# MICCAI 2023 Article Annotation Guide

### Essential Tools and Software

* PDF Reader: Adobe Reader Pro or any capable PDF reader.
* Data Entry Software: Microsoft Excel.

## Resources Provided (2 files and 1 folder)

### File #1: annotation\_guide\_details.doc

This file contains detailed instructions on the annotation process.

### File #2: annotation\_guide\_and\_scheme.xlsx with 4 sheets

**Sheet 1: Annotation guide details**

* 1. Overview of all categories with annotating examples and keywords for searching after relevant sentences

**Sheet 2: All categories**

* 1. Overview of all categories in the annotation scheme

**Sheet 3: Annotation scheme**

* 1. A structured template with predefined categories for you to use when annotating.
  2. The follow categories provide article details such as:
     1. *Title*: Name of the article for reference.
     2. *Paper Number:* Unique identifier for the article.
     3. *Volume Number:* Volume in which the article is published.

**Sheet 4: Keywords organs & image types**

* 1. A list of keywords to assist in locating relevant sections for annotation.

### Article Folder: A directory containing 100 selected PDF articles from MICCAI 2023.

* Naming Convention: Files are titled in the format ‘paper\_<number>\_vol\_<number>’.
* Utilize the **Title** and/or **Paper Number** and **Volume Number** to find articles within the folder.

## 

## Annotation Steps

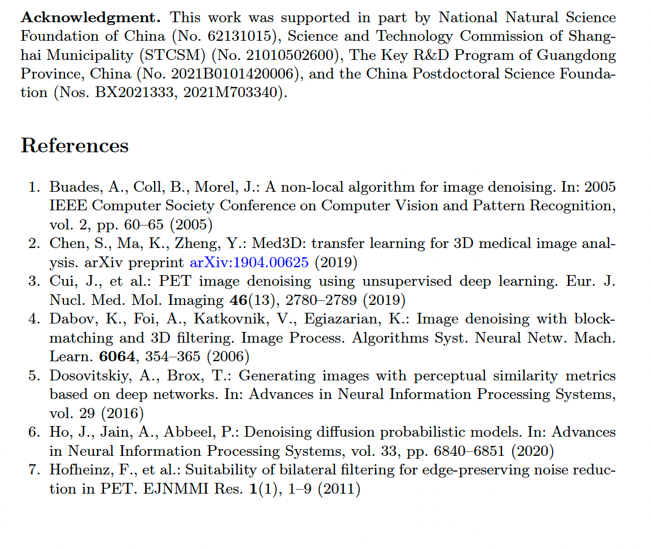
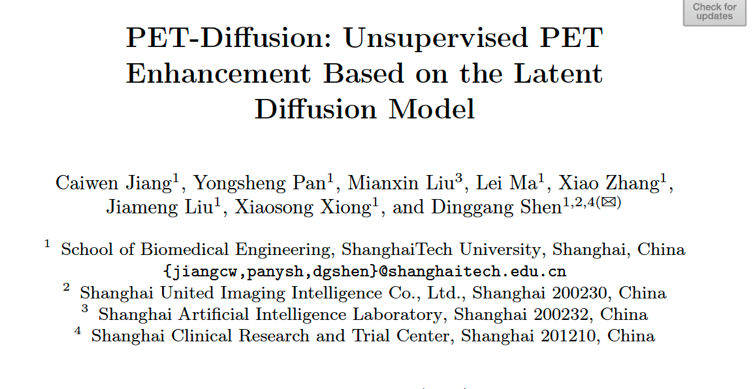
1. *Article Location:* Access the designated folder containing the MICCAI 2023 PDF articles.
2. *Article Identification*: Use the list of titles (under "**title**") and filenames (from "**paper name**" and "**vol**") provided in the Annotation Scheme to locate your assigned article(s) within the folder.
3. *Annotation Process:*
   1. Open the corresponding article.
   2. Begin the annotation process by adhering to the categories outlined in the Annotation Scheme ***(File #2: annotation\_guide\_and\_scheme.xlsx, Sheet 3: Annotation scheme).***
   3. Employ the provided Keywords List to efficiently identify pertinent sentences for each annotation category ***(File #2: annotation\_guide\_and\_scheme.xlsx, Sheet 4: Keywords organs & image types).***

# Annotation Guide - Article information



### Categories

**Does the article contain any of these keywords: cancer/tumor/tumour?**



Excluded sections:

1. Authors, departments, affiliations
2. Acknowledgements, References
3. Only text starting from Abstract and ending with the last line of Conclusion

**Notes**

Here you can add comments and sentences that “justifies” your annotations. For example:

* If you add ‘1’ for a category, please provide the sentence.
* If you add the name of a healthcare facility, please provide the sentence.

**If you find any categories hard to annotate, please add this as a comment. These findings are also important as they show whether the articles provide clear and consistent information or come across as being ambiguous.**

#### Examples:

* It was difficult to determine how many different datasets articles are using.
* Add these sentences for further analysis, or others to back up your annotations.

**Extra categories (not mandatory)**

* Information found outside article: information provided as Supplementary Material and/or by references.

# 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

* Study subject labelled as ‘patient/patients’ in dataset(s)?
* Age (of study subjects)
* Sex/gender (of study subjects)
* Ethnicity (of study subjects)
* Geographical Location (of study subjects)

#### **Follow these steps.**

1. Locate the section that describes the datasets by searching in the PDF for “dataset”, “datasets”, “data”, “data sets”.
2. Determine if the dataset(s) include demographic details like age, gender, race, and geographical location of study subjects.
   1. Search by using the provided keywords.
   2. If demographic data is present, annotate this as ‘1’. If not, annotate as ‘0’.

# Annotation Guide - Dataset information

## 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 dataset’s 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.
  + A dataset is public if the word “public”, “publicly”, “open” or “open access” is used.
* 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.
  + A dataset is private if the word “private”, “privately” or “in-house” is used.
* 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".
* Does the dataset contain sex-specific organ(s)?
  + Annotate ‘1’ if data contains sex specific organ(s). A sex-specific organ is determined by male or female reproductive organs.

#### **Follow these steps.**

1. Locate the section that describes the datasets by searching in the PDF for “dataset”, “datasets”, “data”, “data sets”.
2. Determine if the datasets include information on the categories specified: quantity, public or private status, image types, and organs/body parts.
   1. Search by using the provided keywords.
   2. If you find keywords outside the scope of the article, such as in the section for “References”, which shows that a referred dataset is open/public or private, add this as a comment in “notes”.
3. Use provided keywords and example phrases to assist in identifying relevant information.

# Annotation Guide - Geographical Location of Datasets

## 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 (of dataset/data collection)
* 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 other (center/department/laboratory/university/institution/online)
  + This category encompasses other types of establishments outside the definition of healthcare facility (types of establishments providing medical, surgical, and other forms of health care treatment and services). Name of other is not relevant for this category.
  + Other types of establishments could be a center, institution, department, laboratory, university and/or online database.
* Name of location as other (center/department/laboratory/university/institution/online)
  + Insert name of other
* 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. Locate the section that describes the datasets by searching in the PDF for “dataset”, “datasets”, “data”, “data sets”.
2. Determine if the datasets include details of geographical location.
   1. Search by using the provided keywords.