



Language data and Machine Translation at Swedish Migration Agency (SMA)

Experiences and results of our proof of concepts

Anna Sågvall Hein, Professor Uppsala Universitet and Convertus Torbjörn Ekholm, Innovation Manager, Swedish Migration Agency





Background and five translation scenarios



Large volumes via translation agencies

Manual handling of translation cases

No organization for language data management

Limited access to outsourced translation data

Current process results in quality differences

1 Translation of web pages

2 Translation of documents

3 Translation of e-services

4 Translation in chatbot solution

Translations in meetings

Translations of public web pages - PoC 1



Swedish to English and Arabic

No sensitive data

GNMT as translation engine

Certfied terminology

Translation memory



Example of a postediting view

Swedish to Arabic

3 Det finns också krav på att du ska kunna försörja din familj och ha ett boende av It is also required that you will be able to provide financial support for your family and have housing of a certain standard 4 Kravet gäller inte i alla situationer 5 Vem kan ha rätt till familjeåterförening? Who may have a right to family reunification? 6 Din familj har möjlighet att ansöka om uppehållstillstånd för att flytta till dig i Your family has the possibility of applying for residence permits to move to you in Sweden if you 7 har fått permanent uppehållstillstånd to I have been given a permanent residence permit Jag-har fått permanent uppehålistilistånd 8 har fått ett tidsbegränsat uppehållstillstånd som flykting eller alternativt ps have received a temporary residence permit as a refugee or a person in need of skyddsbehövande och bedöms ha goda möjligheter att få ett permanent subsidiary protection and are deemed to have good chances of receiving a 9 Familjeåterforening för dig som har tildsbegränsat uppehållstillstånd och är flyktling Familj reunification för you if you have a temporary residence permit as a refugee eller alternativt skyddsbehövande Familjeålerförening för dig som har ett-tidsbegränsat uppehållstillstånd och är flykting eller alternativt skyddsbehövande 10 Möjligheten för din familj att komma och bo med dig som har tidsbegränsat uppehälistillstånd som flykting eller alternativt skyddsbehövande i Sverige styrs av flera saker. The possibility for your family to come and live with you in Sweden if you have a temporary residence permit as a refugee or a person in need of subsidiary protection is governed by several things. 11 Först måste du ha ett giltigt uppehållstillstånd. 12 Det betyder att om ditt uppehållstillstånd har slutat gälla och du har skickat in en ansökan om förlängt uppehållstillstånd, måste Migrationsverket först besluta om application for a residence permit extension, the Swedish Migration Agency must first decide if you have a right to a continued residence permit in Sweden. du har rätt till fortsatt uppehållstillstånd i Sverige. 13 Det räcker inte att du har en skyddsstatus och att du tidigare hade ett lt is not enough that you have a protection status and that you previously had a 14 Nar Migrationsverket har kontrollerat att du har ett giltigt uppehållstillstånd, prövar 🧶 Once the Swedish Migration Agency has verified that you have a valid residence vi din familjs ansökan om uppehållstillstånd. permit, we will review your family's residence permit application. 15 Migrationsverket börjar med att titta på om det är trolligt att du skulle få permanent uppehållsstllistånd efter en forlängningsansökan.

6 The Swedish Migration Agency begins by looking at if it is likely for you to receive uppehållsstllistånd efter en forlängningsansökan. 16 Du som är flykting eller alternativt skyddsbehövande har fått en statusförklaring As a refugee or person in need of subsidiary protection, you have received a som visar att du har bedömts vara i behov av skydd. status declaration that shows that you have been deemed to be in need of 17 Som utgångspunkt innebär din skyddsstatus att du har goda möjligheter att få ett 🌑 As a starting point, your protection status means that you have good chances of

Example of a postediting view

Swedish to English



Translations of appeals on decisions – PoC 2



Sensitive data – no public translation tool

Translation Arabic to Swedish

No commercial system available

Lack of sufficent training data for SMT or NMT

Rule-based system in Multra Cube

Modules of Multra Cube

Tokeniser module

for splitting text into words and punctuation; including character normalisation, and splitting of certain words that are written together. Data: List of words commonly written together.

Lexicon Lookup module

for looking up translations and other detailed information for words and phrases.

Data: Arabic–Swedish lexicons – one with Migration Agency terms, one general.

Tagging & Parsing module: NLP Cube

for analysing the grammatical structure of the original sentence.

Transformation module

for combining the lexicon and analysis data into the format that the Multra Transfer module uses.

Multra Transfer module

for transferring the grammatical structure to that of the target language. Data: Arabic–Swedish transfer grammar.

Multra Generation module

for generating a sequence of detailinformation items representing word order and inflection of the translation. Data: Swedish generation grammar.

Target-Lexicon Lookup module

for forming a string of inflected words and punctuation out of the detail information.

Data: Swedish lexicons – one with Migration Agency terms, one general.

Detokeniser module

for removing spaces around punctuation. Data: Swedish spacing rules.

Conclusions





*

MT + TM is a viable option to manual translation

*

In-house TS perform better than raw Google NMT



Certified domain TM is required for translation quality



Language data must be shared and collected



Controlled terminologies have a major impact on quality



Post-editing functions are needed

Collect data nation wide for building Translation Memories

Build domain specific terminology banks

Establish responsibility for language data within the organization

Collaborate with european agencies within your domain for enough data on unofficial text

Define a process for convenient integration of translation agency data into your TM





Thank you!

Anna Sågvall Hein, Professor Uppsala Universitet and Convertus Torbjörn Ekholm, Innovation Manager, Swedish Migration Agency





