# Software Engineering Project Proposal: A Turing Test Game for Machine Translation

## Can you catch the robot translators? And are you better than a robot judge?

#### Date: 2022-9-19 Authors: Chi-kiu Lo Multilingual Text Processing Team Digital Technologies Research Centre National Research Council Canada

### Background

The National Research Council of Canada’s (NRC) Multilingual Text Processing (MTP) team carries out research and development in multilingual natural language processing (NLP). This includes machine translation and other language technologies for multilingual contexts. Machine translation (MT) is making computer models learn to translate between human languages (e.g. translating English into French) by reading through hundreds of thousands of translated documents. Our team works on two areas in machine translation: 1) building and improving MT systems and 2) improving automatic evaluation of translation quality.

In MT system research, the goal is building system that would generate translations which quality is as accurate, fluent and natural as human translators. The older paradigm of MT, statistical MT, was less accurate and fluent than the newer paradigm, neural MT. The NRC MTP team has years of experiences and achieved fruitful results in both paradigms. For instance, statistical MT systems built by the team have been deployed to several federal government departments and neural MT systems built by the team have achieved top rankings in international research competitions.

On the other hand, the goal of automatic MT evaluation research is building metrics that would score translation quality as accurate as human in the lowest cost possible. The automatic evaluation metric, YiSi, built by the NRC MTP team has ranked top in international research competitions for better correlation with human judgment on translation quality.

### Your Challenge

As a federal science department, the NRC has a duty to engage the public in outreach opportunities displaying our science discoveries and inventions. Outreach opportunities include but not limit to “Open Door Ottawa”, “Bring Your Kid to Work Day”, High School Career Day, etc. Participants in these outreach events range from 5 years old to 99, with different education levels. Instead of going through some hard science facts, an interactive game is a better introduction to the results of our research.

Your challenge is to create this interactive game to display the two areas of research of the MTP team: MT results and automatic evaluation of translation quality.

The target players of the game are visitors in outreach events. A set of sentences, including one in the original language and several translations of different quality, is shown to the player. The player is challenged to identify the human translation from the set and compete with YiSi -- the robot judge in doing so.

To accommodate different technical resources available in the outreach events (network may or may not be available), the game should be flexible in set up and plug-n-play/portable for different devices (tablets or laptop; android/iOS/windows platform).

### Preliminary knowledge

A game play animation is attached to this design document. You should go through it to have a rough idea of the game flow.

### Milestones

### Milestone 1: admin panel development

A game admin panel is used by NRC staff to start the game by choosing the game parameters, such as the set of sentence tuples, number of tuples drawn from the set in each game, etc.

### Milestone 2: game development

Game flow in this milestone:

1. Display a sentence tuple: the original sentence, human reference translation and three candidate translations.
2. Allow the player to select potential robot translations and cancel selection.

Allow the player to confirm decision.

1. Display robot judge (YiSi)’s decision. Move on to the next step when player is ready. (e.g. pressing the next button; but open for other implementations)
2. Display true identity of translators (human/SMT/NMT). Move on to the next step when player is ready. (e.g. pressing the next button; but open for other implementations)
3. Display player vs. robot judge scores. Move on to the next step when player is ready. (e.g. pressing the next button; but open for other implementations)
4. Go to step 1 with a new sentence tuple. Game ends when the loaded sets of sentences are exhausted or the game admin chooses to end it.

### Milestone 3: data log and output

Log user decisions in the game play.

When game ends, allow game admin to save the player log for further analysis.