Case study: Experimenting with machine translators

Kaisamari Kuhmonen
26 November 2019

Translation and Language Division

- Translation, language and terminology services to all government ministries

- Government translation needs: ~ 70 000 - 110 000 pages/year and rising

- Staff 70 (+ 14):

  - Swedish Language Unit 44, Foreign Languages Unit 17 (+ 14), Language Services Unit 9

  - Translators, terminologists, senior advisor for legislative affairs, language technology specialists, service coordinators

Need for MT

- Test runs with Commission's statistical engine **MT@EC** (2014-2017) + new neural engine **eTranslation** (2018-2019)

- Would tailored MT give better results?

  ➢ Commission-funded MT project to provide successive Presidencies with web-based MT engine in 2019/2020

  ➢ Meanwhile in Finland: Fiskmö project developing MT for FI-SV-FI

TAS for EU Council Presidency

- TAS = translation automation services

- MT tailored for 3 Presidency countries (using material from each country)

- Objective: facilitate multilingual communications

- Coordinated by Tilde language technology company; Romania, Finland and Croatia as partners

- Finnish Prime Minister's Office decided to participate to gain experience and exchange best practices

Co-financed by the Connecting Europe Facility of the European Union

EU Council Presidency Translator

= Presidency Translator for short

- Trained with some 20 million sentences (including our own translation memories)

- FI-EN-FI translator available in July 2019

- FI-SV-FI translator added as a bonus in September 2019

- Preliminary results encouraging → in daily use

- Accessible via Presidency Translator website (https://fi.presidencymt.eu/#/text) or SDL Studio (plug-in)

Fiskmö project

- Joint project (Universities of Helsinki and Turku + Kites), funded by Swedish Cultural Foundation in Finland

- Goals:

  - create massive parallel corpus in FI and SV to support linguistic research

  - build high-quality, reliable public MT services for FI-SV translation

- Web access + MT plugin for SDL Studio

- https://blogs.helsinki.fi/fiskmo-project/

Share your language data!

- Language data sharing crucial for development of high-quality MT, especially for smaller languages.

- Challenges for translation service:

    - How to distinguish between public and non-public information?

    - How to exclude personal information?

FINNISH GOVERNMENT

Questions and comments:

Kaisamari Kuhmonen
Head of Language Services Unit
Prime Minister's Office
kaisamari.kuhmonen@vnk.fi