

Localizing HTML Web Pages for Francophone Audiences with Machine Translation



John E. Driscoll

Nick Webb (Computer Science) and Charles Batson (French), Advisors



Introduction:

- Internationalization and Localization are the process of adapting computer software to different languages, regional differences and technical requirements of a target locale [1].
- Can Machine Translation be used for Web Page Localization?
- Integrate open-source Machine Translation tools with my own middleware to increase the automation of the localization workflow

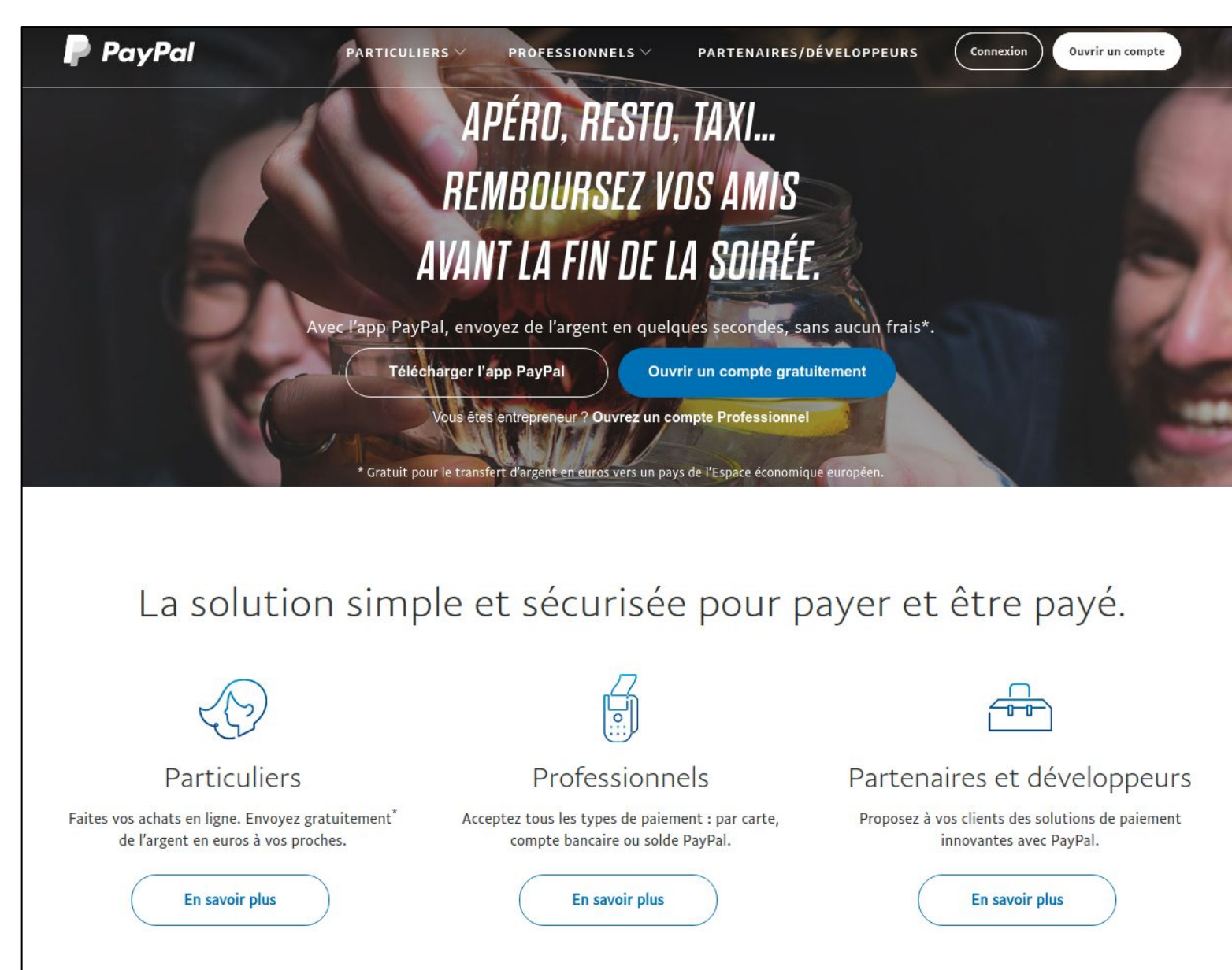


Figure 1. Paypal's home page for France.

Data:

In order to test my hypothesis in a variety of situations, I have selected 3 different types of static HTML web pages:

- 1) The Union College CS Department Homepage
- 2) The Yahoo News about Trump's Tweets
- 3) MIT's Technology Blog, "The Download"

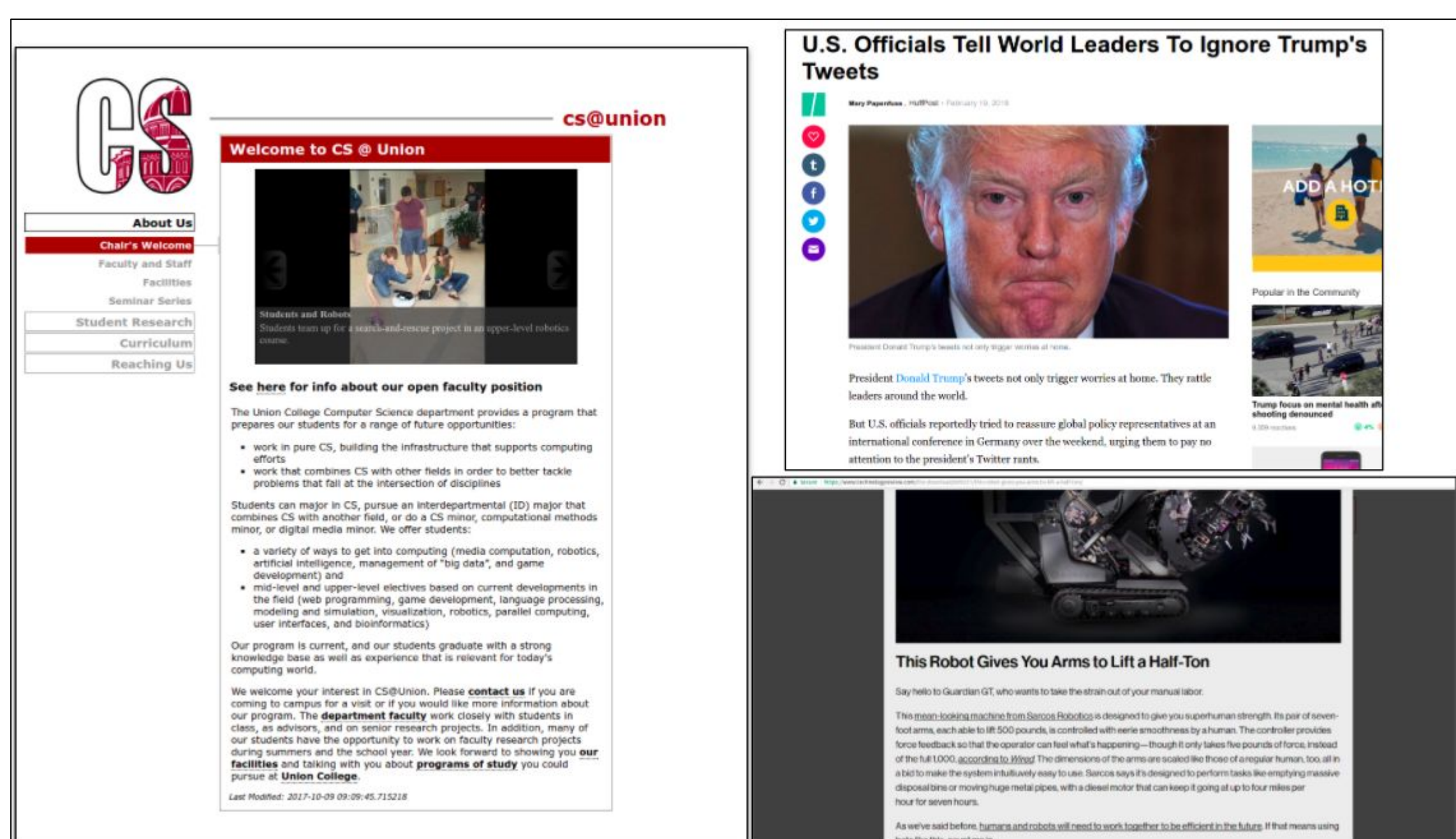


Figure 3. Screenshots of Web Pages.

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

Figure 3. Adequacy and Fluency scoring chart.

Methods:

- Use BeautifulSoup, a Python library, for HTML Scraping
- Parse strings from the HTML document by focusing on specific tags
- Use Python script that sends strings through Google Translate
- Replace English strings with translated French strings and write a new HTML text file with Python
- Evaluate Translation quality via human evaluation metrics (adequacy and fluency) and automated metric BLEU

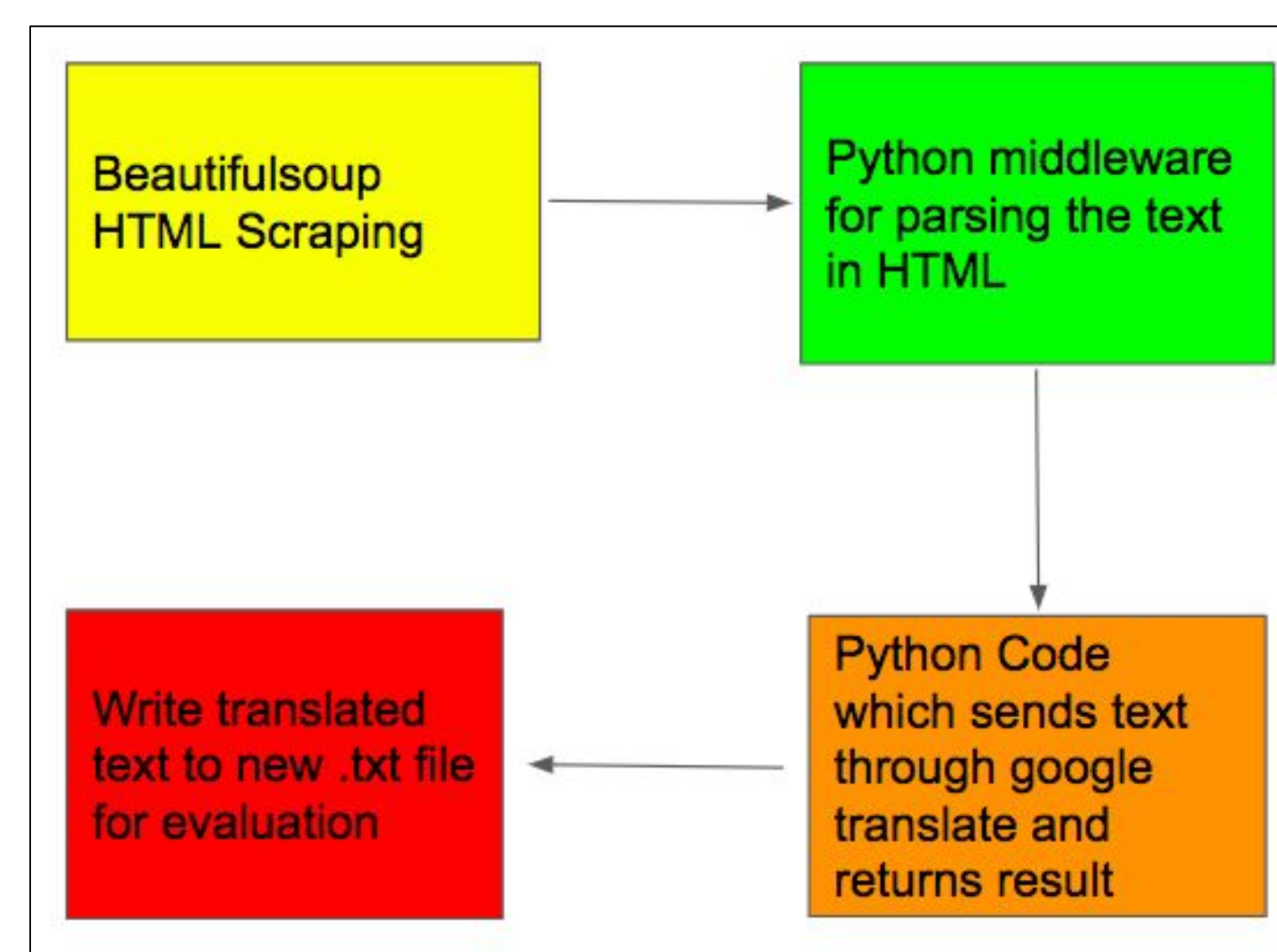


Figure 4. BS4 and Python Translation Pipeline.

Results:

- Strong Adequacy and Fluency Scores
- Participants surprised by quality of translations
- Translations are understandable despite grammatical errors
- BLEU scores did not correlate to human evaluations due to small test corpus of reference translations [2]

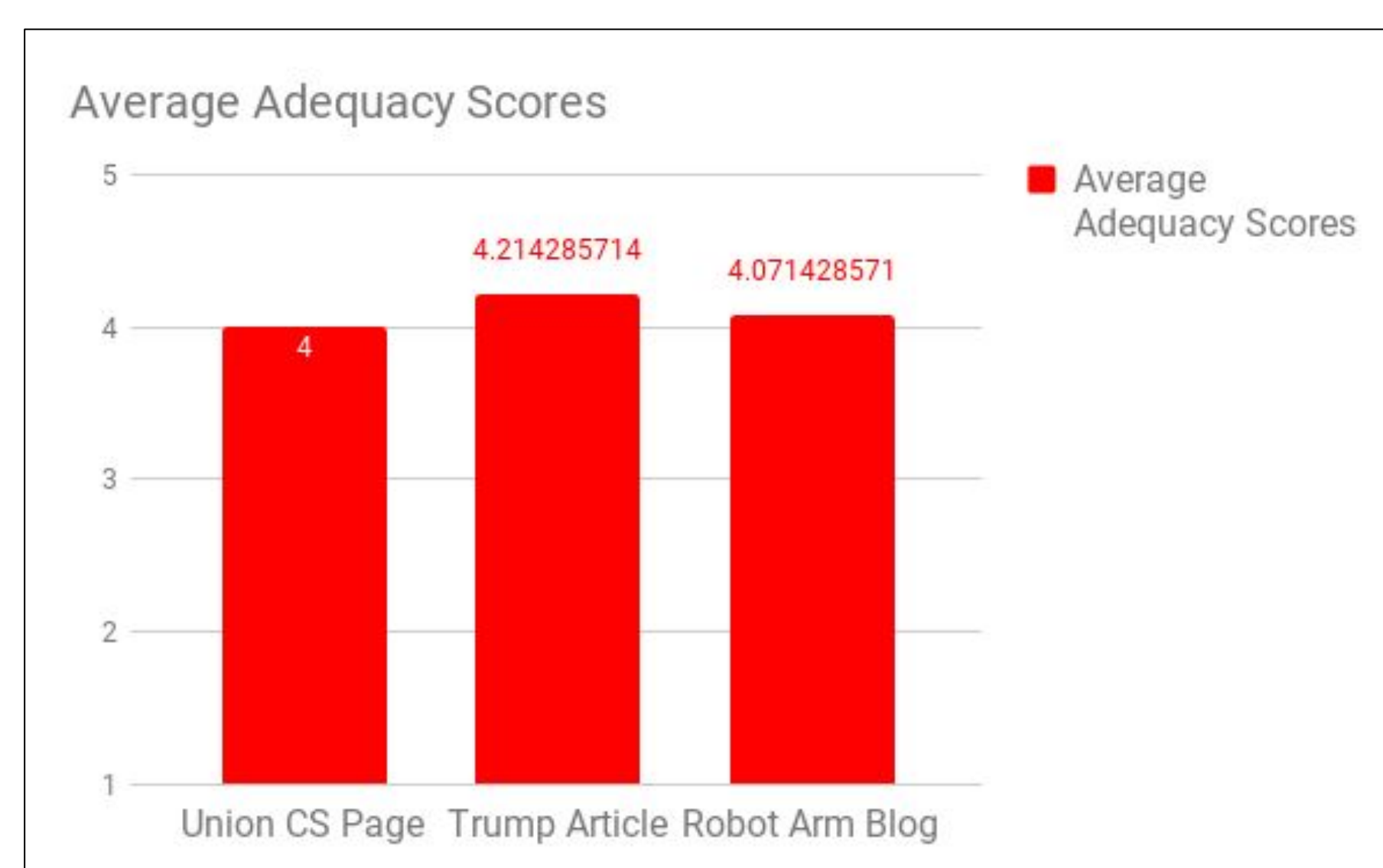


Figure 5. Average Adequacy Scores.

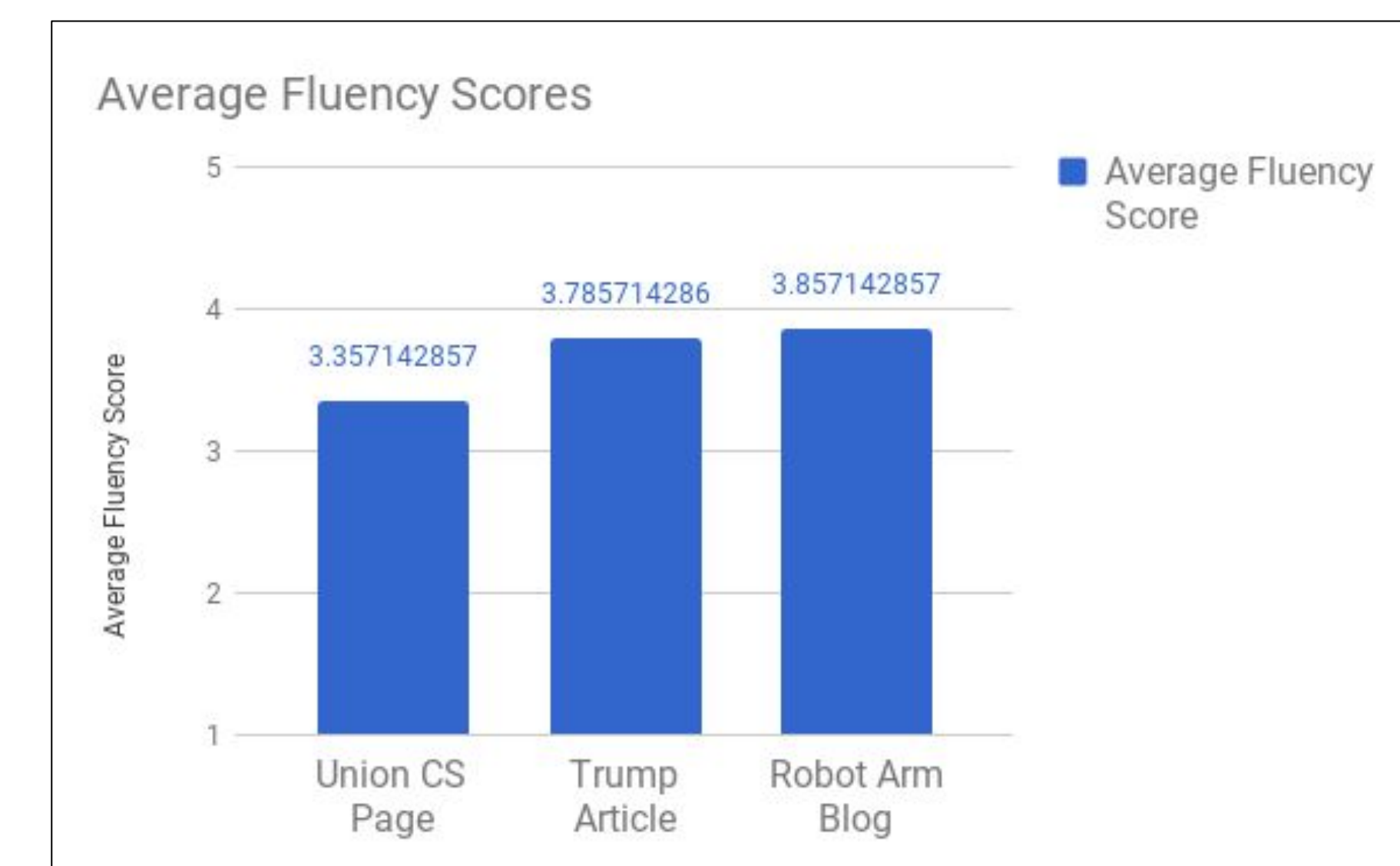


Figure 6. Average Fluency Scores.

Conclusions:

- Based on both the quantitative and qualitative feedback gathered from 14 participants, I believe MT could aid in the translation aspect of the localization process
- BLEU scores did not correlate to human evaluations and were therefore negligible
- The automatically generated translations could act as a preliminary tool in order to reduce the human workload
- Human editors could correct the mistakes in the machine translation output rather than starting the translation from scratch thereby saving organizations both time and money when localizing software and web apps

Future Work:

- Recruit more human evaluators to participate in experiment to generate more feedback
- Test machine translation on more diverse array of web pages
- Create multiple reference translation with the help of other French-speakers to improve BLEU scores for comparison to human evaluation metrics

Acknowledgments:

I would like to thank my parents and siblings for their support and guidance.

References:

- [1] X. Wang, L. Zhang, T. Xie, H. Mei, and J. Sun, "Locating need-to-translate constant strings in web applications," Proceedings of the eighteenth ACM SIGSOFT international symposium on Foundations of software engineering, pp. 87–96, Nov. 2012. .
- [2] Kishore Papinehi et al. "BLEU: a method for automatic evaluation of machine translation". In: *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (July 2002), pp. 311–318.