



SKALNINGSLÖSNINGAR FÖR PROGRAMVARA PÅ FLERA SPRÅK

17 YTTERLIGARE SPRÅK

25 PROJEKT

KOMPONENTER FRÅN GRUNDEN



Vår kund behövde en partner som kunde hjälpa dem att skala sitt implementeringsramverk för programvara till flera språk. Vi levererade – på 17 ytterligare språk inom 25 olika projekt.

Utmaningen

Vår kund har byggt ett ramverk som gör det möjligt att implementera programvara på flera språk. Det här ramverket baseras både på övervakade och icke övervakade maskininlärningsmetoder samt textanalys och procedurer både före och efter bearbetningen, bland annat tokenisering och delar av taltaggar.

Vår kund har utvecklat och testat en uppsättning verktyg och processer, och använder dem för att implementera komponenter för flera språk. Kunden ville ha en partner som kunde dokumentera, utveckla och underhålla ramverket samt implementera komponenterna på ytterligare språk.



Om kunden

Vår kund är en världens största programvaruleverantörer och ledande inom textteknik på naturligt språk. Bland de programvaruprodukter som vår kund utvecklar finns effektiva algoritmer för textbearbetning, vilka sedan kan användas för andra tillämpningar.

Nyckeltermen

Identifiering av namngivna entiteter

Detta innebär att i en given text hitta namngivna entiteter och identifiera vilken fördefinierad kategori de tillhör (till exempel personnamn eller namn på platser).

Exempel: Peter är ingenjör och bor i Toronto
Peter = person, Toronto = plats

Extrahering av nyckelfraser

Detta innebär att extrahera en lista med nyckelfraser eller nyckelord i förhållande till hur relevanta de är i sammanhanget för en specifik text.

Exempel: Utifrån meningarna ovan, 1. Extrahering av nyckelfraser 2. Lista med nyckelfraser 3. Specifik text

Resultaten

Lionbridge hjälpte kunden att utveckla skalningslösningar för 17 ytterligare språk. För de flesta språk har programvaran implementerats från grunden. Vi har skapat all dokumentation, alla handlingsplaner och alla hjälpverktyg som behövs för implementeringen av programvaran – från grunden. Vi har dessutom förbättrat kundens befintliga programvara genom att använda kundens eget ramverk för ett visst språk. Utöver detta har vi utvecklat en struktur för att dokumentera alla steg som vidtas för att programvaran ska uppnå en godtagbar prestandanivå. Slutligen diskuterade vi med kunden om hur ramverket och komponenterna i det kunde förbättras ytterligare.

Vi har lyckats med detta inom ramen för 25 olika projekt. Kunden har integrerat de färdiga lösningarna i sitt molnbaserade API, som gör det möjligt att tillhandahålla tjänster för bearbetning av naturligt språk för grundläggande texter.

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

Lösningen

Lionbridge hjälpte kunden att utveckla skalningslösningar för 17 ytterligare språk. För de flesta språk har programvaran implementerats från grunden.



1. Upprätta ett effektivt arbetsflöde

I samarbete med kunden genomförde Lionbridge en pilotstudie där vi initierade implementeringen av programvaran för ett pilotspråk med hjälp av kundens ramverk. Vi fortsatte sedan internt med implementeringen av programvaran för pilotspråket parallellt med implementeringen av programvaran för flera ytterligare språk. På så sätt kunde vi spara tid samtidigt som processerna förbättrades.



2. Hantera språkspecifika utmaningar

När vårt team gick vidare från ett språk till ett annat lyckades de hantera språkspecifika utmaningar, som:

Identifiering av namngivna entiteter: Problemet med versaler

- I en del språk (till exempel arabiska, japanska och koreanska) används inte versaler – en viktig funktion vid identifiering av namngivna entiteter.

Extrahering av nyckelfraser: Övergenerering

- En nyckelfras kan extraheras korrekt eftersom den innehåller ett visst ord som förekommer ofta i texten. Men en annan nyckelfras kan samtidigt extraheras felaktigt eftersom den innehåller samma ord som förekommer ofta, men nyckelfrasen som extraheras är inte relevant för den specifika texten.

Lionbridge-teamet implementerade språkbaserade regelskript som integrerades i ramverket, och som kunde hantera dessa problem på flera språk. Vårt team implementerade dessutom hjälpverktyg för textbehandling som ett stöd i programvaruutvecklingen, samt analysverktyg som snabbt kunde identifiera förbättringsområden.



3. Skapa en infrastruktur

Lionbridge har världens största team med experter på bearbetning av naturligt språk. Där ingår lingvister, projektledare, dataingenjörer och ett globalt nätverk av språkexperter som arbetar med lösningar och tillgångar för bearbetning av naturligt språk inom praktiskt taget alla språk. Vi tillämpar en molnbaserad plattform som gör det möjligt för vårt globala team att arbeta överallt med Lionbridges verktyg samt få hjälp, vägledning och stöd av Lionbridges projektledare och koordinatörer. Vår globala närvaro innebär att våra team kan arbeta för våra kunder över hela världen, dygnet runt.

Vår kund behövde en partner som inte bara kunde arbeta med de tjänster som krävs vid traditionella lokaliseringprojekt. Kunden behövde en samarbetspartner med djup förståelse för maskininläring – en förståelse som gör det möjligt att utveckla och förbättra procedurer för både för- och efterbehandling av texter vid sidan av att ta fram språkbaserade regler. Traditionella översättare har vanligtvis inte en sådan kompetens. Lionbridge har ett ständigt växande nätverk av experter med kunskaper inom och erfarenhet av maskininläring på över 200 språk.