

ViMMRC 2.0 Dataset User Agreement

1. User Information

- Full name:
- Organization:
- Email:
- Telephone (optional):

Note: If you are a student, this user agreement should be signed by your supervisor.

2. Dataset Information

- Dataset Name: **ViMMRC 2.0**
- Size: 699 reading passages and 5,273 questions with multiple-choice answers.

3. Terms and Conditions

The user of ViMMRC 2.0 developed by the NLP@UIT research group must respect the following terms and conditions:

1. The dataset is only used for non-profit research for natural language processing and education.
2. The dataset is not allowed to be used in commercial systems.
3. Do not redistribute the dataset. This dataset may be modified or improved to serve a research purpose better, but the edited dataset may not be distributed.
4. Summaries, analyses, and interpretations of the properties of the dataset may be derived and published, provided it is not possible to reconstruct the information from these summaries.
5. Published research works that use the dataset must cite the following paper:

Son T. Luu, Khoi Trong Hoang, Tuong Quang Pham, Kiet Van Nguyen, and Ngan Luu-Thuy Nguyen. 2023. *A Multiple Choices Reading Comprehension Corpus for Vietnamese Language Education*. *arXiv preprint arXiv:2303.18162* (2023).

By the User

Signature:

Date:

Name (please print):

**Forward the signed agreement by PDF to this email sonlt@uit.edu.vn*