



LUNDS UNIVERSITET

Ekonomihögskolan

Institutionen för informatik

Maskinöversättningens risker

En studie av hur översättningsbyråer använder maskinöversättning i översättningsprocesserna och hur de hanterar risker som kan uppstå i samband med detta

Kandidatuppsats 15 hp, kurs SYSK16 i Informatik

Författare: Simon Flyman

Handledare: **Björn Svensson**

Rättande lärare: Paul Pierce
Magnus Wärja

Maskinöversättningens risker: en studie av hur översättningsbyråer använder maskinöversättning i översättningsprocesserna och hur de hanterar risker som kan uppstå i samband med detta

ENGELSK TITEL: The risks of machine translation: a study of how translation agencies use machine translation in their translation processes and how they manage risks that may occur in connection with this

FÖRFATTARE: Simon Flyman

UTGIVARE: Institutionen för informatik, Ekonomihögskolan, Lunds universitet

EXAMINATOR: Christina Keller, Professor

FRAMLAGD: maj, 2021

DOKUMENTTYP: Kandidatuppsats

ANTAL SIDOR: 61

NYCKELORD: maskinöversättning, risker, riskhantering, översättningsbyråer

SAMMANFATTNING (MAX. 200 ORD):

Denna uppsats studerar hur översättningsbyråer använder maskinöversättning i översättningsprocesserna och hur de hanterar risker som kan uppstå i samband med detta. Studien har utförts i form av intervjuer med nyckelpersoner på ett antal större översättningsbyråer som är aktiva på den skandinaviska marknaden. Av resultatet framgår att översättningsbyråerna i studien främst betraktar maskinöversättning som ett IT-stöd för mänskliga översättare och att det ofta är kunden som avgör om maskinöversättning ska användas i översättningsprocessen. Byråerna hanterar datasäkerhetsrisker med hjälp av kravspecifikationer i upphandlingsprocessen, slutna interna översättningssystem och förbud mot att använda kostnadsfria maskinöversättningstjänster. För att undvika felöversättningar och förbättra kvaliteten på den slutliga versionen av översättningen genomgår de flesta maskinöversättningar en efterredigeringsprocess. Det finns inga belägg i studien för att översättaryrket nedvärderats genom maskinöversättningens införande utan snarare uppvisar översättningsbyråerna en tendens att använda samma kriterier när de tilldelar efterredigeringsuppdrag som när de tilldelar vanliga översättningsuppdrag.

Innehåll

1	Introduktion.....	1
1.1	Bakgrund	1
1.1.1	Maskinöversättning och datorstödd översättning.....	2
1.2	Problemområde.....	3
1.3	Forskningsfråga	4
1.4	Syfte.....	4
2	Litteraturgenomgång	5
2.1	Översättning och automatisering	5
2.2	Risker kopplade till maskinöversättning	6
2.2.1	Felöversättningar	7
2.2.2	Datasäkerhet	8
2.2.3	Säkring av professionell kompetens.....	8
2.3	Riskhantering	9
3	Metod	12
3.1	Val av metod.....	12
3.2	Urval	13
3.3	Intervjuernas genomförande	14
3.4	Bearbetning och analys av data	14
3.5	Etiska aspekter	15
3.6	Validitet och reliabilitet	16
4	Empiriska resultat	17
4.1	Intervjupersoner.....	17
4.2	Implementering och användning av maskinöversättning	17
4.3	Riskanalys och riskhantering.....	19
4.4	Felöversättningar och översättningskvalitet	19
4.5	Datasäkerhet	20
4.6	Efterredigering och översättares attityd till maskinöversättning.....	21
5	Diskussion.....	23
5.1	Automatisering av översättningsprocessen	23
5.2	Riskhantering vid implementering av maskinöversättning	24
5.3	Felöversättningar och översättningskvalitet	24
5.4	Maskinöversättning och datasäkerhet.....	25
5.5	Maskinöversättning och vikten av professionell kompetens.....	26
6	Slutsats	27
7	Bilagor.....	29

7.1	Bilaga 1 – intervjuguide	29
7.2	Bilaga 2 – intervju 1	31
7.3	Bilaga 3 – intervju 2	40
7.4	Bilaga 3 – intervju 3	49
	Referenser.....	58

Tabeller

Tabell 4.1: Intervjupersonerna i studien..... 17

1 Introduktion

1.1 Bakgrund

Översättningsbranschen är en bransch som normalt inte får särskilt mycket utrymme i media eller andra officiella sammanhang, men det är trots detta en bransch av betydande storlek. Externa språktjänster, dvs. språktjänster som utförs på entreprenad för andra företag, beräknades 2019 omsätta nästan 50 miljarder dollar globalt (CSA Research, 2019). Denna siffra omfattar förvisso inte bara skriftlig översättning utan även näraliggande tjänster som tolkning och programvarulokalisering, men översättning av texter utgör det särklassigt största segmentet inom denna industri. Många stora företag och organisationer har dessutom egna översättningsavdelningar eller egna översättare anställda så siffran ovan täcker inte hela språktjänstesektorns storlek (Esselink, 2003, Schäffner etc., 2014).

Även om vissa översättare arbetar direkt mot en slutkund eller är anställda av en större organisation så utgörs den största delen av översättningsmarknaden av översättningsbyråer (Kelly, 2012). Översättningsbyråer fungerar lite som spindeln i nätet genom hela översättningsprocessen, som omfattar förloppet från det att en kund beställer en översättning till att slutprodukten i form av en färdig översättning levereras (Hofmann, 2012). Översättningsbyråerna hjälper slutkunden att förbereda och anpassa texterna som ska översättas och allokerar lämpliga översättare till det aktuella uppdraget. En översättningsbyrå kan i detta fall använda sig av både anställda översättare, s.k. inhouse-översättare, och en ofta större pool av frilansöversättare som är knutna till byrån på uppdragsbasis (Moorkens, 2017).

Det har skett stora framsteg på språkteknikens område under senare år och utvecklingen av så kallade neurala maskinöversättningssystem framstår alltmer som ett paradigmskifte med långtgående konsekvenser för hela översättningsbranschen (Koehn, 2020). Översättarens yrkesroll sägs vara på väg att förändras till en redaktörliknande roll där maskinöversatta texter korrigeras, bearbetas och anpassas till den aktuella kontexten genom så kallad efterredigering (Olohan, 2021). Denna utveckling innebär en rad fördelar såsom ökad produktivitet och snabbare leveranser, men ger även upphov till nya typer av risker och problem (Canfora och Ottmann, 2020).

Att maskinöversättningssystem inte alltid genererar ett önskat resultat var inte minst tydligt i samband med lanseringen av den svenska versionen av Amazon hösten 2020. Amazon använde sig av maskinöversättning för att översätta namn och beskrivningar för en mängd produkter som utbjuds till försäljning på Amazons webbplats och resultatet visade tydligt att tekniken fortfarande är långt ifrån färdigutvecklad. Vid översättning av begrepp och textfragment utan sammanhang fungerar maskinöversättning fortfarande ofta ganska dåligt och i Amazons fall innebar det att en mängd komiska, för att inte säga stötande, felöversättningar förekom på webbplatsen. Till exempel blev ett pussel med rapsmotiv (*rape*) istället ett pussel med våldtäktsmotiv. (Ewald, 2020)

1.1.1 Maskinöversättning och datorstödd översättning

Som Poibeau (2017) redogjort för i sitt översiktsverk sträcker sig maskinöversättningens historia åtminstone 75 år tillbaka i tiden. Till en början betraktades automatiserad översättning som ett slags avkodningsproblem där det gällde att dekryptera en text på ett främmande språk till det egna språket. Poibeau skildrar hur olika forskargrupper utförde en rad mer eller mindre lyckade försök att utveckla så kallade *regelbaserade maskinöversättningssystem* under en femtioårsperiod från 1940-talet och framåt. Dessa system var oftast uppbyggda med ett domänanpassat lexikon i botten och ovanpå detta tillämpades en mängd grammatiska och syntaktiska regler vid själva översättningen. Poibeau beskriver hur dessa system utvecklades i en ständig anpassnings- och förfiningsprocess som i slutändan kunde ge upphov till relativt goda översättningar inom väl avgränsade områden som t.ex. väderprognoser, medan de fungerade betydligt sämre som översättningssystem för allmän användning (Poibeau, 2017).

På 1980- och 1990-talet utvecklades en ny forskningsinriktning inom maskinöversättning som inte längre byggde på lingvistiska grunder och regelbaserade system. Denna riktning, som kallades *statistisk maskinöversättning*, drog istället nytta av den tillgängliga ökade datorkraften och tillämpade statistiska metoder för att hitta återkommande ordkombinationer i stora, tvåspråkiga textsamlingar och kunde på så sätt återanvända tidigare översättningsfragment för att generera översättningar av nya källtexter (Koehn, 2009). Under 2000-talet började statistisk maskinöversättning uppvisa lovande resultat och fick ett första kommersiellt genomslag i de tidiga versionerna av Google Translate och Microsoft Bing Translator (Fong, 2012).

Under de senaste tio åren har ännu en riktning inom maskinöversättningen fått sitt genombrott och växt fram som den dominerande tekniken på marknaden. Denna nya teknik kallas *neural maskinöversättning* och är baserad på artificiell intelligens och avancerade djupinlärningsmetoder (Koehn, 2020). Inspirationen till modellen bakom neural maskinöversättning kommer från biologin och är tänkt att efterlikna den mänskliga hjärnans komplexa nätverk av neuroner som möjliggör parallella processer i massiv skala (Koehn, 2020). Neurala maskinöversättningssystem består av artificiella neuronnät som tränas upp med hjälp av en stor mängd tvåspråkiga texter och på basis av detta genererar en kraftfull översättningsmotor som översätter texter mellan olika språk (Focada, 2017).

Parallellt med utvecklingen av maskinöversättning har även en rad andra IT-baserade översättningsverktyg utvecklats och integrerats i översättningsprocessen under de senaste decennierna. Som Somers (2003) redogör för, började digitala *översättningsminnen* utvecklas i början av 1980-talet och har använts av översättare i bred utsträckning sedan mitten av 1990-talet. Ett översättningsminne bygger på att tidigare översatta tvåspråkiga segment, motsvarande en sats eller en mening, lagras i en databas. När översättaren ska översätta ett nytt segment göms databasen för att hitta identiska eller likartade källsegment med tillhörande översättning, så kallade ”träffar” (*matches*), som sedan visas för översättaren. Därefter kan översättaren återanvända och eventuellt bearbeta den tidigare översättningen istället för att översätta källtexten från grunden. (Somers, 2003)

Under de senaste decennierna har även datorbaserade *terminologiverktyg* utvecklats och kommit till användning i översättningsprocessen. Dessa verktyg används för att bygga upp digitala termbaser, dvs. specialiserade lexikon med begrepp som används för vissa översättningsområden eller kunder. Med hjälp av termbaserna möjliggörs sedan en konsekvent begreppsanvändning i texterna som översätts. (Bowker, 2003)

Tillsammans med en rad andra stödfunktioner har översättningsminnen och terminologiverktyg på senare år integrerats i så kallade *översättningsverktyg* eller CAT-verktyg (Bowker och Fisher, 2010). Dessa programvaror används av översättaren för att underlätta översättningsarbetet. Även maskinöversättning kan inkluderas i översättningsverktyg genom bland annat insticksprogram. Bowker och Fisher (2010) beskriver översättningsprocesser där översättningsverktyg används som *datorstödd översättning*. Det är viktigt att poängtera att detta är någonting annat än maskinöversättning. Datorstödd översättning innebär alltså att en mänsklig översättare utför översättningen med hjälp av IT-stöd som översättningsminnen och terminologiverktyg, medan maskinöversättning innebär att översättningen utförs helt automatiskt av en dator (Bowker och Fisher, 2010).

Presas et al. (2016) skriver att det finns i princip tre olika former av maskinöversättningslösningar tillgängliga idag. För det första finns det kostnadsfria maskinöversättningstjänster som ofta är tillgängliga online, t.ex. Google Translate och DeepL. I dessa fall handlar det om allmänna maskinöversättningsmotorer som ofta saknar garantier gällande konfidentialiteten för den inmatade informationen. För det andra finns det även en rad olika kommersiella maskinöversättningssystem på marknaden som är anpassade till översättningsbyråers verksamhet. Dessa är regel ganska dyra, men erbjuder å andra sidan relativt hög kvalitet och bättre säkerhet. Det finns även en tredje variant av system, t.ex. Moses, som bygger på open access, men som måste anpassas och utvecklas för att kunna användas i översättningsprocessen, vilket kräver ganska hög kompetens och stora tidsinvesteringar. (Presas et al., 2016)

1.2 Problemområde

Som konstaterats i exemplet med den svenska lanseringen av Amazon ovan ger också moderna maskinöversättningssystem i vissa fall upphov till undermåliga översättningar. Även i situationer där mänskliga översättare är involverade i översättningsförloppet kan det uppstå problem. Det finns en rad risker kopplade till maskinöversättning som uppmärksammats i forskningslitteraturen och som är viktiga att ta hänsyn till när maskinöversättningsteknik implementeras i översättningsprocessen (Nitzke et al., 2019, Canfora och Ottmann, 2020). Till exempel har det rapporterats om flera fall av läckage av känslig information i samband med att personal på företag eller frilansöversättare har utnyttjat kostnadsfria översättningstjänster som Google Translate (Bell och CSO, 2019). Även andra risker och problem, såsom felöversättningar där maskinöversättning är involverad i processen och oklarheter kopplade till ansvarsfrågan, har lyfts fram i forskningslitteraturen i samband med maskinöversättning (Canfora och Ottmann, 2020).

Maskinöversättning kan användas på ett flertal olika sätt i översättningsprocessen. Kastberg (2012) beskriver tre olika tänkbara scenarier. För det första kan maskinöversättning användas för helautomatisk översättning helt utan mänsklig inblandning. Ett andra tänkbart scenario handlar om att maskinöversättning intar den centrala rollen i översättningsprocessen, men att maskinöversättningen tillämpas med mänskligt stöd, t.ex. genom att en översättare förbättrar kvaliteten på maskinöversättningsresultatet. Ett tredje scenario som Kastberg beskriver är maskinstödd mänsklig översättning, där det är den mänskliga översättaren som är central i översättningsprocessen och där maskinöversättningen fungerar som ett av flera IT-stöd för översättaren. (Kastberg, 2012)

Beroende på vilken roll som maskinöversättningen spelar i översättningsprocessen så kan det uppstå olika typer av risker som måste hanteras. Om dessa risker inte hanteras på lämpligt sätt

kan det få svåra följder för översättningsbyråerna och deras kunder. Felöversättningar kan leda till katastrofala skadeverkningar, i synnerhet vid översättningar av säkerhetskritiska texter (Canfora och Ottmann, 2018). Om känslig information läcker ut vid användning av maskinöversättning i översättningsprocessen kan det vålla stora skador för kunden, vilket kan leda till skadeståndsanspråk och andra bekymmer för översättningsbyrån. Det är därför viktigt för översättningsbyråerna att identifiera potentiella risker som kan uppstå genom olika former av maskinöversättningslösningar med hjälp av en lämplig form av riskanalys innan ett maskinöversättningssystem införs i en översättningsorganisation (Nitzke et al., 2019).

Även om forskningen på senare tid har börjat lyfta frågor kring risker kopplade till maskinöversättning så råder det fortfarande brist på empiriska studier på området (jfr. Nitzke et al., 2019, Canfora och Ottmann, 2020). I synnerhet gäller detta studier som empiriskt studerar maskinöversättningens implementering och användning inifrån översättningsorganisationerna. En förståelse utifrån översättningsbyråernas perspektiv är av största vikt eftersom det är här besluten fattas om hur maskinöversättning ska användas i översättningsprocessen och hur risker kopplade till denna användning ska hanteras.

1.3 Forskningsfråga

Avsikten med den föreliggande uppsatsen är att studera hur översättningsbyråer hanterar risker som uppstår i samband med användning av maskinöversättning i översättningsprocessen. Forskningsfrågan går att bryta ned i tre delfrågor kopplade till maskinöversättning och risker:

Hur använder översättningsbyråer maskinöversättning i översättningsprocessen?

Vilken form av riskanalys gör översättningsbyråer i samband med införandet av maskinöversättning?

Hur hanterar översättningsbyråer risker som uppstår i samband med användningen av maskinöversättning?

1.4 Syfte

Uppsatsens huvudsakliga syfte är att beskriva hur översättningsbyråer går tillväga för att hantera de risker som maskinöversättning kan ge upphov till. Detta är i sin tur kopplat till vilken roll som maskinöversättningen spelar i översättningsprocessen. Det innebär att uppsatsen för det första avser att studera hur maskinöversättning används av översättningsbyråer. Därefter undersöks vilken typ av riskanalys översättningsbyråer tillämpar i samband med att maskinöversättningssystem införs i verksamheten. Slutligen studeras den huvudsakliga frågeställningen, dvs. hur risker kopplade till maskinöversättning hanteras av översättningsbyråer.

2 Litteraturgenomgång

I detta kapitel görs en genomgång av forskningslitteratur som är relevant för uppsatsens problemområde och frågeställning. I det inledande avsnittet beskrivs och diskuteras olika sätt att tillämpa maskinöversättning i översättningsprocessen. Därefter följer ett avsnitt som tematiserar ett antal olika risker kopplade till maskinöversättning som diskuterats i forskningslitteraturen under de senaste åren. Slutligen presenteras en modell som kan användas för att förstå hur IT-relaterade riskhanteringsprocesser kan se ut i företag och organisationer.

2.1 Översättning och automatisering

Implementering av maskinöversättningssystem i översättningsprocesser är en del av en mer övergripande utveckling där system baserade på artificiell intelligens (AI) införs i organisationer för att automatisera delar av verksamheten. I takt med att AI implementeras på företag och i organisationer uppstår nya former av hybrider mellan människa och AI. I Rai et al. (2019) beskrivs tre olika former av tänkbara följder vid implementering av AI-baserade system. Enligt Rai et al. kan AI för det första ersätta mänsklig arbetskraft fullt ut (*task substitution*), vilket alltså innebär att det handlar om en total automatisering. Rai et al. skriver vidare att AI även kan användas som en förstärkning av den mänskliga arbetskraften (*task augmentation*), vilket innebär att den fungerar som ett komplement till mänsklig kompetens i arbetsprocessen. Slutligen menar Rai et al. att AI och mänsklig arbetskraft kan sammanfogas i en gemensam process och fungera som en integrerad enhet (*task assemblage*), där det är svårt att fastställa gränsen mellan människa och teknik.

Tillämpat på översättningsprocessernas område går det att föreställa sig en liknande uppdelning av implementeringsformerna för maskinöversättning. Maskinöversättningssystem kan användas som ersättning för mänskliga översättare och tillhandahålla helautomatiska översättningar av texter utan någon form av mänsklig inblandning eller korrektur (Kastberg, 2012). ”Råa” maskinöversättningsresultat utan mänsklig korrekturläsning används till exempel vid så kallad *gisting*, där kunden endast är intresserad av att få en ytlig och övergripande förståelse av innehållet i en text (Nurminen, 2020).

Maskinöversättningssystem kan vidare ha en funktion som påminner mer om den roll traditionella datorbaserade översättningsverktyg, t.ex. digitala översättningsminnen och terminologibanker, har för översättaren. I sådant fall fungerar maskinöversättningen som ett IT-stöd som kan bidra till att öka produktiviteten och kvaliteten i översättningsprocessen (Kastberg, 2012). Vissa branschexperter har även använt begreppet *augmented translation*, eller förstärkt översättning, för att beskriva detta fenomen (DePalma, 2017; Lommel, 2018). Maskinöversättningssystem utgör i detta fall endast en av flera resurser som mänskliga översättare använder sig av i en datorstödd översättningsprocess och bidrar till att befria den mänskliga översättaren från repetitiva och rutinmässiga uppgifter. På så sätt frigörs professionella översättares kapacitet för mer avancerade arbetsuppgifter som terminologihantering och lokalisering. (DePalma, 2017)

Taravella och Villeneuve (2013) föreslår i samma anda att datorstödd översättning med hjälp av översättningsverktyg, maskinöversättning och andra språkliga referensverktyg kan förstås i termer av informationssystem. De föreslår begreppet språkinformationssystem (*language*

information system) för att beskriva ett system som samlar alla relevanta informationsskällor och hjälper översättaren att hantera all tillgänglig kunskap som kan användas som stöd i översättningsarbetet.

Implementering av maskinöversättning i översättningsprocessen skulle för det tredje även kunna betyda att gränsen mellan mänsklig översättning och maskinöversättning alltmer luckras upp. Detta kan t.ex. delvis sägas vara fallet i översättningsprocesser där mänsklig efterredigering tillämpas på resultatet av maskinöversättningen (Olohan, 2021). I detta fall bearbetar och korrigerar en översättare eller annan person maskinöversatta texter och därefter matas oftast dessa bearbetningar tillbaka till maskinöversättningssystemet för att systemet ska kunna lära sig av sina misstag och generera bättre översättningar framöver (O'Brien, 2012).

Den sistnämnda tillämpningen av maskinöversättning innebär att en ny yrkesroll, s.k. efterredigering, har börjat växa fram, där översättare kontrollerar, ändrar och justerar maskinöversatt text så att den blir läslig, korrekt och anpassad till det aktuella sammanhanget (Vieira, 2019). Efterredigering säljs oftast in som en särskild tjänst till slutkunden och skiljer sig därmed från rent mänsklig översättning. Efterredigering innebär att slutprodukten är billigare och kan levereras snabbare än en vanlig översättning, men efterredigering kan även i vissa fall innebära en lägre kvalitet (Dannewitz-Linder, 2018a).

Maskinöversättning med efterredigering har på senare år blivit så utbredd och betydande att den internationella standardiseringsorganisationen (ISO) har utvecklat en särskild standard för efterredigering, *ISO 18587:2017 Translation services — Post-editing of machine translation output* (International Organization for Standardization, 2017). Denna standard innehåller riktlinjer för hur efterredigeringsprocessen ska gestaltas och vilka kvalifikationer och förmågor personer som arbetar med efterredigering bör ha för att den förväntade kvaliteten ska kunna uppfyllas.

Översättningsbyråer erbjuder i många fall två olika typer av efterredigeringsrelaterade tjänster: lätt och full efterredigering. Målet med lätt efterredigering är att den bearbetade texten ska vara innehållsmässigt korrekt och begriplig. Detta innebär att efterredigeraren bör sträva efter att behålla så mycket som möjligt av den maskinöversatta texten och endast ändra och korrigera grava fel. Lätt efterredigering tillämpas främst för texter som inte är tänkta att publiceras utan bara behöver förstås av avnämaren. Vid full efterredigering syftar efterredigeringsprocessen till att slutprodukten ska motsvara en text som översatts av en mänsklig översättare. (Vieira, 2019)

2.2 Risker kopplade till maskinöversättning

Neural maskinöversättning har inte bara potential att öka produktiviteten i översättningsprocessen och göra översättningar av hygglig kvalitet tillgängliga för allmänheten via t.ex. Google Translate (Focada, 2017). Det finns även ett antal risker förknippade med moderna maskinöversättningssystem. I ISO-standarderna för riskhantering, ISO 31000, definieras risk som ”*effect of uncertainty on objectives*” (International Organization for Standardization, 2018, s. 1). Detta riskbegrepp utgår ifrån att det finns ett mått av osäkerhet i alla situationer, vilket gör det omöjligt att förutse konsekvenserna av handlingar och händelser som äger rum i en organisation eller process. En risk kan alltså betraktas som en potentiell avvikelse från det förväntade eller önskade utfallet. I fallet med maskinöversättning kan det t.ex. handla om att översättningsprocessens resultat ger upphov till skadliga följdverkningar eller att känsliga

data läcker ut under översättningsprocessen. Nedan kommer jag att gå närmare in på ett antal risker kopplade till maskinöversättning som diskuterats i forskningslitteraturen.

2.2.1 Felöversättningar

Problem med felöversättningar är inget nytt fenomen i översättningsbranschen. Det förekommer inte bara i samband med maskinöversättning utan har även uppmärksammats för traditionella översättningar (Canfora och Ottman, 2018). Felöversättningar är främst ett problem i samband med översättning av säkerhetskritiska texter där felaktiga formuleringar eller utlämnad information kan leda till allvarliga person- och sakskador eller t.o.m. dödsfall, men även felöversättningar i andra typer av texter kan ha negativa konsekvenser och leda till t.ex. stämningar eller förlorat anseende (Canfora och Ottman, 2018).

I takt med att neurala maskinöversättningssystem implementeras i översättningsprocesser har det visat sig att dessa system genererar vissa typiska fel, t.ex. utelämnade eller tillagda ord, inkonsekvent terminologi och problem med att hantera ovanliga ord (Koehn och Knowles, 2017; Daems et al., 2017). Den gällande ISO-standarden för efterredigering listar vidare en rad andra typiska fel som kan uppstå och som bör uppmärksammas vid maskinöversatt text, t.ex. *”stylistic problems, literal translations, grammar mistakes (e.g. negations, verb translation), translation of names which should not be translated”* (International Organization for Standardization, 2017, s. 9).

Läubli (2017) har uppmärksammat att det också kan uppstå problem som rör kopplingen mellan översättningen av enskilda fraser eller meningar och texten i sin helhet när maskinöversättning används. Även om neurala maskinöversättningssystem har blivit bättre på att fånga upp och beakta kontextuella aspekter så menar han att dessa system saknar en förmåga att ta hänsyn till text som ligger utanför det aktuella textsegmentet, vilket innebär en vidare felkälla som bland annat kan ge upphov till syftningsfel.

Ett sätt att hantera problem med maskingenererade felöversättningar är, som beskrivits ovan, att låta resultatet från maskinöversättningen genomgå en efterredigeringsprocess (Nitzke et al., 2017). Men inte heller detta är en garant för att alla fel hittas och åtgärdas. Det finns studier som visar att några av de vanligaste felen som uppstår vid neural maskinöversättning kan vara svåra att spåra i efterredigeringsprocessen, vilket leder till att slutprodukten riskerar hålla en undermålig kvalitet (Castilho et al., 2017). En jämförande studie av Yamada (2019) där universitetsstudenter i språk utfört efterredigering av maskinöversatta texter från en statistisk och en neural översättningsmotor visar att slutresultatet visserligen förbättras med neural maskinöversättning, men att det fortfarande inte når upp till en professionell standard.

Det faktum att neural maskinöversättning ofta genererar en översättning med bra flyt och ett naturligt språk innebär att många fel kan ”gömma” sig i texten och att det kan vara svårt att hitta dessa fel i efterredigeringsprocessen (Vieira, 2019). Som Koehn skriver är detta problem något som får mycket uppmärksamhet i utvecklingen av neurala maskinöversättningssystem: *”Misinformation due to mistranslation is a significant concern with neutral machine translation, which sometimes prefers fluency over adequacy to the point of completely distorting the output so it has no relation to the input.”* (Koehn, 2020, s. 20)

2.2.2 Datasäkerhet

Många texter som översättningsbyråer översätter innehåller känsliga uppgifter av olika slag. Det kan handla om personuppgifter, affärshemligheter eller annan sekretessbelagd information. För slutkunden är det av yttersta vikt att dessa uppgifter behandlas i enlighet med gällande lagstiftning (t.ex. GDPR) och att de inte läcker ut i offentligheten. Problem med data-läckage och integritet är några av de datasäkerhetsrelaterade risker som uppmärksammats för maskinöversättning i forskningslitteraturen (Kamocki et al., 2015; Canfora och Ottmann, 2020). De flesta av de system och betaltjänster för neural maskinöversättning som används av översättningsbyråer uppges dock ha ett säkert dataskydd som hindrar information från att läcka ut ur systemet (Dannewitz-Linder, 2018b).

Mer problematiskt blir det dock när vissa typer av kostnadsfria maskinöversättningstjänster används i översättningsprocessen. Dessa tjänster innebär i många fall att företaget som tillhandahåller tjänsten har rätt att använda sig av den inmatade texten för att förbättra tjänsten, men även för andra ospecificerade ändamål. Det innebär att denna typ av tjänster ofta inte garanterar en säker datahantering vid översättning av känsliga uppgifter eller konfidentiella texter. (Vashee, 2017)

Under de senaste åren har också flera fall av dataläckage i samband med användning av kostnadsfria översättningstjänster nått ut i offentligheten (Fagerström, 2017). Bland annat har det rapporterats om att personal på norska Statoil använt sig av webbtjänsten Translate.com för att översätta interna meddelanden som sedan läckt ut på webbtjänstens webbplats och därmed gjorts tillgängliga för hela världen (Tomter et al., 2017).

Sakamoto (2019) har rapporterat om att det är vanligt att frilansöversättare på eget bevåg använder sig av olika kostnadsfria maskinöversättningstjänster även vid traditionella översättningsuppdrag. Anledningen till detta är de hjälper till att öka produktiviteten och hastigheten i översättningsprocessen, men det kan även handla om att översättaren behöver förslag till översättningar av vissa begrepp. Ofta görs detta utan att frilansöversättarna informerar sina uppdragsgivare om att maskinöversättningstjänsterna använts, vilket kan vara problematiskt eftersom det kan ge upphov till sådana datasäkerhetsproblem som beskrivits ovan (Sakamoto, 2019).

2.2.3 Säkring av professionell kompetens

Maskinöversättningens utveckling innebär även risker kopplade till professionella språkspécialisters situation och översättningsbyråernas samarbete med dessa. Förbättrade resultat från neurala maskinöversättningssystem har under senare år medfört att vissa kommentatorer förutspått att översättningsyrket kommer förändras i grunden eller kanske till och med försvinna på sikt när maskinöversättningssystemen blir så pass bra att de kan ta över hela översättningsprocessen (Marr, 2018).

Denna utvecklingsprognos kan sägas vara en del av en bredare trend. Framstegen inom artificiell intelligens påverkar inte bara översättares arbetssituation utan förväntas även medföra drastiska förändringar för många andra yrkeskategorier. Susskind och Susskind (2017) förutspår att maskiner baserade på AI kommer att ta över alltfler uppgifter som idag utförs av högutbildade yrkesgrupper och i förlängningen urholka traditionella professioner som läkare, jurister och revisorer.

Ännu har dessa förutsägelser långt ifrån infriats, men maskinöversättning har redan inneburit förändringar för översättares rådande arbetsituation. Den dominerande trenden för tillfället är, som diskuterats ovan, att fler och fler översättare arbetar med efterredigerings tjänster. Frågan är vad denna trend innebär för översättares allmänna situation och därmed i förlängningen för översättningarnas kvalitet. En möjlig risk för det professionella översättarkollektivet är att efterredigering inte betraktas som ett lika kvalificerat och avancerat jobb som översättning och därför istället delegeras till okvalificerad arbetskraft. Susskind och Susskind (2017) benämner detta fenomen paraprofessionalisering. Med detta menar de att arbetskraft med lägre utbildning och mindre erfarenhet kan utföra avancerade arbetsuppgifter som tidigare utförts av professionella yrkesgrupper, t.ex. översättningstjänster, med stöd av teknik som internaliserat expertbaserad kunskap.

Det finns dock anledning att ställa sig tvivlande till om detta perspektiv är applicerbart på efterredigeringsprocessen så som den ser ut idag. Den rådande ISO-standard för efterredigering framhåller tvärtom att efterredigering är en krävande process som fordrar hög kompetens och kvalifikationer motsvarande de en professionell översättare förväntas ha. Dessa krav inkluderar bl.a. en akademisk examen på översättnings-/språkområdet och/eller flerårig erfarenhet inom språkindustrin. (International Organization for Standardization, 2017)

Även forskningslitteratur på området, t.ex. Nitzke et al. (2019), lyfter fram att personer som arbetar med efterredigering, förutom språkliga kvalifikationer och lingvistisk kompetens, också bör besitta en rad andra förmågor. Bland annat kan det handla om teknisk kunskap och förståelse för hur maskinöversättning fungerar, men även en förmåga till riskbedömning och kompetens att avgöra i vilka situationer det är viktigt att lägga ned extra arbete och koncentration för att undvika att det uppstår fel i översättningarna (Nitzke et al., 2019).

Samtidigt finns det farhågor hos många översättare om att inte bara arbetsvillkoren kommer att försämrats utan även att arbetets innehåll riskerar att utarmas och bli mindre kreativt än tidigare i takt med att maskinöversättningen breder ut sig och efterredigering blir en norm i branschen (Moorkens och O'Brien, 2017). Följden av den beskrivna utvecklingen skulle kunna innebära att det blir svårt för översättningsbyråer att hitta kvalificerade personer som är beredda att ta på sig efterredigeringsuppdrag. Om utbildade översättare inte upplever arbetsituationen som tillfredsställande finns det risk för att många lämnar yrket och att unga personer skyr översättarutbildningar med försämrad kvalitet på framtida översättningar som följd (Taravella och Villeneuve, 2013). Ett sätt att råda bot på detta som föreslagits av Cadwell et al. (2018) vore att bättre integrera och ta till vara på översättarnas perspektiv vid implementering av maskinöversättning i översättningsprocesserna.

2.3 Riskhantering

Riskhantering är ett relativt utforskat område inom översättningssektorn (jfr. Canfora och Ottmann, 2015). I mina efterforskningar har jag inte kunnat hitta några akademiska artiklar som berör riskhantering och maskinöversättning, sånär som en artikel som handlar om riskhantering och efterredigering (Nitzke et al., 2019). Men som beskrivits ovan och som Canfora och Ottmann (2020) rapporterat om är det uppenbart att det finns ett antal risker förknippade med maskinöversättning och dessa måste hanteras på något sätt av översättningsbyråerna för att minimera eventuellt skadliga följder. Enligt Stallings och Brown (2018) kan organisationer använda sig av fyra huvudsakliga tillvägagångssätt för att identifiera och hantera risker

kopplade till IT i sin organisation och denna kategorisering lämpar sig väl för att förstå och beskriva hur översättningsbyråer gestaltar sin riskhanteringsprocess vid implementering av maskinöversättning.

Enligt Stallings och Brown (2018) kan företag och organisationer för det första tillämpa en allmän och grundläggande standard för riskhantering. I detta fall handlar det om att använda sig av en uppsättning standardiserade dokument, riktlinjer och säkerhetsanordningar som motsvarar den allmänt fastställda normen i branschen för att på så sätt skydda sig mot de vanligaste hoten och riskerna. Den främsta nackdelen med detta tillvägagångssätt, anser Stallings och Brown, är att det inte tar någon hänsyn till de specifika förutsättningar som gäller i den aktuella organisationen. Det innebär att det finns risk för att en alltför hög och kostsam säkerhetsstandard implementeras i organisationen eller att säkerstandarden inte är tillräckligt omfattande för att täcka alla sårbarheter.

Det andra tillvägagångssättet som Stallings och Brown beskriver, handlar om att utföra en mer informell riskanalys där man drar nytta av den kunskap och expertis som finns i organisationen eller vid behov rådfråga externa konsulter. I detta fall sker ingen strukturerad riskanalys utan här listas istället ett antal risker som lyfts fram av informanterna eller experterna och sedan utförs åtgärder baserat på denna information. Denna metod är relativt billig och kan utföras på ett snabbt och smidigt sätt, men innebär även att det finns en fara för att vissa risker negligeras eller att informationen är påverkad av informanternas åsikter och förutfattade meningar. (Stallings och Brown, 2018)

Det tredje och mest omfattande tillvägagångssättet består enligt Stallings och Brown av en detaljerad och formell riskanalys. Här används ett strukturerat och stegvist förlopp för att fastställa vilka tillgångar som behöver skyddas, identifiera hoten mot dessa tillgångar, uppskatta sannolikheten för att dessa risker ska inträffa och fastslå konsekvenserna för organisationen om detta skulle ske. Utifrån denna information är det sedan möjligt att välja vilka säkerhetsanordningar och andra säkerhetsrutiner som ska implementeras. Detta tillvägagångssätt innebär en omfattande och anpassad riskhantering, men är samtidigt mycket resurskrävande. (Stallings och Brown, 2018)

Internationella standardiseringsorganisationen (ISO) har utvecklat en särskild standard för riskhantering, ISO 31000, som kan ligga till grund för en formell riskhanteringsprocess. Enligt denna standard anses riskhantering vara en integrerad del av alla processer i en verksamhet och den förordar en kontinuerlig riskhantering, där förnyade riskanalyser görs med jämna mellanrum (International Organization for Standardization, 2018). Eftersom ISO 31000 är en horisontell standard kan den tillämpas på alla processer, inklusive maskinöversättning, i en översättningsbyrås verksamhet. Standarden gäller vidare för organisationer av alla storlekar, vilket gör att den omfattar både stora översättningsbyråer och enskilda frilansöversättare (Nitzke et al., 2019).

Stallings och Brown (2018) föreslår slutligen även att ett kombinerat tillvägagångssätt kan tillämpas där element från de tre ovannämnda metoderna kan ingå. I detta fall utgår riskhanteringen oftast från ett antal grundläggande säkerhetsrekommendationer som tillämpas över hela linjen. Därefter identifieras de system eller processer som är kritiska för verksamheten eller där riskerna bedöms vara höga. För dessa system kan sedan en informell riskbedömning eller en detaljerad riskanalys göras. (Stallings och Brown, 2018)

Enligt Nitzke et al. (2019) kan en riskhanteringsprocess även ligga till grund för beslutet om

huruvida maskinöversättning ska användas vid översättning av en viss typ av text eller inte samt om maskinöversättningsresultatet ska efterredigeras. Det är alltså viktigt för översättningsbyråer att analysera vilka risker och negativa följder som kan uppstå om maskinöversättning används innan man kommer överens med kunden om att ett aktuellt översättningsuppdrag ska involvera maskinöversättning (Nitzke et al., 2019).

När väl en riskanalys gjorts finns det enligt Stallings och Brown (2018) en rad olika sätt att behandla de konstaterade riskerna. En första väg som Stallings och Brown nämner är att acceptera den aktuella risken om det finns affärsmässiga skäl att göra detta. En annan strategi går ut på att undvika risken genom att helt enkelt låta bli att använda det IT-system som ger upphov till risken. Det är även möjligt att överföra risker genom att dela risken med en annan aktör, t.ex. genom att försäkra sig mot de negativa konsekvenserna av risk hos ett försäkringsbolag. En fjärde metod går ut på att reducera konsekvenserna av en risk. En sista metod som Stallings och Brown nämner är att minska sannolikheten för att en viss risk ska inträffa genom att använda någon form av säkerhetsrutiner eller säkerhetsanordningar (Stallings och Brown, 2018).

3 Metod

I det följande avsnittet kommer jag redogöra för beslut som tagits och problem som uppkommit under den studie som redovisas i denna uppsats. Allra först kommer jag diskutera varför jag valt att göra semistrukturerade intervjuer. Vidare beskrivs hur urvalet av informanter har gått till. Därefter redogör jag för intervjuernas genomförande. En nyckelaspekt i forskningsprocessen utgörs av hur insamlade data bearbetas och analyseras och detta beskrivs i avsnitt 3.4 i detta kapitel. I ett forskningsprojekt ställs man även inför en rad etiska problem som måste hanteras. Om detta tema handlar avsnitt 3.5. Slutligen diskuterar jag även kort huruvida min studie kan sägas uppfylla kraven på validitet och reliabilitet.

3.1 Val av metod

I denna uppsats är jag främst intresserad av att studera hur översättningsbyråer hanterar risker som uppstår vid användning av maskinöversättning i översättningsprocessen. Under senare år har risker i samband med maskinöversättning alltmer börjat lyftas fram som ett problem i forskningslitteraturen (Canfora och Ottmann, 2020). Men i de litteraturstudier jag gjort i föregående kapitel har jag inte kunnat hitta några mera omfattande empiriska studier på detta område och i synnerhet inte några studier som berör hur riskhanteringen går till på översättningsbyråerna. Eftersom jag alltså rör mig på ett i empirisk mening relativt utforskat område handlar det i mitt fall alltså främst om att göra en explorativ studie som utifrån vissa tematiska områden som jag presenterat i kapitel 2 undersöker hur översättningsbyråer använder maskinöversättningssystem och hanterar risker som kan uppkomma i samband med detta.

För en undersökning av det här slaget anser jag att kvalitativa studier lämpar sig bäst. Kvalitativa intervjuundersökningar bygger på enkla, öppna frågor som ger upphov till komplexa och innehållsrika svar (Trost, 1997). En av fördelarna med kvalitativa metoder jämfört med kvantitativa metoder är att de förra ger ett större utrymme för flexibilitet i forskningsdesignen, dvs. det finns möjlighet att t.ex. ändra och lägga till frågor efter hand som forskningsprojektet utvecklas. Med Jan Trosts ord utmärks kvalitativa intervjuer av ”hög grad av strukturering och låg grad av standardisering” (Trost, 1997, s. 22). Det går alltså att dra lärdomar från de första intervjuerna som sedan eventuellt kan tillämpas vid de senare intervjuerna. Detta är praktiskt när man rör sig på ett relativt nytt forskningsområde där det är svårt att uppställa färdiga hypoteser som t.ex. kan undersökas genom enkätstudier.

Syftet med de kvalitativa intervjuerna är i mitt fall att få en bild av hur maskinöversättning används och hur risker som uppstår i samband med detta hanteras av översättningsbyråerna. Målet är inte att få en representativ bild utan snarare att få en konkret inblick i organisationerna och processerna och lyfta fram ett inifrånperspektiv från översättningsbyråerna. Mer specifikt har jag beslutat mig för att genomföra semistrukturerade intervjuer. Anledningen till detta var att jag ville ha en relativt fast struktur att utgå från vid intervjuerna baserat på de teman som behandlas i litteraturkapitlet, samtidigt som jag ville ha möjlighet att följa nya intressanta trådar som uppkom i intervjuerna och även kunna ställa relevanta följdfrågor när informanterna byter ämne och kommer in på ett nytt spår (jfr. Bryman och Bell, 2005).

3.2 Urval

En viktig del av forskningsprocessen handlar om att välja ut och hitta lämpliga informanter för studien. I den föreliggande studien har jag tillämpat ett tvåstegsurval, likt det som beskrivs i Hallin och Helin (2018). Ett sådant tillvägagångssätt bygger på att man först identifierar en avgränsad grupp av människor/informanter som man är intresserad av och sedan försöker hitta potentiella intervjupersoner i denna grupp.

I det första steget fokuserade jag främst på relativt stora översättningsföretag. Anledningen till detta är att det är tämligen kostsamt att implementera ny avancerad teknik som maskinöversättning och det kräver därför att företagen har en viss storlek och ekonomisk styrka. Jag konstaterade ganska snabbt efter att gjort vissa efterforskningar att de flesta mindre och medelstora översättningsbyråer fortfarande inte har något maskinöversättningssystem på plats eller att maskinöversättning åtminstone ännu inte omfattas av de tjänster som marknadsfördes på respektive översättningsbyrås webbplats.

Vid urvalet av översättningsbyråer utgick jag därför från en lista över de 100 största översättningsföretagen i världen (CSA Research, 2020). Jag valde sedan ut byråer från denna lista som antingen hade sina huvudkontor i Skandinavien eller på annat sätt hade koppling till Skandinavien (t.ex. i form av lokala kontor). Därefter var jag inne och läste på respektive företags webbplats för att se om de erbjöd någon form av maskinöversättningsrelaterade tjänster, vilket det visade sig att de allra flesta gjorde.

I mina inledande kontaktförsök kontaktade jag ett tiotal översättningsbyråer via officiella kanaler, dvs. via ett kontaktformulär på webbplatsen eller en allmän e-postadress. Jag presenterade då kortfattat mig själv och mitt uppsatsprojekt och frågade om det fanns möjlighet att intervjua någon person med kännedom om maskinöversättning och riskhantering på det aktuella företaget. Denna kontaktmetod visade sig inte vara särskilt framgångsrik utan svarsfrekvensen var i princip obefintlig så när som på ett svar, där översättningsbyrån menade sig sakna tid och resurser för att gå mig till mötes.

Efter den första kontaktrundans motgångar bestämde jag mig för att ändra strategi och istället använda mig av LinkedIn för att söka efter personer med en yrkestitel (t.ex. maskinöversättningsspecialist, IT-ansvarig eller teknisk chef) som verkade indikera att personen i fråga arbetar med maskinöversättning på de valda företagen. Därefter skickade jag förfrågningar via e-post direkt till dessa personer, i de fall jag lyckades få tag på deras e-postadresser, eller så vände jag mig till företaget för att höra om det var möjligt göra en intervju med den nämnda personen. Denna väg visade sig vara något mer framgångsrik och ledde fram till att jag hittade tre lämpliga intervjupersoner i de aktuella företagen, men egentligen hade jag nog önskat att kunna göra en eller två intervjuer till.

Trots att jag slutligen lyckades få napp med hjälp av den andra strategin jag använde mig av måste jag ändå konstatera det var svårare än jag hade förväntat mig att hitta informanter som var villiga att ställa upp i min studie. Även i de fall där jag vände mig direkt till personen i fråga, uteblev oftast svaren eller så visade sig personerna ovilliga att ställa upp. Vad detta beror på kan jag bara spekulera i. Möjligen är maskinöversättning fortfarande ett känsligt tema. En av de slutliga informanterna var initialt orolig för att affärshemligheter kopplat till tekniken bakom maskinöversättning skulle kunna röjas vid en intervju. Efter att jag förklarat studiens upplägg så ställde denna person ändå till slut upp. En annan möjlig tolkning av problemet med att hitta informanter är att det rör sig om en relativt ny teknik som delvis fortfarande

håller på att implementeras hos översättningsbyråerna och att många därför kanske inte känner att de har kompetens att prata om dessa frågor.

3.3 Intervjuernas genomförande

Alla intervjuerna i studien genomfördes online, i två fall med hjälp av videokonferenstjänsten Zoom och i ett fall med videokonferenstjänsten Microsoft Teams. Anledningen till detta hade delvis att göra med att vi för tillfället befinner oss mitt i en global pandemi som i princip gör det omöjligt att genomföra intervjuerna i fysisk form, men troligen hade jag ändå utfört intervjuerna online eftersom intervjupersonerna i min studie alla arbetar på företag med kontor i andra delar av Sverige eller i andra europeiska länder.

Jag diskuterade intervjuens upplägg och innehåll med intervjupersonerna per mail eller telefon några dagar i förväg och bestämde därefter en lämplig tid för intervjun. Ett urval av frågorna skickades också till intervjupersonerna i förväg för att ge dem möjlighet att förbereda sig på lämpligt sätt. Jag skickade sedan också ut en påminnelse och inbjudan till ett Zoom- eller Teams-möte strax innan den inbokade intervjun skulle äga rum. Samtliga intervjuer spelades in efter att intervjupersonerna godkänt detta.

För att ha en viss struktur i intervjuerna utarbetade jag en intervjuguide (se bilaga 1) som jag sedan hade till stöd när intervjuerna genomfördes. Intervjuguiden omfattar ett antal olika sektioner och olika teman som i hög grad byggde på resultatet av litteraturstudien som föregick intervjuerna. Den första delen av intervjuguiden innehåller främst ett antal frågor av mer allmän natur om intervjupersonen, översättningsföretaget och vilken typ maskinöversättningssystem som implementerats. Syftet med detta var huvudsakligen att få igång intervjun och få intervjupersonen att börja prata samt ge bakgrundsmaterial till min studie. Därefter följer ett antal frågor kopplade till användningen av maskinöversättning i översättningsprocessen. I den sista delen av intervjuguiden förekommer slutligen ett antal olika teman kopplat till risker och riskhantering vid användningen av maskinöversättning.

De tre intervjuerna förlöpte ungefär som förväntat, förutom i ett fall där förbindelsen var dålig, vilket fick till följd att intervjun gjordes utan att kamerorna var påslagna. Jag följde i stort sett strukturen i intervjuguiden under intervjuerna, men ändrade ordningsföljd på frågorna där det föreföll naturligt och lade även till vissa insticksfrågor som inte fanns med i intervjuguiden. Jag modifierade även intervjuguiden en aning mellan de olika intervjuerna efter att ha upptäckt vissa brister i den ursprungliga versionen.

3.4 Bearbetning och analys av data

Efter att intervjuerna gjorts transkriberades ljudinspelningarna för hand. Transkriberingen skedde i samtliga fall senast ett dygn efter att intervjuerna utförts för att jag skulle kunna hålla dem i färskt minne och återge dem på ett så korrekt sätt som möjligt. Transkriberingen kan också betraktas som en första analys av materialet (Alvehus, 2013). Det är i regel svårt att uppfatta all information under själva intervjun eftersom man ofta har fokus nästa fråga eller för att svaren är väldigt informationsrika. Det innebär att viss information framträder först när intervjun sätts på pränt i samband med transkriberingen.

Det är också viktigt att påpeka att kvalitativa studier ofta har iterativa drag, dvs. det är ibland svårt att hålla isär insamling och analys (Bryman och Bell, 2005). I takt med att nya infallsvinklar framkommer i materialet i samband med transkriberingen kan datainsamlingen justeras efter hand som studien fortskrider. I mitt fall tog det sig till exempel uttryck i att intervjuguiden justerades under undersökningens gång, vilket beskrivits ovan. Vid transkriberingen försökte jag uppnå en balans mellan läslig text och naturtrogen återgivning. Det innebär att jag utlämnade utfyllnadsord (t.ex. ”äh” och ”hmm”), men behöll en uppbruten meningsstruktur. I vissa fall där det varit svårt att höra ett visst ord av den inspelade intervjun så har jag gjort en kvalificerad gissning och kursiverat ordet samt satt det inom parentes. Efter att intervjuerna transkriberats skickades de till intervjupersonerna för genomläsning och eventuella kommentarer och korrigeringar. De transkriberade intervjuernas finns som bilagor i slutet av uppsatsen (se bilaga 2 till 4).

Intervjuer ger ofta upphov till ett rikt och detaljerat, men även något ostrukturerat material (Bryman och Bell, 2005). För att bringa ordning i mitt intervjumaterial använde jag mig av en tematisk strategi där jag utgick från vissa begrepp och nyckelord som var intressanta för studien. Ur dessa utvecklades sedan mer allmänna kategorier och teman som gick att använda för att göra mer generella påståenden relaterade till studiens frågeställning (Hallin och Helin, 2018). Vid analysen använde jag mig av färger som representerade vissa teman som var intressanta för min frågeställning och färgmarkerade avsnitt i intervjuerna som gick att hänföra till dessa teman. Inalles använde jag mig av följande fem färgkategorier i analysen: implementering av maskinöversättning, riskanalys och riskhantering, felöversättningar, datasäkerhet samt efterredigering och kvalifikationer. Med hjälp av dessa färgmarkeringar var det möjligt att göra jämförelser mellan de olika intervjuerna, där gemensamma utmaningar och erfarenheter, men även kontraster i materialet framträdde på ett tydligt och översiktligt sätt. Dessa olika tematiska avsnitt låg sedan till grund för och bildade en ram till resultatkapitlet och det påföljande analyskapitlet

3.5 Etiska aspekter

I sin bok om kvalitativa intervjuer har Hallin och Helin (2018) arbetat ut etiska riktlinjer för intervjuer baserat på fyra krav: informationskravet, samtyckeskravet, konfidentialitetskravet och nyttjandekravet. Enligt Hallin och Helin innebär det första kravet att intervjupersonerna bör vara välinformerade om studiens syfte innan intervjun startar och känna till vilka förutsättningar som gäller. Som jag beskrivit ovan har jag försökt tillgodose detta krav genom att tydligt presentera studien i den inledande mailkontakten och även besvarat frågor kring studien per telefon eller e-post.

Vidare har jag sett till att uppfylla samtyckeskravet, dvs. att intervjupersonerna måste gå med på att bli intervjuade enligt de villkor som angivits, genom att boka in ett separat intervjutillfälle efter de inledande kontakterna och inte ställa några frågor före den inbokade intervjun. Jag var även nog med att begära samtycke för inspelning av intervjun innan detta gjordes.

Kravet på konfidentialitet går ut på att datamaterialet inte får delas fritt och att intervjupersonen har rätt att vara anonym (Hallin och Helin, 2018). Jag valde i konsekvensens namn att pseudonymisera alla de intervjuade samt de företag de arbetade på eftersom en av intervjupersonerna bad om detta. Jag har använt mig av fiktiva namn för intervjupersonerna och företagen för att det ska gå att hålla isär dem. Eftersom jag inte bedömer att intervjupersonernas och översättningsbyråernas riktiga namn är intressanta för studien, då jag främst är intresserad av

att lyfta fram en allmän bild, anser jag inte att detta har påverkat studien i någon nämnvärd omfattning.

För att kunna uppfylla nyttjandekravet, dvs. att det insamlade datamaterialet bara används för det ändamål det är avsett att användas till, har jag varit noga med att påtala för och lova informanterna att de inspelade intervjuerna endast ska transkriberas och användas för analys i denna uppsats för att därefter raderas.

3.6 Validitet och reliabilitet

Validitet beskriver förekomsten av eventuella systematiska fel i undersökningen. Med andra ord handlar det om huruvida den aktuella empiriska studien verkligen undersöker det den är avsedd att undersöka (Lindstedt, 2017). För att en intervjustudie ska kunna uppnå kravet på validitet är det viktigt att datamaterialet speglar det fenomen som studien har för avsikt att studera samt att informanterna är representativa och tillförlitliga (Hallin och Helin, 2018).

För att uppnå kravet på validitet har jag försökt skapa en tydlig koppling mellan studiens syfte och teoretiska utgångspunkter å ena sidan och metodvalet och datainsamlingen å andra sidan. För detta ändamål har intervjuguiden spelat en stor roll. I intervjuguiden har jag skapat en ram för den empiriska undersökningen där frågorna på ett tydligt sätt är kopplade till studiens problemområde och de utgångspunkter för studien som presenteras i litteraturgenomgången.

Intervjupersonerna som ingår i studien har vidare nyckelroller kopplade till maskinöversättning på några av de största översättningsbyråerna på den skandinaviska marknaden, vilket gör dem till viktiga och trovärdiga informanter på området. Dessutom har informanterna delvis olika roller i organisationen, i vissa fall mer tekniska och i andra fall mer processorienterade. Det borgar för att intervjuerna genererat ett mångsidigt och relativt informationsrikt material med tydlig koppling till frågeställningen trots att jag egentligen önskat genomföra någon intervju till i studien.

Reliabilitet handlar om huruvida en upprepning av undersökningen ger upphov till samma resultat (Bryman och Bell, 2005). För att uppnå kraven på reliabilitet måste alltså studien utföras på ett enhetligt och opartiskt sätt. Jag har försökt att sträva efter detta genom att se till att intervjuerna i studien utförts på ett likartat sätt med hjälp av intervjuguiden och genom att uppträda på samma sätt gentemot alla intervjupersoner. Informanterna har vidare få ta del av identiska beskrivningar av studiens upplägg och syfte före intervjuernas genomförande.

4 Empiriska resultat

I detta avsnitt presenteras de huvudsakliga resultaten från intervjustudien strukturerade enligt ett antal relevanta teman. Först beskrivs vilka typer av maskinöversättningssystem som används av översättningsbyråerna och dess implementering i översättningsprocessen. Därefter redogörs för hur risker kopplade till maskinöversättning analyserats och hanterats i samband med att dessa system implementeras. Slutligen redovisas hur översättningsbyråerna hanterar de specifika risker som tematiserats i teorikapitlet.

4.1 Intervjupersoner

I tabellen nedan ges en översikt över informanterna som ingick i studien.

Intervju	Informant (pseudonym)	Företag (pseudonym)	Roll på företaget	Datum och tid	Bilaga (full intervju)
Intervju 1	Andy	X Translation	Technology Manager	2020-04-27, 40 minuter	Bilaga 2
Intervju 2	Jan	Y Translation	IT-ansvarig	2020-04-29, 35 minuter	Bilaga 3
Intervju 3	Jessica	Z Translation	lösningsarkitekt	2020-05-06, 40 minuter	Bilaga 4

Tabell 4.1: Intervjupersonerna i studien

4.2 Implementering och användning av maskinöversättning

Översättningsbyråerna som ingick i studien använder sig av olika typer av maskinöversättningssystem i sina översättningsprocesser. X Translation har köpt in ett färdigt maskinöversättningssystem från en känd leverantör eftersom detta uppfyllde de krav de ställde på systemet och det dessutom var lätt att integrera i det övriga IT-systemet. Även Z Translation använder sig av en färdig lösning. I båda fallen handlar det om system som bygger på neural maskinöversättning. Y Translation, å sin sida, har utvecklat ett eget maskinöversättningssystem på basis av open access-systemet Moses efter att tidigare använt en färdig lösning som de inte var helt nöjda med. Det aktuella maskinöversättningssystemet har sedan anpassats till de olika domäner och översättningsområden som är vanligt förekommande bland översättningsbyråns kunder.

Medan X Translation använt sig av maskinöversättning i cirka nio år, så har maskinöversättningen en kortare historia på både Y Translation och Z Translation, där denna teknik endast varit en del av översättningsprocessen i tre respektive fyra år. Jan på Y Translation ställer sig också aningen skeptisk till att maskinöversättning kan fungera som en universallösning inom översättningsindustrin:

”Så jag säger inte att man inte ska använda det och jag tycker att det är en väldigt bra teknik, ett komplement till mycket annat. Men det är inte lösningen på allt, än så länge.” (Jan, Y Translation)

På Y Translation finns det en relativt lång historia av digitalisering av verksamheten. Här har man använt sig av ett datorbaserat översättningsverktyg med översättningsminnen sedan 1986 för att lagra alla översättningar som gjorts i ett databassystem. Detta innebär att företaget sitter på ett rikt material av tidigare översättningar som kan återanvändas vid t.ex. uppdateringar eller nya utgåvor av dokument som översatts tidigare. I sådana fall uppskattar Jan att det ofta bara är 2–3 procent av hela textmängden som behöver översättas från grunden och då innebär inte maskinöversättningstekniken någon större vinst.

Även om både Y Translation och Z Translation erbjuder sina kunder helautomatisk översättning i vissa avgränsade fall där ledtiderna är korta och kvalitetskraven förhållandevis låga är det vanligaste förfarandet att maskinöversättningsresultaten granskas av mänskliga efterredigerare. På X Translation är mänskliga lingvister eller översättare involverade i alla översättningsprocesser eftersom detta föreskrivs av ISO-standarden 18587 för efterredigering av maskinöversättning.

Av det som framkom i intervjuerna är det tydligt att tekniken, maskinöversättning inbegripet, blir en allt viktigare del av översättningsprocessen. Även om det i flera fall var svårt för informanterna att uppskatta i hur stor del av alla uppdrag maskinöversättning används, verkar intentionen vara att använda det i så stor utsträckning som det är möjligt. På X Translation uppskattar Andy att maskinöversättning varit en del av processen i uppemot 90 procent av alla uppdrag. Även Jessica lyfter framhäver teknikens roll:

”Företaget är en ’language service provider’, kan man säga, som går mer och mer mot att bli ett ’language technology company’. Ja, så att det innebär ju att vi jobbar ju med en integration mellan det lingvistiska, alltså översättning, och olika typer av språktjänster med tekniska lösningar och det är ju dit mycket branschen går.” (Jessica, Z Translation)

Huruvida maskinöversättning ska användas som en del av översättningsprocessen vid ett översättningsuppdrag verkar även i stor utsträckning bero på vad kunden efterfrågar. Även om X Translation utgår från ISO-standarder och egna checklistor när man avgör om maskinöversättning ska användas är det i första hand kunden som bestämmer om detta ska göras. Även hos Y Translation har kunden sista ordet när det bestäms om en text ska maskinöversättas eller inte. Jessica på Z Translation lyfter fram vikten av att förklara tekniken för kunden och ge tydliga rekommendationer om huruvida ett uppdrag lämpar sig för maskinöversättning. Det hänger samman med att kunderna inte alltid är så stora förståelse för maskinöversättningens prestanda och kvalitet. Det är även någonting som Andy understryker:

”They don’t know anything about it. They have heard that it is cheap. But there is a lot of misinformation still to this day going around” (Andy, X Translation)

Jan talar också om att maskinöversättning har blivit ett ”buzz word” på senare tid och att många kunder har orealistiska förväntningar på vad tekniken kan åstadkomma. Han upplever också att detta är också någonting som underblåses av säljare i branschen och att det finns risker med denna utveckling.

4.3 Riskanalys och riskhantering

Andy uppger att X Translation gjorde en ganska omfattande analys i samband upphandlingsprocessen för ett nytt maskinöversättningssystem. Det viktigaste kriteriet här var att systemet hanterade kundens data på ett säkert sätt och man tog därför särskild hänsyn till leverantörernas policy och regler för databehandling. I denna analys använde sig X Translation av en checklista med krav som man själv hade arbetat ut.

Även på Y Translation föregicks implementeringen av maskinöversättningssystemet av en riskanalys. Jan framhåller här vikten av att göra en grundlig analys:

”För att, jag menar, att gå ut och läsa något ”white paper” eller någon som har skrivit någonting som säger det att maskinöversättningen är här nu... nu är det... det är liksom Jesus... här är frälsaren som kommer och det är så här mycket som ni kan köpa eller vi ska förvänta oss förbättringar och såna saker.” (Jan, Y Translation)

På Y Translation tittade man på en rad risker i samband med analysen. Till exempel analyserades säkerhetsrisker såsom lagring av information och dataåtkomst, men det gjordes även utvärderingar av maskinöversättningarnas kvalitet. I samband med detta utfördes en rad olika tester, där systemet testades på olika former av material och olika språkkombinationer. Eftersom Jessica inte jobbade på Z Translation när maskinöversättningssystemet infördes kunde hon inte säkert säga om det gjorts någon riskanalys i samband med detta.

Ingen av översättningsbyråerna som ingick i studien är certifierade enligt ISO-standarden 31000 för riskhantering. Andy på X Translation hävdade dock att de höll på att titta på detta för tillfället och att det eventuellt skulle kunna bli aktuellt för hela verksamheten. Enligt Andy har företaget identifierat ett antal sårbarheter inte bara i maskinöversättningssystemet, utan i organisationen i sin helhet och detta är något som de vill ta tag i.

Jan från Y Translation berättade att de är certifierade enligt ISO 9000 och även i denna standard ingår det ett krav på att risker ska analyseras och hanteras. Han har även noterat ett ökat riskmedvetande hos kunderna och misstänker att ISO-standarden 27000 för informationssäkerhet kommer bli ett krav för översättningsprocessen på sikt.

Jessica å sin sida framhåller vikten av att alla parter är med på tåget när maskinöversättning införs i översättningsprocessen. Hon anser att det är viktigt att både översättare/efterredigerade och kunder är engagerade i processen för att maskinöversättningar ska ge ett önskat resultat och fungera på ett säkert sätt.

4.4 Felöversättningar och översättningskvalitet

Intervjupersonerna i studien har gjort lite olika erfarenheter av vilka typer av felöversättningar som brukar uppstå i samband med maskinöversättning. Detta har möjligen att göra med att de olika översättningsbyråerna använder sig av skilda maskinöversättningssystem. Jan från Y Translation nämnde framför allt att man där upplevt att det statistiska maskinöversättningssystemet som man tidigare använt haft problem med att hantera och hitta rätt översättningar för vissa domänspecifika ord. Detta problem var dock enligt Jan möjligt att åtgärda med hjälp av bättre terminologistöd, t.ex. i form av ”terminology mining” eller ”terminology harvesting”.

Andy menade främst att deras maskinöversättningssystem haft problem med att tillämpa ett konsekvent språkbruk, framför allt vad gäller begreppsanvändningen. Han hade även upplevt att maskinöversättning kan ha problem med att hantera längre meningar. Enligt hans erfarenhet har sådana meningar ofta en tendens att bli ganska styliga när de maskinöversätts. Jessica gav exempel på obegripliga översättningar som hade fått fäste i maskinöversättningssystemet och kunde dyka upp i olika sammanhang. Vidare ansåg hon också att maskinöversättningar kan upplevas som väl bokstavliga.

Det vanligaste sättet att hantera och åtgärda felöversättningar som genereras av maskinöversättningssystem verkar vara att använda sig av kompetenta personer som efterredigerar maskinöversättningsresultaten. Detta var något alla tre översättningsbyråerna tillämpade. Ett problem i samband med efterredigeringsarbetet som Andy nämnde var att många översättare har haft problem att ställa om mentalt från traditionell översättning till efterredigering och därmed haft svårt att skilja mellan mänskliga översättningar från ett översättningsminne, så kallade TM-träffar, och resultat från maskinöversättningen, så kallade MT-träffar. Det kan vara problematiskt eftersom dessa två typer av träffar skiljer sig kvalitativt åt och bör hanteras på olika sätt.

I intervjuerna framkom att översättningsbyråerna även använder sig av en del andra metoder för att hantera risker med felöversättningar. På Translation X används en kombination av automatisk och mänsklig utvärdering för att kontrollera och vid behov åtgärda maskinöversättningarnas kvalitet. Det finns ett antal automatiska variabler tillgängliga för detta ändamål i den plattform som används för maskinöversättning och dessa används för att träna upp systemet till att fungera ännu bättre. Dessutom används en modell där identiska texter som översätts av en mänsklig översättare respektive maskinöversättningssystemet jämförs och utvärderas. På så sätt är det möjligt att hitta och åtgärda återkommande fel i maskinöversättningssystemet.

En annan metod för att undvika maskingenererade felöversättningar är helt enkelt att avstå från att använda maskinöversättning för vissa specifika texter. Jan nämner t.ex. att maskinöversättning inte används för marknadsföringstexter. Det framkom även i intervjuerna att man även i vissa fall accepterar att maskinöversättningen håller en lägre kvalitet, i synnerhet i samband med helt automatiserade översättningstjänster, om översättningen trots detta kan uppfylla sitt syfte. I sådana fall understryker Jan dock att det är viktigt att man då är tydlig med att kommunicera detta till kunden.

4.5 Datasäkerhet

Datasäkerhet i samband med maskinöversättning tycks vara en fråga som tas på stort allvar av alla de översättningsbyråer som ingick i studien. Åtminstone är detta ett tema som alla intervjupersoner återkommer till och lyfter fram som ett nyckelområde. Jessica pratar om betydelsen av att förstå var data som genereras i maskinöversättningsprocessen tar vägen så att man kan kontrollera att den inte används av en tredje part, men eftersom datasäkerhet ligger utanför hennes arbetsområde ville hon inte gå in i detalj på hur detta hanteras på Z Translation.

Andy betonar att för X Translation är det av yttersta vikt att kundens data behandlas på ett säkert sätt. När företaget började intressera sig för maskinöversättning och se sig om efter lämpliga leverantörer förekom det ofta att dessa inte tydligt angav huruvida de lagrade och använde sig av maskinöversättningarna själva. Eftersom X Translation har många

högprofilerade företag och kända varumärken som kunder som lägger stor vikt vid datahanteringen och även reglerar det i avtal var detta någonting som de behövde få klarhet i. Den nuvarande maskinöversättningslösningen är helt krypterad och X Translation har full kontroll över de data som genereras av maskinöversättningssystemet. De har även möjlighet att bestämma var de servrar är placerade där alla data från maskinöversättningssystemet lagras, vilket t.ex. är av betydelse för att kunna uppfylla dataskyddsförordningen (GDPR).

På Y Translation upplever man inte att det förekommer samma datasäkerhetsrisker som hos sådana översättningsbyråer som köpt in maskinöversättningslösningar från externa leverantörer. Som tidigare nämnts har man på Y Translation utvecklat ett eget inhouse-system för maskinöversättning och detta innebär att problemen och utmaningarna som Andy beskriver inte förekommer här eftersom alla data stannar i företaget.

Trots att översättningsbyråerna tagit hänsyn till datasäkerhetsaspekter vid implementationen av maskinöversättningssystemen och anser att dessa fungerar på ett säkert sätt händer det i vissa fall att kunden önskar att känsliga texter inte ska maskinöversättas:

”Some clients are very, very, very strict on the way they want their data used. If they stop us from using any sort of machine translation because of that, that’s fine. We respect that.”
(Andy, X Translation)

Vad gäller de datasäkerhetsproblem som är kopplade till sådana kostnadsfria maskinöversättningstjänster som beskrivits i kapitel 2 är budskapet tydligt från intervjupersonerna. Frilansöversättare får inte använda denna typ av verktyg eller tjänster när de utför översättningsuppdrag för dessa översättningsbyråer. På X Translation är detta uttryckligen förbjudet i de allmänna villkor som ingår i det samarbetsavtal som frilansöversättare sluter med byrån när de påbörjar sitt samarbete med denna. Om detta ändå skulle inträffa avslutar översättningsbyrån normalt samarbetet med översättaren i fråga. Detsamma gäller på Y Translation där detta ingår i sekretessavtalet som sluts mellan byrån och översättaren. Jan varnar dock för att det kan vara svårt att kontrollera att detta efterföljs i praktiken.

På X Translation är man dock inte helt emot att frilansöversättare använder sig av maskinöversättning på egen hand när de utför översättningsuppdrag för byrån. Men i sådant fall måste översättarna vara helt öppna med att detta sker och förklara vad för typ av maskinöversättningssystem det handlar om. Om översättaren till exempel tränat upp en egen maskinöversättningsmotor och denna bedöms vara säker så kan det hända att X Translation tillåter att översättaren använder sig av denna i samband att översättningsuppdrag utförs.

4.6 Efterredigering och översättares attityd till maskinöversättning

Som beskrivits i kapitel 2 ställer ISO-standarderna för efterredigering (ISO 18587) upp ganska stränga villkor för denna typ av arbete. Av de översättningsbyråer som ingår i studien är det bara X Translation som är certifierad enligt ISO 18587. Det innebär att de endast anlitar personer som uppfyller denna standard, dvs. personer med erfarenhet av översättningsyrket och/eller som har genomgått översättarutbildning eller liknande lingvistisk utbildning, i samband med efterredigeringsarbeten och att texten som ska granskas alltid är på personens modersmål. Det sistnämnda kriteriet gäller även på Y Translation.

Jan på Y Translation menar i övrigt att man tillämpar kriterier liknande de som gäller för

översättare när det bestäms vilka personer ska anlitas för efterredigeringsuppdrag. Han hävdar att man i princip skulle kunna anlita vem som helst för att utföra den rent språkliga granskningen av maskinöversatta texter, men liksom vid vanliga översättningar är det av yttersta vikt att personen som utför efterredigering är väl insatt i det aktuella området, i synnerhet vid tekniska översättningar. Här är det viktigt att denna person kan fastställa att innehållet i översättningen verkligen stämmer överens med innehållet i originaltexten.

På X Translation har man upplevt svårigheter med att hitta bra utbildningsmaterial för personer som arbetar med efterredigering i översättningsprocessen. I slutändan ledde det fram till att de utvecklade ett eget kursmaterial som de nu använder sig av. Även externa frilansöversättare erbjuds normalt utbildning på temat efterredigering, men pga. pandemin under det senaste året har det varit svårt att tillhandahålla denna utbildning. Det finns dock en rad andra resurser tillgängliga som externa efterredigerare kan konsultera på detta tema.

Trots att man tillämpar ganska stränga kriterier när man anlitar externa översättare för efterredigeringsarbeten anser ändå Jessica att kvaliteten på efterredigerade texter ofta skiljer sig från kvaliteten på texter som översatts från grunden:

”För att det blir ju annorlunda om man post-redigerar en text... jag vet ibland när jag själv har läst postredigerade texter. Man får en känsla av att om man hade översatt det här från början hade man kanske inte skrivit exakt så här. Man kan inte riktigt peka på att någonting är fel, men att det är någonting som känns lite annorlunda i texten.” (Jessica, Z Translation)

Vad gäller översättares inställning och attityd till maskinöversättning så har Andy märkt skillnader baserat på vilken generation översättaren tillhör. Många äldre och rutinerade översättare har haft problem att anpassa sig till den nya arbetsprocessen och i vissa fall vägrat befatta sig med maskinöversättningar. Unga översättare har en betydligt öppnare attityd, men uppvisar samtidigt allt större svårigheter att översätta en text från grunden:

”And the newer ones, particularly fresh from university, seems to find it very easy to simply start working with it. Often to the point of when you take it away for whatever reason, then they just going like, oh God, I have to think about a sentence from scratch now. I haven't done this in forever.” (Andy, X Translation)

Jan menar att det är viktigt både för översättningsbyråer och översättare att acceptera att maskinöversättningen är här för att stanna och att man måste anpassa sig till denna teknik för att överleva i branschen. Samtidigt anser han liksom Jessica tidigare att det är viktigt att se till att översättarna inte missgynnas utan att bland annat sträva efter att hitta ersättningsmodeller som är rimliga. Andy hävdar vidare att det i slutändan är översättaren som står i centrum för översättningsprocessen. Maskinöversättning är ett verktyg som översättaren ska välja att använda. På X Translation tillämpar de därför principen att om översättarna inte upplever någon nytta med maskinöversättningen för en viss text så används det inte.

5 Diskussion

I det följande kapitlet diskuteras resultaten från intervjuerna i förhållande till tidigare forskning och tillämpliga teorier. Först analyseras hur maskinöversättning används av översättningsbyråerna i studien. Därefter diskuteras hur översättningsbyråerna har gått tillväga för att analysera och hantera risker i samband med implementeringen. Slutligen avhandlas hur de tematiserade riskerna har behandlats och om erfarenheterna på översättningsbyråerna är i linje med de som presenteras i angränsande forskning.

5.1 Automatisering av översättningsprocessen

Som framgått av den empiriska delen utgör språktjänster som omfattar helautomatisk översättning en relativt marginell del av den totala verksamheten på översättningsbyråerna. Två av översättningsbyråerna i studien erbjuder förvisso sina kunder helautomatiska översättningar, men bara i begränsad omfattning för vissa specifika ändamål. Användningsområdet för helautomatiska översättningar verkar ligga i linje med det som beskrivits i forskningslitteraturen på området, dvs. det är främst någonting som används när kunden snabbt behöver tillgodogöra sig innehållet på ett översiktligt sätt eller för att kunna göra en bedömning av vad materialet innehåller (jfr. Nurminen, 2020).

Utifrån intervjuerna framstår det som att den huvudsakliga tillämpningen av maskinöversättning motsvarar *task augmentation*, dvs. att maskinöversättning fungerar som en extra stödfunktion åt översättaren, snarare än *task assemblage*, där maskinöversättningssystemet och översättarens efterredigeringsfunktion inte går att skilja åt i en helt integrerad översättningsprocess (jfr. Rai et al., 2019).

Som Jan är inne på fungerar maskinöversättningssystemet på Y Translation främst som ett komplement till de interna översättningsminnena som byggts upp i flera decennier och därför innehåller ett rikt översättningsmaterial som kan återanvändas i översättningsprocessen. I Y Translations fall är det fortfarande dessa översättningsminnen som utgör kärnan i den datorstödda översättningsprocessen.

Intervjupersonerna framhåller också att det primärt är tänkt att maskinöversättningssystemet ska underlätta arbetet för översättarna, men att det fortfarande är de mänskliga översättarna som står i centrum i översättningsprocessen. Det verkar även i stor utsträckning gälla vid efterredigeringsarbeten. Jessica pekar på att maskinöversättningen främst är tänkt att fungera som ett stöd i översättningsprocessen och det är ”*översättaren som avgör hur pass mycket man vill använda av den.*” Andy är inne på ett liknande spår där han menar att företaget endast använder maskinöversättning i de fall där det är till gagn för översättaren:

”As I said, if we don’t think that particular domain is helped by machine translation and it might make our translators’ life harder, we refrain from it” (Andy, X Translation).

Dessa idéer ligger väl i linje med de som har lyfts fram av företrädare för *augmented translation*, där maskinöversättning syftar till automatisera vissa rutinartade aspekter av översättningsprocessen för att översättarna ska kunna öka produktiviteten och fokusera mer på terminologiarbete och efterforskningar (DePalma, 2017).

5.2 Riskhantering vid implementering av maskinöversättning

I intervjuerna med informanterna från de tre översättningsbyråerna framkom att ingen av dessa att gjort någon egentlig formell riskanalys innan maskinöversättning infördes i översättningsprocessen. Det visade sig också att inte heller någon av översättningsbyråerna var certifierad i enlighet med ISO-standarden för riskhantering, ISO 31000. Detta resultat ligger i linje med andra studier som påvisat att formella och detaljerade riskhanteringsprocesser kopplade till översättningsprocessen i stort är ovanliga på översättningsbyråer eftersom de upplevs som dyra och tidskrävande (jfr. Canfora och Ottmann, 2018).

Stallings och Brown (2018) skriver att ett sätt att hantera risker kopplade till informationsteknik är att tillämpa någon form av *baseline approach* eller *industry best practice* som visat sig uppfylla vissa grundkrav gällande säkerheten. Detta skulle kunna sägas gälla för den typ av färdiga och välbeprövade maskinöversättningslösningar som används på X Translation och Z Translation. När t.ex. Andy motiverade valet av det aktuella maskinöversättningssystemet anförde han det faktum att denna lösning hade tydliga rutiner för datahantering som stämde överens med de krav som efterfrågades på detta område.

På både X Translation och Y Translation gjordes relativt grundliga analyser innan ett maskinöversättningssystem infördes i översättningsprocessen. Dessa analyser var informella i den bemärkelsen att de inte följde något standardiserat förfarande utan snarare utgick från egenutvecklade metoder eller drog nytta av den egna tekniska kompetensen på företaget. I fallet med X Translation skulle man dessutom kunna hävda att de tillämpade en kombination av en baseline-lösning och en informell analys (jfr. Stallings och Brown, 2018).

Informella riskanalyser behöver dock inte nödvändigtvis innebära att riskanalysen inte varit djupgående, utan snarare tvärtom. Som beskrivits i föregående kapitel gjordes det t.ex. på Y Translation en relativt genomgripande analys av maskinöversättningssystemet och dess output innan det togs i drift.

5.3 Felöversättningar och översättningskvalitet

Många av de felöversättningar som intervjupersonerna haft erfarenhet av i sina respektive maskinöversättningssystem har även tematiserats och lyfts fram av tidigare forskning som vanligt förekommande fel i samband med maskinöversättning. Till exempel så framhäver Koehn och Knowles (2017) att neural maskinöversättning har problem med att hantera ovanliga ord samt att denna teknik uppvisar sämre kvalitet för längre meningar, båda saker som även framkom i intervjuerna. Intervjupersonerna nämnde också förekomsten av bokstavliga översättningar, vilket lyfts fram som ett problem i ISO-standarden för efterredigering (International Organization for Standardization, 2017), samt en oförmåga att hålla sig till en konsekvent begrepps användning, något som Daems etc. (2017) också framhållit som ett frekvent problem.

Jessica från Z Translation pratade i intervjun om att två huvudsakliga dimensioner, *fluency* och *accuracy*, används i bedömningen av maskinöversatta texter som efterredigerats. Enligt hennes erfarenhet, kunde maskinöversatta texter ofta uppvisa ett bra flyt (*fluency*), men istället ha problem med begrepps användning och innehåll (*accuracy*). Detta har även uppmärksamats i forskningslitteraturen där det konstaterats att texter som producerats av neurala maskinöversättningssystem, i synnerhet om de efterredigerats, ofta uppvisar ett naturligt språk med bra flyt, men att det i många fall kvarstår problem med innehållet (Castilho et al., 2017).

Översättningsbyråerna använder sig av en rad olika åtgärder för att minska risken för felöversättningar eller dålig kvalitet i maskinöversättningarna. Som framkom i intervjuerna används i princip genomgående efterredigering för att minska risken för fel i översättningarna och öka kvaliteten på de maskinöversatta texterna. På X Translation efterredigeras alla maskinöversättningar utan undantag. Det framkom dock inte i intervjuerna om det förekom några särskilda kvalitetsproblem avseende de efterredigerade maskinöversättningarna, något som forskning på området pekat ut som ett problem (Castilho et al., 2017; Yamada, 2019). Däremot talade bland annat Andy på X Translation om att en del översättare hade haft vissa omställningsproblem i samband med att de börjat utföra efterredigeringsarbeten och haft svårigheter att hålla isär maskinöversättningar och mänskliga översättningar från översättningsminnet. Detta är ett problem som uppmärksammats även i litteraturen i takt med att båda dessa typer av input och stöd för översättaren integrerats i moderna översättningsverktyg (Vieira, 2019).

Även andra sätt att behandla risken för felöversättningar, t.ex. att helt avstå från maskinöversättning eller acceptera risken i form av översättningar av lägre kvalitet, tillämpades av översättningsbyråerna (jfr. Stallings och Brown, 2018). Till exempel nämner Jan på Y Translation att man avstår från att översätta kreativa texter som marknadsinformation med hjälp av maskinöversättning något som även rekommenderas av Nitzke et al. (2019). Det förekommer också att man accepterar en högre risk för felöversättningar i vissa situationer i samband med helautomatisk översättning och lätt efterredigering om kunden önskar detta (jfr. Stallings och Brown, 2018).

5.4 Maskinöversättning och datasäkerhet

Datasäkerhetsproblem i samband med maskinöversättning verkar främst vara en fråga som är kopplad till kostnadsfria maskinöversättningstjänster (Vashee, 2017), medan kostnadspliktiga lösningar och maskinöversättningssystem som bjuds ut på marknaden oftast uppfyller grundläggande krav på datasäkerhet, åtminstone om man kan lita på leverantörernas uppgifter (Dannewitz-Linder, 2018b). Den förstnämnda typen av maskinöversättningssystem används inte av översättningsbyråerna i studien, utan de system som implementeras hos dessa är antingen av det andra slaget eller egenutvecklade.

Att dessa system uppfyller de önskade säkerhetskraven är också något som informanterna i studien vill göra gällande. På både X Translation och Z Translation har grundläggande datasäkerhetskriterier som datasekretess och dataintegritet (jfr. Kamocki et al., 2015) varit avgörande när översättningsbyrån valt vilket maskinöversättningssystem som ska köpas in. X Translation valde sin nuvarande leverantör av maskinöversättningsteknik eftersom denne kunde erbjuda översättningsbyrån krypterad och säker datahantering samt garantera att översättningsbyrån själva äger och kontrollerar alla data som genereras av maskinöversättningssystemet. Just denna sista punkt var helt central för X Translation och en anledning till att man bytt leverantör eftersom den tidigare leverantören inte fullt ut kunnat garantera detta.

På Y Translation har man valt en annan väg för att se till att garantera datasäkerheten i maskinöversättningssystemet. Här har istället en egenutvecklad programvara använts och man har byggt upp ett slutet, internt system som innebär att de data som genereras av maskinöversättningen inte lämnar företaget. Denna typ av slutna översättningssystem som frilansöversättare loggar in på när de utför översättnings- eller efterredigeringsarbeten innebär att hela

översättningsprocessen utförs företagsinternt, vilket garanterar en säker datahantering i processen (Canfora och Ottmann, 2020).

Vad gäller datasäkerhetsproblem kopplade till frilansöversättares användning av kostnadsfria översättningstjänster och verktyg (jfr. Sakamoto, 2019), så var policyn tydlig på översättningsbyråerna som ingick i studien. Detta är någonting som inte tolereras överhuvudtaget och det regleras i regel genom avtal med de externa frilansöversättare som anlitas av företaget.

5.5 Maskinöversättning och vikten av professionell kompetens

Det finns inget stöd i intervjumaterialet för att maskinöversättning med efterredigering har givit upphov till en paraprofessionalisering av språkbranschen, dvs. att översättningsbyråerna använder sig okvalificerad arbetskraft för denna typ av arbeten (jfr. Susskind och Susskind, 2017). Översättningsbyråerna tillämpar i princip samma urvalsmetoder och ställer samma krav när de anlitar personer för dessa tjänster som när de anlitar frilansare för vanliga översättningsuppdrag.

Översättningsbyråerna verkar alltså vara ganska eniga om att efterredigering i princip kräver samma kvalifikationer och egenskaper som gäller vid professionell översättning. Andy på X Translation refererar till ISO-standarden för efterredigering ISO 18587 (International Organization for Standardization, 2017) som han menar innehåller krav på kompetenser och kvalifikationer som motsvarar dem som anges för översättare i ISO-standarden för översättning ISO 17100 (International Organization for Standardization, 2015). Även Jan och Jessica är inne på samma linje och Jan understryker särskilt att fackmässiga ämneskunskaper är mycket viktiga både vid översättning och efterredigering.

Viss litteratur på området har förordat att efterredigering kräver en rad ytterligare förmågor utöver de som efterfrågas vid traditionell översättning, t.ex. kunskap om hur maskinöversättning fungerar och en förmåga att göra riskbedömningar (Nitzke et al. 2019). Detta var dock inget som någon tog upp i intervjuerna utan alla intervjupersonerna verkade snarast vara av uppfattningen av likheterna mellan översättning och efterredigering var större än skillnaderna.

Yamada (2019) hävdar att personer som utför efterredigering behöver en särskild utbildning för att kunna fokusera på problem som är specifika för dessa jobb och utföra arbetet på ett effektivt och korrekt sätt. Enligt Cespedes (2018) förväntas ofta frilansöversättare själva införskaffa den kunskap som är nödvändig för att kunna utföra efterredigeringsarbeten. Detta gäller dock inte för alla översättningsbyråerna som ingick i studien. Till exempel så har X Translation arbetat ut en utbildning på temat efterredigering och denna har även erbjudits till externa frilansöversättare som utför denna typ av arbete.

Översättningsbyråerna verkar inte ha mött något allmänt motstånd från sina översättare vid införandet av maskinöversättning och efterredigering utan reaktionerna kan snarare sägas ha varit blandade. Alla intervjupersonerna uppger att det funnits vissa översättare som har varit negativa till maskinöversättning eller till och med vägrat att befatta sig med det, men att andra översättare, i synnerhet från den yngre generationen, har välkomnat maskinöversättning som ett hjälpmedel. Cadwell et al. (2018) rekommenderar att det är viktigt att integrera översättarna i förändringsprocessen från första början för att de ska acceptera tekniken. Intervjupersonerna är inne på liknande tankegångar där de betonar vikten av rimliga villkor och att alla parter är med på taget.

6 Slutsats

Syftet med denna uppsats har dels varit att studera hur översättningsbyråer använder sig av maskinöversättning i sina översättningsprocesser och dels hur de hanterar de risker som uppstår i samband med detta. Dessa två dimensioner är delvis sammankopplade eftersom det är svårt att säga något generellt om vilka risker som uppstår i samband med maskinöversättning. Snarare är det så att olika typer av användning kan ge upphov till olika typer av risker.

Studien visar att implementeringen av maskinöversättning i översättningsbyråernas översättningsprocesser inte följer någon lagbunden ordning. Istället är olika användningsscenarier tänkbara och det framstår som att det är kunden och kundens behov som i stor utsträckning avgör hur maskinöversättning används. Helautomatisk översättning förekommer än så länge endast i mindre skala på översättningsbyråerna som ingår i studien. Den vanligaste tillämpningen av maskinöversättning i översättningsprocessen verkar vara att den spelar rollen av ett kompletterande stöd till tidigare hjälpmedel som digitala översättningsminnen och terminologiverktyg.

Ingen av de tillfrågade översättningsbyråerna gjorde någon formell, detaljerad riskanalys i samband med att maskinöversättning infördes i organisationen. I de fall riskanalyser gjordes var de snarare beteckna som informella då de inte byggde på någon särskild standard för riskanalys. Detta innebär dock inte översättningsbyråerna ignorerade riskerna med maskinöversättning utan informanterna framhävde i synnerhet att man lade stor vikt vid datasäkerhetsaspekter när man valde eller utvecklade en maskinöversättningslösning för översättningsprocessen.

Att maskinöversättningssystem kan ge upphov till vissa typer av felöversättningar var något som även framkom i studien, där informanterna hade erfarenhet av olika former av återkommande fel i maskinöversättningarna. Problemet med felöversättningar kan hanteras på olika sätt. Den vanligaste åtgärden är att låta kvalificerade översättare efterredigera texten. Ett annat alternativ är undvika att använda maskinöversättning i översättningsprocessen för vissa typer av texter där den anses vara dåligt lämpad. I vissa fall, t.ex. vid snabba leveranser av helautomatiska översättningar, kan även en högre felfrekvens tolereras om kunden är med på detta.

Översättningsbyråerna i studien har valt olika vägar för att garantera att maskinöversättningssystemet uppfyller de krav som ställs på datasäkerhet. I de fall där färdiga lösningar köpts in har datasäkerhetsaspekter spelat en viktig roll i upphandlingsprocessen. En av översättningsbyråerna har utvecklat en egen intern maskinöversättningslösning som bygger på att översättaren utför översättningen i ett slutet system, vilket förhindrar dataläckage. Det framkom även att samtliga översättningsbyråer tillämpar en strikt förbudspolicy gällande frilansöversättares användning av kostnadsfria översättningstjänster eftersom dessa visat sig ha vissa problem med säkerheten.

Det finns inga tecken i studien på att översättningsprofessionen som sådan nedvärderats efter att maskinöversättning införts. Översättningsbyråerna tillämpar i princip samma kriterier när de tilldelar efterredigeringsuppdrag som när de tilldelar översättningsuppdrag. Intervjupersonerna säger sig ha upplevt att maskinöversättning mottagits på ett blandat sätt av översättarna. Dock uppvisar informanterna en samstämmighet om att en lyckad implementering av maskinöversättning förutsätter att man tar hänsyn till översättarnas intresse i samband med detta.

Slutligen är det väl på sin plats att konstatera att denna uppsats bygger på en relativt begränsad studie som förhoppningsvis bidragit med en viss insikt i hur maskinöversättning används och hur risker förknippade med denna användning hanteras av översättningsbyråer, men som endast kan betraktas som en översiktlig empirisk introduktion till problemområdet. Framtida forskning bör rikta in sig på fördjupade studier av de olika delområdena, t.ex. olika tillämpningar av maskinöversättning i översättningsprocesserna och mer grundliga studier av de enskilda riskerna kopplade till maskinöversättning.

7 Bilagor

7.1 Bilaga 1 – intervjuguide

Inledning

- Är det ok om jag spelar in intervjun?
- Kan du beskriva din roll på företaget?
- Skulle du kunna ge en kort beskrivning av företaget (storlek, position, inriktning)?

Maskinöversättning

- Hur länge har ni använt maskinöversättning i översättningsprocessen på ert företag?
- Vad för typ av maskinöversättningssystem använder ni er av (statistiskt/neuralt)?
- Använder ni er av ett egenutvecklat system eller en färdig lösning?
- Ungefär hur stor andel av alla era översättningsuppdrag involverar maskinöversättning?

Användning av maskinöversättning i översättningsprocessen

- Utför ni några översättningsuppdrag baserat på helautomatisk översättning utan mänsklig inblandning?
- Erbjuder ni lätt respektive full efterredigering som särskilda tjänster?
- Vilka typer av texter anser du lämpar sig väl respektive mindre väl för maskinöversättning?
- Finns det någon typ av texter eller dokument där maskinöversättning aldrig används?
- Hur bestämmer ni om maskinöversättning ska användas för ett uppdrag (kriterier)?
- Hur bestämmer ni om uppdrag ska utföras med maskinöversättning och efterredigering?

Felöversättningar

- Har ni noterat att maskinöversättningssystem ger upphov till vissa typer av felöversättningar?
- Vad för typ av fel kan detta handla om?
- Vidtar ni några särskilda åtgärder för att motverka dessa fel?

Maskinöversättning och datasäkerhet

- Har ni uppmärksammat några datasäkerhetsproblem vid användning av maskinöversättning?
- Om ja, vad vidtar ni för åtgärder för att åtgärda dessa problem?
- Vad har ni för policy rörande användning av externa, kostnadsfria översättningstjänster (t.ex. Google Translate)?
- Tillåter ni att frilansöversättare använder sig av maskinöversättning vid vanliga översättningsuppdrag?
- Om inte, vad gör ni för att motverka detta?

Efterredigeringsuppdrag och kompetenser

- Hur bestämmer ni vilka personer som ska anlitas för ett efterredigeringsuppdrag (kriterier)?
- Anser du att en person som utför ett efterredigeringsuppdrag bör ha några särskilda

kvalifikationer?

- Hur ställer sig översättare till maskinöversättning enligt dina erfarenheter?
- Har ni upplevt att det är svårt att hitta kvalificerade översättare som vill ta på sig efterredigeringsuppdrag?
- Erbjuder ni någon form av utbildning för personer som jobbar med efterredigering?
- Måste man ha genomgått denna utbildning för att få utföra efterredigeringsuppdrag?

Riskhantering och kvalitetssäkring

- Har ni gjort någon form av riskanalys i samband med implementeringen av maskinöversättning i översättningsprocessen?
- Om ja, kan du beskriva hur denna såg ut?
- Är företaget certifierat för riskhantering enligt ISO 31000?
- Är företaget certifierat enligt andra ISO-standarder (t.ex. ISO 18587 för efterredigering)?
- Har ni vidtagit några andra åtgärder för att minska riskerna i samband med maskinöversättning?
- Har du stött på några andra risker i samband med maskinöversättning som vi inte tagit upp?

Avslutning

- Är det något mer du vill ta upp?

7.2 Bilaga 2 – intervju 1

Andy (pseudonym), Technology Manager, X Translation (pseudonym)

Simon: Could you start by telling me about yourself in a few words? About your history in the translation industry and so on.

Andy: Yes, so I am a translator by training. I have a BA in translation and interpreting. I am originally Italian, but I have been living in the UK for a good third of my life. Mostly because I started my career in translation here and then I sort of stayed and developed. I have worked mostly as a project manager/translator for the first couple of years of my working life, in a smaller company and then I moved into my current company, X Translation, to then do more project management and then continuing into a more focused role which was essentially account management and localization management, specifically for clients that wanted software localization and more technically minded clients, let's put it this way. A few years back, due to the needs of the company and due to positions that became available, I moved into the sort of technology manager role that I have now, in the sense that I manage technology. I am not the manager of the technology team. I report to someone. But I took over the bulk of the machine translation strategy from the person that left. And ever since 2016, more or less, I have been dealing with the strategical development, training, staff training, translator training and research on the machine translation resources that the company employ.

Simon: OK, can you say something about the history of machine translation in your company. How long have you been using machine translation in the translation process and so on?

Andy: Yeah. So, as I said, I took over from someone that sort of brought forward the idea of machine translation here in around 2013. So, we are about nine years in, more or less. At the time, it was very much still a budding thing. We were sort of dipping our toes into the wider topic. So, the lady, that was there in charge of it, kind of laid the groundwork for the foundation of the programme. I kind of owe her a lot, particularly when it comes to metadata and the classification of the data, that we produce and the translation data that we have. She did a lot of work that way, that I could then simply reuse and (*tweak*), rather than having to start from scratch. We started with statistical machine translation at first, because that was the trend of the time. It was new, well, it wasn't new, but the commercially available option. And then coming 2017, more or less, due to very much data problems, we had to switch to a different provider. At the time, the provider started offering neural machine translation, so we decided to switch to that. And, we are still using that. And we are being, well... basically we have been employing neural machine translation in our daily workflows for about four years now with the current state being the most reliable that we found so far.

Simon: Maybe you can say something about the percentage of translations made with help of machine translation and in how many cases you don't use machine translation, so to speak.

Andy: Ok, so this also a very topical topic (laughter)... sorry for the word salad... to discuss. Because it is kind of changing. We are in a bit of a particular situation in the sense that we do a lot of work sub-contracting from other agencies, which means that a lot of the processes sometimes are not necessarily within our control. We do the work, rather than actually organizing it. That said, the policy that we have company-wide is to use machine translation

whenever we have the option to do it. So, I think the highest we ever had was about 80 percent of tasks done with machine translation. These days it's lower in a bit, mostly because our clients want to be in charge of it and often, they have their own machine translation systems. So, we cannot exactly use ours because they have theirs. So, if you are talking about just using any machine translation system, ours or the client's, it is probably around 90 percent anyway. When it comes to using our own systems, it is getting to around 70 percent, more or less.

Simon: In terms of translation services, are you providing any fully automated services without any human intervention?

Andy: No, we are qualified by the ISO standards 17000 and 18587. So, both in terms of quality management systems and the post-editing standards that we comply with we have to have at least one person involved in every step. Even if it is just a one-step process, which is translation only or post-editing, there is at least a qualified native speaker with translation experience doing that step.

Simon: And are you offering both full post-editing services and light post-editing services?

Andy: Yes, we are. I have to say that, in my experience, a lot of people offer light post-editing service, but very few people buy it for what we have seen. It is something we offer very happily, but the vast majority of the work we are being asked to do machine translation wise is always being full post-editing.

Simon: What is your experience with customers, when they are coming to you and wanting to have a translation made? Have they a good awareness of machine translation and its performance or rather not?

Andy: In the nicest way possible, absolutely not. They don't know anything about it. They have heard that it is cheap. But there is a lot of misinformation still to this day going around, not just in the sense of clients being new to the translation industry and buying localization services for the first time, but also with even agencies that haven't dealt with machine translation much and are trying to understand it any better. We've been active for almost ten years now, so I feel that we are, you know, also certified... so I feel that we have a good understanding of it. Most of our staff have been trained on it, when it comes to translation and project management as well, but with clients there is always a very, very lengthy conversation that we have to have. Generally, between me and our sales people to try to educate them as much as we can.

Simon: Can you say something about what kinds of texts that are well suited for machine translation? Are there some specific kinds of texts or areas that you think are better suited?

Andy: Yeah, in our experience we find that anything that is extremely repetitive works really well. Obviously, because machine translation works with pattern recognition so anything that tends to be iterative and reiterative tends to work well. More specifically, anything like end user license agreements, code of conducts, human resource policies, contracts, terms and conditions, and the world has those in spades, so there is always a need for that. Whereas, on the other side of the spectrum, there is usually very carefully crafted texts like marketing campaigns, slogans and branding, documents that tend to rely on, you know, play on words, puns,

subtexts, contexts, memes, anything these days. Machines don't have any sense of humour so you cannot really expect it to express itself. Also, from a more dull point of view, other texts that are really, in our experience, not very suitable tend to be pharmaceutical and some types of financial documentation that are not very repetitive, mostly letters to shareholders for example. Pharmaceutical texts are a bit of a special case, because they are so reliant on templates and rules and regulations that you have to follow. And there is a very, very strict quality process, that generally adding machine translation as another variable does not seem do it much good. Also, there are legal implications when something goes wrong in the process so a lot of people tend to prefer to refrain from it and only use specialists, trained people, who have only experience with those kinds of texts and that is it. I would say that, even in pharmaceutical domains, I have seen a lot of progress being done recently, also due to the situation we are in at the moment, where a lot of work has been put into automating a lot of pandemic-related documents. But there have been entities like TAUS for instance, and the EU has been pouring a lot of work into solving this. But I still think, particularly from a buyer's perspective, when you are dealing with heavily regulated texts from a financial or legal point of view, even patents for example, you have to be very careful about what you are using.

Simon: I think you already answered this question partly, but are you using some certain criteria when you are deciding not to use machine translation?

Andy: Yes. We have a checklist so to speak. It is partly also via ISO standards, but the first thing that we check is if the client is ok with it. Some clients have very strong views on machine translation based, I am sorry to say, purely on prejudice. Probably bad experiences because they were trying to be cheap on something and then they got, you know, scammed on some provider that was not really honest with them and then they believe now that everything machine translation-related is terrible. That said, we still want to collaborate with them. So, if they express a strong dislike for it and don't want us to use it, we don't. And we respect their wishes. Then we make a sort of informed choice, when it comes to domains and text types. As I said, if we don't think that that particular domain is helped by machine translation and it might make our translators' life harder, we refrain from it. And the last thing is data security considerations. Some clients are very, very, very strict on the way they want their data used. If they stop us from using any sort of machine translation because of that, that's fine. We respect that. Technically that is not a concern from us, because our systems are secure. But people have opinions and ultimately, we have to sign contracts with them so we have to respect that.

Simon: Let us talk a bit about possible or potential translation errors that can occur when you are using machine translation. Have you noticed in your machine translation systems that there have been some certain kinds of errors that have been coming up when using machine translation?

Andy: Yes, this is actually the very focus of one of our training sessions when we've got new translation staff starting. The main point there being that translators have been trained for years on using translation memories, which are essentially human translations repurposed for whatever texts we are using. Whenever they deal with machine translation, if they had previous experience as translators, they tend to find it really difficult mentally, even though they know it is a different source, to differentiate between the translation memory and a machine translation match. It's much easier for newer translators that come straight from university or into the market now to adopt it directly to machine translation because they've never known anything else. In terms of straight error types, when dealing with machine translation, consistency is an issue, as in maintaining the same sort of terminology throughout the text. But it

really depends on the engine that you're using too. Then, some language standards but it also depends on the level of training that you've done. We found this most of our clients' machine translation rather than our own cause put in quite a lot of work into that, but you can have anything like spaces or punctuation, marks not applied correctly, formatting, often even if it's like maintaining bold or italics in a text. It's not much, but still in one more click that you have to do. So, it does take time over the translation process. And then, there is the sort of... the fact that machine translation seems to handle longer sentences in a much more stilted way than it does shorter sentences. It's much more geared up to a subject, verb, object kind of clause that is self-sustained and doesn't rely on subjective clauses rather than a very long-winded phrasing that has indentations and, you know, inserts, quotes or anything like that in between. This becomes particularly important for some of the languages we do, like German or Danish, where... I mean I'm sure you've seen enough Danish yourself to see that, you know, it tends to be much more long-winded. Swedish, on the other hand, is ideal for that because it's mirrors the structure of English very well, and the only thing you have to really worry about is the word order and the position of certain words. But Danish and German, for example, given how comma-based they tend to be, tend to struggle a bit more.

Simon: Do you use any special measures to deal with this kind of errors?

Andy: Yes, so we do a combination of automatic measurement and human evaluation. The platform that we use, I checked with the managers so I am free to say that we do... the provider which we are using is Microsoft. And they do have a series of automatic metrics that are available in the platform, when you're training engines. And the human evaluations that we use is basically a model that we have created internally to do an evaluation of a sample between generally the source text, the raw MT output and the human translation of it to compare the two and see how it would end up and assigning a weight-based evaluation to each of the errors based on, you know, a critical major or minor sort of category.

Simon: Okay, let's move on to data security, maybe. You've already brushed upon it a bit, but maybe you can tell me a bit more about if you have experienced any data security issues when implementing or using machine translation in your company.

Andy: Yeah. So well, there is plenty to say. The major issue that we have personally, at X Translation at least, is the importance of keeping our client data safe. Back in the day when we initially started, it wasn't very clearly stated by a lot of providers of machine translation services, whether or not any sentences translated with their technology would be retained by them or not. Because we work with a lot of brands and high-profile clients, often it is stated in their service level of agreements and their contracts whether or not this is a something that they care about. Generally, they do. They're getting more and more savvy each day. So, one of the fundamental things that we did when we switched providers was exactly that. Finding one that was completely encrypted, secured and of which we were in complete control. We tried a couple and then we ended up with Microsoft for, I would say for two main reasons. The first one is that if at any point we decided to delete everything and remove a profile and destroy our engines if we wanted to, they retain nothing. So at least I know that anything we've trained our engineers with, anything that have been sent through it is not stored, is not kept, and it's simply forgotten. The second one is that you can choose the location in which your servers are based, which makes them GDPR compliant, should we need to. And given that most of our market is Europe based, we had to and it helps us bypass our problems, especially the headache that was Brexit and still is. On the other hand, I would say, I mentioned that we had to switch providers for a question of data security. The reason was that the

provider we were using initially for statistical engines got acquired by a client of ours, which meant that all of our training data and all of our resources were now in their hands. That was not acceptable, so we had to essentially ask them to remove everything and delete everything, which they did. But it meant that we had to rather quickly try to find an alternative before reassessing our entire data structure, basically. I would say the last thing that we are very, very careful with is the fact that whenever we are using our own machine translation or the clients' machine translation as well that our external collaborators, the freelancers that we employ for translation work when our internal translators are not available, also do not use machine translation of their own. We're not against it per se. In fact, sometimes we allow it, provided they explain to us everything about it. We encourage people to be autonomous and well-trained on machine translation. It's all well and good. The problem is that we have to still be in control of our process because we have a responsibility to our clients. We have had cases in the past where some freelancers didn't admit initially to using their own machine translation, which turned out to be a secure one as well, so it wasn't really a problem per se, but we need transparency, otherwise that becomes a problem.

Simon: There have been cases, as you know, with external translation services like Google Translate and so on, where data are leaked into the public sphere of the Internet. What is your policy towards freelance translators using, say, Google Translate or other free external translation services?

Andy: Yeah. So, they are explicitly forbidden from using any of that in the terms and conditions that they have to agree with when I start working for us. And should they be found doing this, generally we end up terminating our collaboration with them because it's simply not acceptable. We've had cases where people have come to us and told us: Listen, I have my own personally tried and secure MT engine that I use and I've been using it for a couple of years and it is very good and it lets me work much faster. I would like to keep on using it. In general, we are open to the idea, as long as people are honest with us and they tell us straight away, and if the systems they are using are secure and personal.

Simon: Okay, I would like to ask you a few questions about how you're handling post-editing jobs and how you assign that kind of jobs. You talked a little bit about it previously and you're only using a certain group of people for that, mostly like linguists and translators, right? And you are also certified according to the... there is a special ISO standard for post-editing ... are you certified according to that standard as well?

Andy: Yeah, we've got a certification for the ISO 17100 and the 18587, is it called? Yes. Both for quality management purposes and specifically post-editing and machine translation quality management. This is because before being a company that employs machine translation, we were a translation company and we still are, so we've always only employed native speakers of the target language that we ask them to translate into. And we only employ people that meet the ISO quality standards requirements, which means they have a certain amount of years of experience in translation, or at least a relevant title linked to it, a degree or a master or anything like that. And even more so if they got a proven track record of the work they've done, and particular certain specializations. Generally, we have a very strict database of freelancers that we categorize by domains and text types, where people can decide what kind of texts they would like to be contacted about. For certain more specific domains like legal, financial and medical, we also do tests, meaning that if they claim to be specialists of those fields, we assess that. For the reasons I mentioned earlier, there are legal repercussions if one of those projects goes wrong. So, we need to be sure that we are providing quality that is up to

par. And generally, there's a testing period of about six or seven real projects to see if we can collaborate properly and if the quality is up to par. If it's not, if they're open to feedback, and if we see an improvement going on, we have a large in-house team of translators that are being trained extensively into this. So, we rely on them specifically being the sort of official reviewers and assessors of this. And the entire database is divided into categories based on, I would say, the experience we have with them. So, for example, if you got a certain number of translator names preferred because we worked with them for years and years, they are reliable, they are good, they produce good quality and then you might have backups for whatever reason, or maybe having like a red flag because a recent project hasn't gone well. So maybe to add an extra layer of revision is necessary, you know. Basically, this is something that we've always taken very seriously, and when we switched into a more machine translation-oriented workflow, we carried it over specifically on that. I would say we have really, really struggled to begin with post-editing training. Simply because there isn't really much around. We tried to find resources online. We were TAUS members for the longest time, but to be completely honest, I don't really find the TAUS material to be very useful. It's a bit too vague. And we very much believe that people learn by doing real life examples. So...

Simon: Sorry, but you do offer some kind of education or training for potential post-editors...

Andy: Yes, that was what I was getting at. Basically, we had to develop our own because we couldn't find anything online, so we ended up creating a bit of an extensive introduction programme for the first two weeks and then continues for three months when someone joins the company as a translator.

Simon: As an in-house translator?

Andy: Yeah, and we essentially first have them work on a few practice jobs and get revision and then feedback and then we start with real life projects and then they also get feedback on that. Usually someone is considered junior for the first year here and based on the on the productivity metrics that they manage to achieve and the quality they manage to produce then they get promoted as, you know, translator then senior and so on. When it comes to freelancers, we have done quite a lot of extensive post-editing training in the past. I'm sorry to say that we haven't been able to keep it up much over the last year or so due to the general chaos that's been going on with the pandemic. So, we have had other things to worry about. But we do recommend whenever we have a freelancer that maybe is new or wants to get into this: go to certain resources that we have found or we've produced that we believe to be useful for their own reference. And we are in a very close collaboration with the ATC, the association of translation companies in the UK, to develop this kind of framework and help it make it available to translators when possible.

Simon: But is it a requirement from your side for post-editors, external freelance post-editors, that they should have gone through some kind of training or education, or?

Andy: So, the ISO standards for machine translation and the ISO standard for translation actually have the exact same requirements, so there isn't a difference in that because after very, very lengthy discussions between the certification body, the ATC and a number of experts in the industry. I was only listening to that one, but one of my colleagues was participating in the discussion as well. It was basically accepted that the skill set for a translator and the skill set for a post-editor are by and large the same, with some minor differences in the kind of tasks

they employ, but ultimately attention to detail, language knowledge and ability to write are the three key components for everyone. So, as long as someone fits the requirement to be a professional translator in our own standard, then they generally comply with all the requirements that we have.

Simon: The main reason I ask is that there has been reports of certain kinds of errors, primarily by neural machine translation systems, that are quite difficult to recognize for translators or proof-readers or post-editors, and I'm a bit interested in... you're obviously aware of that... how you handle that kind of problem in your company.

Andy: Um, so for internal staff we have training sessions specifically on that because it is quite a big topic. It's about like an hour and a half session, but then we got plenty of material that they consult every time. We have a wiki where we keep a lot of material, including some sort of frequently found errors, if you want, just so people can keep an eye on this kind of instances. I also use it a lot on feedback when it comes to a report on the quality of the engine, seeing if I can fix something with further trainings and when it comes to freelancers, when we start onboarding them and we start doing the first few projects with them, we generally give them feedback from our own senior */*kort avbrott i förbindelsen*/* check if they have some bits and pieces they could improve on, or some errors that have recognized that they, you know, could have done better.

Simon: Regarding the attitude from the side of translators towards machine translation, have you experienced some kind of resistance, or? What's your experience?

Andy: Plenty. Or, a bit of both really. It really... not to sound disparaging of this... but it really comes with age. The more seasoned translators seem to find it very difficult to adopt. Extraordinarily so in some cases. And the newer ones, particularly fresh from university, seems to find it very easy to simply start working with it. Often to the point of when you take it away for whatever reason, then they just going like, oh God, I have to think about a sentence from scratch now. I haven't done this in forever. The main thing is that ultimately, we believe machine translation to be a tool that people need to choose to use. We have a few translators, there's one particular sample that comes to mind, there is a lady that has been working for us for 30 years. She's very good and very skilled. She only deals with a couple of very specialized domains, so mostly financial and medical. And she has had experiences with machine translation and then simply decided she's not going to use it. She's vehemently against it, and that's fine. She has her own client base. She is aware of the fact that this means a reduction in work that we can offer her because most of our jobs are machine translated. She's okay with it. She's got her established client base so it doesn't affect her financially too much. This means, we continue cooperating, so we offer only projects where machine translation isn't necessarily a component. There are plenty of examples like that. Ultimately, it's a question of cooperating and talking and agreeing on the finer points. As I said, it depends on the person, on the text, on the situation, on the context, on a number of variables.

Simon: Okay, so just finally I want to ask you a few questions more explicitly about risk management in connection to machine translation. Have you used some kind of explicit risk analysis in the process of implementing machine translation?

Andy: It depends on what you mean. Meaning like a service or an audit or...

Simon: Yeah, or some kind of explicit process... or how did you deal with risks

connected to machine translation when you implemented the machine translation system in your company, so to speak.

Andy: Okay, so we did a fairly thorough... when we were essentially trying to find a new system to move into and even with the first one to begin with ...we did a fairly thorough review of the offerings that these providers were trying to give us and decided on the ones that were meeting our criteria. The main one was client security client, client data security basically, privacy of the overall environment and data retention logics and processing. It was a checklist that we created ourselves there were, there wasn't really anyone else want to contact, and ultimately because we weren't creating engines ourselves directly, but we're relying on out-of-the-box solutions, it became fairly easy to simply ask them for their policies and their data regulation content and rules. And pick and choose, essentially. We ended up with Microsoft for a number of reasons. Data compliance is one. The fact that we are a very Microsoft heavy company when it comes to the management of our own accounts and systems. So, it was already easily integrated in all our systems where we already use multi-factor authentication. All sorts of cyber security protection. There's the question of accountability, meaning that we can rely on them to fix their own problems when it happens. And that it requires a lot less input on our side, because we're just buying the solutions from them already made.

Simon: There is a specific standard for risk management as well, as you may be aware of, 31000. Are you certified according to this standard as well?

Andy: No, but we're looking into it. It's something that's... It's... With the move to digital, when we have to move everything into a remote sort of environment. We had quite a lot of challenges. We managed, but in the end, you know, it was a learning opportunity for everyone and we have identified a few points of vulnerabilities that we definitely want to address, not just related to machine translation, but to the company as a whole. So, in the last few months we've actually done quite a lot of work... We have actually got an audit from our risk assessment analyst on our entire structure, not just machine translation. So, we are working towards that eventually, yes.

Simon: OK, are there any other risks connected to machine translation that we haven't discussed so far, that you have experienced in the translation process or in your company in some way?

Andy: I wouldn't say so. To me, the biggest ones are always, you know, the ownership of the data and what happens with it. We have specific clauses when it comes to... because there is one key thing that has always been a bit difficult to define... So, we have a specific clause in our terms and conditions when working with everyone, clients and translators which means that when you are carrying out a task, who is the owner of the content that is being translated. Because technically we did it, so we are the owners of it. That said, obviously clients may want to retain ownership of it, but we found that that's not the case. Often, they only care for the finished product, not anything that has been ongoing. So, we have explicitly said that, in a lot of our terms and conditions, that unless otherwise stated, we retain ownership of our own translations. Nothing to do with, you know, translation memory contents, that the client might have, that's completely separate and nothing to do with their own machine translation. But if something originates from our own machine translation and our own translations, we consider it, I forgot the exact legal expression, that is used in the terms and conditions, but is basically a by-product of the task. It's something that we created as part of the request they have made. So that retains... that we retain the ownership of that. Which means we can also use it

to retrain our own engines and we use it when we have to. In that cases where clients have not been OK with it, which means that I simply scrapped their data from our engines. Again, because I can, so Microsoft enables me to do this. So, it's very simple, which is also one of the reasons why we went with Microsoft. Data handling is extremely easy.

Simon: Ok, I think I have already asked the questions I have. Is there anything you want to add that you think is relevant for this topic that we haven't really discussed so far?

Andy: No, I think you covered quite a broad amount of topics.

7.3 Bilaga 3 – intervju 2

Jan (pseudonym), IT-ansvarig, Y Translation (pseudonym)

Simon: OK, skulle du kanske bara kunna börja med att berätta lite kort om din roll i företaget i allmänna ordalag?

Jan: Yes, okej, jag heter Jan Palm, jobbar som IT-chef eller, ja, IT-ansvarig på företaget. Jag har varit på företaget sedan 2001 och har jobbat både med teknisk säljsupport, kan man säga, när vi tecknar nya kunder och försöker hjälpa dem att förstå problematiken med översättning och hur man på ett mer maskinellt sätt ska hjälpa översättare att få en bra kvalitet och andra typer av delar då också. Samtidigt som jag har en IT-avdelning på ungefär 36 personer, varav en stor del är utvecklare och jobbar med att ta fram program och den typen av språkinriktade program då som hjälper översättare, som till exempel CAT-verktyg eller maskinöversättning eller liknande saker då, eller även administration kring det, som terminologi i systemet.

Simon: Skulle du kunna bara säga några ord om företaget också allmänt sett? Om inriktningen kanske, om ni främst använder er av inhouse-översättare eller frilansöversättare och så vidare.

Jan: Vi använder oss mestadels av frilansöversättare. Vi har ett stort frilansnätverk eftersom volymerna kommer och går lite uppåt och nedåt. Det är väldigt svårt att hålla det på det sättet, men man kan säga, även fast vi jobbar med frilansöversättare ganska mycket, så har vi ju ett mindre, eller mindre antal, men om vi säger att vi har ett nätverk på kanske 5000–6000 översättare runt om i världen så kanske vi jobbar regelbundet med 500–600 utav dem som vi likasom, mer eller mindre som inhouse-personal blir det ju, för att de får ju hela tiden uppdrag av oss, sen nu under coronan så har det kanske varit lite si och så beroende på hur kunder och allting runt omkring har varit. Men det är väl det som är det normala, kan man säga.

Simon: Är ni inriktade på några specifika språkområden, eller?

Jan: Vi håller på rätt mycket med tekniska översättningar. Vi är ganska stora mot fordonsindustrin och så, men har även mot institutionella delar som Europeiska unionen och andra typer av institutioner inom Sverige och myndigheter och så också. Men mestadels skulle jag säga att det är teknisk översättning.

Simon: Om vi då går in på det här med maskinöversättning lite mera. Hur länge har ni använt er av maskinöversättning i översättningsprocessen på ditt företag?

Jan: Man kan väl säga...aktivt skulle jag säga i 3 år lite drygt. Vi har varit, kan man säga, initialt försiktiga till själva tekniken. Och det har väl egentligen varit av två anledningar, dels har... när man håller på med översättningar som vi har gjort under ganska lång tid då så använder man sig utav CAT-verktyg med översättningsminnen och när man har en ganska tight relation med en kund och man har hållit på länge, då har man ju byggt upp en ganska stor mängd av översättningsminnen som genererar en ganska bra återanvändning egentligen utav det, framför allt inom tekniska översättningar då, vilket gör att man, kan man säga, har man hållit på och översatt åt en kund i 10 år och det är en liten förändring i en produkt eller någonting, då kanske det är 2–3 procent som behöver uppdateras utav hela volymen av informationen egentligen eftersom allting finns ju redan översatt sedan tidigare och kvalitetssäkrat et cetera och det är utav de där 2–3 procenten som man eventuellt skulle kunna göra en vinning

utav att gå in och göra maskinöversättning för att resten hämtas ju redan från det tidigare. Och då var det en sån liten del i förhållandevis till det andra, så frågan är hur mycket investeringen och nyttan utav det skulle bli förhållandevis till det andra. Så att det har ju varit, kan man säga, incitamentet för att kanske inte satsa så mycket på det initialt. Men nu i och med att tekniken har mognat... att det har kommit... att de har blivit bättre med översättningsmotorerna. Det är "closer time to market" när man ska försöka leverera någonting. Man försöker göra allting nu för att både spara pengar och skapa en högre kvalitet och försöka få det snabbare, självklart. Då måste man ju ta till alla knep som man kan för att man ska kunna göra det annars blir man omsprungnen av konkurrenter också självklart. Och det efterfrågas av kunder för att konkurrenten är där och knackar på och säger: vi kan sänka kostnaderna med 30 procent. Och det är 30 procent utav de här 3 procenten vi pratar om. Och det är liksom det där som är svårt då, det att pedagogiskt förklara för en kund...

Simon: Och vad är det för typ av maskinöversättningssystem ni primärt använder då alltså. Du behöver inte gå in i detalj, men handlar det primärt om statistiskt baserade eller neurala system, eller?

Jan: Man kan säga så här. Först så köpte vi ett system som heter SYSTRAN. Men det var väldigt, väldigt dyrt och fungerade, tyckte jag, sådär, just när vi höll på med det då. Och då var det en hybrid kan man säga. Det var både statistiskt och, kan man säga, en språkmodell som låg i grunden för det då, som var själva huvudmotorn i det då. Sedan valde vi då, när vi tyckte att det inte funkade... det funkade, vi lärde oss ganska mycket utav det när vi höll på att testa och utvärdera det här systemet och så kände vi att, nej, men för att vi ska kunna vara konkurrenskraftiga och göra någonting bra utav det här så valde vi att själva utveckla en motor baserad på Moses då som är den här motorn som är... finns gratis kan man säga... runt om, men som var något EU-finansierat projekt, har jag för mig, som låg bakom det där. Så vi började att titta på den och började titta på hur man skulle kunna göra med det och tog fram en egen motor kring det. Så vi har utvecklat en egen maskinöversättningsmotor som vi själva kan skruva och "tweaka" hur vi vill. Och vi har väl, kan man... alltså om man tittar på den akademiska världen så tittar man ju ganska mycket då på så här... vi ska översätta en text. Hur gör man det generellt sett och sedan så jobbar man på det. Men vi försökte på något sätt titta på det istället och säga, okej, vi nischar oss mer mot en domän. Motorn på en bil eller det här på någonting annat och försöker träna motorerna så att de ska förstå just det och ingenting annat för att öka kvaliteten och då såg vi att vi fick en bättre träffbild då. Självklart så finns det massa utmaningar med det här. Om en författare inte har ett bra terminologistöd eller vet hur man ska skriva en text, en originaltext, så påverkar det ju självklart väldigt mycket hur resultatet blir utav en maskinöversättning. Så att det finns också vissa förutsättningar som måste uppfyllas för att man ska kunna säga att man uppnår en god kvalitet eller inte en god kvalitet på det här och hur det här slår är extremt svårt att förutspå. Alltså. Det är det är lite lotto, liksom varje gång. Man vet inte. Och sedan är det också vilka språkkombinationer... vissa språk är lättare att göra det till, medans andra är svårare att göra det till beroende på hur det ser ut.

Simon: Det hänger kanske samman med min nästa fråga här då som handlar lite om... finns det någon särskild typ av texter där ni märker att det fungerar bättre med maskinöversättning respektive finns det vissa andra typer av texter där det inte fungerar alls så att säga?

Jan: Man kan säga så här. Marknadstexter som ska ha en ganska... jag vet inte vad man ska kalla det för... en ganska kommersiell ton om man säger så... det kanske ska vara mer lokaliserat eller mer marknadsanpassad information... skulle jag säga är inte riktigt lika lämpat för

det, för att det är inte då en copywriter eller människa som sätter sig ned och försöker skriva den här typen av text. Medan om man tittar på mer en verkstadsmanual eller en mer... kan man kalla lite mer torrare information då med lite mer maskinellt så här: ”Montera fästelement med 25 newtonmeter” eller någonting liknande... något sådant där... repetitiva texter på det sättet. Det är lämpar sig väldigt mycket mer. Det är enklare att få till då. Men samtidigt så kan man säga så här. Med AI-tekniken nu och med de datorerna som vi också nu har utvecklat och tittar på lite grann så ser vi också att ibland så blir man faktiskt positivt överraskad när man ser en text som en maskin har lyckats producera. Om den är mänskligt hanterad eller om den inte är det, därom tvista de lärde. Men det är ju så... men generellt sett om man liksom ska försöka svara på frågan. Då skulle jag säga det att det som är lite mer copywriting-hanterat eller som den typen av material är, tycker jag, sämre lämpat då. Men det där kan ju ändras till nästa dag. Ja, det beror på och sedan är det då självklart också olika språkvarianter. Vissa språk är väldigt mycket svårare att göra det till i förhållande till det och det kan ju också vara att det finns mindre material att träna på också när man försöker göra så.

Simon: Har du någon uppfattning om i vilken utsträckning ni använder maskinöversättning? Ungefär för hur stor andel av uppdragen? Är det mer än hälften eller är det 25 % eller 75 %?

Jan: Man försöker göra det... alltså i början så gjorde man det på tester på en mindre population av kunderna då och tittade på de uppdragen som var lämpliga. Men ju mer tekniken mognar desto mer vill man ju försöka få in det på det mesta. För det handlar ju om, som jag pratade om innan, måste man vara konkurrenskraftig mot resten av den här branschen då. Det finns incitament hos kunden att de vill ha det här för det har blivit ett ”buzz word” på sista tiden, också, att det måste finnas och det ska vara där... man ska liksom ha det så där... man efterfrågar det för att konkurrenter vill att det ska vara sådär. Det är lite ibland ”bells and whistles” tycker jag. Men ibland är det också en nödvändighet för att man ska göra... framförallt om det är ett nytt material. Om man inte har någonting sedan tidigare. Då kan det finnas stora vinster med det.

Simon: Hur upplever ni att kunderna... vad har de för insikt i maskinöversättning när de beställer den här tjänsten?

Jan: Oj! Det är jättevarierande. En del är väldigt intresserade. Jätteintresserade och vill verkligen veta. Andra vill bara beställa och få ett resultat. De skiter hur man gör det egentligen. Så, det där är extremt varierande. Jag kan inte svara på den frågan.

Simon: Utför ni några översättningsuppdrag baserade på helautomatisk översättning, så att säga, utan någon mänsklig inblandning alls?

Jan: Man kan väl säga så här. Det beror på lite grann vad kunden vill ha som kvalitet och vad de har för förväntningar. Och det är egentligen vad man betalar för... förväntningar och ledtider egentligen som man måste förhandla om i sådana fall. Det kan finnas vissa typer av dokumentation som man måste ha väldigt snabb ledtid på. Man kanske inte hinner att få till, men man behöver också på... att deras förväntningar då att det då inte blir en perfekt översättning utan du får ett hum om vad det ska stå där. Lite grann kanske som vi gemene man använder sig utav Google Translate. Man vet inte vad det står på någonting, man slänger in det så får man åtminstone ett hum om det inte. Det kanske inte är helt korrekt, men man kan i alla fall bilda sig en uppfattning om vad det handlar om. Och det efterfrågas ju av vissa kunder för olika typer av produkter. Om man tar inom fordonsindustrin till exempel så finns det

någonting som heter servicebulletiner. Det är när en bil går sönder och sedan så ska man försöka få ut någonting väldigt, väldigt fort till en kund eller till en verkstad och det kanske handlar om 3–4 timmar för att man ska liksom kunna få ut en sådan snabbt. Det blir rätt kostsamt då att ha en bas av översättare som bara sitter och väntar på att det ska komma ett uppdrag och så hoppas de på att nu kommer det och så skulle vi översätta det så fort som möjligt och sedan så sitter man och väntar på nästa och det kommer inte så frekvent. Så där finns det ett användningsområde, skulle jag säga då, om det är den typen av uppdrag. Men det måste också kommuniceras tydligt mot slutkund att det här är det du kan tänkas få. Det här är resultatet. Det här är förväntningen. Ni kan inte förvänta er att ni ska få någonting som är helt perfekt på den ledtiden. Då måste det finnas en större ledtid för då måste en människa verifiera att det här faktiskt är korrekt.

Simon: Den normala processen är alltså då att ni har någon form av efterredigering kopplat till maskinöversättning?

Jan: Ja precis, ”post-editors” som går in och gör det.

Simon: Och erbjuder ni då både det man kallar full efterredigering och lätt efterredigering, så att säga?

Jan: Ja, alltså, det är återigen... Vi kan göra båda och, men det är också lite grann... Det är den här kända pyramiden med kostnad, ledtid och kvalitet och så där och hur man liksom vill få till det så att... Om förväntningen hos kunden är av en viss del och att det är tydligt kommunicerat så kan vi självklart leverera båda och, liksom. Det beror på vad man har och skulle det vara något medicintekniskt så ska det väl granskas tre gånger eller fyra gånger innan man vågar skicka iväg det för att det är de standarderna som gäller då. Så det är helt och hållet vad uppdraget... Ja, vad det handlar om egentligen, det man gör.

Simon: Finns det någon typ av uppdrag eller texter eller domäner där ni aldrig använder maskinöversättning överhuvudtaget?

Jan: Jag skulle säga marknadsinformation eller om kunden specifikt säger att nej, det här vill vi inte få maskinöversatt. Det här måste översättas från scratch utav en människa egentligen. Men man ska ju aldrig säga aldrig... att man inte kommer göra någonting baserat på hur teknikutvecklingen går heller. Så man måste hela tiden vara öppen för att saker och ting förändras så att det blir bättre... att man kan göra saker på ett annorlunda sätt. Det är farligt om man inte skulle vara där heller. Men självklart om en kund kommer till oss och säger så här: ”Vi vill inte att det här ska maskinöversättas. Det här är så känsligt. Eller det här är en pressrelease som ska gå ut i hela världen. Det här vill vi inte att en maskin eller någonting rör överhuvudtaget för tillfället”. Så, då gör vi inte det, självklart. Så att det är ju det handlar ju väldigt mycket om uppdragets natur, kan man säga.

Simon: Och om vi går in på det här lite med risker då och kanske eventuella problem som kan uppstå i maskinöversättningsprocessen. Har ni noterat att maskinöversättningssystemen ger upphov till vissa specifika typer av fel som är vanligt förekommande?

Jan: Nej, jag vet inte faktiskt, det där. Att det blir fel, det blir det ju och ibland blir det ju konstigheter. Det är väl framför allt, tycker jag, när det var den statistiska delen, om man inte hade ett bra terminologistöd... där den inte riktigt visste var vissa ord skulle vara då eller det

inte fanns någonting som man hade lärt... att det var... framförallt specifika ord som var väldigt, kanske, domänspecifika. Någonting som var väldigt...kanske bara en gång som det förekom, men det var väldigt viktigt för det sammanhanget. Då såg man ju...om det var på engelska från början, då bara lyfter den över det ordet på engelska för den visste inte vad den ska göra med det. Den hade ingen direkt översättning utav det. Den visste inte vad det var, så att säga. Och då får man ju försöka hitta den typen av problematik då och bygga upp terminologistöd kring det. Och det finns ju väldigt många olika varianter med "terminology mining" och "harvesting" och allt möjligt om man försöker hitta nya ord och hur man ska göra det och så. Men det är en vetenskap i sig att försöka få ut det, så det kan man väl säga var väl en utav problemen som vi såg det då. Och det är väl återkommande om maskinen inte vet vad den ska göra och om man inte har något bra förslag för det då blir det ju inte bra på just den biten. Medans andra paragrafer som man som översätter då eller man tittar på kan vara extremt bra och man funderar på det om det inte är en människa som har gjort det här, liksom.

Simon: Det har ju till exempel uppmärksammats i litteraturen kring neurala maskinöversättningssystem att de genererar väldigt bra texter som är väldigt läsbara och som ser väldigt bra ut. Men där det till exempel kan förekomma att ett ord har utelämnats eller lagts till som inte hör dit, va? Och att det kan vara väldigt svårt att upptäcka den typen av fel. Är det någonting ni känner igen eller har stött på?

Jan: Ja, och det är ju därför vi kör det här med post-editors då... att man försöker ha någon människa som alltid... innan vi släpper iväg någonting till våra kunder, generellt sett, kan man säga, om inte de har begärt någonting annat då som vi pratade om innan... deras förväntningar och deras ledtider och såna grejer beroende på vad vi har kommit överens om. Men i det generella fallet så är det alltid så att vi använder oss utav post-editorer som får gå igenom materialet. Och det kan väl jämföras lite grann med en "fuzzy match" egentligen som man får när man tittar... när man pratar om vanliga CAT-verktyg då... att det krävs en genomläsning för att förstå att en del "fuzzy matches" är korrekta faktiskt för att det är små grejer som har förändrats. Men det är mer eller mindre korrekt som man måste läsa. Och man försöker då göra... utveckla systemen så att de ska försöka förstå vad för någonting man borde titta på och göra grejer. Men det är svårt. Det är liksom inte... det är språk man har att göra med. Det är inte någon logisk grej ibland hur det ska vara, utan det är det här... det här hantverket som sitter i huvudet hos folk som man ska försöka sätta en maskin på. Det är inte alltid så jävla enkelt.

Simon: Det var en annan intervju jag hade här i veckan där informanten pratade om det här med att många efterredigerare har problem att hålla isär det här med "fuzzy matches" och maskinöversättningsträffar. Det är ändå olika typer av träffar man får där som man kanske bör betrakta eller åtgärda på olika sätt?

Jan: Absolut, och det gör ju vi också. Vi håller ju isär dem. Så, jag menar en "fuzzy match" är en "fuzzy match" och en maskinöversättningsdel är ju en maskinöversättningsdel. För jag menar när vi gör en maskinöversättning så börjar vi alltid med att göra en normal matchning med "fuzzys" och allt möjligt. Sedan den delen som blir kvar, det maskinöversätts. Så att det primära är ju fortfarande att använda översättningsminnen enligt den klassiska översättningsmodellen, eller vad man vill kalla den för då.

Simon: Om vi går in på det här lite mer då med efterredigering kanske. Har du någon uppfattning om hur ni bestämmer vilka ni anlitar för efterregeringsuppdrag? Är det vissa kriterier som gäller där, eller?

Jan: Alltså kriterierna skulle jag säga är snarlika till dem som översätter. Däremot så är det ju så att man måste... alltså det viktigaste jag skulle säga... det är att man förstår ämnet som man översätter, framför allt inom tekniska översättningar då. Så att om vi håller på med borrarriggar för gruvinstrument till exempel eller någonting liknande... så att de vet den typen av... hur en sådan maskin fungerar, vad man gör med den alltså... vad de olika delarna heter och att det är den informationen som kanske är viktigare så att det som står där sedan... att man kan ta del utav det så att man inte missuppfattar det som står i originalet, egentligen. Och jämför mot det. Så att det är väl ett utav kriterierna då, men det gäller ju generellt sett också mot en vanlig översättare eller granskare för den sakens skull. Har man inte kunskap om ämnet så är det väldigt svårt som vanlig lekman, även fast man kan språket, att förstå hur det här funkar om man inte... om det är det som ska granskas då. Men om det är en språklig granskning som ska göras... Är det rätt grammatik? Är det rätt punktuation? Har man rätt stavning? Och såna här grejer. Om det är det man ska titta på då kan det göras utav vem som helst egentligen, men då har man ju inte läst igenom vad innehållet är. Är det ett-till-ett-förhållande mellan originalet och det som är översatt. Eller står det någonting annat? Är det någonting som man syftar på som inte alls har med det att göra eller någonting liknande? Så att det gäller ju... så det är väl de två delarna som är viktigt då. Och sedan har vi ju som kriterier för oss också att man granskar bara till sitt modersmål, kan man säga. Så att... det är ju... att man översätter till sitt modersmål och man granskar det som står i sitt modersmål och inte någonting annat, med vissa undantag, med översättare som har flera modersmål. Men det är ju inte så jättevanligt.

Simon: Så i princip använder ni er av vanliga översättare för de här efterregeringsuppdragen också? Ni har inga specifika personer för detta?

Jan: Ja, jag tror att vi har några specifika som går ut och säljer sig som post-editors som är liksom rent krasst en post-editor. Men jag tror att en post-editor kan också sälja sig som en översättare beroende på vilket sammanhang och vilka glasögon man har på samma person då egentligen. Så att jag skulle inte säga att vi har exklusivt en person som är en post-editor bara. Men jag låter det vara osagt för att det... ja, jag vet inte faktiskt.

Hur har du upplevt att översättarkollektivet, så att säga, ställer sig till... eller översättare i allmänhet ställer sig till det här med maskinöversättningar? Märker du några skillnader mellan... kanske mellan inhouse-översättare och frilansöversättare eller yngre och äldre översättare och så vidare?

Jan: Det är ju självklart en förändring och branschen förändras ju hela tiden. Det kommer nya saker och, jag menar, man måste hela tiden hänga med. Samma sak hände ju när CAT-verktygen kom in i 80-, 90-talet någon gång... när de kom in då. Jag menar alla satt ju och översatte liksom i sin WordPerfect eller vad man hade då och satt och meckade ihop alltihopa därigenom då eller på papper för den sakens skull. Och sedan så kom ju Internet och hela den revolutionen då, med allt som har hänt... så att det gäller att hänga med i tekniken... så visst det knorras väl en del i leden ibland och att folk tycker vad ska jag göra nu... blir jag ersatt av en maskin och liksom hur ska det se ut och så? Och det är ju den naturliga utvecklingen skulle jag säga, egentligen, kring det här. Och hänger man inte med då är man ju ute och då blir man Facit räknemaskiner eller någonting annat så småningom då, så att det måste ju på något sätt ske en förändring. Sedan gäller det då självklart att hitta ersättningsmodeller som är rimliga... som funkar så att det liksom inte... att man utnyttjar någon bara för att det kommer en ny teknik utan att man får en rimlig ersättning för någonting ändå, men att man utnyttjar tekniken för att kunna bli konkurrenskraftig så att det där är ju en väldigt svår balansgång hur man ska... hur man ska göra det då, faktiskt, skulle jag säga. Men det är väl egentligen det som är

det stora problemet, men sedan självklart, det är också en ganska pedagogisk övning att prata med någon att förklara hur saker och ting funkar och göra och så där. Men lyckligtvis är inte jag den personen som ska göra det utan det får Vendor Management och de som jobbar med de frågorna hos oss göra istället.

Simon: Jag tänkte höra lite om det här med om ni har stött på några datasäkerhetsproblem kopplade till användningen av maskinöversättning hos er?

Jan: Ja, man kan säga så här, vi har ju löst det på det sättet att vi använder ju egentligen inga kommersiella produkter. Vi använder oss utav egenutvecklad programvara så vi har ett eget CAT-verktyg som jag använder utav. Vi har vår egen maskinöversättningsmotor och sånt där, så att vi har kanske inte riktigt de problemen eftersom allting är mer eller mindre in-house hos oss. Men säg att vi skulle använda oss utav en mer kommersiell produkt... för att ta den vanligaste då som är Trados då... till exempel om man skulle använda den. Den har ju möjlighet att till exempel att kunna gå ut mot Google Translate och göra grejer. Om man då har ett väldigt känsligt material och man lägger upp det hos Google, då står det ju i det finstiltade att då äger Google det som man har lagt upp och det är deras, liksom, rättighet att få göra i stort sett vad fan de vill med det och så. Och det är inte så många av våra kunder sugna på då. Så därför har vi valt den strategin att mycket läggs internt då. Eller att vi använder oss av externa motorer då, där vi, kan man säga, köper den tjänsten och säger att det här är vårt material. Det här får inte spridas. Det här får inte göras någonting annat med om det då är ett sådant krav egentligen. Sedan vet jag ju också att många kunder har efterfrågat någonting som de kallar för "gisting"... att de vill ha motorer som kanske ligger på ett intranät eller någonting liknande med deras tränade material, där de snabbt ska kunna skriva in en text och få ett svar tillbaks och leverera den typen av tjänster, då. Och det kan vi göra idag med både externa motorer som är dedikerade för dem... att man köper en sådan tjänst som är liksom bara är för dem... alternativt att vi använder våra egna (*delar*) då. Och det tror jag eliminerar en stor del utav den här säkerhetsproblematiken som är att en översättare som man säger åt, kanske säger så här, jamen du får inte använda Google Translate och sedan så gör det i alla fall för att det är enkelt att göra det och så... och att man slipper det. Våra verktyg möjliggör inte den delen då. Vi kan i och för sig aldrig kontrollera om de klipper och klistrar och sitter gör det, men det är inte i alla fall inte en automatisk knapp inifrån systemet som bara gör det åt dem liksom.

Simon: Men ni använder bara den typen av verktyg, så att säga. Så ni använder inte Trados och andra vanliga verktyg?

Jan: Vi gör det när kunder efterfrågar det. Om de säger att jag har en egen avdelning som håller på med översättning och de vill ha den typen av leveranser tillbaks så att leveransformatet är, kan man säga, ett Trados-paket eller ett Déjà Vu-paket eller ett Star... vad heter de... ett Star Transit-paket eller någonting liknande. Om det finns den typen av krav då. Men annars så använder oss av våra egna verktyg för att göra det och det är ingenting som vi säljer egentligen, utan vi säljer det som en tjänst till våra kunder istället, så att säga. När ni köper en tjänst av oss så får ni det här på köpet och vi *utnyttjar* det för oss för att dels slippa den problematiken med just säkerhetsdelen.

Simon: Men har ni en uttalad policy gentemot frilansöversättare vad gäller till exempel användning av Google Translate om de klipper och klistrar och sådana grejer?

Jan: Jo, ja, absolut. Vi har ju NDA med dem där det står tydligt, liksom, vad man får göra och sådana saker. Sedan hur man kan kontrollera den efterlevnaden...det vet jag inte om någon

kan till 100 procent säga att det verkligen är på det sättet. Det... jag menar... människan är väldigt kreativ när det kommer till att gå runt barriärer om man säger så, men det är ju utmaningen, kan man säga, kring det. Men vi har vi har den typen av avtal och det finns sekretessklausuler och allt möjligt som säger att det här är det man måste förhålla sig till.

Simon: Jag tänkte ställa några frågor här mer explicit om riskhantering i slutet och tänkte höra om ni gjorde någon form av riskanalys i samband med att ni implementerade maskinöversättning i översättningsprocessen?

Jan: Ja absolut. Det gjordes. Vi alltså... när vi började titta på det så fanns det fanns ju många typer av risker. Säkerhetsriskerna självklart då med var man lagrar information, hur datan kan accessas, från vem, vem kan se data... sådant där... hur ska vi göra det? Så det togs ju självklart upp som en del. Andra risker var ju kvalitet, självklart, som hur kommer det här påverka kvaliteten? Risker som kostnadsbilder också. Vi har en ganska hög uppstartskostnad, initial investering, för att göra någonting och träna och få igång det. När kan en sådan investering börja betala sig? Hur lång tid måste man vänta? Tar det ett år för en kund? Tar det ett halvår? Tar det en månad, liksom? Var någonstans är break-even för det här... den här investeringen och så här. Så det gjordes ju ett väldigt gediget arbete när vi testade dem här och även testade olika typer av material... olika typer av språkkombinationer... försökte förstå liksom hur mycket kan man faktiskt återanvända. För att, jag menar, att gå ut och läsa något "white paper" eller någon som har skrivit någonting som säger det att maskinöversättningen är här nu... nu är det... det är liksom Jesus... här är frälsaren som kommer och det är så här mycket som ni kan köpa eller vi ska förvänta oss förbättringar och såna saker. Det är vet man inte riktigt hur pass partisk eller opartisk den är egentligen. Så vi har gjort väldigt mycket egna tester kring allt det här då. Och delvis för att vi har vi hållit på sedan 1986 med maskin... eller alltså med översättningsminnen. Vi har ju extremt mycket inhouse-material egentligen som vi har då för att kunna träna de här motorerna själva och se hur det funkar och då kan vi få ut ganska mycket statistik också. Och bara testa alla möjliga olika grejer, plus att vi vet ju också vad resultatet blir med vanliga översättningsminnen och såna grejer så att det går ju att få ganska mycket statistik som vi har gjort. Så att det har varit... jag tycker ändå att vi har gjort ett ganska gediget säkerhets... eller kan man säga, riskhanteringsarbete kring själva implementeringarna och försökt analysera vad som är bra, vad som är dåligt och hur vi ska använda det här på bästa möjliga sätt och i olika situationer.

Simon: Det finns ju en specifik ISO-standard, 31000, för riskhantering. Är det någonting ni är certifierade för eller har tillämpat?

Jan: Nej, vi har en ISO 9000 där det också ingår, kan man säga, att man ska ta upp risker och sådant och jobba enligt det då. Och sedan är vi också certifierade enligt... ska se vi vad den heter nu då... 17000 någonting... som har med översättningar att göra. Det har vi tittat på och sedan har jag på IT-sidan då börjat snegla lite grann... eller vi påbörjade ett arbete med ISO 27000 också, informationssäkerhet. Och det är en ganska omfattande standard och var ännu mer omfattande när jag tittade på den för några år sedan, så de har gjort den lite enklare nu, men det är väl någonting som kommer mer eller mindre att bli ett krav och man ser det också att våra kunder och de som vill köpa tjänster utav oss börjar bli mer och mer medvetna om den här typen av risker och nu under senaste tiden med alla ransomware och allting som har kommit och snurrat runt så blir det väldigt påtagligt att det blir mer och mer av den delen. Så jag tror att det kan vara en viktig del för oss att satsa mer på faktiskt.

Simon: Det finns ju även en ISO-standard för efterredigering... jag kommer inte ihåg exakt numret på den... vet du om ni tillämpar den också?

Jan: Den tillämpar vi inte vad jag vet i dagsläget, men det får vi väl titta på också. Jag menar det finns väldigt många standarder. Det tar väldigt mycket tid implementera också så att vi försöker ta en sak i taget kan säga, men absolut.

Simon: Jag tänkte lite mer på det här med efterredigering också. Erbjuder någon form av utbildning för folk som börjar arbeta med efterredigeringar antingen internt eller liksom vad gäller externa frilansare?

Jan: Oj, det är faktiskt mer än vad jag vet. Jag tror att det är nog en fråga för Vendor Management. Jag tror att de har... vi har ju någon form av utbildning för alla våra översättare. Sedan om de har utvecklat... jag tror att de har någonting för dem också, men det är inte någonting som jag tagit del av. Det är ingenting som IT varit involverade i alla fall vad jag vet.

Simon: Ja, jag tror att jag har kommit ganska långt med mina frågor här. Jag kanske skulle kunna höra lite om det om det är några andra risker kopplade till maskinöversättning som vi inte har diskuterat som du kan komma ihåg att du har stött på i samband med att ni har implementerat det här på företaget?

Jan: Jag tror att en stor risk... och det har väl kanske mer med den kommersiella delen... det är ju kunders enorma förväntningar ibland över hur den här tekniken ska kunna hjälpa dem. Man hör väldigt mycket... att det blir ett "buzz word" som fluktuerar i den här branschen och sedan så är det någon säljare som lovar för mycket som ringer och säger att vi kan sänka era kostnader med 30 procent eller någonting och de vet inte ens vad man har för några kostnader. Så att den typen av risker ser vi idag som ganska påtagliga... att man målar upp en bild som är väldigt, väldigt grön och i själva verket kanske någonstans rödgult kanske, jag vet inte, som det faktiskt kan hjälpa mig. Så jag säger inte att man inte ska använda det och jag tycker att det är en väldigt bra teknik, ett komplement till mycket annat. Men det är inte lösningen på allt, än så länge. Det kanske kommer imorgon, jag vet inte. Men än så länge så är språk för komplext för att en maskin än så länge ska kunna klara av att göra det som en människa klarar att göra. Det är väl ungefär det som är slutsatsen utav det och den risken är väl påtaglig hela tiden, skulle jag vilja säga.

Simon: Ja, okej, jag kan kanske bara avsluta med att fråga om det är något mer du känner att du vill ta upp i det här sammanhanget som du tycker att vi inte har diskuterat så mycket, eller?

Jan: Säkert, men jag kommer inte på någonting (skratt).

Simon: Nej, men jättebra. Jag tror jag avslutar inspelningen här då. Du ska ha ett stort tack för din medverkan.

7.4 Bilaga 3 – intervju 3

Jessica (pseudonym), lösningsarkitekt, Z Translation (pseudonym)

Simon: Kan vi kanske börja med om du bara lite kort kan berätta vad din roll på firman är och kanske lite kort om din bakgrund utan att bli för personlig?

Jessica: Jag jobbar som lösningsarkitekt, Solutions Architect. Så att det är på företagets teknologiavdelning. Och det är en roll där man jobbar med specifika kundcase och kundimplementationer där det finns ett "scope" för en teknisk lösning.

Simon: Skulle du också kanske bara lite kort kunna säga någonting om ditt företag? Jag tänker mest på om ni främst använder er av inhouse-översättare eller externa översättare och lite sådana saker?

Jessica: Precis. Företaget är en "language service provider", kan man säga, som går mer och mer mot att bli ett "language technology company". Ja, så att det innebär ju att vi jobbar ju med en integration mellan det lingvistiska, alltså översättning, och olika typer av språktjänster med tekniska lösningar och det är ju dit mycket branschen går. Vad det gäller översättningsresurser så jobbar vi mycket med frilansare, men också med olika andra, liksom byråer, som tillhandahåller översättningsresurser. Så vi har väl en del inhouse-översättare, men framförallt med externa resurser.

Simon: Har du en uppfattning om hur länge ni har använt er av maskinöversättning i översättningsprocessen?

Jessica: Maskinöversättning i översättningsprocessen? Jag började på företaget... vad kan det varit... 2017. Och då tror jag man hade gjort en del... alltså, man hade liksom jobbat med implementation tidigare, men det var väl kanske därefter som det faktiskt... som det har implementerats som del i olika produktionsflöden. Så jag skulle säga att där någonstans runt 2017, kanske, att det börjar hända saker.

Simon: Vet du vad det är för typ av maskinöversättning ni primärt använder? Är det statistiskt baserat eller ett neuralt maskinöversättningssystem?

Jessica: Neuralt. Och det är ju i linje med hur branschen har gått, så vi har ju även där flyttat oss från statistiska metoder, då, till de här neurala motorerna. Vi kanske möjligtvis använder någon statistisk motor i något specifikt fall, men övergripande så är det ju mot det neutrala vi går.

Simon: Använder ni er främst av ett egenutvecklat maskinöversättningssystem eller har ni köpt in en färdig lösning?

Jessica: Vi köper in en färdig lösning för det.

Simon: Har du någon uppfattning på ett ungefär i hur stor andel av alla översättningsuppdrag som ni använder maskinöversättning i processen? Alltså, är det mer än hälften eller 75 % eller ja?

Jessica: Oj, den är jättesvår att svara på. Jag tror man behöver fundera på vad man menar med när det används. Man kan säga att det finns ju som en del integrerat i våra produktionsflöden, men där är det på något sätt översättaren som avgör hur pass mycket man vill använda av den. Men sedan så finns det ju liksom rena kundcase och där kan det vara så att det finns ett intresse av att mer förutbestämma hur pass mycket man ska använda maskinöversättning och då är ju det kopplat till... vad ska man säga... kopplat till en kvalitetskravspec och en kostnadskravspec och det kan vara kopplat till (*ohörbart*) turnaround times, alltså hur snabbt man vill ha leveransen. Så att det finns en... vad ska man säga... maskinöversättningar legitimeras av andra aspekter eller andra krav kopplat till översättningen. Så det behöver man kanske skilja på lite... att vi integrerar i maskinöversättning i produktionen, men då är det mer som ett stöd i översättningsprocessen. Men sedan har vi då de specifika kundcasen där man applicerar... liksom man bestämmer hur man ska använda det helt enkelt och, som sagt, det är kopplat då till en kravspec som har många olika parametrar i sig.

Simon: Du nämnde där att i vissa fall får översättaren själv avgöra i vilken utsträckning den vill använda sig av maskinöversättningssystemet, så att säga, och att det fungerar lite som ett IT-stöd eller ett stöd i översättningsprocessen. Kan du kanske beskriva det lite mer? Vad det går ut på, så att säga?

Jessica: Ja, absolut. Jag kan göra ett försök. Nu jobbar jag ju inte som (*översättare*)... men vi använder ju olika CAT-verktyg och där kopplas ju maskinöversättningen på så att översättaren kan se förslag från maskinöversättningen egentligen och beroende på hur mycket eller om de vill använda den så går det att använda. Men sedan kan det vara att man gör en bedömning... alltså det ska ju bli en översättning så att översättaren är ju ansvarig för den kvaliteten som levereras. Så bedömer översättaren att nej, det här var inte alls en bra träff, då behöver man inte använda den så att det finns inget som styr hur pass mycket det ska appliceras utan det är upp till den individuella som jobbar med det så att säga. Ja då syns de här träffarna, kan man säga, i CAT-verktyget då som man jobbar med.

Simon: Så det fungerar lite som om du får en träff från översättningsminnet... att det kommer upp som en träff som du kan använda dig av och bearbeta?

Jessica: Ja, precis. Den är ju flaggad som en MT-träff.

Simon: Precis, jag diskuterade det här i en annan intervju med en annan informant som menade på att de har haft lite problem med vissa översättare... att de hade svårt att hålla isär det här med maskinöversättningsträffar och träffar från ett översättningsminne då som är kvalitativt olika saker och att det var lite bekymmer. Har ni upplevt något liknande?

Jessica: Jag kan inte riktigt svara på det för att jag jobbar inte riktigt med den delen, men det jag har hört i alla fall, det är att det blir... i början när det här implementerades så fanns det ett stort motstånd. Man tyckte väl att... ja, det var dålig kvalitet, att det störde processen. Men nu så blir det ju mer och mer normerande, alltså, att man vill ha det här stödet så att säga. Men sedan om det finns ett bekymmer i att hålla isär en MT-träff från en TM-träff, det kan jag inte riktigt svara på. Kan du mer gå in på vad du menar med att det är svårt att hålla isär.

Simon: Vi kanske kan fortsätta på det andra spåret där som jag tyckte var ganska intressant... som du är inne på... det här med att för många översättare är det lite av en tillvänjningsprocess, som det låter, att använda sig av maskinöversättning. Har ni mött

något... generellt kan man säga någonting om översättares attityd eller inställning till det här med att använda maskinöversättning. Finns det några skillnader mellan olika grupper av översättare, till exempel gamla och unga?

Jessica: Ja, jag kan inte säga någonting om åldrarna, men däremot så vet jag att det inom vissa språkkombinationer kan det vara mer av en tradition att man kanske inte vill jobba med MT och utföra den här typen av granskning som det ofta blir om man pratar post-editing-jobb till exempel. Så det kan nog skilja sig mellan olika språkkombinationer. Det är väl det jag vet i alla fall. Ja att det i vissa språkkombinationer är liksom mer normerat än i andra. Och sedan kanske det finns en... jag vet inte om det är någon slags fördom, men att de som jobbat som översättare väldigt, väldigt länge och kommer från en tradition av att översätta allting från scratch och att det kanske är svårare att ta till sig den nya tekniken, men det är bara så här en... det vet jag faktiskt ingenting om utan det känns bara som en sådan fördom man kan ha, kanske.

Simon: Jag funderar lite på... om vi går tillbaka här med vilka typer av språktjänster ni erbjuder. Erbjuder ni även översättningsuppdrag som är baserade på helautomatisk maskinöversättning utan mänsklig inblandning?

Jessica: Ja, det gör vi. Och återigen, det är så viktigt att koppla det till själva caset eller användarcaset här... vad är det kunden efterfrågar... för att någonstans bedöma vilken typ av översättning och kanske översättningskvalitet och liksom process ska man implementera. Och vi har alltså... inte bara själva tjänsten, maskinöversättning finns, men det finns också att tillgå från olika plattformar. Du kan lägga en beställning med oss som på något sätt hanteras i vårt normala beställningsflöde, men som är en ren maskinöversättning eller så har ju vi en plattformstjänst där kunden i sitt eget användargränssnitt, kanske... ja alltså inte ens in i Z Translation liksom, deras system, men att man där kan... genom vår tjänst då egentligen skicka på maskinöversättning och få tillbaka ett direkt svar eller vad man ska säga. Så då går det inte in i våra normala översättningsprocesser och då ser ju fakturering och hur produkten är paketerad helt annorlunda ut. Så då kanske man inte faktureras per projekt eller per beställning utan man faktureras kanske mer på sin licens, baserat på konsumtion till exempel. Så att det finns ju lite olika affärsmodeller med det här beroende på hur man vill applicera den här automatiseringen, eller vad man ska säga.

Simon: Kan du kanske ge något exempel på situationer där kunder efterfrågar helautomatisk översättning?

Jessica: Absolut. Ja menar maskinöversättning det... jag ska säga så här... du hade en fråga om det finns fall där man aldrig använder maskinöversättning? Och jag skulle vilja svara nej på den frågan för hela tiden måste man återkomma till vad är det kunden efterfrågar. Och även om man kanske tänker så här att en juridisk text, kanske, eller en... att det tänker man spontant kanske inte är så anpassat för maskinöversättning för att det kan komma mycket tveksamheter i responsen. Så återigen måste man ställa sig frågan: vad är det kunden efterfrågar? Det kanske inte spelar någon roll... eller att de ska göra någon egen redigering och så vidare... så att hela tiden usercaset är väldigt viktigt. Så tillbaka till din fråga som jag ville koppla samman till det där. Då ska vi se vad det var du frågade... kan du återupprepa frågan?

Simon: Ja, det var kanske om du hade något exempel på där kunder efterfrågar helautomatisk översättning.

Jessica: Ja precis. Och det kan verkligen variera, men ofta kan det vara för typ intern kommunikation. Alltså om man som företag har global spridning och man vill enkelt kunna kommunicera mellan olika grenar i olika länder... att man snabbt ska kunna förstå varandra, till exempel, det kan ju vara en sådan sak och då det kanske är väldigt... det är inte så noga om det blir exakt rätt bara man skapar sig en förståelse... så sådan typ av intern kommunikation. Sedan kan det ju också vara om man sitter och vill göra...säg att du har väldigt mycket material och du vill kunna göra en snabb bedömning på vad som finns i det här materialet... då är det inte heller så noga exakt vilken respons du får. Du kanske hellre vill få en gissning, en mer generell förståelse för vad som är i dokumenten. Så när man gör någon slags bakgrundsarbete... någon slags research kanske det kan vara relevant med ren maskinöversättning. Även om man tittar... och det kanske inte är supervanligt... men det blir ju mer och mer vanligt om man tänker sig e-commerce eller e-commerce-handelsplattformar, där man har stor omsättning av produkter som säljs och där... jag ska inte säga att det är supervanligt, men där kan det också i vissa fall vara så att man måste snabbt få ruljans på översättningen och få tillbaka översättningen snabbt och där kanske kvaliteten är lite mindre viktig och så. Ja, det är väl några stycken case... ska se vad vi mer kan tänka oss.

Simon: Eller vi kanske ska fortsätta med att gå in på det vi var inne på tidigare... vi pratade om efterredigering eller post-editing och i de fall där ni erbjuder den tjänsten... erbjuder ni både lätt efterredigering och full efterredigering om du är med på vad jag menar?

Jessica: Ja, ja precis. Tidigare så tror jag vi hade det paketerat som att det fanns två olika sådana tjänster. Det var light och full eller vad det nu kallas för? Men i dagsläget så har vi nog bara paketerat... det finns bara en enda tjänst och där återigen handlar någonstans om att hitta den redigeringsnivå som är utifrån kundens behov. Så om man vill kalla det lätt eller full... det vet jag inte egentligen om det är relevant utan det gäller nånstans att hitta en överenskommelse med kunden. Vad är nivån av efterredigering som krävs?

Simon: När ni avgör om eller i vilken utsträckning maskinöversättning ska användas för ett visst uppdrag... är det främst kunden som avgör det eller har ni andra kriterier som ni själva använder er av för att titta på om det kan vara lämpligt att använda sig av maskinöversättning?

Jessica: Ofta, det vi gör är ju att ge kunden en rekommendation, kan man ju säga, att baserat på det du säger att du vill ha så erbjuder vi det här. Det rekommenderar vi. Vi kan titta på de här sakerna och så vidare. Man har liksom en dialog kring det. Alltså, det beror väldigt mycket usercasen, men är det ett sådant användarfall där man verkligen vill säkerställa att det verkligen blir rätt nivå av applikation på maskinöversättningen, då handlar ju någonstans om att göra en ganska kvalitativ utvärdering, alltså att göra en del tester helt enkelt med olika språkkombinationer ifall kunden har det. Kanske titta på olika typer av texter om kunden också har det och att ha en lingvist som på något sätt bedömer den här outputen. Så att man får en kvalitativ utvärdering och återigen att göra den kvalitativa utvärderingen i förhållande till vad kunden någonstans är intresserad av i kvalitetsmått, kanske. Sedan finns det ju andra typer av mera tekniska kvalitetsparametrar. Man kan väl prata om t.ex. BLEU score och det finns väl x antal sådana, men de jobbar vi nu inte supermycket med egentligen för att... nu ska vi se... dels de här BLEU score-parametrarna är ju framför allt om man har gjort en helt custom... att man på något sätt anpassat den här... man har tränat motorn. Och det gör vi ju kanske inte jätteofta. Och sen någonstans så är det väl inte så att i grund och botten... ja, man kan ha ett numeriskt mått, men det kommer alltid ned till vad är det kunden och användaren

ser som kvalitet och inte. Är det mer orienterat mot terminologi, alltså det här accuracy som man kan prata om. Man kan prata om de här dimensionerna... accuracy-dimensionen kontra fluency, alltså hur stringent och lätt att läsa det är... alltså den till dimensionen. Det är ju två helt olika dimensioner, egentligen, och det kan ju skilja sig från kund till kund... vad som egentligen är viktigt. Så att jag skulle säga att den här mänskliga utvärderingen är jätteviktig.

Simon: Och om man ser det från kundens sida... hur pass medvetna är kunderna om maskinöversättningens prestanda eller hur pass bra maskinöversättning fungerar?

Jessica: Oj, det var en väldigt svår fråga. Det kan säkert skilja sig jättemycket. Vi jobbar ju med massor av olika typer av kunder och användarfall. Jag tror oftast att det är en viktig uppgift för oss att också beskriva vad den här tekniken gör och inte gör... att verkligen komma med den här typen av rekommendationer. Ja, men vill ni ha det, då rekommenderar vi det här... att hela tiden... det upplever jag är väldigt viktigt och det tror jag att kunderna kanske inte alltid tänker på. Så där tror jag att vi fyller en viktig roll att kunna problematisera kring vad det är de efterfrågar kontra vilken typ av teknik och vilken nivå av maskinöversättning man bör applicera.

Simon: Har du märkt av att kunden kommer in med orealistiska förväntningar på det här med maskinöversättningar och att det har blivit ett problem?

Jessica: Ja, det jag nog märkt av. Jo, men återigen, det är kopplat till att man har en orealistisk förväntning och att man kanske inte har gått igenom förväntningarna riktigt... att liksom det kvalitetskrav som kunden slutligen hade, det var kanske inte tydligt från början. Så att vi behöver ofta kanske stödja kunden i det här att definiera kraven... att man behöver liksom verkligen utmana kunden lite i att beskriva de här kvalitetsparametrarna för att också kunna utvärdera och efter ett tag... när man har kört maskinöversättning ett tag... okej hur uppnår vi det här och uppnår vi det inte. För annars är det som du säger att... men det kan ju vara så att maskinöversättningen är bra från ett fluency-perspektiv, men det visar sig att det var väldigt viktigt för kunden att vissa termer eller vissa namn blev rätt och det kanske alltid slår fel på. Så jag har inget specifikt case som sådant, annat än att det är alltid är så viktigt...

*(*Avbrott i förbindelsen och inspelningen*)*

Jessica: Jag, då vet jag inte var du försvann riktigt. Men ja, du frågade om det här med om det fanns specifika fall där kunden på något sätt har haft andra förväntningar på maskinöversättningen.

Simon: Ja, precis. Eller om de hade orealistiska förväntningar.

Jessica: Ja och det finns det. Jag vet själv att... jag jobbade som projektledare en gång i tiden och projektledare hos oss det är en person som hanterar beställningen helt enkelt och där var det ju någonstans... om man ska ta det specifika exemplet... så var det att kunden hade kort om tid och man ville liksom på något sätt snabba på processen och det legitimerade användning av maskinöversättning. Men att den här post-redigeringen då, eller vad man ska säga, gjorde i det här fallet att man ändå inte nådde upp till förväntan hos kunden på kvaliteten. Och där finns det ju ett tvådelat ansvar någonstans. Återigen att kunden behöver ställa kvalitetskrav i början för att man ska kunna utvärdera om maskinöversättning passar eller inte, men också för oss som företag... att på något sätt säkerställa det här... att vara mer utmanande kanske på att definiera kravställningen. För att det blir ju annorlunda om man post-redigerar

en text... jag vet ibland när jag själv har läst postredigerade texter. Man får en känsla av att om man hade översatt det här från början hade man kanske inte skrivit exakt så här. Man kan inte riktigt peka på att någonting är fel, men att det är någonting som känns lite annorlunda i texten. Så att det är klart det blir säkert en skillnad i olika grader.

Simon: Kopplat till det här, har ni märkt att det finns vissa typer av texter som lämpar sig sämre för maskinöversättning?

Jessica: Ja, och vi var inne på det lite tidigare. Jag ska säga att det finns väl ingenting som man så här absolut aldrig skulle... alltså att det absolut inte passar för det... återigen kundens krav. Men som vi sa juridiska texter, där det är mycket terminologi, där det kanske inte får finnas tvetydighet, utan det är verkligen viktigt med någon med juridisk kompetens... alltså en människa med juridisk kompetens som verkligen kan avgöra vad som är rätt terminologi. Här är ju accuracy, tänker jag, väldigt, väldigt viktigt... att det är korrekt begrepp som är översatt. Så juridiska texter... men sedan har du ju även sådana här marknadstexter som är svävande på något vis eller en kreativ text där det också finns tvetydigheter många gånger. Man kanske använder metaforer och symboler och så vidare, alltså kanske väldigt långa segment och sådär. Så kreativa texter kanske inte heller lämpar sig alltid jättebra med den här råa maskinöversättningen. Så att det är väl de två fallen kanske. Och sedan så är det så klart... jag menar maskinöversättningen kräver ju också att du har ett format... alltså att texten på något sätt ligger i ett format som går att processa genom maskinöversättningen. Det kanske finns vissa PDF eller inskannade filer och så där, där man kanske inte ens jobbar med ett CAT-verktyg utan jobbar utanför det.

Simon: Om vi går in på det här lite med efterredigering eller post-editing då, alltså. Är det ofta översättare ni anlitar för de här uppdragen eller kan ni anlita andra typer av personer också eller?

Jessica: Ja, det var den här frågan som var lite svår. Men, vad ska man säga? De lingvister som finns i vår databas, de har ju... alltså det är olika inköpsprodukter, helt enkelt. Så att om vi köper en översättning och då köper vi en översättning och köper vi en post-redigering så köper vi en sådan. Så det beror helt enkelt på vad lingvisten är registrerad hos oss och så... och vilka kompetenser den har?

Simon: Anser du att en person som utför efterredigeringsuppdrag behöver ha några särskilda kvalifikationer eller några särskilda kompetenser?

Jessica: Den där frågan skulle jag vilja kanske komplettera, om du vill, på mail för jag tänker att jag skulle behöva prata med vår inköpsavdelning och Vendor Management-ansvarige för att se om det finns några specifika...

Simon: Ja, det får du gärna göra om du vill. Vi kan hoppa över den så länge. Vi kanske ska gå in lite på det här med problem som kan uppstå i maskinöversättningsprocessen, till exempel översättningsfel eller felöversättningar då. Har ni noterat några särskilda typer av felöversättningar som förekommer i samband med maskinöversättning?

Jessica: Alltså, det finns ju såklart tänker jag lite så här roliga exempel. Jag vet något exempel där en översättare noterade att det med jämna mellanrum kom in i maskinöversättningen en text där det stod: "Jag har inte tid med det här"... alltså som återkommer på olika ställen. Och det kan man ju bli så här... vart kommer det ifrån, liksom? Hur har det fått ett sådant fäste? Så

sådana exempel finns det säkert gott om. Det här var bara ett sådant som jag vet om. Men annars är det väl, som jag sa, att det kan bli så väldigt bokstavligt ibland... som inte passar in i sammanhanget och då återigen det där med juridiska texter eller copy-texter eller kreativa texter. Och sedan vet vi att det är vissa språkkombinationer där MT kanske presterar sämre. Där vi i alla fall får mer klagomål, möjligtvis. Så det är det jag vet.

Simon: Har ni några särskilda åtgärder, vad du vet, för att hantera översättningsfel som uppstår i maskinöversättningssystemet?

Jessica: Nej, och det där är ju en väldigt bra fråga för att när det är någonting som vi... när vi liksom inte riktigt ansvarar för outputen... och vad jag menar då, det är ju när kunden kanske köper helt rå maskinöversättning, då kan ju vi aldrig... vi vet ju inte riktigt vad som kommer ut då, så att säga, och där gör ju vi ingen... Då köper de egentligen den rena outputen så. Men där... om det uppkommer fel i det... då är det en bra fråga hur hanterar man det? För det är inte vi som... vi köper ju in, vad ska man säga, algoritmen eller modellen någonstans. Så då bör man ju, tänker jag, egentligen rapportera det tillbaka till den leverantören, alltså den som man köper maskinöversättningsmöjligheten från. För den är ju svår för oss att påverka. Men annars när vi på något sätt ansvarar mer för leveransen, då menar jag att vi har... alltså en människa som tittar på texten, då behöver man ju någonstans gå tillbaka till den människan eller den leverantören och se vad det är som är fel helt enkelt. Om det är så att det är något jättteuppenbart som borde ha redigerats. Så det är ju två helt olika fall på något vis. Men sedan har vi fall, som jag säger, där vi tränar motorer. Och vad jag menar när vi tränar motorer, det är ju att kunden har liksom en "bulk of linguistic", assets, eller ja språktillgångar då som man provar lägga ovanpå den här baselinen helt enkelt och ser om det kan, på något sätt, flytta kvaliteten eller göra det lite mer anpassat för kunden. Och där så tänker jag ju spontant att... nu har jag inget sådant exempel... men om man upptäcker att där är något som är jättetokigt i den outputen så kan man ju kanske mixtra med den datan och se om man får någon annan effekt.

Simon: Om vi går in på det här lite med datasäkerhet. Har ni märkt av några särskilda datasäkerhetsproblem i samband med användningen av maskinöversättning?

Jessica: Nej, men här är ju så här viktigt att förstå, tycker jag, vart data flyttas runt... hur processen ser ut när man applicerar maskinöversättning för att förstå vart hamnar vår data och sedan också förstå hur den här datan används. För den ska ju inte användas av någon annan än oss, liksom... så att den inte används av en tredje part på något vis och sedan också... om man i de fallen där man tränar en sådan där maskinöversättningsmotor för en kunds räkning att också förstå vad det är i det här materialet och att kunden verkligen också förstår det så att man känner att det inte tränas med data som kanske kan vara problematisk ur ett GDPR-perspektiv eller att det finns känslig personlig data och så. Det är väl sådana frågeställningar som man tittar på.

Simon: Har ni vidtagit några särskilda åtgärder för att till exempel undvika dataläckage som du var inne på... att data läcker ut ur systemet?

Jessica: Den frågan är också... den kan jag inte svara på. Vi har ju såklart en sådan ansvarig på företaget, men jag känner att jag kan inte svara på den.

Simon: Om vi går vidare på det här... det finns ju även sådana här kostnadsfria översättningstjänster på nätet, till exempel Google Translate är den mest kända. Vad har ni

för policy rörande den typen av tjänster gentemot översättare? Alltså, får de använda det, eller?

Jessica: Nej, det får de inte göra. Det är mitt spontana svar alltså att det... alltså du tänker att översättaren skulle ta liksom texten och köra igenom...

Simon: Ja, antingen att man klipper och klistrar eller jag tror till och med man i vissa fall kan lägga in den som en plug-in i CAT-verktyget och köra texten mot till exempel Google Translate eller något annat sådant kostnadsfritt verktyg.

Jessica: Jaha, nej det tänker jag spontant att de inte får göra... eller det får de inte göra så klart.

Simon: Om vi går vidare lite mer, kanske i slutet här, rent explicit med er riskhanteringsprocess. I samband med att ni har infört maskinöversättning i översättningsprocessen... vet du om ni gjorde någon form av riskanalys i samband med det här?

Jessica: Det kan jag egentligen inte heller riktigt svara på vad som gjordes i början. Det är väl några år sedan säkert, men jag tänker så här... nu om man tänker framåt i alla fall, det finns ju risker från ett affärsmässigt perspektiv... att kanske inte använda maskinöversättning för det är ju mycket dit branschen går. Så att det gäller ju också att utvärdera... det är ju inte bara att applicera det, men det gäller att utvärdera så här... finns det fördelar med att applicera det... att liksom kunna mäta det... på vilka språk? I vilka typer av branscher exempelvis eller vilka typer av motorer? Var finns det liksom fördelar med att använda maskinöversättning för alla parter någonstans? Jag tänker att det finns en risk... alltså en affärsmässig risk med att inte utforska det här. Ja, för att det är som du säger. Jag tänker att det är dit branschen går väldigt mycket.

Simon: Det finns ju en särskild ISO-standard för riskhantering som heter 31000. Vet du om det någonting ni har tillämpat?

Jessica: Ja, då ska vi se. Jag kollade lite vad vi hade för ISO-certifieringar och det står på vår hemsida också. Och där noterade jag... och jag kan egentligen inte svara så mycket på det här. Men jag bara noterade att vi är ju certifierade enligt lite olika ISO-standarder. Men vilken var det du sa som det var?

Simon: Den här jag nämnde var först då... det finns en särskild riskhanteringsstandard som heter 31000.

Jessica: 31000? Nej. Den kanske inte fanns. Jag vet inte. ISO 9001:2015. Sedan har vi ISO 14001:2015 och ISO 27001:2013 och sedan så... oj, vad blir det här då? 17100:2015. De ISO-certifieringarna hittade jag.

Simon: Det finns ju en ISO-standard även för efterredigering eller post-editing då som jag tror heter 18587, men den fanns inte med i din lista?

Jessica: Nej, inte vad som står på hemsidan i alla fall. Men sedan är frågan om det finns någon annan information internt, men jag gissar att det som är på hemsidan är det som gäller. Men jag kan kolla upp det också... om det så att vi behöver... jag kan komplettera med det. 18587? Ja, men jag kan kika om det är något som vi har.

Simon: Sedan funderade jag på det här lite med de här personerna då som jobbar med efterredigering, antingen in-house eller externt... erbjuder ni någon form av utbildning för dem, liksom kopplat till efterredigeringen?

Jessica: Nej, men det är väl en jättebra fråga också, men det är också något som jag skulle vilja prata med vår Vendor Management-avdelning om. Men vi har ju för alla... vad jag vet i alla fall för alla tjänster så finns det liksom... ja men instruktioner eller på något sätt beskrivning av vad som förväntas av den lingvist som utför tjänsten. Så det vet jag finns. Men sedan vid själva onboardingen av en ny lingvist om vi har något speciellt så här utbildningspaket, det kan jag också kolla upp. Jag ska kolla det.

Simon: Ja, jättebra. Vi är nästan framme vid slutet här nu. Jag funderade bara på om det finns några andra risker eller problem kopplade till maskinöversättning som vi inte har tagit upp här hittills i intervjun. Du var inne på det här med lite med affärsmässiga risker om man låter bli att använda... finns det några andra risker, kanske mer kopplade till själva maskinöversättningen och dess output som ni har stött på?

Jessica: Ja, men det var väl det som vi pratade lite om... med datasäkerheten där... att vill man verkligen förstå det där hur data flyttas och hur det sparas också... den delen är ju viktig att ha med sig. Och sedan även, som vi sa, det här affärsmässiga är ju också väldigt viktigt när man går in i framtiden. Men jag tänker så här i alla led att ha alla med sig på tåget någonstans, både internt... interna processer, men även lingvisterna på något sätt som jobbar för oss och som ska utföra det här... att de också förstår vad som förväntas om man pratar, som du säger, lite efterredigeringsperspektiv. Och det här som vi också var inne på... jag tänker att en annan risk är ju det här att förväntningarna på något sätt med en kund och ett specifikt case är att alltid försöka "aligna" kring förväntningarna helt enkelt på den nivå av översättning man beställer. Ja, det är väl kanske det?

Simon: Nej, men jättebra. Det är nog de frågor jag har helt enkelt. Jag tänkte bara avsluta med att fråga om det är något mer du känner att du skulle vilja ta upp kopplat till det här temat som du känner att vi inte riktigt har hunnit med att diskutera?

Jessica: Nej, men jag tänker så här. Nästa steg väldigt mycket det är ju... den här tekniken finns. Men jag tänker att en svårighet är att hitta hållbara affärsmodeller kring den och att på något sätt... utöver tekniken vilka typer av tjänster ska och bör erbjudas kring det här. Vilka problem och utmaningar har kunderna att förstå de här problemen någonstans för att kunna rekommendera inte bara så här... okej, använd maskinöversättning eller inte, men att så här... i vilka fall? för vilka språk? för vilka typer av texter? vilken motor lämpar sig bäst? Det finns ju jättemånga leverantörer av den här typen av teknik på marknaden. Men att hitta någon slags intelligens kring det här, det är ju väldigt viktiga och väldigt intressanta frågor där jag tror att vi som erbjuder tjänster kring det här, att där finns ett stort arbete att göra och mycket att utforska om man tänker sig om man går in i framtiden, men det kanske inte är kopplat till specifika risker som du nu var intresserad av, utan det är mer den här affärsmässiga biten. Hur man på bästa sätt gör affärer kring det här och, som jag säger, skapar hållbara affärsmodeller.

Simon: Ja, men jättebra. Då har vi nog nått fram till slutet här. Du ska ha stort tack för att du ställde upp på det.

Referenser

- Alvehus, J. (2013). *Skriva uppsats med kvalitativ metod: en handbok*. Stockholm: Liber.
- Bell, T. och CSO Staff (2019). Data breached in translation. *CSO Online*. 27 november. Tillgänglig under <https://www.csoonline.com/article/3236348/data-breached-in-translation.html>. [Hämtad 2021-04-13].
- Bryman, A. och Bell, E. (2005). *Företagsekonomiska forskningsmetoder*. Malmö: Liber Ekonomi.
- Bowker, L. (2003). Terminology tools for translators. i H. Somers (ed.), *Computers and Translation: A Translator's Guide*. Amsterdam/Philadelphia: John Benjamins, s. 49–66.
- Bowker, L. och Fisher, D. (2010). Computer-aided translation. i Y. Gambier och L. Doorslaer (eds) *Handbook of Translation Studies, Volume 1*. Amsterdam/Philadelphia: John Benjamins, s. 60–65.
- Cadwell, P., O'Brien, S., och Teixeira, C.S.C. (2018). Resistance and accommodation: factors for the (non-)adoption of machine translation among professional translators. *Perspectives*, 26:3, s. 301–321.
- Canfora, C. and Ottmann, A. (2015). Risikomanagement für Übersetzungen. *trans-kom* 8:2, s. 314–346.
- Canfora, C. och Ottmann, A. (2018). Of ostriches, pyramids, and Swiss cheese. Risks in safety-critical translations. *Translation Spaces*, 7: 2, s. 167–201.
- Canfora, C. och Ottmann, A. (2020). Risks in neural machine translation. *Translation Spaces*, 9:1, s. 58–77.
- Castilho, S., Moorkens, J., Gaspari, F., Calixto, I., Tinsley, I. och Way, A. (2017). Is Neural Machine Translation the New State of the Art? *The Prague Bulletin of Mathematical Linguistics*, 108, s. 109-120.
- Céspedes, B. (2018). Mind the gap: Language Service Providers' perceptions of the technological training of professional translators. i Postigo, E. (ed). *Nuevas tecnologías, procesos cognitivos y estrategias para la optimización de las competencias del traductor e intérprete*. Berlin: Frank and Timme, s. 5–25.
- CSA Research (2019). Global Market for Outsourced Translation and Interpreting Services and Technology to Reach US\$49.60 Billion in 2019. Pressrelease, 17 maj. Tillgänglig under <https://csa-research.com/More/Media/Press-Releases/ArticleID/546/Global-Market-for-Outsourced-Translation-and-Interpreting-Services-and-Technology-to-Reach-US-49-60-Billion-in-2019>. [Hämtad 2021-03-22].
- CSA Research (2020). The Top 100 Language Service Providers: 2020. Tillgänglig under <https://csa-research.com/Featured-Content/Global-Market-Study/Top-100-LSPs>. [Hämtad 2021-04-10].
- Daems, J., Vandepitte S., Hartsuiker, R. J. och Macken, L. (2017). Identifying the Machine Translation Error Types with the Greatest Impact on Post-Editing Effort. *Frontiers in Psychology*, 8:1282.
- Dannewitz-Linder, M. (2018a). Efterredigerings olika ansikten. *Facköversättaren*, nr 2, s. 10–12.
- Dannewitz-Linder, M. (2018b). Hur (o)säker är maskinöversättning? *Facköversättaren*, nr 2, s. 6–8.

- DePalma, D. (2017). Augmented Translation Powers up Language Services. CSA Research. [Blogg]. 15 februari. Tillgänglig under <https://csa-research.com/Blog/ArticleID/140/Augmented-Translation-Powers-up-Language-Services> [Hämtad 2021-04-12].
- Esselink, B. (2003). Localisation and translation. i H. Somers (ed.), *Computers and Translation: A Translator's Guide*. Amsterdam/Philadelphia: John Benjamins, s. 67–86.
- Ewald, H. (2020). Amazons Sverigelansering hånas för översättningsmissar. *Dagens Nyheter*, 28 oktober, Tillgänglig under <https://www.dn.se/ekonomi/amazons-sverigelansering-hanas-for-oversattningmissar> [Hämtad 2021-03-24].
- Fagerström, N. (2017). Populär översättningstjänst publicerar i hemlighet användares privata meddelanden och företags interna mejl. *Yle*, 5 september, Tillgänglig under <https://svenska.yle.fi/artikel/2017/09/05/popular-oversattningstjanst-publicerar-i-hemlighet-anvandares-privata-meddelanden> [Hämtad 2021-04-25].
- Focada, M. L. (2017). Making sense of neural machine translation. *Translation Spaces*, 6(2), s. 291–309.
- Fong, D. (2012). Online translation breaks language barriers. *Deutsche Welle*. 20 december. Tillgänglig under <https://www.dw.com/en/online-translation-breaks-language-barriers/a-16457840>. [Hämtad 2021-04-23].
- Guerberof Arenas, A. (2013). What do professional translators think about post-editing? *The Journal of Specialised Translation*, 19, s. 75–95.
- Hallin, A. och Helin, J. (2018). Intervjuer. Lund: Studentlitteratur.
- Hofmann, S. (2012). Prozessgestütztes Übersetzen. Vom funktionsorientierten Übersetzungsprozess zum GP-Modell für die Dienstleistung Übersetzen. Lichtenberg: Harland Media.
- International Organization for Standardization. (2015). Translation services — Requirements for Translation Services (ISO Standard No. 17100:2015).
- International Organization for Standardization. (2017). Translation services — Post-editing of machine translation output — Requirements (ISO Standard No. 18587:2017).
- International Organization for Standardization. (2018). Risk management – Guidelines (ISO Standard No. 31000:2018).
- Kelly, N. (2012). Clearing up the Top 10 Myths about Translation. *Huffington Post* [hämtad 2021-03-22]. Tillgänglig under http://www.huffingtonpost.com/nataly-kelly/clearing-up-the-top-10-my_b_1590360.html?view=print&comm_ref=false.
- Koehn, P. (2009). *Statistical Machine Translation*. Cambridge: Cambridge University Press.
- Koehn, P. (2020). *Neural Machine Translation*. Cambridge: Cambridge University Press.
- Koehn, P. och Knowles, R. (2017). Six Challenges for Neural Machine Translation. *Proceedings of the First Workshop on Neural Machine Translation*, s. 28–39.
- Kamocki, P., O'Regan, J. och Stauch, M. (2015). All Your Data Are Belong to us. European Perspectives on Privacy Issues in Free Online Machine Translation Services. i Aspinall, D., Camenisch, J., Hansen, M, Fischer-Hübner, S. och Raab, C. (eds) *Privacy and Identity 2015: Privacy and Identity Management. Time for a Revolution?* Springer: Cham.
- Kastberg, P. (2012). Machine Translation Tools – Tools of the Translator's Trade. *Communication & Language at work*, 1, s. 34–45.
- Läubli, S. (2017). 3 Reasons Why Neural Machine Translation Is a Breakthrough. SlatorCon 2017 Zürich.
- Lindstedt, I. (2017). *Forskingens hantverk*. Lund: Studentlitteratur.

- Lommel, A. (2018). Augmented translation: A new approach to combining human and machine capabilities. *Proceedings of the 13th Conference of the Association for Machine Translation in the Americas (Volume 2: User Papers)*, s. 5–12.
- Marr, B. (2018) Will Machine Learning AI Make Human Translators An Endangered Species? *Forbes*, 24 augusti, Tillgänglig under <https://www.forbes.com/sites/bernardmarr/2018/08/24/will-machine-learning-ai-make-human-translators-an-endangered-species/?sh=679f8d853902> [Hämtad 2021-04-05].
- Moorkens, J. (2017). Under Pressure: Translation in Times of Austerity. *Perspectives*, 25: 3, s. 464–477.
- Moorkens, J., och O'Brien, S. (2017). Assessing user interface needs of post-editors of Machine Translation. i Kenny, C. (ed.) *Human issues in translation technology*. Abingdon: Routledge. s. 109–130.
- Nitzke, J., Hansen-Schirra, S. och Canfora, C. (2019). Risk management and post-editing competence. *The Journal of Specialised Translation*. 31, s. 239–259.
- Nurminen, M. (2020). Raw Machine Translation Use by Patent Professionals. A case of distributed cognition. *Translation, Cognition & Behavior* 3:1, s. 100–121.
- O'Brien, S. (2012). Translation as human–computer interaction. *Translation Spaces*, 1, s. 101–122.
- Olohan, M. (2021). Post-editing: A Genealogical Perspective on Translation Practice. i M. Bisiada (ed.) *Empirical Studies in Translation and Discourse*. Berlin: Language Science Press, s. 1–25.
- Poibeau, T. (2017). *Machine Translation*. Cambridge: The MIT Press.
- Presas, M., Cid-Leal, P. och Torres-Hostench, O. (2016). Machine Translation Implementation among Language Service Providers in Spain: A Mixed Methods Study. *Journal of Research Design and Statistics in Linguistics and Communication Science*, 3:1, s. 126–144
- Rai, A., Constantinides, P., och Sarker, S. (2019). Next-generation digital platforms: Toward human-AI hybrids. *MIS Quarterly*, 43, s. iii–ix.
- Sakamoto, A. (2019). Unintended consequences of translation technologies: from project managers' perspectives. *Perspectives*, 27:1, s. 58–73.
- Schäffner, C., Tcaciuc, L. S. och Tesseur, W. (2014). Translation Practices in Political Institutions: A Comparison of National, Supranational, and Non-Governmental Organisations. *Perspectives* 22:4, s. 493–510.
- Somers, H. (2003). Translation Memory Systems. i H. Somers (ed.), *Computers and Translation: A Translator's Guide*. Amsterdam/Philadelphia: John Benjamins, s. 31–46.
- Stallings, W. och Brown, L. (2018). *Computer Security. Principles and Practice*. Harlow: Pearson Global Edition.
- Susskind, R. och Susskind, D. (2017). *Professionernas framtid: hur teknologin kommer att förändra experters arbete*. Göteborg: Daidalos.
- Taravella, A. och Villeneuve, A. O. (2013). Acknowledging the Needs of Computer-Assisted Translation Tools Users: The Human Perspective in Human-Machine Translation. *The Journal of Specialised Translation* 19, s. 62–74.
- Tomter, L., Zondag, M. H. W. och Bye Skille, Ö. (2016). Warning about translation web site: Passwords and contracts accessible on the Internet. *NRK*, 3 september, Tillgänglig: https://www.nrk.no/urix/warning-about-translation-web-site_-passwords-and-contracts-accessible-on-the-internet-1.13670874 [Hämtad 2021-03-22].
- Trost, J. (1997). *Kvalitativa Intervjuer*. Lund: Studentlitteratur.

- Vashee, K. (2017). Data Security Risks with Generic and Free Machine Translation. *Empty Pages Blog*. Tillgänglig under <http://kv-emptypages.blogspot.com/2017/09/data-security-risks-with-generic-and.html> [Hämtad 2021-03-22].
- Vieira, L. N. (2019). Post-Editing of Machine Translation. i O'Hagan, M. (ed.) *The Routledge Handbook of Translation and Technology*. London: Routledge, s. 319–335.
- Vieira, L. N. (2020). Automation anxiety and translators. *Translation Studies*, 13:1, s. 1–21.
- Yamada, M. (2019). The impact of Google Neural Machine Translation on post-editing by student translators. *The Journal of Specialised Translation*, 31, s. 87–106.