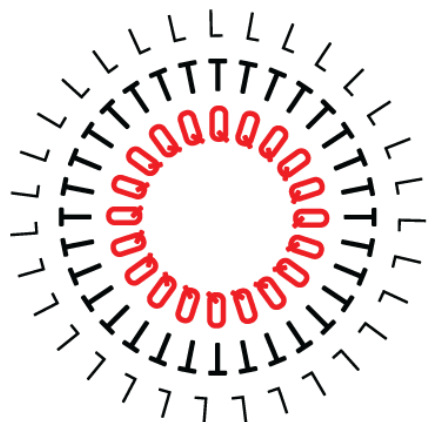# IT-domain translation task

@WMT 2016, Berlin, August 11

Martin Popel

IT-task overview

- 7 language pairs (**4 new in WMT**), English to:
  - **Basque**, **Bulgarian**, Czech, **Dutch**, German, Spanish, **Portuguese**
- 31 systems submitted by 10 teams
- 82k manual-evaluation pairwise comparisons
- Test data: 1000 answers (QTLeap Corpus, Batch 3)
- In-domain training data
  - 2000 answers (QTLeap Corpus, Batches 1+2)
  - localization files (LibreOffice, KDE, VLC), Wikipedia terms
- Out-of-domain training data
  - all corpora from News task
  - PaCo2-EuEn (Basque), SETimes (Bulgarian)

IT-task specifics

- Adaptation to IT domain
  - technical jargon
  - frequent named entities: menu items, product names

- Adaptation to translation of help-desk answers
  - Instructions with imperatives

Example

Source:

Click and drag the file onto the Trash icon **on the Dock**. Alternatively, you can select the file and press Command-Delete.

Czech translation by the QtlMoses system:

Přetáhněte soubor do koše ikonu **na lavici obžalovaných**. Případně můžete vybrat soubor a tisk parametry – odstranit.

"bench of the accused ones"

Results (1)

English→Bulgarian

#	score	range	system
1	5.26	1	QTL-MOSES
2	−5.26	2	QTL-DEEPFMOSES

English→Spanish

#	score	range	system
1	3.53	1	PROMT-HYBRID
2	−0.80	2–3	QTL-CHIMERA
	−0.81	2–3	QTL-TECTOMT
3	−1.93	4	QTL-MOSES

English→Czech

#	score	range	system
1	0.53	1–2	QTL-CHIMERA-PURE
	0.43	1–2	ILLC-UVA-DS
2	0.13	3	QTL-TECTOMT
3	−0.47	4–5	QTL-CHIMERA-PLUS
	−0.62	4–5	QTL-MOSES

English→Basque

#	score	range	system
1	1.57	1	QTL-TECTOMT
2	−1.57	2	QTL-MOSES

Results (2)

English→German

#	score	range	system
1	1.61	1	PROMT-RULE-BASED
2	−0.04	2–5	UHBS-LMI
	−0.06	2–6	UHDS-DOC2VEC
	−0.06	2–6	QTL-RBMT-SMTMENUS
	−0.09	3–6	RBMT
	−0.10	3–6	QTL-RBMT-MENUS
3	−0.19	7–8	DFKI-SYNTAX
	−0.19	7–8	JU-USAAR
4	−0.38	9	QTL-SELECTION
5	−0.49	10	QTL-MOSES

English→Dutch

#	score	range	system
1	1.95	1	ILLC-UVA-SCORPIO
2	0.36	2	QTL-CHIMERA
3	0.15	3	QTL-TECTOMT
4	−2.46	4	QTL-MOSES

English→Portuguese

#	score	range	system
1	4.61	1	PROMT-HYBRID
2	−1.06	2	QTL-TECTOMT
3	−1.27	3	QTL-CHIMERA
4	−2.28	4	QTL-MOSES