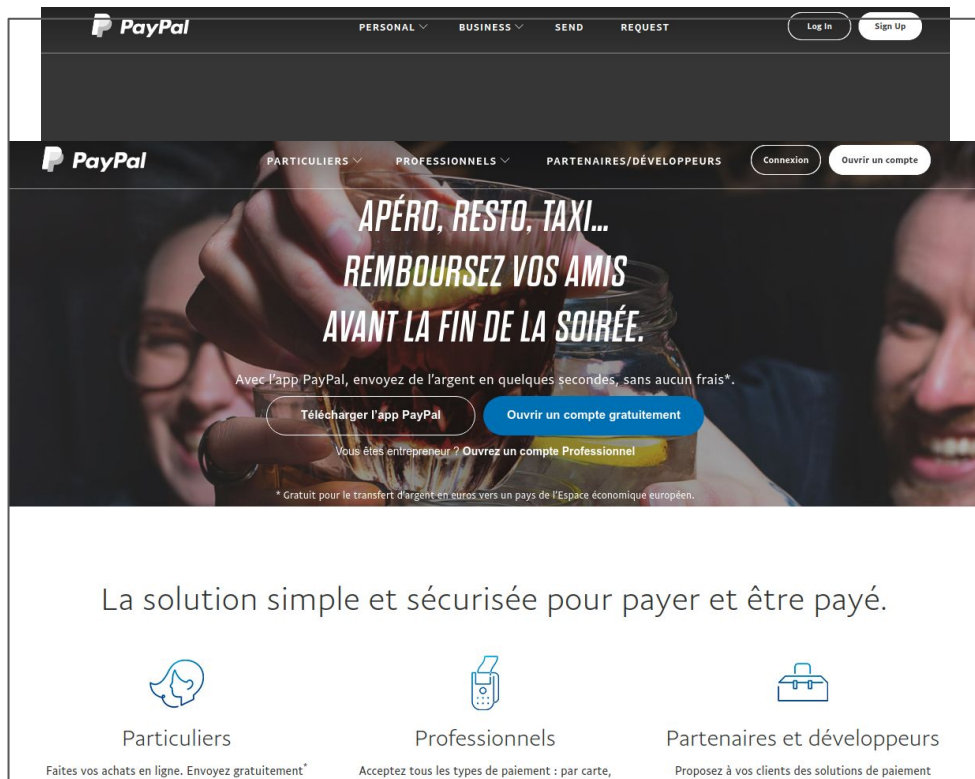


# Localizing HTML Web Pages for Francophone Audiences with Machine Translation

By Johnny Driscoll

# Introduction

- What is Localization?
- Why Localize?



The screenshot displays the PayPal website interface in French. The top navigation bar includes the PayPal logo, links for 'PERSONAL', 'BUSINESS', 'SEND', and 'REQUEST', and buttons for 'Log In' and 'Sign Up'. Below this, a secondary navigation bar features 'PARTICULIERS', 'PROFESSIONNELS', and 'PARTENAIRES/DÉVELOPPEURS', along with 'Connexion' and 'Ouvrir un compte' buttons.

The main banner features a background image of two people smiling and clinking glasses. The text on the banner reads: 'APÉRO, RESTO, TAXI... REMBOURSEZ VOS AMIS AVANT LA FIN DE LA SOIRÉE.' Below this, it states: 'Avec l'app PayPal, envoyez de l'argent en quelques secondes, sans aucun frais\*.' Two prominent buttons are present: 'Télécharger l'app PayPal' and 'Ouvrir un compte gratuitement'. A smaller link 'Vous êtes entrepreneur ? Ouvrez un compte Professionnel' is also visible.

At the bottom, a section titled 'La solution simple et sécurisée pour payer et être payé.' lists three user types with corresponding icons: 'Particuliers' (person icon), 'Professionnels' (POS terminal icon), and 'Partenaires et développeurs' (briefcase icon). Each category has a brief description of the service offered.

**Particuliers**  
Faites vos achats en ligne. Envoyez gratuitement\*

**Professionnels**  
Acceptez tous les types de paiement : par carte,

**Partenaires et développeurs**  
Proposez à vos clients des solutions de paiement

# Machine Translation (MT)

- Research began in the 1950's
- Statistical MT
- Human evaluation
- Automated evaluation (BLEU)
- Google Translate (2006)
- Neural MT (2016)



1954

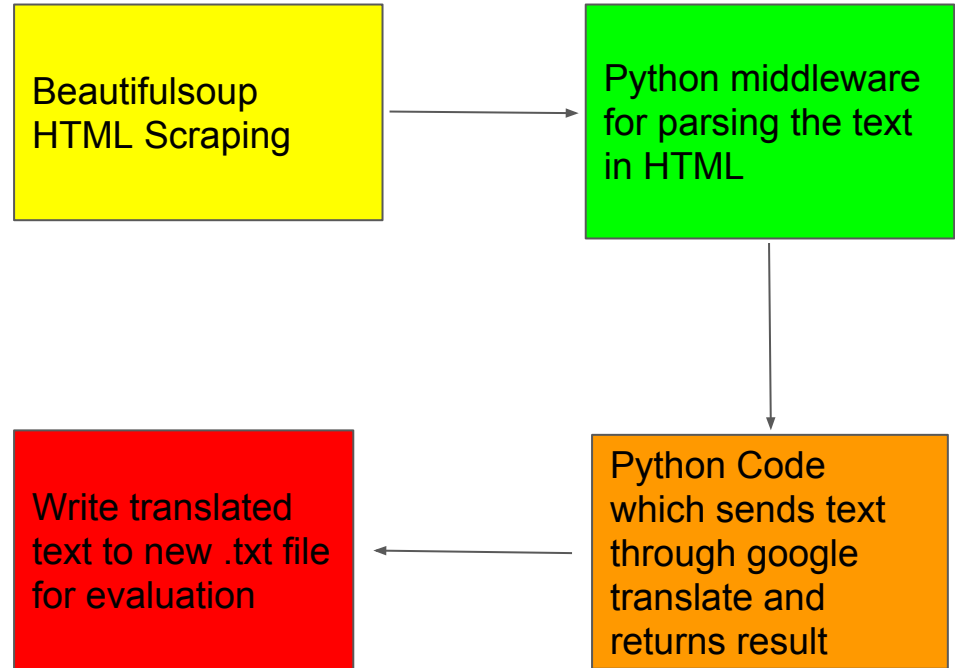


Sentences in Russian are punched into standard cards for feeding into the electronic data processing machine for translation into English

Georgetown-IBM experiment


# Can MT provide high-quality automated Localization?

- Python Implementation
- Translation pipeline



# Data


- 3 different static web pages
  - a. Union College CS Homepage
  - b. Political article from Yahoo News
  - c. Blog post from MIT's Technology Review
- Reference Translations



Chair  
Faculty  
Sem  
Student  
C  
Re

## U.S. Officials Tell World Leaders To Ignore Trump's Tweets


Mary Papenfuss · HuffPost · February 10, 2018




President Donald Trump's tweets not only trigger worries at home.

President **Donald Trump's** tweets not only trigger worries at home. They rattle leaders around the world.


But U.S. officials reportedly tried to reassure global policy representatives at an international conference in Germany over the weekend, urging them to pay no attention to the president's Twitter rants.



Popular in the Community



Trump focus on mental health after shooting denounced  
9,359 reactions 4%



### This Robot Gives You Arms to Lift a Half-Ton

Say hello to Guardian GT, who wants to take the strain out of your manual labor.

This mean-looking machine from Sarcos Robotics is designed to give you superhuman strength. Its pair of seven-foot arms, each able to lift 500 pounds, is controlled with eerie smoothness by a human. The controller provides force feedback so that the operator can feel what's happening—though it only takes five pounds of force, instead of the full 1,000, according to Wired. The dimensions of the arms are scaled like those of a regular human, too, all in a bid to make the system intuitively easy to use. Sarcos says it's designed to perform tasks like emptying massive disposal bins or moving huge metal pipes, with a diesel motor that can keep it going at up to four miles per hour for seven hours.

As we've said before, humans and robots will need to work together to be efficient in the future. If that means using bots like this, count me in.

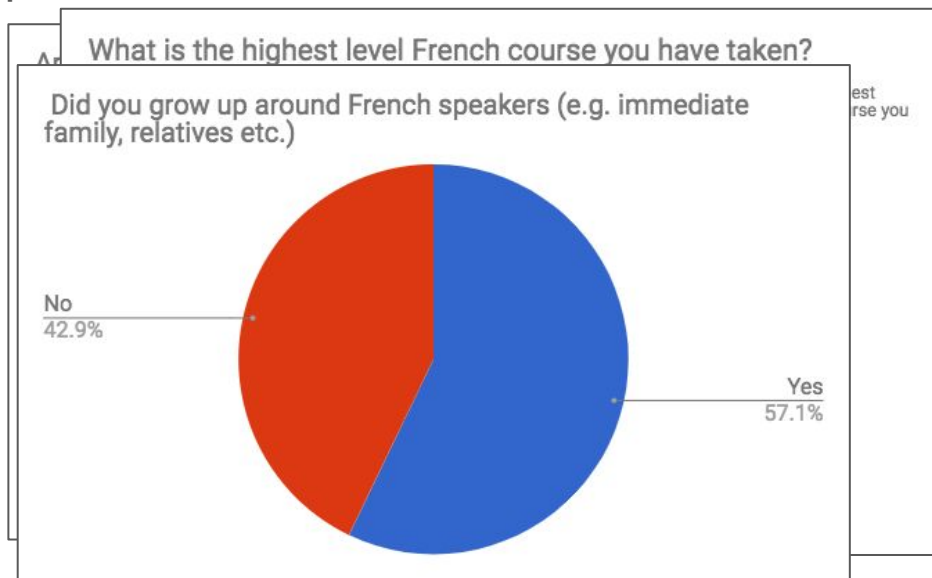
# Outline of Methods

- 14 Participants
- Participants' Background in French
- Participants evaluated each translation
- No access to reference translations
- Quantitative and Qualitative feedback
  - Adequacy + Fluency Scores
  - Post-Editing Marks

Metric	Score	Definition
ADEQUACY	1	None of the meaning is preserved
	2	Little of the meaning is preserved
	3	Much of the meaning is preserved
	4	Most of the meaning is preserved
	5	All the meaning is preserved
FLUENCY	1	Incomprehensible target language
	2	Error-filled but understandable
	3	Fewer errors and better comprehension
	4	Good quality target language
	5	Flawless quality target language

# Preliminary Questions

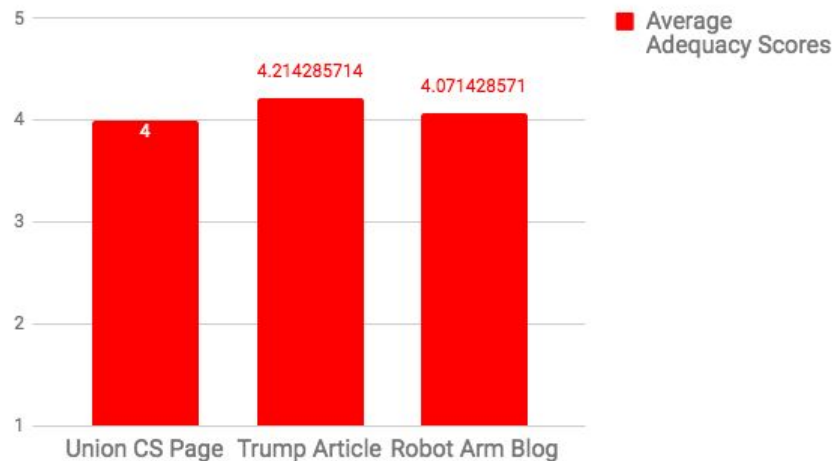
- Native French speaker?
- Did you grow up around French speakers?
- Highest level French course?



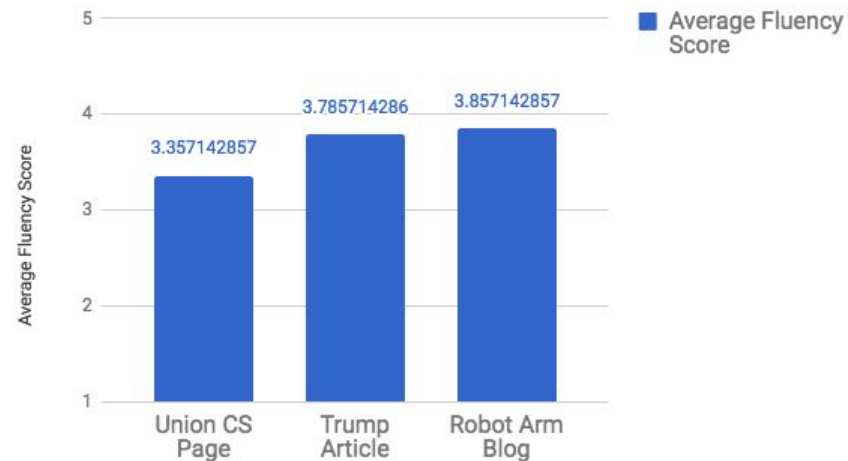
# Quantitative Results

- Adequacy
- Fluency

Average Adequacy Scores



Average Fluency Scores





# Qualitative Results

- Word choice errors
- Verb tense errors
- Word Placement errors
- Examples from translations

REFERENCE: Last year, Secretary of State Rex Tillerson

TRAN

REFERENCE: Rewriting Life

Rex Tillerson

TRANSLATION: récrire la vie '



Ilyas El Arif  
4:23 PM Feb 20



Delete: "CS"



Ilyas El Arif  
4:30 PM Feb 20



Replace: "La dernière année" with "l"



Omar Lebbar  
2:09 PM Feb 24

Resolve

Reconstitution de la vie

Reply...



Marie Marceau  
8:52 PM Feb 21

Resolve

What do you mean by "computing effort"? "Les efforts de calcul" doesn't mean anything in this sentence

Reply...

# Conclusion

- French-Speakers surprised with quality of translations
- Use NMT as preliminary translation service
- Reduce work for human editors

# Future Work

- Get results for automated metric (BLEU score)
- Evaluate on more diverse set of websites/translations

# Acknowledgements

I would like to thank my advisors Nick Webb and Charles Batson, as well as my parents and siblings for their support and guidance throughout this project.

I would also like to thank David Frey for always being a helpful and positive presence in the department.