Steps for tagging the domain of packages within TensorFlow-based SSC

- Step 1. Domain tags exploration. To facilitate the tagging of package domains, our initial step involves identifying available tags. We refer to Gao et al. labeling of package domains in PyPI and explore additional labels. To accomplish this, we employ two labelers who each independently select and tag the domain of 100 packages within TensorFlow SSC without any pre-existing tags. Subsequently, we conduct a comprehensive analysis and discussion of the labeler's assessments, resulting in the identification of 26 distinct domain tags.
- Step 2. Implementing tagging. Following the establishment of domain tags, we label the packages within TensorFlow SSC, this task is also undertaken by the two initial labelers. To ensure accuracy, we randomly select a sample of 140 packages, representing approximately 10% of the 1,397 packages, and assign this data for independent tagging by the two labelers. After careful analysis, Krippendorff's α coefficient for the tagging reached a value of 0.89, meeting the reliability standard. Subsequently, we divide the remaining packages into two subsets, comprising 628 and 629 packages, and allocate them to each labeler.