

Lokaliseringens framtid



AI Business
e-bokserie



*Så kan du använda språk och kultur som
verktyg i din egen digitaliseringsprocess*

I samarbete med:

LIONBRIDGE

2

Kärnan i företagets expansion

Lokalisering handlar om att anpassa produkter, kommunikation, tjänster med mera, så att de stämmer överens med kundens förväntningar på en viss plats.

I en globaliserad värld använder företag lokalisering för att ta sig in på internationella marknader för att öka intäkterna och samtidigt uppnå stordriftsfördelar. Lokalisering utgör därmed kärnan i alla affärsstrategier som höjer blicken bortom en lokal produkt. Även företag som är knutna till en viss region genom sitt varumärke – exempelvis Apples löfte om produkter som är ”Designed by Apple in California” – måste arbeta med lokalisering för att tala kunder i andra länder.

Översättning är naturligtvis fortfarande en central pusselbit i lokaliseringsarbetet. Som vi ska diskutera i den här e-boken utförs delar av det här arbetet manuellt av språkexperter. Andra aspekter anförtros banbrytande maskinöversättningsverktyg (MT). Åter andra utförs med hjälp av en hybridmetod, så kallad förstärkt översättning (se sidan 5), som är en blandning av dessa.

Verklig lokalisering sträcker sig dock bortom språket. För att framgångsrikt kunna överföra en tjänst eller produkt från en marknad till en annan krävs fingertoppskänsla för de små nyanser som är unika för varje plats. ”I princip återskapar vi en upplevelse på ett nytt språk”, berättar Inés Rubio, Senior Manager for Language Services på NI (National Instruments), för AI Business. ”Det handlar inte så mycket om språk, utan om kultur.”

I den här e-boken, som är sponsrad av översättnings- och lokaliseringsexperten Lionbridge, tittar vi närmare på några av de utmaningar som internationella företag ställs inför och fördelarna med väl utförd lokalisering. Vi intervjuar experterna och lämnar plats för andra perspektiv från organisationer som Zynga, Royal Caribbean, Volvo och Wikimedia Foundation. Vi hoppas att den här e-boken ska förse dig med kunskaperna du behöver för att utveckla verkligt globala produkter och tjänster.

Luke Dormehl | Associate Editor | AI Business

e-bok i samarbete med:

LIONBRIDGE

Lokaliseringens framtid | www.lionbridge.com



3

Introduktion till moderna lokaliseringstjänster
Språk och kultur

4

Marknadsöversikt
Lokalisering blir standard

5

Sätta AI i arbete
Vad är förstärkt översättning?

7

Fallstudie
Så hjälpte Lionbridge Royal Caribbean med deras lokaliseringsprocesser

8

Arbeta för en global publik
Zynga använder lokalisering för att vinna miljontals spelares kärlek

10

Tänka lokalt
Använda språktjänster för att förbättra Wikipedia

12

Framsteg inom maskinöversättning
Brytpunkter

14

Fallstudie
Så hjälpte Lionbridge Volvo Cars att skapa en enhetlig onlinemarknadsplats

15

Vad kommer sen?
Språkteknikens framtid

17

Ett akademiskt perspektiv
En fråga om enhetlig kvalitet

Språk och kultur

Det handlar inte så mycket om språk, utan om kultur. När vi till exempel pratar om spanska, vad riktar vi in oss på? Menar vi global spanska? Eller europeisk spanska? Hur ser vi till att målgruppen kan relatera? Det handlar definitivt inte bara om att översätta ord. Utan går mycket, mycket [djupare än så].

Översättning är viktigt – men det är ännu viktigare att förstå målgruppen

Det allra vanligaste är att lokalisering avser språklig översättning av material som webbplatser, instruktionshandböcker och programvara, så att de kan läsas eller förstås på annat sätt.

”Syftet med lokalisering är egentligen att göra produkten eller erbjudandet lokalt och en stor del i det handlar om att tala slutkonsumentens språk”, förklarar Kajetan Malinowski, produktledare och strateg på Lionbridge, för AI Business. Men språk är bara en del av sanningen – och några av världens mest lukrativa marknader kräver kulturell kompetens.

Bortom översättning

Har du någonsin hört uttrycket ”Det är inte vad du sa, utan hur du sa det?”. Det är inte utan anledning som man skiljer på lokalisering och översättning. Vid lokalisering tar man hänsyn till fler aspekter inom lokal kommunikation. Lingvistik är en del i det, men också regionala skillnader. Måttenheter är ett enkelt exempel på en standard som varierar mellan olika platser.

Andra inslag som kan variera beroende på plats är juridiska och tekniska aspekter samt ton och tilltal.

”I princip återskapar vi en upplevelse på ett nytt språk”, berättar Inés Rubio, Senior Manager for Language Services på NI (National Instruments), för AI Business. ”Det handlar inte så mycket om språk, utan om kultur. När vi till exempel pratar om spanska, vad riktar vi in oss på? Menar vi global spanska? Eller europeisk spanska? Hur ser vi till att målgruppen kan relatera? Det handlar definitivt inte bara om att översätta ord. Utan går mycket, mycket [djupare än så]”.

Lokalisering kan också handla om att översätta grafik, exempelvis gränssnittsikoner, för att tillgodose målgruppens preferenser. Sådana ikoner har inte nödvändigtvis universell eller otvetydig betydelse. Till exempel är det vanligt att använda en glödlampa för att symbolisera en idé på den amerikanska eller brittiska marknaden, men symbolen har inte samma betydelse på andra platser. En ”tummen upp”-emoji har positiva konnotationer i USA och Storbritannien, men är stötande i Latinamerika, Mellanöstern och Västafrika. Färgers betydelse skiljer sig också åt. I Indien betyder orange kärlek och mod, men i mellanösternkulturer förknippas färgen snarare med förlust och sorg.

Lyckad lokalisering innebär enorma fördelar. Men om lokaliseringen går snett eller man inte sätter sig in i lokala sedvänjor ordentligt, kan effekten bli extremt allvarlig. Det här ett av de viktigaste verktygen i moderna företags arsenal på 2020-talet.



Lokalisering blir standard

När man sätter igång finns lokalisering inte med på kartan”, säger hon. ”Det är något man kommer på i sista minuten. Det diskuteras inte alls tidigare i processen. När man lär sig av misstag och inser hur mycket dyrare det kommer att bli att göra ändringarna i efterhand lär sig företaget och hela teamet [att uppskatta det].

En titt på lokaliseringsmarknadens omfattning – och vad som driver på tillväxten

Modern lokalisering – som ofta förkortas ”l10n” där siffran 10 står för antalet bokstäver mellan l och n i localization – såg dagens ljus på 1980-talet. Det är ingen slump att detta även var årtiondet då datorprogram blev ett självklart inslag i kontor och hem världen över. För att dra nytta av den här växande marknadsmöjligheten var företag tvungna att använda lokalisering, oavsett om det handlade om att stödja en ny teckenuppställning eller att lägga till eller anpassa funktioner för vissa marknader. Till att börja med gjordes dessa internt, men omfattningen och betydelsen av sådana arbetsinsatser ledde snabbt till utvecklingen av en fullfjädrad lokaliseringsindustri.

Större än någonsin

De senaste tio åren har språktjänstmarknaden – som består av vägledning, översättning och en bredare uppsättning lokaliseringsinsatser – nästan fördubblats i storlek. Enligt en aktuell rapport från Slator, en ledande publikation i språktjänstbranschen, var den tillgängliga översättnings-, lokaliserings- och tolkningsmarknaden år 2020 värd 23,8 miljarder US-dollar. Teknik, spel, media och life sciences var

de kundsegment som växte snabbast och branschen som helhet förväntades växa med upp till 10 procent under 2021.

Statisa uppger att Europa utgör ungefär hälften av den globala språktjänstmarknaden, trots att kontinenten bara är hem för mindre än 10 % av världens befolkning. Det beror på att så många språk talas i regionen. När det gäller marknadsstorlek behåller USA dock förstaplatsen.

De här skillnaderna syns inte bara i jämförelser av olika länder. Lokalisering är en etablerad disciplin i vissa sektorer, men håller fortfarande på att befästa sin betydelse inom andra (framför allt nyare) branscher och företag.

Inés Rubio, Senior Manager for Language Services på NI, anser att intresset för lokalisering kan ses som ett tecken på mognad. ”När man sätter igång finns lokalisering inte med på kartan”, säger hon. ”Det är något man kommer på i sista minuten. Det diskuteras inte alls tidigare i processen. När man lär sig av misstag och inser hur mycket dyrare det kommer att bli att göra ändringarna i efterhand lär sig företaget och hela teamet [att uppskatta det]. Sedan börjar de lägga till sådana saker i processen.”

Tack vare tekniska framsteg kopplade till lokalisering blir verktygen som driver den här industrin hela tiden bättre. Nya innovationer, exempelvis neural maskinöversättning (NMT), har dykt upp först de senaste åren – men vi ser redan att den får stor effekt. Den här sortens genombrott kommer att fortsätta förändra lokaliseringsvärlden i grunden för både kunder och slutanvändare.





Vad är förstärkt översättning?

Ett samtal om förhållandet mellan människa och maskin med Will Rowlands-Rees och Jaime Punishill från Lionbridge

”Maskinöversättning är ett av de grundläggande användningsområdena för AI”, berättade Will Rowlands-Rees, Chief Product Officer på Lionbridge, för AI Business. ”Den första maskinöversättningen gjordes i början på 1950-talet. Redan då förutspådde man att manuell översättning skulle försvinna. De där förutsägelseerna stämde visst inte helt.”

Närmare 75 år senare spelar människan fortfarande en viktig roll för översättningsarbetet. Det gäller även när maskinöversättning används. Översättningsmotorer som är tillgängliga för allmänheten är visserligen gratis och kan fungera bra i vissa sammanhang, men de lämpar sig sällan för professionell användning.

”Vi känner alla till roliga historier om hur man klistrar in lite text och fått väldigt konstiga resultat, eller hur?”, säger Jaime Punishill, marknadsdirektör på Lionbridge, till AI Business. ”De här översättningsmotorerna är inte tränade. Det fungerar något bättre i en kommersiell miljö, men det uppstår

ändå knäppa, språkliga problem mellan två språk. Det är inte lätt att försöka översätta engelska till koreanska, än mindre tyska till koreanska. Maskinen kan hitta på mycket kul.”

Vikten av bra översättningar

Det kan vara okej att läsa en översatt nyhetsartikel om allt du vill ha är kontentan av historien. Men i andra lägen är ”bra nog” helt enkelt inte bra nog. Ett prospekt till en investeringsprodukt, en medicinsk text eller en juridisk text är bara några exempel på områden där översättningen måste göras helt perfekt.

”Om jag producerar en handbok till en ledande sportbilstillverkare och ger fel instruktioner kan det påverka både hälsa och säkerhet”, förklarar Will. ”Eller tänk dig klinisk forskning eller ett annat område där ord kan vara [en fråga om] liv och död. Ur ett företagsperspektiv påverkar allting också varumärket. Du vill vara säker på att innehållet översätts rätt, eftersom det annars får negativa följder för varumärket.” Traditionellt har den här processen kallats för efterredigering, vilket innebär att maskinöversättningsverktyg används

för att utföra den första översättningen. De råöversatta MT-segmenten lämnas sedan över till språkexperter som rättar till eventuella fel som datorerna har gjort. I teorin sparar man tid med den här metoden, men den har också sina begränsningar, till exempel det faktum att språkvetaren tar över först i slutet av processen.

Förstärkt översättning kommer in i bilden

Förstärkt översättning är ett begrepp som ursprungligen myntades av CSA Research. Här används en rad olika tekniker för att ge språkexperter möjlighet att översätta snabbare, mer konsekvent och med högre precision. Förstärkt översättning lägger grunden till en bättre översättningsupplevelse genom att bättre kombinera människa och teknik. Många av de här metoderna har utforskats tidigare, men genom att koppla samman dem går det att skapa ett mer heltäckande ramverk. Många arbetsuppgifter som anses mindre viktiga men tar mycket tid i anspråk automatiseras, så att språkexperterna kan fokusera sin kompetens på de områden där den bäst behövs. ➤

6

► Förstärkt översättning omfattar sju huvudområden:

Översättningsminne

Adaptiv neural maskinöversättning

Kvalitetsbedömning

Automatisk innehållsberäkning (ACE)

Terminologihantering

Obemannad projekthantering

Translation Management System (TMS)

Alla dessa arbetar sida vid sida. Till exempel kan översättningsminnen hjälpa till med återkommande översättningsuppdrag. ”Vi får kanske i samma produkthandbok till översättning varje år”, säger Jaime Punishill. ”Större delen av handboken ser likadan ut som förra året, men i den tidigare, analoga världen betalade företaget för att översätta den två gånger – eller ännu fler. Översättningsminnen är utformade för att titta på källtexten för att se om det finns en matchning.”

Om det finns en exakt träff kan texten köras igenom utan några ytterligare ändringar. Om det å andra sidan finns avvikelser i vissa avsnitt, kan de markeras så att en språkexpert kan uppdatera dem. På så sätt får man också möjlighet att stämma av omgivande ord och meningar för att se till att tillägget inte påverkar textsammanhanget.

En annan aspekt av förstärkt översättning är att automatiskt kunna avgöra vilka lingvist- och maskinresurser som passar bäst till ett visst översättningsjobb. Det kan innebära att jobbet tilldelas en översättare eller att man får veta vilken maskinöversättningsmotor som ger bäst resultat.

Utnyttja styrkorna

”En sak som man upptäcker på det här området är att alla maskiner är bra på olika saker”, fortsätter Jaime. ”En är bra på att översätta engelska till tyska, en annan från spanska till koreanska. En är bra på innehåll från läkemedelsindustrin, en annan bättre på marknadsföring. Man

måste blanda och kombinera olika MT-motorer för att få en kraftfull uppsättning som kan driva översättningssystemet.

”På Lionbridge hjälper vi företag att upprätta sådana och driver [systemen] åt dem. Det omfattar allt man måste göra, från att träna motorerna till att underhålla riktlinjerna för varumärke och röst, terminologordlistor och alla lingvistiska tillgångar som får allt att rulla på – helt friktionsfritt. Det handlar inte bara om att åstadkomma bra översättningar, utan också om att översätta med exempelvis Apples röst eller Microsofts röst.”

Will Rowlands-Rees påpekar: ”Vi kan säga till användaren att vi tror att det blir bäst om den här texten översätts manuellt av följande skäl eller att vi vill skicka det här vidare till MT och det här är anledningen till att vi tror att MT blir bra. Vi har förstått vissa av anledningarna till att manuell översättning är bättre, eller när MT är bättre, beroende på situation. Vi har programmatiskt insett det när vi tar emot jobb och kan sedan skicka innehållet vidare på lämpligaste sätt, vilket är effektivt för oss och framför allt för våra kunder”.

Förstärkt översättning kan underlätta översättningsarbetet även på andra sätt. Det gäller exempelvis möjligheten att använda prediktiv analys för att mäta vilka texter som kan översättas med gott resultat, vilket bidrar till förbättringar både tidigare och senare i skapandet av innehåll. Det går också att genomföra omfattande bedömningar av sökmotoroptimering (SEO) på andra språk, vilket kan leda till betydligt bättre synlighet online.

Förstärkt översättning utgör det allra senaste inom maskinöversättning och kan leda till stora kvalitetsförbättringar. Genom att kombinera tekniker som översättning och kvalitetsberäkning (för att göra en statistisk uppskattning av hur hög kvalitet man kan förväntas uppnå med maskinöversättning) med de senaste adaptiva neurala maskinöversättningsverktygen, kan förstärkt översättning bli ett värdefullt vapen i varje lokaliseringsteams arsenal. Räkna med att det här blir ett begrepp vi får höra mer om de närmaste åren.



Så hjälpte Lionbridge Royal Caribbean med deras lokaliseringsprocesser

Den ledande kryssningsarrangören ville hitta sätt att locka en mer global publik

Ett kryssningsfartyg som seglar jorden runt är, per definition, internationellt. Varje år packar omkring 5,3 miljoner resenärer från jordens alla hörn sina resväskor för att ge sig av till olika exotiska resmål tillsammans med Royal Caribbean. För att se till att de både får den upplevelse de drömmer om och värna om integriteten hos Royal Caribbeans varumärke, krävs översättningar av hög kvalitet på företagets webbplats – royalcaribbean.com. Här kan besökarna inte bara boka sin semester, utan också utforska Royal Caribbeans fartygsflotta, aktiviteter och resmål.

Företagets problem

För ett ansett varumärke som Royal Caribbean kan ett översättningsarbetsflöde som leder till ett tungrott och ogenomtänkt innehåll få negativ inverkan på både innehållets korrekthet och förmågan att kommunicera. När företaget bestämde sig för att uppdatera sin innehållsplattform vände man sig till etablerade lokaliseringsexperten för

att få hjälp. Royal Caribbean valde att anlita Lionbridge, ett företag med mer än 25 års erfarenhet av språktjänster. Målet var att driva både trafik och försäljning med en centraliserad webbplats för innehåll, förbättra arbetsflöden för översättning och leverera en förstklassig användarupplevelse.

Den valda lösningen

Lösningen som valdes bestod av två verktyg: Lionbridges Connector™ och Adobe Experience Manager™. På så sätt fick man tillgång till både de effektiva arbetsflöden och den flerspråkiga expertis som behövdes för att se till att Royal Caribbeans innehåll inte bara uppdaterades på rätt sätt utan också optimerades för lokala sökningar. Materialen gjordes tillgängliga på önskade språk, inom godtagbar tid, till rimlig kostnad och med hög kompetens. På bara fyra månader

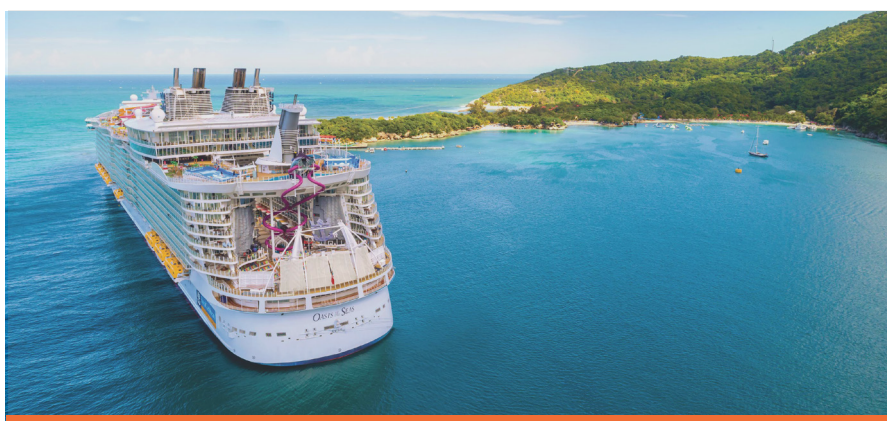
översattes över 80 procent av Royal Caribbeans befintliga innehåll – motsvarande ungefär 1 500 sidor – till sju språk.

”Genom att använda Lionbridges konnektor för Adobe Experience Manager för att automatisera webböversättningsprocessen har vi enkelt översatt tusentals webbsidor”, berättar Ariadna Castro, Product Content Optimization Manager på Royal Caribbean Cruise Lines. ”Tekniken som Lionbridge tillhandahåller har hjälpt oss att implementera ett effektivt arbetsflöde för översättningar. Teamet har varit mycket hjälpsamma. De är alltid redo att göra allt för att stödja våra mål.”

1 500
sidor, översatta till 7 språk
på 4 månader av Lionbridge

Lokaliseringsprojekt bör först och främst skapa en bättre användarupplevelse för kunderna. De måste dock även kunna omsättas i konkreta resultat som visar varför investeringen är så fördelaktig för företagen. I Royal Caribbeans fall var effekterna märkbara nästan omedelbart. De första åren efter införandet rapporterade företaget en 35-procentig ökning av antalet organiska klick, hela 37 procent fler organiska sökvisningar samt en imponerande tredubbling av antalet klick globalt.

Med en mer personlig upplevelse för kunder från hela världen kunde alla involverade skörda framgångarna av en väl genomförd och omvälvande lokalisering.



8

Zynga använder lokalisering för att vinna miljontals spelares kärlek

Vi försöker undvika begreppet lokalisering, i stället försöker vi använda 'kulturisering', säger Bernard Kim, Zyngas President of Publishing. "Vi förändrar alltså inte bara hur produkten läses sett till språket, utan försöker anpassa spelet dynamiskt till den lokala marknaden med fokus på spelupplevelse och engagerande funktioner.

President of Publishing Bernard Kim och Head of Strategic Partnerships Vineet Rajosi Sharma diskuterar utveckling av spel som kan locka en global publik

Spel är universella och intresset känner inga gränser. För att blomstra på dagens spelmarknad behöver stora spelutvecklare röra sig lika obehindrat på spelplanen. Zynga grundades 2007 och är i dag en av världens ledande utvecklare av spel för sociala medier som spelas av miljontals människor världen över, varje dag. En av företagets mest välkända titlar, FarmVille, lanserades på Facebook i juni 2009 och samlade tio miljoner dagligen aktiva användare på bara några veckor. Mer än tio år senare har Zynga mestadels gått över till mobila enheter, men fortsätter locka tiotals miljoner spelare i fler än 150 länder varje månad, med titlar som Zynga Poker, Words With Friends 2, Game of Thrones Slot Casino och Harry Potter Puzzles and Spells.

En av förklaringarna till att man har nått sådana framgångar med sina "evighetsspel" (som företaget kallar sin beständiga egendom) är det tydliga

fokuset på att anpassa sina produkter till marknader i olika delar av världen. De som har engelska som första- eller andraspråk utgör ju faktiskt mindre än en tolfedel av världens befolkning. Spelupplevelser som bara erbjuds på engelska går miste om större delen av den potentiella användarbasen. Därför är det sunt affärsförnuft att lansera spel på flera språk. Men, som Zyngas chefer förklarade för AI Business, handlar framgångsrik lokalisering om mycket mer än att bara översätta text i ett spel.

Kalla det "kulturisering"

"Vi försöker undvika begreppet lokalisering, i stället försöker vi använda 'kulturisering', säger Bernard Kim, Zyngas President of Publishing. "Vi förändrar alltså inte bara hur produkten läses sett till språket, utan försöker anpassa spelet dynamiskt till den lokala marknaden med fokus på spelupplevelse och engagerande funktioner. Till exempel kanske koreanska spelare engagerar sig fler gånger per dag och blir mer uppslukade av ett av våra spel. Vi försöker skapa erbjudanden och utmaningar för de här spelarna som är kulturiserade för den lokala marknaden."

Vineet Rajosi Sharma, Head of Strategic Partnerships på Zynga, berättar för AI Business att det var viktigt att spelarna känner att spelen har designats speciellt för dem. "Om jag är en användare och interagerar med en app ska det kännas som om den är gjord för mig", säger han. ➤



► Det är inte lätt att skapa ett spel som tilltalar miljontals användare i flera länder och samtidigt känns personligt. Bernard är veteran i spelbranschen och berättar att han har lång erfarenhet av att hantera problem med produkter som är svåra att översätta till nya regioner. En läxa som han har lärt sig är att lokaliseringen av titel till en viss marknad inte får kännas ”påklitrad”.

”[Spelarna vet om] ett spel är ett som de skulle kunna spela länge”, förklarar Bernard. ”Känslan, språket, hur saker och ting presenteras – och om detta görs i sista minuten för vissa marknader ger det [negativa resultat] när det gäller engagemangstalen.”

För att motverka det här lägger Zynga upp en strategi för lokalisering (eller kulturisering) redan i början av alla nya projekt. Allt handlar om att ”breda kanalerna” sett till antalet potentiella spelare.

”Om vi tar våra eviga franchisespel, så tilltalar de en global publik”, säger Vineet Rajosi Sharma. När de väljer vilka koncept som ska få grönt ljus ”får kulturiseringen högsta prioritet”, tillägger han. ”Det är viktigt att sätta sig in i alla aspekter av att lansera appen på marknaden.”

Språket – viktigt men inte den enda faktorn

Av förklarliga skäl är språk en viktig fråga för Zynga, som har studior utspridda över hela världen och kontor i USA, Kanada, Indien, Irland och Turkiet. Vineet Rajosi delade med sig av ett tips som kan användas för att förenkla lokaliseringsprocessen: minska textvolymen. ”Om du lägger in för mycket text i spelet gör di princip bara jobbet svårare”, förklarar han. ”Och ingen gillar ändå att behöva läsa så där mycket text.” Det gäller framför allt mobila enheter med mindre skärmar.

För att göra det hela ännu knepigare måste all text översättas på ett sätt som förmedlar originaltitelns ”röst”. Tillsammans med text och

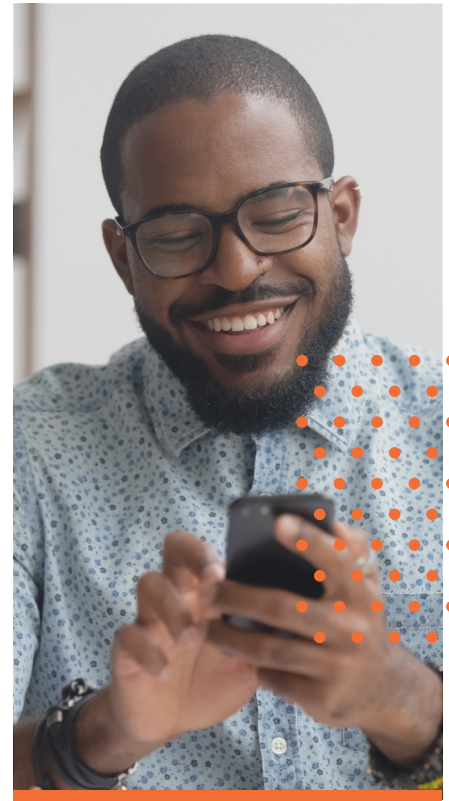
kommentarer i spelet kan det även vara inslag i fiktiva världar som valutor, måttenheter, datumformat med mera. Sedan har vi beskrivningar till appbutiker, marknadsföringsmaterial och annan tillhörande text som behöver översätta. ”Det finns en intern eller extern översättare”, förklarar Vineet Rajosi Sharma. ”Det vi verkligen inriktar oss på är [översättningskvaliteten]. Vi använder kvalitetssäkring för lokalisering, eller ”internationell QA” som vi kallar det, för att se till att översättningarna som kommer in är sammanhängande och stämmer.”

Tal och skrift är bara toppen av isberget i det större kulturiseringsarbetet. Det finns till exempel vissa symboler, sätt att hantera kön och titlar som man behöver tänka på. Detsamma gäller integrering med vanliga plattformar för sociala medier på olika platser, vare sig det gäller marknadsföring, autentisering eller delning med andra användare. Bernard säger att en av de detaljer som tydligt skiljer sig åt mellan olika marknader är ”formen på köp i app-paket” i spel, som ofta anpassas för att tilltala spelare i en viss region. Tidigare har företaget undvikit att lansera vissa spel – exempelvis sitt Harry Potter-spel på vissa asiatiska marknader – för att säkerställa en optimal upplevelse.

”Det är precis därför vi anlitar översättare. De specialiserar sig på våra spel och bidrar och arbetar med spelen under väldigt lång tid”, förklarar Bernard Kim.

Vikten av lokalisering i gratis spelvärlden

Som Z yngas chefer noterar är lokalisering viktigare än någonsin i dagens spelindustri med många gratis spel. Till skillnad från premiumspelen som säljs till ett fast pris tjänar ”freemium”-spelen pengar på att locka tillbaka spelarna gång på gång. Det innebär att det inte räcker att fokusera på ytlig lokalisering, exempelvis genom att köra en bra, lokal marknadsföringskampanj men göra minimala ändringar i själva spelet. Enligt Bernard leder freemium automatiskt



till hög kvalitet, för på marknader med många aktörer återvänder spelarna bara till miljöer där de känner sig hemma. Det betyder vanligtvis spel som utvecklas med lokalisering i åtanke. Det finns i dag runt 3,8 miljarder smartphoneanvändare i världen, enligt Statista – vilket motsvarar ca 48,33 % av jordens befolkning – så lokalisering är en smart investering för alla mobilspelestudier.

”Rent allmänt ser vi kulturisering som en kreativ resa”, säger Bernard. ”Det omfattar hela vägen från det att användaren först hör talas om vårt spel – genom en annons eller en lokal marknadsföringskampanj – till det att hen blir spelare och relationen hanteras med CRM (kundrelationshantering). Görs det via e-post eller meddelanden? Hur mycket är klockan? Du vill nämligen inte skicka push-meddelanden till en person som sover, eller skicka dem på ett olämpligt språk? Alla de här faktorerna ingår i kulturisering.”

Använda språktjänster för att förbättra Wikipedia

När maskininlärning tillämpas ... används modeller för att förbättra upplevelsen för läsare och redigerare”, säger Chris. ”Till exempel har vi modeller som analyserar, och försöker förutse, kvaliteten på en redigerad text för att ta reda på om vi anser att en ändring är skadlig eller inte.

Chris Albon, Director of Machine Learning på Wikimedia Foundation, diskuterar projektets målsättning och betydelsen av lokala data.

Världens största användargenererade encyklopedi ska vara tillgänglig på så många platser som möjligt. Kan lokaliseringstjänster hjälpa till att göra en värdefull resurs ännu bättre för världens användare? Det var en fråga som skaparna bakom Wikipedia ställde sig, en tjänst som i dag erbjuder fler än 40 miljoner artiklar på 300 språk och tillgodoser en publik på 500 miljoner användare varje månad. Svaret löd: självklart.

Det mest uppenbara sättet att se till att alla Wikipedia-versioner är så omfattande som möjligt vore att samla dem alla och automatiskt översätta till exempel engelska artiklar (eftersom den största Wiki-sidan är skriven på engelska) till andra språk med färre användare.

The Wikimedia Foundation har till viss del gjort just detta. Med hjälp av ett översättningsverktyg som stöds av stiftelsen produceras en rå maskinöversättning av Wikipedia-artiklar som sedan granskas manuellt och redigeras före publicering, om så behövs. Det här steget är viktigt, eftersom oöversatta maskinöversättningar kan ställa till problem. Ett anmärkningsvärt exempel är när orden ”village pump” på engelska översattes till ”bomba byn” på portugisiska. Genom att inkludera ett manuellt granskningssteg kan sådana fel minimeras. Hittills har det automatiska verktyget använts för att hjälpa till att översätta över en halv miljon artiklar.

Men som Chris Albon, Director of Machine Learning på Wikipedia, berättade för AI Business har organisation undvikit helautomatisk

maskinöversättning och i stället infört en metod ”som tar mycket större hänsyn till lokala aspekter”.

Lokaliserad maskininlärning

Wikipedia använder maskininlärningsmodeller i mycket större omfattning än många användare kanske tror. ”När maskininlärning tillämpas ... används modeller för att förbättra upplevelsen för läsare och redigerare”, säger Chris. ”Till exempel har vi modeller som analyserar, och försöker förutse, kvaliteten på en redigerad text för att ta reda på om vi anser att en ändring är skadlig eller inte.”

Wikipedia använder hundratals ”botar” med skiftande komplexitet som har tagits fram för att göra allt från att föreslå tänkbara artikelämnen till att förhindra skadegörelse, så att sidorna förblir tillgängliga. Men även om Chris och hans kolleger (som inte personligen skriver innehåll till Wikipedia) bara vill den kostnadsfria encyklopedins bästa vill de inte heller riskera att bädda in sina egna biaser i de AI-modeller de bygger.

Vad kan bias betyda när det gäller Wikipedia? Enkelt uttryckt kan inte en San Francisco-bos uppfattning om vad Wikipedia behöver, eller vad som utgör en bra Wikipedia-artikel, automatiskt tillämpas på en Wikipedia-användare i Kenya med helt andra kulturella förväntningar.

”Den Wikipedia-version som jag har läst klart mest är naturligtvis den på engelska förklarar Chris. ”Så min uppfattning om vad som skulle vara en bra artikel, eller vilket ämne en viss >



► artikel kan innehålla, eller hur den artikel ska vara uppbyggd, eller viken ton som bör användas i den artikeln, allt detta bygger på min erfarenhet av en enda community i Wikipedia-projektet. Och den kan inte automatiskt appliceras på alla andra communities. Vi delar alla ett vist samförstånd genom att vara en del av Wikipedia-gemenskapen, men när det handlar om vad som utgör ändringar av hög kvalitet, eller skadliga ändringar, eller hur man hanterar [ett visst ämne], så är det väldigt lokalt. Det är något som bör styras av den lokala communityn på det lokala språket.”

Lokalisering ingår i Wikipedia-uppdraget

Wikipedia driver för närvarande omkring 100 maskininlärningsmodeller som tränas helt och hållet med hjälp av lokala communitydata. De distribueras till omkring 30 Wikipedia-sidor.

”Anledningen till att lokalisering är så viktig för oss på stiftelsen är att det är en central del i hela projektet”, säger Chris Albon, och hänvisar till att målet är att alla människor i världen ska ha tillgång till en kostnadsfri encyklopedi på sitt modersmål.

Skälen till att varje modell tränas med lokala data är att användare i olika länder kan ha helt olika förväntningar på vad som utgör en bra Wikipedia-artikel.

”Det här blir viktigt om man betänker att vad swahili-communityn tycker är viktigt, eller oväsentligt, i artiklar, eller viktiga saker att länka till, eller utforskade ämnen, skiljer sig åt [från] till exempel engelska Wikipedia på grund av det lokala sammanhanget”, förklarar Chris Albon. ”På engelska Wikipedia anses exempelvis något som en matatu (en privatägd minibuss) inte vara [särskilt viktigt]. Men på Wikipedia-sidan på swahili ägnas ordet stor uppmärksamhet. Det är ett fenomen som är grundläggande för det dagliga livet för människor i östra, södra och västra Afrika ... Vi vill inte tvinga på dem [vår uppfattning]



genom att skapa en modell baserad på engelska Wikipedia och sedan tillämpa den på Wikipedia-sidan på swahili.”

En av maskininlärningsmodellerna som används av Wikipedia är konstruerad för att ge förslag på när ett ord i en artikel borde länka till en annan sida. I teorin kan merparten av alla substantiv på Wikipedia ha egna sidor, vilket riskerar att leda till att artiklar består av enbart hyperlänkar. Det kan få sidan att kännas överbelamad och dra uppmärksamhet från länkar som kan vara relevanta för en potentiell läsare. Därför har Wikipedia beslutat att man inte vill länka till sådant som anses vara ”allmänt känt”, det vill säga information som betraktas som så grundläggande att besökarna nästan garanterat känner till den. Men det fanns ett problem: Definitionen av allmänt känt information kan skilja sig åt i olika regioner. Därför är det viktigt att AI-modeller som används till den här uppgiften antar ett lokalt perspektiv på vad som betraktas som allmänt känt information.

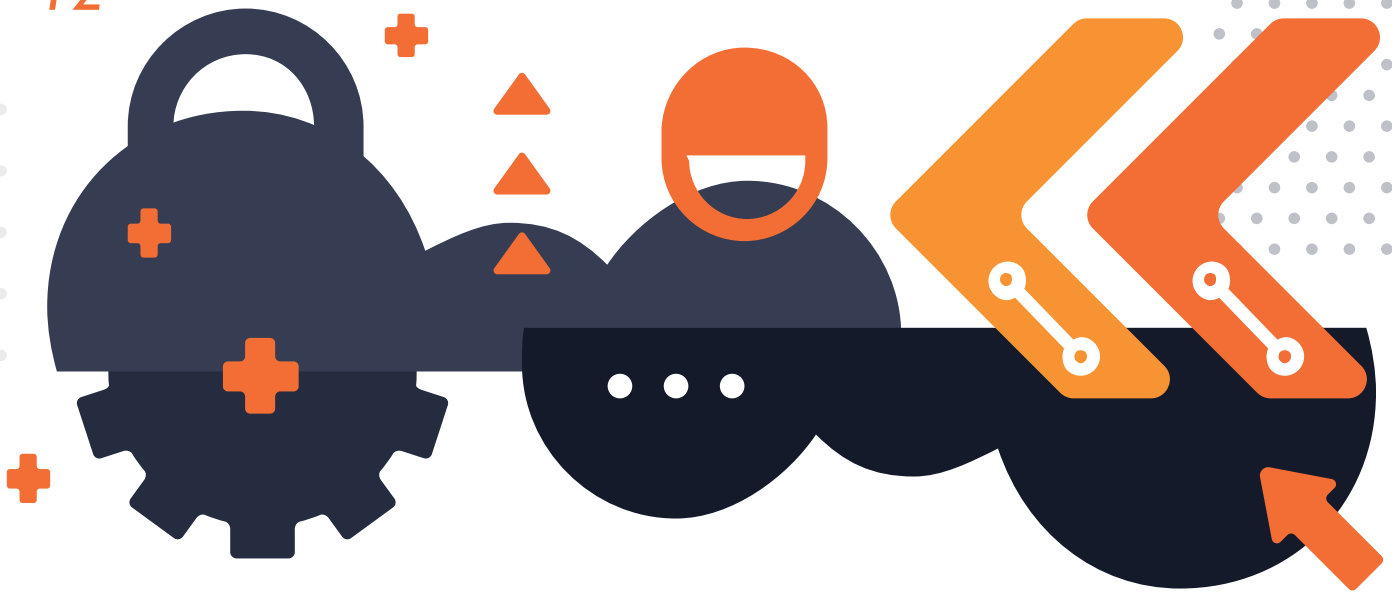
Den gyllene standarden

”Vår gyllene standard för modellproduktion är att när vi tränar en maskininlärningsmodell använder vi data från den community där modellen ska tillämpas”, förklarar Chris Albon. ”Om till exempel Wikipedia på swahili kommer och säger att de vill ha en modell för

att bedöma redigeringskvalitet, så ser vår gyllene standard till att vi samlar in träningsdata om hur en bra artikel kan se ut, från den communityn. Sedan tränar vi modellen och erbjuder den till dem.”

När kvaliteten på olika Wikipedia-sidor höjs blir de mer användbara för användarna, vilket i sin tur lockar till sig fler användare. De här nya användarna kan sedan ge feedback som gör de lokalt tränade maskininlärningsmodellerna smartare och ännu mer lokaliserade. Den här lokaliseringsmetoden kan beskrivas som en ”unik, positiv cirkel” som gör en värdefull onlineresurs ännu bättre.

”Det är därför det är så viktigt att engagera communityn i sådana här saker”, säger Chris. ”För om vi inte gör det blir det en grupp människor från en viss region med en viss uppfattning om hur Wikipedia ska se ut som pådyvlar den [på alla andra]. Man brukar prata om att bias kan tvättas bort med maskininlärningsmodeller. Det är precis vad vi inte vill göra. Vi vill att communityn ska få göra sin röst hörd om sina modeller, vara en del i arbetet med att ta fram modellerna och ingå i processen för att övervaka dem. När allt kommer omkring handlar det ju om att använda [AI]-modellerna för att skapa en bättre upplevelse för communityn.”



Brytpunkter

Diskussion om förhållandet mellan maskininlärning och språk med Rafa Moral och Kajetan Malionowski från Lionbridge

Översättning har länge varit en del av artificiell intelligens. Redan 1933, mer än två decennier innan AI fick status som egen disciplin, presenterade den ryske vetenskapsmannen Peter Troyanskii sin "maskin för val och utskrift av ord vid översättning mellan språk" för Vetenskapliga akademien i Sovjetunionen. På världsmässan i New York 1964 visade IBM upp ett maskinöversättningsverktyg som kunde leverera råöversättningar mellan engelska och ryska.

Det kan vara enkelt att betrakta maskinöversättning som ett av många områden som ryms under AI-paraplyet, men Rafa Moral, VP of innovation på Lionbridge, hävdar att det här är ett av teknikens centrala teman. "Själva grundtanken med AI är att imitera mänsklig intelligens och det finns inget mer mänskligt än språk", säger Rafa till AI Business. "Så om en AI-forskare försöker skapa något som betar sig som en människa innebär det rent begreppsmässigt att språk blir inblandat. Jag tror det är därför språk har varit på tapeten – ja kanske till och själva huvudfrågan – för AI ända från början."

Trimmad maskinöversättning

I likhet med många andra AI-inriktningar har maskinöversättning genomgått många förändringar genom åren. De har banat väg för de senaste årens extraordinära framsteg som, samtidigt som tekniken äntligen har mognat och gått från lovande F&U-projekt till fullfjädrade verktyg.

En banbrytande utveckling är ordinbäddning. Det här begreppet ingår i verktygslådan för maskin- och djupinlärning och avser en ordvisningsmetod där ord med likartad betydelse kan visas på liknande sätt. Vid ordinbäddning fångas ordens betydelse på ett sätt som kan förbättra översättningsprocessen avsevärt. Den skiljer sig också åt från tidigare metoder där olika ord visades på helt olika sätt, även om de betydde nästan samma sak.

"Innan inbäddning fanns var ord textsträngar", förklarar Rafa Moral. "När inbäddningskonceptet lanserades förde det också med sig vektoriseringskonceptet och hur vi kan omvandla ord till siffror där siffran eller vektorn kan ha semantisk betydelse. Inbäddning handlar om att

identifiera vilken siffra eller vektor du ska associera med ett visst ord, utifrån sammanhang."

Neural MT: en tydlig brytpunkt

Ett av de största genombrotten de senaste åren har varit införandet av neural maskinöversättning, något som Rafa kallar för en "tydlig brytpunkt" för hela branschen. Neural maskinöversättning (NMT) skiljer sig från tidigare använd, regelbaserad MT (RbMT) respektive statistisk MT (SMT). Det är en helautomatisk metod som förlitar sig på neurala nätverk, de verktyg som skapats med hjärnan som förebild och som utgör grunden för en stor del av den moderna maskininlärningen. NMT erbjuder bättre och mer korrekta översättningar genom att ta hänsyn till i vilket sammanhang ord används i stället för att analysera varje ord för sig. Den tittar inte bara på ord, utan också på relationen mellan dem.

Teknik som NMT har bidragit till att trimma de lokaliseringstjänster som erbjuds av företag som Lionbridge, som använder den som en del i sin arsenal för att tillhandahålla förstklassig översättning av dokument och andra texter. >

► ”Det finns mer innehåll än vi någonsin skulle kunna översätta, flera hundra gånger mer än alla översättare i världen klarar av”, säger Kajetan Malinowski, Product Leader och strateg på Lionbridge, till AI Business. ”Det betyder att maskinöversättning hjälper oss att revolutionera industrin och hantera mer innehåll och ger språkexperterna möjlighet att fokusera på områden där de kan göra skillnad.”

Precis som Rafa Moral anser Kajetan Malinowski att införandet av metoder som NMT har varit avgörande för automatiserad översättning. Men en viktig del i all förstklassig översättning är att veta vilken metod man ska välja – vissa passar bättre till vissa typer av innehåll än andra.

Välja rätt metod för optimala resultat

”Det finns olika typer av innehåll för olika ändamål”, säger Kajetan. ”En del innehåll är väldigt kortlivat och ses av ett fåtal. Då behöver inte översättningen hålla högsta kvalitet för att texten ska vara användbar. Sedan finns det annat innehåll, som reklamtext eller annonser, eller väldigt specifika dokument inom juridik, sjukvård eller regelverk, där översättningen måste hålla högsta klass. Det vi kan göra med hjälp av maskinöversättning är att hantera mer innehåll och att låta översättarna ägna sig åt uppgifter som tillför ett 'mervärde'.”

På Lionbridge använder företaget till exempel AI-baserade verktyg för att avgöra vilken metod som ska väljas för att få ett optimalt resultat – innan ett enda ord har översatts.

”Vi har infört en hybridprocess som tar det bästa av varje metod och sedan försöker fastställa samtliga av följande problem: språklig tyngd och ordförråd, komplexitet och variation samt kvalitet”, tillägger Kajetan Malinowski. ”Det här är vår utgångspunkt. När vi väl har definierat innehållet tydligt för kunden kan vi välja vilken process som passar bäst och sedan välja arbetsflöde för den aktuella typen av innehåll. Det handlar till



När vi väl har definierat innehållet tydligt för kunden kan vi välja vilken process som passar bäst och sedan välja arbetsflöde för den aktuella typen av innehåll. Det handlar till exempel om vilket översättningsminne som är bäst att använda, vilken maskinöversättningsmotor som passar till språket eller området och vilken verktygsuppsättning som bör användas till den aktuella innehållstypen.

exempel om vilket översättningsminne som är bäst att använda, vilken maskinöversättningsmotor som passar till språket eller området och vilken verktygsuppsättning som bör användas till den aktuella innehållstypen.”

Människan är kvar i leken

Som vi beskriver i artikeln om förstärkt översättning på [sidan 5](#) är det bästa alternativet ofta att kombinera maskinöversättning med mänsklig språkexpertis.

”[Maskinöversättningsmotorerna] har blivit så bra att de kan hantera alla former av innehåll med låg synlighet, där kvaliteten bara behöver vara bra nog”, säger Kajetan. ”Om den sedan behöver vara något bättre kan människan kliva in och snygga till det. På så sätt kan språkexperterna fokusera på viktigare innehåll som kräver mer omsorg. Texten behöver kanske vara mer lättläst, ha ett

visst flyt eller kunna väcka vissa känslor hos målgruppen. Maskinöversättningar fixar inte det – ännu.”

Den subjektiva tolkningen av språk, och ambivalensen kring vad som anses utgöra en bra översättning, är fortfarande en utmaning även för de allra bästa maskinöversättningsverktygen, men den moderna översättningen har dragit enorm fördel av framsteg på det här området. ”Mängden innehåll som människan genererar har ökat exponentiellt jämfört med för bara ett eller två år sedan, och blir bara större”, säger Kajetan Malinowski. ”Maskinöversättning och alla AI- och maskininlärningsverktyg som vi har tillgång till, hjälper oss att skapa mer innehåll, kommunicera med fler målgrupper och bli bättre på att hantera innehållet.”

Så hjälpte Lionbridge Volvo Cars att skapa en enhetlig onlinemarknadsplats

Vi ville gå ifrån den splittrade onlineupplevelsen och bygga ett enhetligt varumärke och en enhetlig plattform.

Företaget ansvarade tidigare för omkring 100 olika webbplatser

När Volvo Cars bestämde sig för att gå över till helt digital onlineförsäljning och elbilar senast 2030 inledde de ett samarbete med Lionbridge och Avanade, ett företag som levererar affärstjänster och molnlösningar, för att genomföra nödvändiga förändringar. Man ville bland annat tänka helt nytt när det gällde företagets onlinenärvaro och förvandla webbplatsen till en digital marknadsplats, snarare än enbart ett marknadsföringsverktyg. Det innebar att man behövde skapa en användarupplevelse som kunde informera och tilltala potentiella kunder över hela världen.

Företagets problem

Innan förändringsarbetet inleddes hade Volvo Cars en stark webbnärvaro, men den var extremt fragmenterad. Företaget ansvarade för omkring 100 webbplatser på sammanlagt 45 språk. Webbssidorna skapades av ett centralt team, men skickades sedan vidare till teamen på respektive marknad för att översättas. Det här systemet påverkade varumärkets helhetsintryck, eftersom olika marknader valde att utforma sina webbplatser och försäljningsstrategier på helt olika sätt.

Genom att centralisera webbplatsen hoppades Volvo Cars uppnå styrning av lokaliseringsarbetet från ledningen, med central överblick över översättningar och direkt kontakt med dem som utför det här otroligt viktiga arbetet. "När vi tar steget över till onlineförsäljning behöver vi ta ett förnyat grepp om vårt budskap",

förklarar Cecilia Ernby, produktionsansvarig för Global Online Digital på Volvo. "Vi ville röra oss bort från det splittrade utseendet och vår fragmenterade varumärkeshistoria och i stället skapa ett varumärke och en plattform som redigerades och översattes centralt."

Den valda lösningen

För att optimera översättningsprocessen använde Volvo Cars Smartling Translation Management System och kunde på så sätt minska den tid som ägnades åt att hantera översättningsuppdrag med runt 1 000 timmar. En del av detta handlade om att centralisera användningen av översättningsminnen som tidigare var unika för varje marknad. Företaget gjorde också om arbetsflödet för översättning och lade till ett steg för att se till att innehållet blir ordentligt optimerat för att uppfylla lokala krav på sökmotoroptimering (SEO).

Centralisering och lokalisering: En kraft att räkna med

Till följd av förändringarna kommer Volvo Cars webbplats nu att se likadan ut på samtliga marknader. Det betyder att kunder som tittar på produktsidor i Storbritannien, Tyskland eller Belgien alla tar del av samma användarupplevelse, fast numera med översatt innehåll av hög kvalitet. På så sätt blir det enklare för Volvo Cars att lansera kampanjer och produkter i flera regioner.

"Vi kan nu lansera massor av kampanjer på en gång, det hade helt enkelt inte varit möjligt med våra tidigare lösningar", säger Cecilia Ernby.

Den här fallstudien är ett talande exempel på hur lokalisering och centralisering tillsammans kan skapa en sammanhängande användarupplevelse, utan att försaka några av de fördelar som lokala kunder efterfrågar.

1 000

timmars arbete undveks med nya system

Språkteknikens framtid

En närmare titt på möjligheterna med kroppsnära enheter, förstärkt verklighet och kolossala språkmodeller

Det finns en klassisk, eventuellt svårbegriplig, historia om maskinöversättningens begränsningar. Det sägs att ett tidigt MT-verktyg fick i uppdrag att översätta den engelska meningen: "The spirit is willing, but the flesh is weak" till ryska. Resultatet blev något i stil med "Whiskeyn är stark, men köttet är ruttet".

Maskinöversättning och språkteknik har kommit en lång väg sedan industrins vagna, men språktekniken har fortfarande en bra bit kvar innan den kan bli den centrala del i våra liv som förespråkarna alltid har trott är möjligt.

Några av de mest imponerande uppvisningarna i nästa generations språkteknik har sin grund i hårdvara. De här innovationerna, som ofta involverar den snabbväxande kategorin med kroppsnära teknik, bär med sig ett löfte om att översättning ska bli sömlös och otvetydig på ett sätt som gör att slutanvändare inte ens behöver tänka på den. Vi närmar oss då tanken på den "universella översättare" som länge har förekommit som hypotes inom science fiction och ser till att man omedelbart förstår alla världens språk.

Vägen till en universell översättare

Ett praktiskt exempel på den här tekniken är hörlurar med översättningsfunktion, som kombinerar ordbehandling med maskinöversättning för att tillhandahålla realtidsöversättning av det som sägs. Till de mer intressanta exemplen hör Pilot Smart Ear Buds (som även transkriberar samtal), Google Pixel Buds och WT2 Plus AI Real-time Translator Ear Buds.

I den visuella änden av spektrumet återfinns Google Translates funktion för förstärkt verklighet (AR). Den kallas

Google Lens och kan användas för att rikta en smartphonekamera mot en skylt, meny eller annan text på ett främmande språk och se en översättning ovanpå bilden på mobilens display – ofta med samma teckensnitt som originaltexten. Allt detta utförs med hjälp av integrerade maskininlärningsverktyg, vilket innebär att Lens fungerar även utan internetanslutning. I januari 2021 avslöjade företaget att den fristående Google Lens-appen hämtats fler än 500 miljoner gånger i appbutiken Google Play. Google Lens stöder de flesta Google Translate-språk, för närvarande fler än 100 stycken.

Google Lens kan tyckas avancerat, men utgör ändå första generationens AR-baserade översättningsverktyg. Lens-appen lanserades i slutet av 2017 och ett mer utbrett stöd kom först senare. Men situationen kan förändras dramatiskt genom intaget

av populära head-up-displayer (HUD), exempelvis smarta glasögon. De här bärbara enheterna kommer att frigöra AR från de begränsningar som mobiler och surfplattor innebär och göra det betydligt smidigare och vanligare att förstärka användarens omgivning. Sett ur ett lokaliseringsspektiv innebär det att översättningar kan visas på mycket enklare sätt.

Framsteg under huven

Demonstrationer för allmänheten fungerar visserligen som imponerande påminnelser om framstegen inom språktekniken. Men för dem som efterfrågar översättningar av hög kvalitet återfinns de verkliga innovationerna under huven, väl inbakade i programvara, snarare än i smarthörlurar eller smartglasögon.

Som vi förklarar i avsnittet "Vad är förstärkt översättning?" på [sidan 5](#) >



16

► kommer översättningstekniken i den närmaste framtiden nästan garanterat inbegripa en hybrid, förstärkt metod som omfattar både AI och språkexperter. Domenico Lombardini, CEO för Absolute Scientific Translation (ASTV), en språktjänstleverantör specialiserad på immateriell egendom, juridik och life sciences, ser den vanligt förekommande konceptualiseringen av människa mot maskin som ett stort misstag.

”Ny teknik kan visserligen i viss mån minska eller undanröja behovet av manuella ingrepp, men det är också sant att de skapar nya professionella roller och nya arbetssätt”, säger Domenico till AI Business. ”Allt fler översättare går över till att efterredigera maskinöversättning, men deras bidrag spelar fortfarande en avgörande roll eftersom maskinen fortfarande kan göra fel. Träning utförd av lingvister är därför en nödvändig förutsättning för översättningarnas slutliga kvalitet.”

Människan har fortfarande sista ordet när det gäller översättningarnas kvalitet, men AI-verktyg kan ändå vara otroligt användbara vid uppgifter där man använder förstärkt översättning – till exempel för att skapa rekommendationssystem som föreslår den bästa översättaren till

ett visst projekt, baserat på specifik kompetensuppsättning. Snarare än att lämna över hela översättningsuppdraget till en algoritm kan medarbetare använda AI för att analysera dokumentets innehåll och särskilja nyckelord och stilistiska krav. De kan också märka och kategorisera texten innan de skapar korsreferenser till informationen.

Precis som nästa generations grammatik- och stavningskontroller kommer maskinerna att kunna erbjuda allt bättre stöd till språkexperter som arbetar med översättningar. Det kan handla om något så enkelt som att rätta till ett stavfel – men de kan också flagga för stilistiska avvikelser eller variationer och SEO-argument eller varna dem för kulturellt känsliga frågor som de (och företaget som har anlitat dem) kan vilja att de är medvetna om.

Kolossala språkmodeller

Sedan har vi användningen av AI-språkmodeller för andra uppgifter som kan underlätta kraven på lokalisering. Ett av de mest betydande AI-framstegen under 2020 var skapandet av GPT-3, en enorm språkmodell skapad av OpenAI. GPT-3:s neurala nätverk rymmer hela 175 miljoner parametrar, samtidigt som själva modellen har tränats med 45 terabyte data – vilket motsvarar en avsevärd del av allt tillgängligt innehåll

på internet. Genom att fyllas med så mycket kan data kan GPT-3 utföra jobb med otroligt hög precision när det handlar om att förutse vilka ord som rent statistiskt bör följa på andra – exempelvis att veta att ”apelsin” ofta följs av ”juice”.

Tack vare sitt enorma förråd av träningsdata kan GPT-3 användas till ett stort antal tillämpningsområden, inklusive översättningsuppgifter som kan växla språk mellan olika ”röster” för att förvandla juridisk text till vanlig svenska, eller tvärtom.

Verktyg som GPT-3 öppnar upp helt nya möjligheter för språktekniken, inte bara för att översätta befintlig text utan också för att generera helt nytt innehåll. ”AI kan hjälpa oss att göra mycket mer med språk än enbart översättning”, säger Kajetan Malinowski, Product Leader och strateg på Lionbridge. ”Den kan även hjälpa oss att författa innehåll. Vi har testat den aktuella GPT-3-modellen, som möjliggör flera användningsfall och där innehållsgenerering är en av de mest intressanta.”

I en värld där användare förväntar sig mer sofistikerade och personliga budskap kan innovationer som stora språkmodeller användas för att lokalisera text inte bara för enskilda länder eller regioner utan potentiellt även för individuella användare.



Språkteknikens framtid

Hur nära är egentligen maskiner att kunna imitera den mänskliga språkförmågan?

I en intressant forskningsrapport från 2020, "A Set of Recommendations for Assessing Human-Machine Parity in Language Translation, ([webblänk](#))" hävdade författarna att "kvaliteten på maskinöversättning har förbättrats markant de senaste åren, till den grad att den numera inte kan skiljas från översättningar utförda av språkexperter i flera empiriska studier".

Två av författarna, Samuel Läubli från universitetet i Zürich och Sheila Castilho på Dublin City-universitetet, tog ett samtal med AI Business om kraften och potentialen hos maskinöversättning.

AIB: I rapporten påpekar ni att maskinöversättning har gjort betydande framsteg de senaste åren. Vad beror de här framstegen främst på?

Samuel Läubli (SL): Jag ser tre huvudsakliga faktorer: en demokratisering av resurser, ökad datorkraft och växande kommersiellt intresse. Projekt som Paracrawl har gjort stora mängder träningsdata tillgängliga för i princip vem som helst som är intresserad av att skapa maskinöversättningssystem. Eftersom neurala maskinöversättningsmetoder rent begreppsmässigt är enklare än tidigare metoder, till exempel statistiska metoder, och flera implementeringar är tillgängliga utan kostnad, är det enklare än tidigare att bygga ett maskinöversättningssystem för akademiskt, eller till och med kommersiellt, bruk.

Arbetet med att träna neurala modeller med stora volymer träningsdata kräver massor av datorkraft, men eftersom man numera får betydligt mer kraft för samma pris jämfört med för bara några år sedan, är det här inte längre något hinder för andra organisationer eller individer än stora aktörer som Google, Microsoft eller Amazon. Sedan har vi det kommersiella intresset. Även om

man redan har kunnat visa att det går att spara massor av tid – och därmed pengar – i översättningsprocesser med föregångstekniken, har neural modellering för första gången möjliggjort skapandet av helt flytande meningar, vilket gav en "wow-effekt" hos allmänheten, och även bland beslutsfattare.

AIB: Vilket är det bästa måttet vid bedömning av översättningsprecision?

Sheila Castilho (SC): Den frågan är nästan omöjlig att besvara. Det finns tusentals arbeten om översättningsbedömning och många författare är inte bara oense om valet av metod utan också om vad översättningskvalitet egentligen är. För mig är det bästa sättet att mäta översättningskvalitet att först definiera vad kvalitet är för det aktuella översättningsdokumentet. Sedan utformar man metodiken därefter.

SL: Jag håller med Sheila. Om syftet med en översättning är att användare ska lösa ett problem med sin bärbara dator är det kanske inte meningsfullt att låta språkexperter hitta och kategorisera språkliga fel i översättningen. Men jag skulle vilja tillägga en sak: Oavsett vad du gör, inkludera ett manuellt steg någonstans. Om förbättringar i översättningskvalitet har visat något så är det att metoder för att automatiskt bedöma översättningskvalitet, till exempel **BLEU** ([webblänk](#)), blir allt mindre tillförlitliga.

AIB: Vilka är de mest spännande och lovande översättningstekniker som skymtar vid horisonten?

SC: Jag tycker att allt som blandar olika metoder är intressant, till exempel interaktiv efterredigering och tal.

SL: Förutom översättning med blandade metoder ser jag två viktiga förändringar som är på gång. För det första kan maskinöversättningssystem snart ta större hänsyn till sammanhang, dvs. översätta hela dokument i stället för att klippa isär dem i meningar och översätta varje mening isolerat, för att sedan [klistra] ihop dem igen. Det problemet är i princip redan löst ([webblänk](#)) i forskningsmiljö och kommer, när det väl släpps på marknaden, att leda till ökad samstämmighet.

För det andra kommer vi att få ökad kontroll över vad maskinöversättningssystem producerar. Maskiner producerar redan översättningar med imponerande kvalitet, men beslut om stil, till exempel, kan användarna inte påverka. I framtiden kommer du inte bara att kunna styra om du vill ha en formell snarare än informell översättning ([webblänk](#)) av texten, utan även många andra funktioner som ordriktedom och könsanpassning.

Den här intervjun har förtydligats och förkortats.





AI Business
e-bokserie



LIONBRIDGE

Lionbridge samarbetar med olika varumärken för att riva barriärer och bygga broar över hela världen. I 25 år har vi hjälpt företag att komma i kontakt med globala kunder och medarbetare genom att leverera översättnings- och lokaliseringslösningar på mer än 350 språk. Via vår plattform hanterar vi ett nätverk med passionerade experter över hela världen som samarbetar med olika varumärken för att skapa kulturellt anpassade upplevelser. Eftersom vi har en obehaglig kärlek till språk använder vi det bästa av mänsklig kompetens och maskinintelligens för att kommunicera budskap som är anpassade till våra kunders kunder. Lionbridge har sitt huvudkontor i Waltham, Massachusetts i USA och lösningscenter i 23 länder.

Läs mer på www.lionbridge.com



Läs mer här

