

# The GaitLU-1M Dataset Ethical Requirements

## Introduction

Gathering web data on a large scale has become a common practice that significantly promotes AI research and applications. However, what cannot be ignored is that large-scale web crawling inevitably involves personal privacy since most Internet data is created by and related to individuals. This is still an unsolved ethical deal between privacy protection and AI advances. On this background, the distributor of the GaitLU-1M dataset would like to try its best to protect personal data by taking the following measures.

- Data Minimization.** The data of GaitLU-1M is exhibited as the sequence of binary silhouettes and body skeletons. No visual RGB information will be available.
- Expiration Date.** GaitLU-1M can only be used for 20 years since the corresponding paper publishes. After this date, all data will be deleted and not allowed to be used.
- Opt-out.** For each gait sequence in GaitLU-1M, its data will be deleted once identified by some algorithms or people with providing confirmed evidence.
- Unattended Pre-treatment.** The pre-processing of GaitLU-1M is based on automatic algorithms, meaning that nobody has manually processed personal visual data.
- Anonymization.** The raw data of GaitLU-1M, including video links and other intermediate results, have been deleted. In addition, weak annotations such as the date and city information have been erased by randomly numbering the sequences.
- Purpose Limitation.** GaitLU-1M can be distributed only for non-commercial research purposes with a case-by-case access application.
- Ethical Requirements.** Applicants are required to sign this Ethical Requirements.

## Code of Ethics

Users of the GaitLU-1M dataset **MUST** obey the following rules.

- Any form of personal identity inferring is **PROHIBITED**.
- Any form of additional manual labeling is **PROHIBITED**.
- Any form of private distribution is **PROHIBITED**.
- Any form of model / data inversion attack is **PROHIBITED**.
- Any form of biased analysis about personal attributes like gender, race, nationality, and health is **PROHIBITED**.
- Any form of conducts potentially cause lawfulness, fairness, and transparency concerns is **PROHIBITED**.
- Try your best to be a responsible AI researcher.

## Indemnification

Researchers agree to indemnify, defend, and hold harmless Southern University of Science and Technology, its officers, employees, and agents from any and all losses, expenses, damages, demands, and claims arising from or related to the researcher's use of the GaitLU-1M dataset. Researchers shall be responsible for paying any damages, claims, judgments, or expenses resulting from their use of the dataset.

**Note:** Please fill in the form below in printed form, do not use handwriting

Information	Legal representative, e.g., your supervisor if you are a student	Point of contact, i.e., student or person severing as coordinator (if any)
Name		
Organization		
Position		
Email address (Institute domain is needed, e.g. "sustech.edu.cn". Free e-mail domain is not acceptable, e.g. gmail.com, qq.com)		
The <b>Homepage</b> of the legal representative, or research group, laboratory, department, etc., including basic information of legal representative.		

**With my signature**, I hereby request to obtain and use the GaitLU-1M dataset, and I state that I have the authority to accept the terms stated in this Requirements.

\_\_\_\_\_  
SIGNATURE OF LEGAL REPRESENTATIVE

\_\_\_\_\_  
SIGNATURE OF POINT OF CONTACT (if any)

\_\_\_\_\_  
DATE

Send a scanned PDF file by email. The update to date email address can be found at the GaitLU-1M homepage: <https://github.com/ChaoFan996/GaitSSB>