# Annotation Template Setup Guide

## Complete Instructions for Creating Your Annotation Workbook

**Version:** 1.0  
**Date:** October 2025  
**For:** Violent Event Annotation Project

## Table of Contents

1. Quick Start
2. Excel Template Structure
3. Google Sheets Setup (Alternative)
4. Data Validation Rules
5. Installation Instructions
6. Usage Workflow
7. Troubleshooting

## 1. Quick Start

### What You’ll Create

A multi-sheet Excel workbook with: - **Sheet 1:** Article Metadata - **Sheet 2:** Event Records (main annotation sheet) - **Sheet 3:** Entity Annotations (detailed) - **Sheet 4:** Relationships - **Sheet 5:** Reference Lists (taxonomy, vocabularies)

### Time Required

* Template setup: 30-45 minutes
* Python tools installation: 10 minutes

## 2. Excel Template Structure

### SHEET 1: Article Metadata

**Purpose:** Store information about each article

**Columns:**

| Column Name | Data Type | Width | Validation | Example |
| --- | --- | --- | --- | --- |
| Article\_ID | Text | 15 | Required, Format: ART\_XXX | ART\_001 |
| Article\_URL | Text | 50 | Required | https://… |
| Article\_Title | Text | 60 | Required | “Attack in Mali” |
| News\_Source | Text | 25 | Required | Reuters |
| Publication\_Date | Date | 15 | Required, Format: YYYY-MM-DD | 2024-03-20 |
| Annotator\_Name | Text | 25 | Required, Dropdown | John Doe |
| Annotation\_Date | Date | 15 | Auto-fill with TODAY() | 2024-10-20 |
| Article\_Text | Long Text | 100 | Required | [Full article text] |
| Notes | Text | 50 | Optional | Any notes |

**Excel Formula for Annotation\_Date:**

=TODAY()

**Data Validation for Annotator\_Name:** - Go to Data > Data Validation - Allow: List - Source: John Doe, Jane Smith, Ahmad Hassan, Mary Johnson - (Customize with your actual annotator names)

### SHEET 2: Event Records (MAIN SHEET)

**Purpose:** Main annotation sheet - one row per event

**Column Specifications:**

#### A. Identification Columns

| Column | Type | Width | Format | Validation | Formula/Note |
| --- | --- | --- | --- | --- | --- |
| Event\_ID | Text | 20 | ART\_XXX\_EVT\_YY | Required | Manual entry |
| Article\_ID | Text | 15 | ART\_XXX | Required, Dropdown from Sheet 1 |  |
| Event\_Trigger\_Text | Text | 30 |  | Required | Word like “attacked” |
| Sentence\_Numbers | Text | 15 |  | Optional | “3,4,5” |
| Event\_Description | Text | 80 |  | Required | Brief summary |

**Data Validation for Article\_ID:**

=ArticleMetadata!$A$2:$A$1000

#### B. Actor (Who) Columns

| Column | Type | Width | Validation | Example |
| --- | --- | --- | --- | --- |
| Actor\_Text | Text | 50 | Required | “Al-Shabaab militants” |
| Actor\_Normalized | Text | 40 | Required | “Al-Shabaab” |
| Actor\_Type | Dropdown | 25 | Required, List | Non-state armed group |
| Actor\_Confidence | Number | 15 | 0.0-1.0 | 0.85 |

**Data Validation for Actor\_Type:** - Source: Reference\_Lists!$A$2:$A$7 - Options: State forces, Non-state armed group, Criminal organization, Communal group, Unknown, Multiple

**Data Validation for Actor\_Confidence:**

Custom: =AND(A2>=0, A2<=1)  
Error Alert: "Must be between 0.0 and 1.0"

#### C. Victim (Whom) Columns

| Column | Type | Width | Validation | Example |
| --- | --- | --- | --- | --- |
| Victim\_Text | Text | 50 | Required | “at least 15 civilians” |
| Victim\_Normalized | Text | 40 | Required | “15 civilians” |
| Victim\_Type | Dropdown | 20 | Required, List | Civilian |
| Deaths | Number | 10 | ≥0 | 15 |
| Injuries | Number | 10 | ≥0 | 30 |
| Victim\_Confidence | Number | 15 | 0.0-1.0 | 0.9 |

**Data Validation for Victim\_Type:** - Source: Reference\_Lists!$B$2:$B$6 - Options: Civilian, Combatant, Mixed, Infrastructure, Unknown

**Data Validation for Deaths/Injuries:**

Custom: =A2>=0  
Error Alert: "Must be non-negative"

#### D. Location (Where) Columns

| Column | Type | Width | Validation | Example |
| --- | --- | --- | --- | --- |
| Location\_Text | Text | 50 | Required | “Bama market, Borno” |
| Location\_Normalized | Text | 50 | Required | “Bama, Borno State, Nigeria” |
| Location\_Specific | Text | 40 | Optional | “Bama market” |
| Location\_City | Text | 25 | Optional | “Bama” |
| Location\_State | Text | 30 | Optional | “Borno State” |
| Location\_Country | Text | 25 | Required | Nigeria |
| Location\_Coordinates | Text | 25 | Format: “Lat,Long” | “11.5204, 13.6891” |
| Location\_Confidence | Number | 15 | 0.0-1.0 | 0.85 |

#### E. Date/Time (When) Columns

| Column | Type | Width | Validation | Example |
| --- | --- | --- | --- | --- |
| Date\_Text | Text | 30 | Required | “yesterday” |
| Date\_Normalized | Date | 15 | Format: YYYY-MM-DD | 2024-03-19 |
| Time\_of\_Day | Dropdown | 20 | Optional, List | Morning |
| Date\_Confidence | Number | 15 | 0.0-1.0 | 0.9 |

**Data Validation for Time\_of\_Day:** - Options: Early morning, Morning, Afternoon, Evening, Night, Unknown

#### F. Weapon/Method (How) Columns

| Column | Type | Width | Validation | Example |
| --- | --- | --- | --- | --- |
| Weapon\_Text | Text | 50 | Required | “guns and explosives” |
| Weapon\_Normalized | Text | 40 | Required | “Firearms; Explosives” |
| Weapon\_Category | Dropdown | 20 | Required, List | Multiple |
| Tactic | Text | 25 | Optional | “Ambush” |

**Data Validation for Weapon\_Category:** - Source: Reference\_Lists!$C$2:$C$8 - Options: Firearms, Explosives, Edged weapons, Fire/Arson, Heavy weapons, Multiple, Unknown

#### G. Taxonomy Columns

| Column | Type | Width | Validation | Example |
| --- | --- | --- | --- | --- |
| Taxonomy\_L1 | Dropdown | 35 | Required, List | Political Violence |
| Taxonomy\_L2 | Dropdown | 35 | Required, Conditional | Terrorism |
| Taxonomy\_L3 | Dropdown | 35 | Required, Conditional | Bombing/Explosive Attack |
| Taxonomy\_L4 | Text | 35 | Optional | Suicide Bombing |
| Classification\_Confidence | Number | 20 | 0.0-1.0 | 0.85 |
| Alternative\_Classification | Text | 50 | Optional | Other valid options |

**Data Validation for Taxonomy\_L1:** - Source: Reference\_Lists!$D$2:$D$5 - Options: Political Violence, Criminal Violence, Communal Violence, State Violence Against Civilians

**IMPORTANT:** Taxonomy\_L2 should be conditional on L1. See “Conditional Dropdowns” section below.

#### H. Quality Control Columns

| Column | Type | Width | Format | Example |
| --- | --- | --- | --- | --- |
| Multi\_Label | Checkbox | 12 | TRUE/FALSE | FALSE |
| Flagged\_for\_Review | Checkbox | 18 | TRUE/FALSE | FALSE |
| Notes | Text | 80 | Optional | Any notes or questions |

### SHEET 3: Entity Annotations (Detailed)

**Purpose:** Detailed entity-level annotations for advanced analysis

| Column | Type | Width | Example |
| --- | --- | --- | --- |
| Entity\_ID | Text | 25 | ART\_001\_EVT\_01\_ENT\_01 |
| Event\_ID | Text | 20 | ART\_001\_EVT\_01 |
| Entity\_Text | Text | 60 | “Al-Shabaab militants” |
| Entity\_Type | Dropdown | 20 | Actor |
| Start\_Char | Number | 12 | 245 |
| End\_Char | Number | 12 | 265 |
| Sentence\_Number | Number | 15 | 3 |
| Normalized\_Form | Text | 40 | “Al-Shabaab” |
| Coreference\_Group\_ID | Text | 25 | ART\_001\_EVT\_01\_ACTOR\_1 |
| All\_Mentions | Text | 80 | “Al-Shabaab militants; The attackers; They” |

**Data Validation for Entity\_Type:** - Options: Actor, Victim, Location, Date, Weapon, Other

### SHEET 4: Relationships

**Purpose:** Document relationships between events

| Column | Type | Width | Example |
| --- | --- | --- | --- |
| Relationship\_ID | Text | 20 | REL\_001 |
| Event\_1\_ID | Text | 20 | ART\_001\_EVT\_01 |
| Relationship\_Type | Dropdown | 25 | Retaliation |
| Event\_2\_ID | Text | 20 | ART\_001\_EVT\_02 |
| Description | Text | 80 | “Event 2 is retaliation for Event 1” |
| Confidence | Number | 15 | 0.8 |

**Data Validation for Relationship\_Type:** - Options: Temporal\_Sequence, Causality, Retaliation, Coordination, Related\_Context

### SHEET 5: Reference Lists

**Purpose:** Store controlled vocabulary lists for dropdowns

#### Column A: Actor Types

State forces  
Non-state armed group  
Criminal organization  
Communal group  
Unknown  
Multiple

#### Column B: Victim Types

Civilian  
Combatant  
Mixed  
Infrastructure  
Unknown

#### Column C: Weapon Categories

Firearms  
Explosives  
Edged weapons  
Fire/Arson  
Heavy weapons  
Multiple  
Unknown

#### Column D: Taxonomy L1

Political Violence  
Criminal Violence  
Communal Violence  
State Violence Against Civilians

#### Columns E-H: Taxonomy L2 (by L1 category)

**Column E: Political Violence L2**

Rebellion/Armed Insurgency  
Terrorism  
Coup and Regime Change Violence  
Election Violence  
Political Repression

**Column F: Criminal Violence L2**

Organized Crime Violence  
Armed Robbery/Banditry  
Kidnapping for Ransom  
Criminal Gang Violence

**Column G: Communal Violence L2**

Ethnic/Tribal Conflict  
Religious Violence  
Resource-Based Conflict  
Pastoralist-Farmer Clashes

**Column H: State Violence L2**

Extrajudicial Killings  
State Repression of Protests  
Mass Atrocities by State Forces  
Forced Displacement by State  
Arbitrary Detention with Violence

## 3. Conditional Dropdowns (Advanced)

### Making Taxonomy\_L2 Conditional on L1

**Problem:** L2 options should change based on what L1 category is selected.

**Solution: Using INDIRECT function**

1. **Name your L2 ranges:**
   * Select Column E (Political Violence L2) → Name it: PoliticalViolence\_L2
   * Select Column F (Criminal Violence L2) → Name it: CriminalViolence\_L2
   * Select Column G (Communal Violence L2) → Name it: CommunalViolence\_L2
   * Select Column H (State Violence L2) → Name it: StateViolence\_L2
2. **Create helper column (can be hidden):**
   * Add column after Taxonomy\_L1
   * Name it: L2\_List\_Name
   * Formula:

* =SUBSTITUTE(SUBSTITUTE([Taxonomy\_L1\_Cell]," ","\_"),"Against","")&"\_L2"
  + This converts “Political Violence” → “PoliticalViolence\_L2”

1. **Set Data Validation for Taxonomy\_L2:**
   * Select Taxonomy\_L2 column
   * Data > Data Validation
   * Allow: List
   * Source: =INDIRECT([L2\_List\_Name\_Cell])

**Alternative (Simpler but manual):** - Don’t use conditional dropdowns - Use all L2 categories in one list - Rely on Python validator to catch mismatches - Annotators manually ensure correct pairing

## 4. Formatting Recommendations

### Color Coding

**Header Row:** - Background: Dark Blue (RGB: 68, 114, 196) - Font: White, Bold, Size 11

**Identification Section (Columns A-E):** - Light Gray background (RGB: 217, 217, 217)

**Actor Section:** - Light Blue background (RGB: 221, 235, 247)

**Victim Section:** - Light Red background (RGB: 252, 228, 214)

**Location Section:** - Light Green background (RGB: 226, 239, 218)

**Date Section:** - Light Yellow background (RGB: 255, 242, 204)

**Taxonomy Section:** - Light Purple background (RGB: 237, 230, 246)

**Quality Control:** - Light Orange background (RGB: 252, 245, 229)

### Cell Borders

* All cells: Thin borders
* Header row: Medium bottom border
* Every 5 rows: Slightly thicker border for readability

### Text Wrapping

* Enable text wrap for: Notes, Event\_Description, Actor\_Text, Victim\_Text, Location\_Text
* Adjust row height: Auto-fit or minimum 30 pixels

### Freeze Panes

* Freeze: Row 1 (header) and Columns A-C (IDs)
* View > Freeze Panes > Freeze Panes

### Column Widths

Set as specified in column tables above for optimal display.

## 5. Google Sheets Setup (Alternative)

If using Google Sheets instead of Excel:

### Advantages

* Cloud-based (multiple annotators can work simultaneously)
* Auto-save
* Easy sharing and collaboration
* Version history

### Setup Steps

1. **Create new Google Sheet**
   * Go to sheets.google.com
   * Click “Blank”
   * Rename: “Violent\_Event\_Annotations”
2. **Create sheets:**
   * Rename Sheet1 → “Article Metadata”
   * Add sheets: Event Records, Entity Annotations, Relationships, Reference Lists
3. **Set up columns:**
   * Same structure as Excel version
   * Copy column headers
4. **Data validation:**
   * Select cells
   * Data > Data validation
   * Criteria: List from a range
   * Example: =Reference\_Lists!A2:A7
5. **Conditional formatting (optional):**
   * Format > Conditional formatting
   * Add rules for color coding sections
6. **Sharing:**
   * Click “Share” button
   * Add annotators with “Editor” permissions
   * Enable “Notify people” to alert them

### Google Sheets Specific Features

**Add comments:** - Right-click cell > Comment - Tag annotator: Type @ and name - Use for questions/discussions

**Version history:** - File > Version history > See version history - Restore earlier versions if needed

**Apps Script (optional automation):** - Extensions > Apps Script - Can add custom functions for validation

## 6. Installation Instructions

### Step 1: Install Python (if not installed)

**Windows:** 1. Download Python 3.9+ from python.org 2. Run installer 3. ✅ Check “Add Python to PATH” 4. Click “Install Now”

**Mac:**

brew install python3

**Linux:**

sudo apt-get install python3 python3-pip

### Step 2: Install Required Packages

Open terminal/command prompt:

pip install pandas numpy scikit-learn matplotlib seaborn openpyxl

Or create requirements.txt:

pandas>=1.3.0  
numpy>=1.21.0  
scikit-learn>=0.24.0  
matplotlib>=3.4.0  
seaborn>=0.11.0  
openpyxl>=3.0.0

Then install:

pip install -r requirements.txt

### Step 3: Download Tools

1. Save the Python code artifact as: annotation\_tools.py
2. Place in your project folder
3. Make executable (Unix/Mac):

* chmod +x annotation\_tools.py

### Step 4: Test Installation

python annotation\_tools.py --help

Should display usage instructions.

## 7. Usage Workflow

### Daily Annotation Workflow

#### For Annotators:

1. **Open template** (Excel or Google Sheets)
2. **Select next article** to annotate
3. **Fill Article Metadata** sheet
4. **Annotate events** in Event Records sheet
5. **Fill entity details** (optional, if needed)
6. **Mark relationships** (if multiple events)
7. **Self-review** using checklist
8. **Save** with filename: annotations\_[YourName]\_[Date].xlsx
9. **Submit** to project lead

#### For Project Lead (Quality Control):

**Daily:** 1. **Collect** submitted annotation files 2. **Run validator:** bash python annotation\_tools.py validate --file annotations\_John\_20241020.xlsx --output validation\_report\_John\_20241020.txt 3. **Review** validation report 4. **Provide feedback** to annotator if errors found

**Weekly:** 1. **Check inter-annotator agreement:** bash python annotation\_tools.py iaa --file annotator1.xlsx --file2 annotator2.xlsx --output iaa\_report\_week4.txt 2. **Review disagreements:** python from annotation\_tools import InterAnnotatorAgreement iaa = InterAnnotatorAgreement() disagreements = iaa.identify\_disagreements('ann1.xlsx', 'ann2.xlsx', 'Taxonomy\_L1') disagreements.to\_excel('disagreements\_to\_adjudicate.xlsx') 3. **Hold adjudication meeting** to resolve disagreements

**Monthly:** 1. **Generate statistics:** bash python annotation\_tools.py stats --file all\_annotations\_merged.xlsx --output monthly\_stats\_report.txt 2. **Review progress** toward annotation goals 3. **Identify** problematic categories or patterns 4. **Refine** guidelines if needed

## 8. File Organization

### Recommended Folder Structure

violent\_event\_annotation/  
│  
├── templates/  
│ ├── annotation\_template.xlsx  
│ └── annotation\_template\_blank.xlsx  
│  
├── guidelines/  
│ ├── annotation\_guidelines.pdf  
│ └── taxonomy.pdf  
│  
├── annotations/  
│ ├── in\_progress/  
│ │ ├── annotator1/  
│ │ ├── annotator2/  
│ │ └── annotator3/  
│ │  
│ ├── completed/  
│ │ └── [submitted files]  
│ │  
│ └── validated/  
│ └── [quality-checked files]  
│  
├── articles/  
│ ├── to\_annotate/  
│ └── completed/  
│  
├── tools/  
│ ├── annotation\_tools.py  
│ └── requirements.txt  
│  
├── reports/  
│ ├── validation/  
│ ├── iaa/  
│ ├── statistics/  
│ └── charts/  
│  
└── adjudication/  
 ├── disagreements/  
 └── resolved/

## 9. Troubleshooting

### Problem: Data validation not working

**Solution:** - Check that Reference Lists sheet exists - Verify range names are correct - Ensure no typos in formulas

### Problem: Conditional dropdowns not working

**Solution:** - Verify named ranges are created - Check INDIRECT formula syntax - Test with simple list first, then add complexity

### Problem: Python tools give error

**Solution:** - Check Python version: python --version (should be 3.7+) - Reinstall packages: pip install --upgrade [package name] - Check file path is correct - Verify Excel file has correct sheet names

### Problem: File too large / slow

**Solution:** - Split into multiple files (one per week or 100 events) - Remove formatting from data rows (keep header only) - Use CSV instead of XLSX for large files - Close other programs

### Problem: Merge conflicts (Google Sheets)

**Solution:** - Assign non-overlapping articles to annotators - Use “Version history” to see changes - Communicate about who’s working on what - Consider locking cells: Data > Protect sheets and ranges

## 10. Quick Reference Card

### Validation Command

python annotation\_tools.py validate --file [FILE] --output [REPORT]

### IAA Command

python annotation\_tools.py iaa --file [FILE1] --file2 [FILE2] --output [REPORT]

### Statistics Command

python annotation\_tools.py stats --file [FILE] --output [REPORT]

### Essential Checks Before Submission

* All required fields filled
* Event\_IDs unique and properly formatted
* Dates normalized to YYYY-MM-DD
* Confidence scores between 0.0-1.0
* Taxonomy hierarchy consistent
* No obvious typos
* Notes explain low confidence cases

## 11. Next Steps

### For Annotators:

1. ✅ Read Annotation Guidelines document
2. ✅ Set up Excel template or request Google Sheets link
3. ✅ Complete training exercises
4. ✅ Take qualification test
5. ✅ Begin annotation with first 3 articles (supervised)
6. ✅ Get feedback and refine
7. ✅ Continue with production annotation

### For Project Lead:

1. ✅ Create master template
2. ✅ Distribute to annotators
3. ✅ Set up shared folder structure
4. ✅ Test validation tools
5. ✅ Schedule weekly check-ins
6. ✅ Monitor progress daily
7. ✅ Run quality checks weekly

## 12. Support & Questions

**For technical issues with tools:** - Email: [your email] - Response time: 24 hours

**For annotation questions:** - Check Guidelines document first - Email: [advisor email] - Weekly office hours: [time/day]

**For urgent issues:** - Contact: [emergency contact]

**Document Status:** Ready for use  
**Last Updated:** October 2025  
**Version:** 1.0

**Good luck with your annotation project!** 🎯