridge effectively helped our customer develop scaling solutions for 17 additional languages. For most languages, the software has been implemented from scratch. We created all necessary documentation, roadmaps, and auxiliary tools for implementing software from scratch, as well as improving an existing software, using the client's framework for any language. We've additionally developed a structure to record all steps taken for the software to reach a satisfactory performance level. Finally, we consulted our client on how it could further improve the framework and its components.

We've made these achievements in the context of 25 projects, and our client has integrated the resulting solutions into its cloud-based API, providing Natural Language Processing services over basic text.

LEARN MORE AT

[LIONBRIDGE.COM](https://www.lionbridge.com)

The Solution

Lionbridge effectively helped our customer develop scaling solutions for 17 additional languages. For most languages, the software has been implemented from scratch.



1. Establish Efficient Workflow

In collaboration with the client, Lionbridge underwent a mentoring period whereby we initiated the implementation of a software for a pilot language using the client's framework. We then continued, in house, the implementation of software for the pilot language in parallel with the implementation of software for several additional languages. This resulted in timely execution and process improvements.



2. Overcome Language-Specific Challenges

While moving from language to language, our team was able to successfully address language-specific challenges, including:

Named Entity Recognition: Capitalization Challenges

- Some languages (e.g., Arabic, Japanese, Korean) have no word capitalization—a main feature in Named Entity Recognition.

Key Phrase Extraction: Overgeneration

- A key phrase may be correctly extracted because it contains a certain word that occurs frequently in the text; however, another key phrase may simultaneously be incorrectly extracted because it contains the same frequent word, but the key phrase extracted is not relevant to the given text.

The Lionbridge team implemented linguistic-based rule-writing scripts, which we integrated into the framework to address these challenges across languages. Our team also implemented auxiliary tools for text processing to support the software development, as well as analytical tools that could easily identify areas for improvement.



3. Provide Infrastructure

Lionbridge has built the world's largest Natural Language Processing team, which comprises linguists, project managers, data engineers, and a global network of language experts dedicated to the creation of Natural Language Processing solutions and assets for potentially any language. We employ a cloud-based platform that enables our global team to work from any location, using Lionbridge's tools and benefiting from guidance and oversight by Lionbridge project managers and coordinators. Our global reach means our teams can be productive for our clients around the world, around the clock.

Our client required a company that could go beyond the services required for traditional localization projects. It needed a partner with a deep understanding of machine learning methods—an understanding that develops and enhances text pre- and post-processing procedures along with linguistic-based rule-writing. These skills are not part of a typical translator's background. Lionbridge cultivates and deploys a network of experts with machine learning understanding and facility in over 200 languages.