

PRINCIPLE

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



The Principle project in MT

Co-operation of four institutions and one private company: Dublin City University, University of Zagreb, University of Iceland, The Norwegian National Library, and Iconic Translation Machines

<https://principleproject.eu/>

CEF project

PRINCIPLE is a 2 year Connecting Europe Facility (CEF) funded project (Action 2018-EU-IA-0050, Grant Agreement No. INEA/ CEF/ICT/A2018/1761837) whose main aim is to identify, collect and process high-quality Language Resources (LRs) for four under-resourced European languages:

- Croatian
- Icelandic
- Irish
- Norwegian (Bokmål and Nynorsk)

Outline

Language Resources (LRs) are collected from data holders and 'early adopters' from each of the 4 countries involved (in the specific domains of eJustice and eProcurement)

Machine Translation (MT) systems are produced from these LRs and evaluated to ensure high quality output

The MT systems are provided to 'early adopters' for free for the duration of the project

'Early adopters' use the MT systems and provide feedback

Based on MT output evaluation and early adopter feedback, high quality LRs are identified

Parallel LRs of high quality are uploaded to the ELRC-SHARE portal in order to improve the automated translation system eTranslation.

Stakeholders in Iceland

Arni Magnusson Institute

Ministry for Foreign Affairs

Icelandic Standards

Icelandic Met Office

University of Iceland

External Evaluation: ADAPT Centre, Dublin City University

- Automatic Evaluation

EN>IS	BLEU	METEOR	TER
Iconic	35.9	30.5	53.4
Google	23.6	24.1	63.1
Bing	18.4	21.4	69.6
eTranslation	23.2	24.3	69.5

IS>EN	BLEU	METEOR	TER
Iconic	45.3	35.7	43.0
Google	31.9	30.1	53.7
Bing	26.0	27.7	57.2
eTranslation	31.6	29.5	54.5

TUs collected
814,000

- Human Evaluation (Post-editing)

EN>IS

	Mode of operation	No. of segments	Segment number	Overall time
Evaluator 1	Translation	500	1-250	7h 45m
Evaluator 2	Post-editing	500	1-250	3h 28m
Evaluator 1	Post-editing	500	251-500	8h 12m
Evaluator 2	Translation	500	251-500	6h 10m

External Evaluation

- Automatic Evaluation

EN>IS	BLEU	METEOR	TER
Iconic	29.0	27.4	54.1
Google	21.7	24.0	60.1
Bing	19.4	22.4	63.4

TUs collected
17,000

External Evaluation

- Automatic Evaluation

IS>EN	BLEU	METEOR	TER
Iconic	59.7	44.4	22.7
Google	01.8	15.5	107.3
Bing	02.3	14.9	96.8

“SE 18-23 AND RAIN. S 13-20 TONIGHT, STRG WND IN E. LGT WND LATE TONIGHT.”

TUs collected
212,000

PRINCIPLE

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

