# Annotation guide for selected MICCAI 2023 papers

## Annotation Guide: Demographic Information

**Purse of annotation:**

*To identify and record specific characteristics of the population included in study datasets. These characteristics typically include age, gender, race, ethnicity, and geographical location.*

When examining papers, your task is to identify and annotate demographic information provided in the 'Datasets' or 'Experiments' sections, or sections with a similar title. This information helps us understand the diversity and representativeness of the datasets used in the research.

**Categories:**

* Age
* Sex/gender
* Ethnicity
* Geographical Location

**Follow these steps:**

1. Open the file called “….”
2. Locate the section that describes the datasets.
3. Determine if the datasets include demographic details like age, gender, race, and geographical location.
4. If demographic data is present, annotate this as ‘1’. If not, annotate as ‘0’.

****

# Annotation Guide: Geographical Location details

**Purse of annotation:**

*To identify dataset sources, examining definitions of 'location', understanding that 'location' can extend beyond geographical areas to specific establishments, and recognizing implicit references to location.*

When annotating for geographical location information, our goal is to capture details about where each dataset originates.

1. **Source Annotation:**

Check if the dataset source is specified. Annotate any mention of where the dataset was gathered or sourced. This could be a country, city, institution, or any other place of origin provided in the documentation.

1. **Location Definition Annotation:**

Look at how each paper defines 'location.' Remember, 'location' isn't just about maps and coordinates; it can also refer to types of places. For example, if a paper talks about data from a 'hospital,' 'clinic,' or 'health center,' annotate this as the location, even though it doesn't give an exact address or city.

Always pay attention to the context — sometimes, the type of place (like a healthcare facility) tells us about the location, even when specific geographical details are not given.

**Categories:**

* Location
* Location as healthcare facility
  + This category encompasses various types of establishments dedicated to providing medical, surgical, and other forms of health care treatment and services. Name of healthcare facility is not relevant for this category.
* Name of healthcare facility
  + Insert name of healthcare facility only
* Location as large-scale geographical entity
  + A location as a large-scale geographical entity includes the largest territorial divisions, typically recognized on an international scale.
* Name of location as large-scale geographical entity
  + Insert name of country, countries and/or continents
* Location as subnational geographical entity
  + A location as a subnational geographical entity includes entities that are subdivisions of a country, ranging from larger areas like states and provinces to smaller localities like towns and villages.
* Name of location as subnational geographical entity
  + Insert name of city, cities, province, state, region, town, village, area and/or district

**Follow these steps:**

1. Open the file called “….”
2. Locate the section that describes the datasets.
3. Determine if the datasets include details of geographical location.





# Annotation Guide: Dataset details

**Purpose of annotation:**

*To concisely catalogue the data used in studies, including the number of datasets, their public or private status, the types of images, and any mentioned organs. These annotations help understand the datasets' scope, critical for assessing research replicability, potential bias, and results applicability. Essentially, it ensures transparency and promotes further research.*

Annotators will review research papers to collect specific information about the datasets used by the authors. This includes details about dataset quantity, access status (public or private), image types, and organs/body parts featured.

**Categories**

* Dataset Quantity
  + Count and annotate the number of datasets mentioned in the paper. Record as a whole integer. Annotate ‘0’ if no datasets are mentioned.
* Public Datasets
  + Identify and annotate the number of datasets described as public or publicly available. Use whole numbers and ‘0’ if none are public.
* Private Datasets:
  + Count and annotate the number of datasets described as private. Record this information as a whole integer, using ‘0’ if there are no private datasets.
* Image Types
  + Annotate the specific types of images described in the datasets (e.g., MRI, CT, X-ray). If not specified, annotate as "Unknown".
* Organs/Body Parts
  + Annotate the names of organs or body parts mentioned in relation to the datasets. If none are mentioned, annotate as "Unknown".
* Sex-specific cancer
  + Annotate ‘1’ if the cancer is sex specific. A sex-specific cancer is determined by male or female reproductive organs.
  + Cancer can be sex-specific in two ways:
    - There is a higher occurrence of the illness in one sex compared to the other, or
    - The illness is determined by/related to underlying biological factors or one sex only.
    - In this case we only annotate '1' if cancer is determined/related to one sex only: use the column "organ" as guideline.

**Follow these steps:**

1. Open the file “…..”
2. Locate the section in the paper where the datasets are described.
3. Determine if the datasets include information on the categories specified: quantity, public or private status, image types, and organs/body parts.
4. Use provided keywords and example phrases to assist in identifying relevant information.

