# Literature review

## Introduction

Due to the extraordinary globalization, cross-language plagiarism detection tools become necessary. For example, some international students who have other languages as mother tongue rather than English might translate some paper without stating the origin.

In this literature review, we discuss 3 following questions:

Q1: What is plagiarism (include cross-language situation)?

Q2: How can the detection tools deal with different languages?

Q3: How can our team have a success in building the tool?

# Research and Discussion

In this section, we are discussing 3 above questions.

### Q1: What is plagiarism (include cross-language situation)?

According to the Oxford Dictionary and the Longman Dictionary, basically, plagiarisms are somebodies take others' work without declaration.

This question involves copy, translation, references, bibliographies, paraphrase and machine translation. We use a decision matrix to illustrate this:

|  |  |  |
| --- | --- | --- |
|  | References/Bibliographies | Without Declaration |
| Copy | × | × |
| Machin translation | ? | × |
| Human translation | √ | × |
| Paraphrase | √ | × |

There is one case should be confirmed with the uses, which is using the machine translators with references or bibliographies.

### Q2: How can the detection tools deal with different languages?

Machine translators should be used instead of human assessors because of the increasing costs of human resource. Choosing commercial translate services would be reasonable because of the astonished difficulties, include technologies and budget.

According to the article, "A Monte Carlo Method for Metamorphic Testing of Machine Translation Services", Google translation might be a good choice. We may use some metrics, such as Levenshtein Distance, BLEU and Cosine Similarity, to access the similarity.

There is one confusion should be clarified with the uses, which are the translate directions, namely, translate from English to the target language and translate from the target language to English. Should the detection tool deal with these two directions?

In addition, we assume that the cross-language detection tool might be better with the natural sciences area than humanity area. We could prove or disprove this hypothesis at the test stage.

### Q3: How can our team have a success in building the tool?

Building a tool with limited time and budget, the functions of the tool should be restricted.

We may use the wiki of GitHub as a Kanban to guide the team, to show the progress of development and balance team members' workload.

To reduce the risk of failure, we prefer agile rather than a waterfall. We are using a walking skeleton, which is a tiny runnable model to develop the tool. We are going to deploy and confirm the model frequently.

To deal with emergencies, address changes, improve efficiency and balance workload, we use cellular manufacturing system. It means though we divide the team members into different roles, we persuade a person to learn different skills and different work.

## Conclusion

To fulfil the task, we will search continually and use flexible development methods.

By Peng Tian, 11 Aug. 18

References:

(Daniel Pesu, 2018, RADIGAN, Davis, 2014, Inman, Xiongyi LI, 2012)

*plagiarism* [Online]. Available: <https://en.oxforddictionaries.com/definition/plagiarism> [Accessed 11 Aug 2018].

DANIEL PESU, Z. Q. Z., JINGFENG ZHEN, DAVE TOWEY. A Monte Carlo Method for Metamorphic Testing

of Machine Translation Services. MET '18 Proceedings of the 3rd International Workshop on Metamorphic Testing, May 27 - 27, 2018 2018 Gothenburg, Sweden. New York, NY, USA: ACM, 38-45.

DAVIS, N. 2014. *Kickstart your next project with a Walking Skeleton* [Online]. Available: <https://codeclimate.com/blog/kickstart-your-next-project-with-a-walking-skeleton/> [Accessed 11 Aug 2018].

INMAN, R. A. *CELLULAR MANUFACTURING* [Online]. Available: <http://www.referenceforbusiness.com/management/Bun-Comp/Cellular-Manufacturing.html> [Accessed 11 Aug 2018].

RADIGAN, D. *What is kanban?* [Online]. Available: <https://www.atlassian.com/agile/kanban> [Accessed 11 Aug 2018].

XIONGYI LI, S. Z. 2012. A Study of Cell Production Mode from the Perspective of

Labor Division. *2012 International Conference on Engineering and Business Management.*