



## SKALIERUNGSLÖSUNGEN FÜR SOFTWARE IN MEHREREN SPRACHEN

17 ZUSÄTZLICHE  
SPRACHEN

25 PROJEKTE

NEU ENTWICKELTE  
KOMPONENTEN



Unser Kunde hat einen Partner gebraucht, um sein Framework für die Softwareimplementierung in mehrere Sprachen zu skalieren. Im Rahmen von 25 Projekten haben wir 17 zusätzliche Sprachen geliefert.

### Herausforderung

Unser Kunde hat ein Framework entwickelt, mit dem eine Software in verschiedenen Sprachen implementiert werden kann. Dieses Framework basiert auf beaufsichtigten und unbeaufsichtigten Methoden für maschinelles Lernen sowie auf Textanalysen und Textvor- und Nachbereitung, unter anderem Tokenisierung und Part-of-Speech-Tagging.

Unser Kunde entwickelte und testete Tools und Prozesse, um damit Komponenten für mehrere Sprachen zu implementieren. Er war auf der Suche nach einem Partner, der das Framework dokumentierte, entwickelte und wartete sowie die Komponenten in zusätzlichen Sprachen implementierte.



### Über den Kunden

Unser Kunde, einer der weltweit größten Anbieter von Enterprise-Software, ist führend im Bereich Natural Language Text Technology. Im Rahmen seiner Produktsuite entwickelt unser Kunde effiziente Algorithmen zur Textverarbeitung, die dann in anderen Anwendungen eingesetzt werden können.

## Zentrale Begriffe

### Eigennamenerkennung

Hierbei werden in einem Text Eigennamen gefunden und es wird ermittelt, in welche vordefinierte Kategorie sie gehören (z.B. Personennamen, Ortsnamen).

Z.B. Peter ist Ingenieur und lebt in Toronto.  
Peter = Person, Toronto = Ort

### Extrahierung von zentralen Begriffen

Damit wird eine Liste mit wichtigen Begriffen oder Keywords auf Basis ihrer Relevanz für den Kontext eines Texts extrahiert.

Z.B. basierend auf dem Satz oben: 1. Extrahierung von zentralen Begriffen, 2. Liste zentraler Begriffe, 3. Text

## Das Ergebnis

Lionbridge unterstützte den Kunden effektiv dabei, Skalierungslösungen für 17 zusätzliche Sprachen zu entwickeln. In den meisten Sprachen wurde die Software dabei ganz neu implementiert. Wir erarbeiteten alle nötigen Dokumentationen, Projektpläne und Hilfstools für die Implementierung der Software. Außerdem optimierten wir mithilfe des Frameworks des Kunden eine bestehende Software. Wir entwickelten zudem eine Struktur zur Erfassung aller Schritte, die nötig waren, um die Software so weit zu optimieren, dass eine zufriedenstellende Leistung erzielt wurde. Schließlich zeigten wir unserem Kunden auf, wie er das Framework und die Komponenten weiter verbessern konnte.

Wir haben diese Leistungen im Rahmen von 25 Projekten erzielt. Unser Kunde hat die daraus entstandenen Lösungen in seine cloudbasierte API integriert, um Natural Language Processing zu ermöglichen.

ERFAHREN SIE MEHR AUF  
**LIONBRIDGE.COM**

## Die Lösung

Lionbridge hat den Kunden effektiv dabei unterstützt, Skalierungslösungen für 17 zusätzliche Sprachen zu entwickeln. In den meisten Sprachen wurde die Software dabei ganz neu implementiert.



### 1. Effiziente Workflows etablieren

In Zusammenarbeit mit dem Kunden hat Lionbridge eine Mentoringphase durchlaufen, bei der wir die Implementierung einer Software mithilfe des Frameworks des Kunden für eine Pilotsprache initiierten. Anschließend arbeiteten wir inhouse an der Implementierung der Software für die Pilotsprache und parallel an der Implementierung der Software für mehrere zusätzliche Sprachen. Dadurch waren eine schnelle Umsetzung und Prozessverbesserungen möglich.



### 2. Sprachspezifische Herausforderungen meistern

Beim Wechsel von einer Sprache zu einer anderen meisterte unser Team sprachspezifische Herausforderungen:

#### Eigennamenerkennung: Probleme mit der Großschreibung

- In einigen Sprachen (z.B. im Arabischen, Japanischen und Koreanischen) gibt es keine Großschreibung. Die Großschreibung ist eine zentrale Funktion der Eigennamenerkennung.

#### Extrahierung von zentralen Begriffen: Übergenerierung

- Ein zentraler Begriff kann korrekt extrahiert werden, weil er ein bestimmtes Wort enthält, das im Text häufig vorkommt. Allerdings wird vielleicht fälschlicherweise gleichzeitig noch ein anderer, für den vorliegenden Text nicht relevanter Begriff extrahiert, weil er dasselbe häufige Wort enthält.

Das Lionbridge-Team hat Scripts zur sprachbasierten Regelerstellung implementiert. Diese Scripts haben wir in das Framework integriert, um diese Herausforderungen sprachübergreifend zu lösen. Unser Team hat außerdem Hilfstools für die Textverarbeitung implementiert, um die Softwareentwicklung zu unterstützen, sowie Analysetools, um Optimierungspotenzial leicht zu erkennen.



### 3. Infrastruktur bereitstellen

Lionbridge hat das weltweit größte Team für Natural Language Processing aufgebaut. Es besteht aus Linguisten, Projektmanagern, Datentechnikern und einem globalen Netzwerk aus Sprachexperten, die Lösungen und Assets für Natural Language Processing in praktisch jeder Sprache entwickeln. Wir nutzen eine cloudbasierte Plattform, die es unserem Team ermöglicht, an jedem Standort der Welt zu arbeiten. Dabei profitiert unser Team von den Lionbridge-Tools sowie von der Unterstützung und Anleitung der Lionbridge-Projektmanager und -Koordinatoren. Dank unserer globalen Reichweite ist unser Teams auf der ganzen Welt rund um die Uhr für unsere Kunden einsatzbereit.

Unser Kunde brauchte ein Unternehmen, das weit mehr bieten konnte als herkömmliche Lokalisierungsanbieter. Er brauchte einen Partner mit umfassenden Kenntnissen über Methoden für maschinelles Lernen – um die Textvor- und -nachbearbeitung sowie die sprachbasierte Regelerstellung zu entwickeln und zu optimieren. Ein typischer Übersetzungsanbieter verfügt nicht über diese Fertigkeiten. Lionbridge greift auf ein Netzwerk aus Experten für maschinelles Lernen zurück, die in über 200 Sprachen bewandert sind.