

# Twitter Crowd Translation

Eduard Šubert

Faculty of Nuclear Sciences and Physical Engineering  
Czech Technical University in Prague

Ondřej Bojar

Institute of Formal and Applied Linguistics  
Faculty of Mathematics and Physics  
Charles University, Prague, Czech Republic

November 16, 2014

# Twitter Usage

284 million monthly active users

500 million Tweets are sent per day

80% of Twitter active users are on mobile

<https://about.twitter.com/company>

# Twitter Usage

Twitter supports 35+ languages

<https://about.twitter.com/company>

# Goals

- ▶ provide translations of tweets
- ▶ accurate
- ▶ fast (aiming at “instant”)
- ▶ stay within the Twitter platform

# Machine Translated Tweet

@Philae2014:

Some problems with my cold gas thruster could mean that it might be up to harpoons & ice screws to make sure I stay on  
#67P #CometLanding

některé problémy se zimou plynových trysek by mohlo znamenat, že by to mohlo být až harpuny & mnaga; led šrouby, aby se ujistil, že zůstane na 67 P # # cometlanding

some problems with cold gas jets could it be to harpoon & mnaga; ice screws to assure that stays on 67 p # # cometlanding

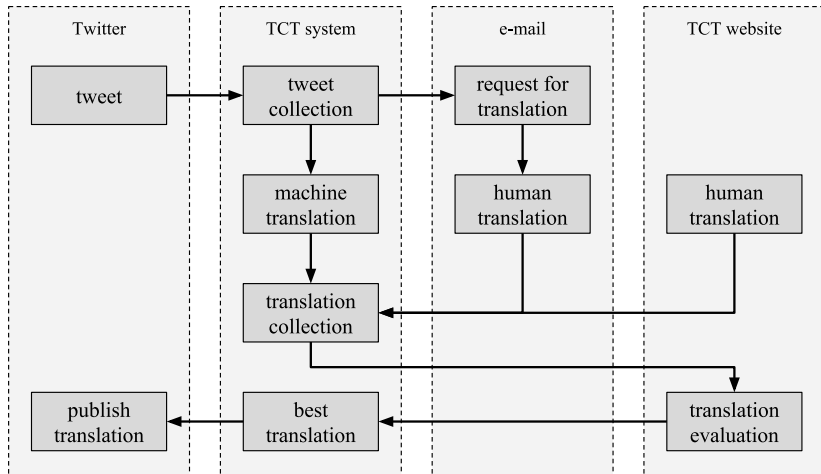
# Goals

- ▶ collect appropriate parallel corpus
- ▶ provide translations of tweets
- ▶ accurate
- ▶ fast (aiming at “instant”)
- ▶ stay within the Twitter platform

# TCT - Twitter Crowd Translation

- ▶ human translation of tweets
- ▶ scoring of translations
- ▶ autonomous tweet collection and publishing

# TCT Workflow





# Tweet Example



tct retweeted



**Stefan Füle** @StefanFuleEU · Sep 26

Practical overview of EU financial help to [#Ukraine](#) - how much we provide, for what & under what conditions; read here: [bit.ly/1utqdbf](https://bit.ly/1utqdbf)



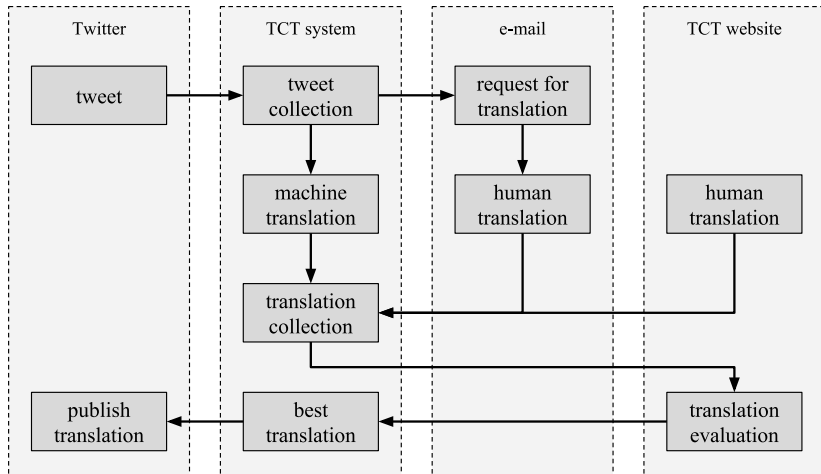
37



10



# TCT Workflow



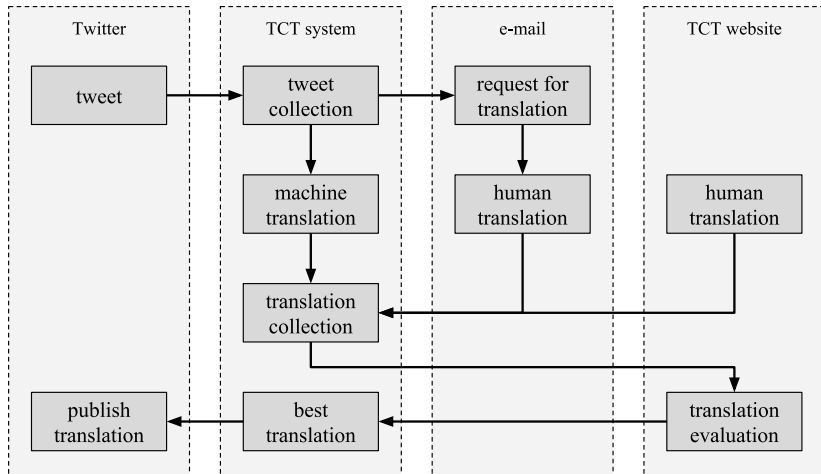
# E-mail Example

Please translate the following post to language: Czech

Practical overview of EU financial help to #Ukraine - how much we provide, for what & under what conditions; read here: [bit.ly/1utqdbf](https://bit.ly/1utqdbf)

ID:51d46e56c720c0a20e98e7e628b5806a

# TCT Workflow



# Translation Evaluation

Practical overview of EU financial help to #Ukraine - how much we provide, for what & under what conditions; read here: [bit.ly/1utqdbf](https://bit.ly/1utqdbf)

Praktický přehled finanční pomoci EU Ukrajině. Kolik poskytujeme, za co a za jakých podmínek. Čtete zde

hate it

dislike it

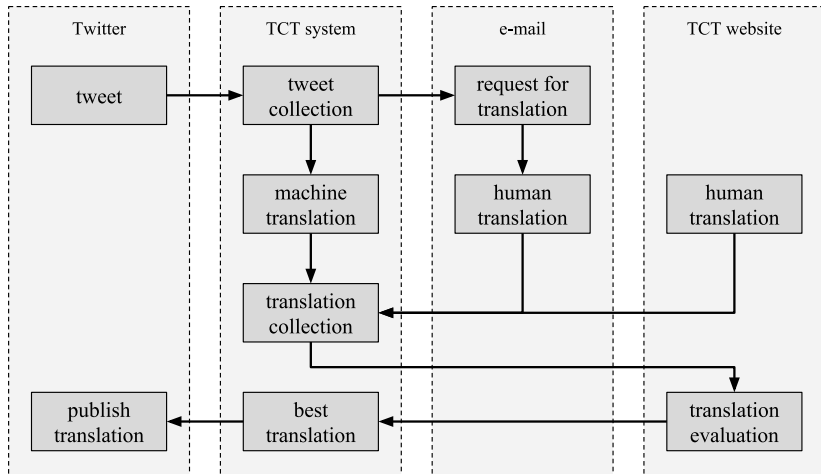
it's OK

like it

love it

Skip

# TCT Workflow



# Best Translation

tct @tctranslation · Sep 27

.@StefanFuleEU Praktický přehled finanční pomoci EU Ukrajině. Kolik poskytujeme, za co a za jakých podmínek. Čtete zde <http://t.co/ZHHOwJrL...>



[View conversation](#)



tct retweeted



**Stefan Füle** @StefanFuleEU · Sep 26

Practical overview of EU financial help to [#Ukraine](#) - how much we provide, for what & under what conditions; read here: [bit.ly/1utqdbf](http://bit.ly/1utqdbf)



37



10





Tweets



Score



Translators



About



@tcttranslation

quest.ms.mff.cuni.cz/tweeslate

 @tcttranslation

powered by: [CakePHP 2.5.3](#), [Zurb Foundation 5](#) and [Moses Decoder](#)

© Charles University in  
Prague, ÚFAL



# Tweet Translation

- ▶ context
- ▶ abbreviation

Terrorists used #Russian-supplied Smerch MRLs against  
#Ukraine forces in the conflict zone <http://t.co/s9ZKP2zsVk>  
EMPR <http://t.co/9ElqFP7qlg>

Mykolaiv armored plant handed over 10 APCs to the border  
guards <http://t.co/ANOtAKcUdw> via @HromadskeTV  
<http://t.co/mUmfbYKuwf>

Terrorists used #Russian-supplied Smerch **MRLs** against #Ukraine forces in the conflict zone <http://t.co/s9ZKP2zsVk>  
EMPR <http://t.co/9ElqFP7qlg>

Mykolaiv armored plant handed over 10 **APCS** to the border guards <http://t.co/ANOtAKcUdw> via @HromadskeTV  
<http://t.co/mUmfbYKuwf>

MRLs = multiple rocket launcher  
APC = armoured personnel carrier

Russian SOF in Ukraine with "Polite People" patch and new gear #CrisisUcrania <http://t.co/gqLvuhCof>

Russian SOF in Ukraine with "Polite People" **patch** and new **gear** #CrisisUcrania <http://t.co/gqLvuehCof>



Terrorists used #Russian-supplied Smerch MRLs against  
#Ukraine forces in th conflict zone <http://t.co/s9ZKP2zsVk>  
EMPR <http://t.co/9ElqFP7qlg>

Mykolaiv armored plant handed over 10 APCs to the border  
guards <http://t.co/ANOtAKcUdw> via @HromadskeTV  
<http://t.co/mUmfYKuwf>

# Summary

- ▶ human translation of tweets
- ▶ translation evaluation
- ▶ parallel corpus
- ▶ understanding of tweet structure