# Starting Annotation - Week 2 Launch Plan

## Complete Guide to Launching Your Annotation Team

**Version:** 1.0  
**Date:** October 2025  
**Status:** Ready to Execute

## Table of Contents

1. Status Check
2. Immediate Action Plan (7 Days)
3. Training Session Plans
4. Qualification Testing
5. Production Launch (Week 3)
6. Ongoing Workflow
7. Quality Control System
8. Team Management
9. Troubleshooting
10. Checklists

## 1. Status Check ✅

### What You’ve Completed:

* ✅ Taxonomy defined
* ✅ Annotation guidelines written
* ✅ Excel/Sheets template created
* ✅ Python validation tools ready
* ✅ Quality control system in place

### What’s Next:

**Get 3-4 annotators trained and working**

## 2. Immediate Action Plan (Next 7 Days)

### Day 1-2: Recruit Annotators

#### Who to Recruit:

**Ideal Candidates:** - Graduate students (your department or related fields) - Research assistants - Upper-level undergrads (computer science, linguistics, political science) - Detail-oriented, reliable, available 10-15 hours/week

**Red Flags:** - ❌ Can’t commit to 6-8 weeks - ❌ Poor attention to detail - ❌ Limited English proficiency - ❌ Not available for weekly meetings

#### How to Recruit:

**Option 1: Department Posting**

RESEARCH ASSISTANT POSITION  
Violent Event Annotation for Master's Thesis  
  
Duration: 8 weeks (Oct-Dec 2025)  
Hours: 10-15 hours/week, flexible schedule  
Rate: [Your rate] per hour OR academic credit  
Location: Remote work possible  
  
Requirements:  
- Excellent English reading comprehension  
- Strong attention to detail  
- Reliable and punctual  
- Available for weekly team meetings  
- Interest in conflict studies/NLP (helpful but not required)  
  
Tasks:  
- Annotate news articles about violent events  
- Extract structured information (who, what, where, when)  
- Classify events using provided taxonomy  
- Attend weekly team meetings  
  
Training: Fully provided (6-8 hours paid training)  
Start Date: [Next week]  
  
Contact: [Your email]  
Subject: Annotation Assistant Application  
Include: CV, availability, why interested

**Option 2: Direct Recruitment** - Ask professors for recommendations - Post in graduate student groups - Reach out to students who’ve worked on research projects - Contact students who’ve taken relevant courses

**Option 3: Your Advisor’s Network** - Ask advisor to recommend students - Often knows reliable workers - Can provide academic credit if needed

#### Interview Questions (15 min each):

1. “Are you available 10-15 hours/week for the next 8 weeks?”
2. “Can you attend weekly team meetings [proposed time]?”
3. “Tell me about a time you did detailed, repetitive work. How did you stay accurate?”
4. “This requires reading news about violence. Are you comfortable with that?”
5. “Rate your English reading comprehension 1-10”
6. “Have you done research work before? Tell me about it.”

**Select:** 3-4 best candidates (ideally 4 for redundancy)

### Day 3: Prepare Training Materials

#### Create Training Package:

**Folder 1: Essential Reading (send 48 hours before training)**

📁 Annotation\_Training/  
├─ 📄 Welcome\_Letter.pdf  
├─ 📄 Annotation\_Guidelines.pdf (the full document)  
├─ 📄 Taxonomy.pdf (sections 1-4, 7)  
├─ 📄 Quick\_Reference\_Card.pdf (1 page - create this)  
└─ 📄 Schedule.pdf (training session times)

**Welcome Letter Template:**

Dear [Annotator Name],  
  
Welcome to the Violent Event Annotation Project!   
  
You'll be helping extract structured information about violent events from   
news articles to train a machine learning system for the African Union's   
early warning system.  
  
BEFORE TRAINING (Monday):  
Please read these documents (4-6 hours):  
1. Annotation Guidelines (pages 1-25 essential)  
2. Taxonomy document (sections 1-4 and 7)  
3. Quick Reference Card (keep handy during annotation)  
  
TRAINING SCHEDULE:  
Session 1: [Date/Time] - 3 hours - Introduction & Practice  
Session 2: [Date/Time] - 2 hours - Qualification Test  
Location: [Room or Zoom link]  
  
WHAT TO BRING:  
- Laptop (with Excel or access to Google Sheets)  
- Questions from reading materials  
- Notebook for notes  
  
COMPENSATION:  
- Training: [X] hours at [rate]  
- Annotation: [X] hours/week at [rate] for 6-8 weeks  
- Payment: [Bi-weekly / Monthly]  
  
Questions? Email: [your email] or Slack: [channel]  
  
Looking forward to working with you!  
  
[Your name]  
Project Lead

**Quick Reference Card (1 page):**

# VIOLENT EVENT ANNOTATION - QUICK REFERENCE  
  
## 5W1H TO EXTRACT:  
✓ WHO (Actor): Perpetrator - mark complete noun phrase  
✓ WHAT (Event): Type of violence - use taxonomy  
✓ WHOM (Victim): Target - mark casualties if mentioned   
✓ WHERE (Location): Most specific place + coordinates  
✓ WHEN (Date): Normalize to YYYY-MM-DD format  
✓ HOW (Method): Weapons/tactics used  
  
## TAXONOMY L1 (4 categories):  
□ Political Violence - Political objectives, state/opposition  
□ Criminal Violence - Economic gain, criminal actors  
□ Communal Violence - Identity groups, resources  
□ State Violence Against Civilians - State forces vs civilians  
  
## CONFIDENCE SCORING:  
0.9-1.0 = Very high (clear, unambiguous)  
0.7-0.8 = High (minor uncertainty)  
0.5-0.6 = Medium (some ambiguity)  
0.3-0.4 = Low (significant ambiguity)  
0.0-0.2 = Very low (FLAG FOR REVIEW)  
  
## RED FLAGS (always flag):  
⚠ Can't determine if 1 or multiple events  
⚠ Confidence < 0.5 on any field  
⚠ Contradictory information in article  
⚠ Can't confidently classify  
  
## GOLDEN RULES:  
1. Annotate ONLY what the article says  
2. NO external knowledge  
3. Mark "Unknown" if not stated  
4. Resolve coreference (He/They/The militants = ?)  
5. When unsure: document both options + note why  
  
## DAILY WORKFLOW:  
1. Read article fully once  
2. Identify # of events   
3. For each event: Extract entities → 5W1H → Classify  
4. Self-review checklist  
5. Submit  
  
Questions? Check Guidelines Section X or ask in Slack!

**Folder 2: Practice Materials**

📁 Practice\_Articles/  
├─ 📰 Article\_001\_Simple.pdf (single event, clear)  
├─ 📰 Article\_002\_Medium.pdf (single event, some ambiguity)  
├─ 📰 Article\_003\_Complex.pdf (multiple events)  
├─ 📊 Answer\_Keys/ (for comparison after practice)  
└─ 📝 Blank\_Template.xlsx (for practice)

### Day 4: Training Session 1 (3 hours)

#### Session 1 Agenda:

**Hour 1: Introduction & Overview (Lecture)**

**0:00-0:15: Welcome & Project Overview** - Introduce yourself - Project goals and importance - Their role and contribution - Timeline and expectations - Q&A logistics (Slack, email, meetings)

**0:15-0:45: Taxonomy Deep Dive** - Walk through Level 1 categories with examples - Show how events fit into hierarchy - Discuss edge cases - Answer questions

**0:45-1:00: Break**

**Hour 2: Annotation Process Demo**

**1:00-1:30: Live Annotation Demo** - Open Article\_001\_Simple on screen - Walk through annotation step-by-step: - Read article aloud - Identify event(s) - Mark entities - Extract 5W1H - Classify using taxonomy - Fill template fields - Self-review - Explain every decision out loud - Show how you handle uncertainty

**1:30-2:00: Common Mistakes & How to Avoid** - Show examples of common errors - Demonstrate corrections - Discuss why errors happen - Prevention strategies

**Hour 3: Hands-On Practice**

**2:00-2:30: Guided Practice** - Everyone annotates Article\_002\_Medium - You stay on Zoom/in room for questions - No talking/comparing until done - Submit when finished

**2:30-2:50: Group Review** - Share your annotation on screen - Compare with theirs - Discuss differences - Identify learning points - Answer questions

**2:50-3:00: Homework & Next Session** - Homework: Annotate Article\_003\_Complex alone - Bring completed annotation to Session 2 - Read Guidelines Section 10 (Common Mistakes) - Prepare questions for next session

### Day 5-6: Between Training Sessions

**Your Tasks:** 1. Review everyone’s Article\_002 annotations 2. Note common mistakes for each person 3. Prepare individual feedback 4. Create answer key for Article\_003 5. Prepare qualification test articles (2 articles)

**Annotators’ Tasks:** - Complete Article\_003 (homework) - Re-read annotation guidelines - Prepare questions

### Day 7: Training Session 2 (2-3 hours)

#### Session 2 Agenda:

**Hour 1: Review & Feedback**

**0:00-0:20: Homework Review** - Discuss Article\_003 annotations - Show answer key - Identify common issues - General feedback

**0:20-0:40: Individual Feedback** - Meet with each annotator 5 minutes - Point out their specific strengths/weaknesses - Answer their questions - Clarify confusion

**0:40-1:00: Advanced Topics** - Handling very ambiguous cases - When to flag for adjudication - Coreference resolution tricks - Speed vs accuracy balance

**Hour 2: Qualification Test**

**1:00-1:05: Test Instructions** - You’ll annotate 2 articles independently - No communication with others - Use all your resources (guidelines, taxonomy) - Take your time (quality matters) - Aim for 30-40 minutes per article - Submit when done

**1:05-2:00: Qualification Test** - Everyone annotates same 2 articles - You monitor but don’t help - They work silently - Submit via email/shared folder when done

**2:00-2:15: Debrief** - How did it feel? - What was hardest? - Any remaining questions?

**2:15-2:30: Next Steps** - Results in 24-48 hours - Qualification threshold: 75% agreement with your annotations - If qualified: Start production next week - If not: Additional practice + re-test - Production workflow explained - Questions?

### Day 8-9: Evaluate Qualification Tests

#### Your Tasks:

**1. Annotate the 2 test articles yourself** (gold standard) - Your annotations = ground truth - Be extra careful and consistent

**2. Compare each annotator’s work to yours:**

Run validation:

# For each annotator  
python annotation\_tools.py validate --file annotator1\_test.xlsx  
  
# Calculate agreement with your gold standard  
python annotation\_tools.py iaa --file gold\_standard.xlsx --file2 annotator1\_test.xlsx

**3. Calculate scores:**

For each annotator, score: - Event identification (did they find all events?) - Entity extraction accuracy (actors, victims, locations correct?) - Taxonomy classification (L1, L2, L3 agreement) - Overall agreement (Cohen’s Kappa)

**Passing Criteria:** - ✅ Event count: 100% agreement - ✅ Taxonomy L1: Kappa ≥ 0.75 - ✅ Taxonomy L2: Kappa ≥ 0.70 - ✅ Major entities: ≥ 80% agreement - ✅ Overall: No critical errors

**4. Provide feedback:**

**If PASS (expected for 3-4 out of 4):**

Subject: Qualification Test Results - PASSED ✓  
  
Hi [Name],  
  
Congratulations! You passed the qualification test with strong performance:  
  
- Event identification: 100%  
- Taxonomy L1: Kappa = 0.82 (Excellent)  
- Taxonomy L2: Kappa = 0.74 (Good)  
- Entity extraction: 85% agreement  
  
Strengths:  
- [Specific positives]  
  
Watch out for:  
- [Minor issues to improve]  
  
NEXT STEPS:  
- Start production annotation Monday  
- You'll receive 25 articles this week  
- Daily submissions expected  
- Weekly team meeting: [Time]  
  
Welcome to the team!  
[Your name]

**If FAIL (maybe 1 out of 4):**

Subject: Qualification Test Results - Additional Training Needed  
  
Hi [Name],  
  
Thank you for completing the qualification test. You showed good effort,   
but we need to work on a few areas before production annotation:  
  
Results:  
- Event identification: [%]  
- Taxonomy classification: Kappa = [X]  
- Main issues: [Specific problems]  
  
NEXT STEPS:  
- Schedule 1-hour one-on-one session  
- Practice 3 more articles together  
- Re-test next week  
  
This is normal - annotation takes practice! Let's schedule a time to work   
on this together.  
  
[Your name]

## 3. Week 3: Start Production Annotation 🚀

### First Week Production Schedule:

#### Monday: Launch

**Your Tasks (Morning):**

**1. Prepare article batches:** - Create folders: annotator\_1, annotator\_2, annotator\_3, annotator\_4 - Assign 25 articles each (non-overlapping) - Plus 5 overlap articles (all 4 annotators do same ones) - Label clearly: “ART\_001 to ART\_025” and “OVERLAP\_001 to OVERLAP\_005”

**2. Send launch email:**

Subject: Production Annotation Starts Today! 🚀  
  
Team,  
  
You're officially qualified and ready to start! Here's your Week 1 assignment:  
  
ARTICLES FOR THIS WEEK:  
- Your folder: [link to their folder]  
- Main batch: 25 articles (ART\_XXX series)  
- Overlap batch: 5 articles (OVERLAP\_XXX) - everyone does these  
- Goal: Complete all 30 by Friday 5 PM  
  
DAILY WORKFLOW:  
1. Annotate 5-6 articles per day (1-2 hours)  
2. Self-review before submitting  
3. Save as: annotations\_[YourName]\_[Date].xlsx  
4. Upload to: [submission folder]  
5. I'll validate and provide feedback next morning  
  
OVERLAP ARTICLES:  
- These are for quality control  
- Everyone annotates the same ones  
- I'll calculate inter-annotator agreement  
- Helps us improve consistency  
  
QUALITY > SPEED:  
- Take your time the first week  
- Aim for accuracy, speed comes with practice  
- Flag anything uncertain  
- Ask questions in Slack anytime!  
  
WEEKLY MEETING:  
- Friday 2 PM: Team check-in (30 min)  
- Discuss challenges, questions, improvements  
  
You've got this! Remember: we're building something that matters.  
  
[Your name]

**Annotators’ Tasks (All Week):** - Annotate 5-6 articles/day - Submit daily - Attend Friday meeting

**Your Tasks (Daily, 15-30 min):**

**Each morning:**

# Validate yesterday's submissions  
for file in submissions/\*.xlsx; do  
 python annotation\_tools.py validate --file "$file" --output "reports/$(basename $file .xlsx)\_report.txt"  
done  
  
# Review reports  
# Email feedback to each annotator if issues found

**Each evening:** - Check Slack for questions - Respond within a few hours - Be encouraging!

#### Friday: Week 1 Review

**Before Meeting (1 hour):**

**1. Calculate IAA on overlap articles:**

# Compare all pairs  
python annotation\_tools.py iaa --file ann1\_overlap.xlsx --file2 ann2\_overlap.xlsx  
python annotation\_tools.py iaa --file ann1\_overlap.xlsx --file2 ann3\_overlap.xlsx  
# ... all combinations

**2. Generate week 1 statistics:**

# Merge all completed annotations  
python annotation\_tools.py stats --file week1\_all\_annotations.xlsx --output week1\_report.txt

**3. Prepare meeting agenda:** - Statistics overview - IAA results - Common issues - Questions from team - Week 2 assignment

**Meeting Agenda (30 min):**

**0:00-0:10: Week 1 Wins! 🎉** - “We completed [X] articles and extracted [Y] events!” - Celebrate milestones - Recognize good work

**0:10-0:20: Learning & Improvements** - IAA results: “We’re at Kappa 0.73, target is 0.75” - Common mistakes this week - How to improve - Clarify confusing taxonomy cases

**0:20-0:25: Questions & Discussion** - Open floor - Team helps each other - Share tips and tricks

**0:25-0:30: Week 2 Preview** - Next batch assignments (30 articles each) - New goal: increase speed slightly - Reminder of daily workflow - Motivational close

## 4. Ongoing Workflow (Weeks 3-8)

### Daily Routine:

**Annotators (1.5-2 hours/day):** - Annotate 5-6 articles - Self-review - Submit by end of day

**You (15-30 min/day):** - Validate submissions - Provide quick feedback - Answer Slack questions - Track progress

### Weekly Routine:

**Monday:** - Assign new batch (30-35 articles each) - Include 5 overlap articles - Send motivational update

**Friday:** - Calculate weekly IAA - Generate statistics - Team meeting (30 min) - Celebrate progress

### Monthly Routine:

**End of Month:** - Comprehensive statistics report - Quality trends analysis - Adjust training if needed - Report to advisor - Team celebration/appreciation

## 5. Quality Control System

### Red Flag System:

**🟢 GREEN (all good):** - Validation passes - Confidence scores reasonable - Few flagged events → Action: None, keep going!

**🟡 YELLOW (watch closely):** - Validation warnings - Low confidence scores (avg < 0.6) - Many flagged events (>20%) - IAA slightly below target (Kappa 0.60-0.70) → Action: One-on-one check-in, targeted feedback

**🔴 RED (intervention needed):** - Validation failures - IAA well below target (Kappa < 0.60) - Repeated same mistakes - Quality declining → Action: Stop, re-train, don’t continue until fixed

### Intervention Protocol:

**If annotator hits RED:**

1. **Immediate:** Stop their work, don’t assign new articles
2. **Diagnose:** Review last 10 annotations carefully, identify pattern
3. **Meet:** 1-hour one-on-one retraining session
4. **Practice:** 5 articles together, giving real-time feedback
5. **Test:** 2 new test articles
6. **Decide:** Pass → resume work; Fail → consider replacement

## 6. Team Management

### Communication:

**Slack Workspace (recommended):**

Channels:  
#general - Team announcements  
#questions - Quick questions  
#wins - Celebrate milestones  
#random - Team bonding  
  
DMs:  
- Individual feedback  
- Sensitive issues

**Or Email + WhatsApp Group:** - Email: Formal communication - WhatsApp: Quick questions, updates

### Progress Tracking:

**Simple Spreadsheet:**

| Annotator | Week 1 | Week 2 | Week 3 | … | Total | Status |
| --- | --- | --- | --- | --- | --- | --- |
| Ann 1 | 30 | 35 | 35 | … | 250 | ✓ |
| Ann 2 | 28 | 35 | 32 | … | 243 | ✓ |
| Ann 3 | 30 | 30 | 35 | … | 237 | ⚠ |
| Ann 4 | 25 | 28 | 30 | … | 215 | ⚠ |
| ———– | ——– | ——– | ——– | —– | ——- | ——– |
| TOTAL | 113 | 128 | 132 | … | 945 |  |
| TARGET | 120 | 240 | 360 | … | 2500 |  |

## 7. Motivation & Retention

### Keep Morale High:

**Weekly Wins Email Template:**

Subject: Week 3 Wins! 🎉  
  
Team,  
  
Another great week! Here's what we accomplished:  
  
📊 THIS WEEK:  
- 132 articles annotated  
- 198 events extracted  
- IAA improved to 0.76 (up from 0.73!)  
- 945 total events so far (38% of goal!)  
  
🌟 SHOUTOUTS:  
- [Ann 1]: Perfect validation rate this week!  
- [Ann 2]: Best IAA scores, great consistency!  
- [Ann 3]: Excellent question in Friday meeting!  
- [Ann 4]: 35 articles - productivity champion!  
  
📈 PROGRESS:  
- Week 1: 113 articles  
- Week 2: 128 articles   
- Week 3: 132 articles  
- We're gaining speed! 🚀  
  
🎯 NEXT WEEK:  
- Goal: 140 articles (35 each)  
- We'll hit 1,000 events milestone!   
- Pizza party when we hit 1,000! 🍕  
  
Keep up the amazing work!  
[Your name]

**Milestone Celebrations:** - 500 events: Coffee/lunch together - 1,000 events: Pizza party - 2,000 events: Team dinner - 2,500 events: Completion celebration + certificates

## 8. Troubleshooting

### Problem: Annotator Drops Out

**Backup Plan:** - Keep list of backup candidates - Have 4 annotators (can continue with 3) - Recruit replacement immediately - Fast-track training (1 day intensive)

### Problem: IAA Too Low

**Solutions:** 1. Additional training session 2. Adjudication meeting (resolve disagreements together) 3. Refine guidelines based on common disagreements 4. More frequent check-ins

### Problem: Too Slow

**Solutions:** 1. Extend timeline (acceptable) 2. Add one more annotator 3. Work with advisor to reduce target (2,000 instead of 2,500) 4. You help annotate some articles

### Problem: Quality Issues

**Solutions:** - More frequent validation (twice daily) - Lower daily quota (quality over speed) - Pair annotators (work together on hard articles) - More training on problem areas

## 9. What If Things Go Wrong?

### Common Issues & Solutions:

**Issue: Low motivation after 3-4 weeks** - Solution: Change routine, add variety, celebrate small wins, remind of purpose

**Issue: Disagreements between annotators** - Solution: Adjudication meeting, clarify guidelines, update taxonomy examples

**Issue: Technical problems (Excel crashes, file corruption)** - Solution: Regular backups, Google Sheets alternative, technical support

**Issue: Personal emergencies** - Solution: Flexibility, backup annotators, redistribute workload

**Issue: You’re overwhelmed** - Solution: Automate more (scripts), ask advisor for help, reduce other commitments

## 10. Checklists

### Pre-Launch Checklist

Before Monday morning, confirm:

* 3-4 annotators recruited
* Training completed (all passed qualification)
* Articles prepared (150-200 for first 2 weeks)
* Folders organized
* Templates distributed
* Slack/communication set up
* First batch assigned
* Weekly meeting scheduled
* Validation routine ready
* You’re excited! 😊

### Daily Checklist (For You)

**Morning (15 min):** - [ ] Check submissions folder - [ ] Run validation on new files - [ ] Review validation reports - [ ] Send feedback emails if needed - [ ] Update progress tracker

**Evening (10 min):** - [ ] Check Slack/email for questions - [ ] Respond to questions - [ ] Plan tomorrow’s tasks - [ ] Mental check: is everyone on track?

### Weekly Checklist (For You)

**Monday:** - [ ] Prepare new article batch - [ ] Assign to annotators - [ ] Send motivational email - [ ] Update progress dashboard

**Friday Before Meeting:** - [ ] Calculate IAA on overlap articles - [ ] Generate weekly statistics - [ ] Prepare meeting agenda - [ ] Review individual progress

**Friday After Meeting:** - [ ] Send meeting notes - [ ] Update advisor on progress - [ ] Plan next week’s batches - [ ] Celebrate wins!

### Monthly Checklist

**End of Each Month:** - [ ] Comprehensive statistics report - [ ] Quality trend analysis - [ ] Financial tracking (payments) - [ ] Advisor meeting/report - [ ] Team appreciation event - [ ] Reflect and adjust process

## 11. Success Metrics

### Target Goals (8 Weeks)

| Metric | Week 1 | Week 2 | Week 4 | Week 8 (End) |
| --- | --- | --- | --- | --- |
| **Articles Annotated** | 50 | 150 | 500 | 1,500-2,000 |
| **Events Extracted** | 75 | 225 | 750 | 2,500-3,500 |
| **Annotators Trained** | 3-4 | 3-4 | 3-4 | 3-4 |
| **IAA (L1)** | >0.70 | >0.75 | >0.75 | >0.75 |
| **IAA (L2)** | >0.65 | >0.70 | >0.70 | >0.70 |
| **Error Rate** | <20% | <15% | <10% | <5% |

### Quality Indicators

**✅ Good signs:** - Validation passes most files (>80%) - IAA scores meet targets - Confidence scores reasonable (0.7-0.9 average) - <10% events flagged for review - Annotators getting faster (15-20 min per article by Week 4)

**⚠️ Warning signs:** - Validation failures >30% - IAA below targets consistently - Many events flagged (>20%) - Annotators still taking 45+ min per article after Week 3 - High disagreement on same categories repeatedly

**Action:** If warning signs appear, hold emergency training session and refine guidelines.

## 12. File Organization

### Recommended Structure

violent\_event\_annotation/  
│  
├── 00\_setup/  
│ ├── annotation\_tools.py  
│ ├── requirements.txt  
│ ├── template\_master.xlsx  
│ └── setup\_guide.md  
│  
├── 01\_guidelines/  
│ ├── annotation\_guidelines.pdf  
│ ├── taxonomy.pdf  
│ └── quick\_reference.pdf  
│  
├── 02\_articles/  
│ ├── batch\_1/ (articles 001-200)  
│ ├── batch\_2/ (articles 201-400)  
│ └── ...  
│  
├── 03\_annotations/  
│ ├── annotator\_1/  
│ │ ├── in\_progress/  
│ │ ├── submitted/  
│ │ └── validated/  
│ ├── annotator\_2/  
│ ├── annotator\_3/  
│ └── annotator\_4/  
│  
├── 04\_quality\_control/  
│ ├── validation\_reports/  
│ ├── iaa\_reports/  
│ └── disagreements\_to\_adjudicate/  
│  
├── 05\_statistics/  
│ ├── weekly\_reports/  
│ ├── monthly\_reports/  
│ └── charts/  
│  
└── 06\_final\_dataset/  
 └── (merged, validated, ready for ML)

## 13. Quick Command Reference

### Daily Commands

# Validate submissions  
python annotation\_tools.py validate --file annotator1\_2024-10-21.xlsx --output report.txt  
  
# Quick check all files  
for file in submissions/\*.xlsx; do  
 python annotation\_tools.py validate --file "$file"  
done

### Weekly Commands

# Calculate IAA  
python annotation\_tools.py iaa --file ann1.xlsx --file2 ann2.xlsx --output iaa\_report.txt  
  
# Generate statistics  
python annotation\_tools.py stats --file week1\_merged.xlsx --output week1\_stats.txt  
  
# Generate charts (auto-created with stats)  
# Charts saved to ./annotation\_charts/

### Monthly Commands

# Comprehensive statistics with visualizations  
python annotation\_tools.py stats --file month1\_all.xlsx --output monthly\_report.txt  
  
# Merge all weekly files  
# (Use pandas or Excel to combine files)

## 14. Timeline Overview

### Week-by-Week Breakdown

**Week 1 (Current):** - Day 1-2: Recruit annotators - Day 3: Prepare training materials - Day 4: Training Session 1 - Day 5-6: Homework & prep - Day 7: Training Session 2 & Qualification - Day 8-9: Evaluate & provide feedback

**Week 2:** - Monday: Production launch (25-30 articles/person) - Daily: Annotation + validation - Friday: First team meeting, Week 1 review

**Weeks 3-4:** - Ramp up: 30-35 articles/person/week - Focus on speed while maintaining quality - Should have 500-750 events by end of Week 4

**Weeks 5-6:** - Peak production: 35-40 articles/person/week - Quality should be stable (IAA >0.75) - Should have 1,200-1,500 events by end of Week 6

**Weeks 7-8:** - Final push: Maintain pace - Complete any remaining articles - Final validation and cleanup - **Target: 2,500-3,500 events by end of Week 8**

## 15. Critical Success Factors

### 1. Start Small, Scale Fast

* Week 1: 2 annotators, 50 articles (learn and refine)
* Week 2: Add 1-2 more annotators if Week 1 went well
* Week 3+: Full production with all 4 annotators

### 2. Daily Validation

* NEVER let errors accumulate
* Validate same day or next morning
* Quick feedback loop keeps quality high

### 3. Weekly IAA Checks

* Schedule Friday afternoon for IAA analysis
* Discuss Monday morning
* Adjust guidelines as needed

### 4. Regular Communication

* Daily Slack/email for questions
* Weekly 30-min meeting (all annotators)
* Monthly check-in with advisor

### 5. Celebrate Milestones

* 500 events: Small celebration
* 1,000 events: Team lunch
* 2,500 events: Annotation phase complete! 🎉

## 16. Common Pitfalls to Avoid

### ❌ Don’t:

* Skip validation to save time (costs more later)
* Let annotators work in isolation too long
* Ignore low IAA scores
* Change guidelines mid-stream without retraining
* Forget to backup data regularly

### ✅ Do:

* Validate everything
* Communicate constantly
* Address issues immediately
* Keep guidelines stable (document changes)
* Backup to cloud daily (Google Drive, Dropbox)

## 17. Support & Resources

### For Technical Issues (Python, Excel)

* First: Check troubleshooting section in setup guide
* Email: [your email]
* Response: 24 hours

### For Annotation Questions

* First: Check annotation guidelines
* Check taxonomy document
* Email: [advisor email]
* Weekly office hours available

### For Urgent Issues

* Dataset corrupted
* Major disagreement between annotators
* Technical failure preventing work
* Contact: [emergency contact]

## 18. Motivation & Mindset

### Remember Why This Matters

This isn’t just an academic exercise. Your work will: - Help the African Union prevent conflicts - Save lives through early warning - Advance research in NLP and conflict analysis - Create valuable open data for researchers - Contribute to peace and security in Africa

### For Difficult Days

**When you’re tired:** - Remember your progress (look at charts!) - Take a break, come back fresh - Ask for help from the team - Celebrate small wins

**When quality dips:** - It’s normal - annotation is hard - Address it early - Retrain if needed - Quality > quantity always

**When behind schedule:** - Adjust timeline if needed (talk to advisor) - Add resources (annotators) - Reduce scope slightly - Stay calm, stay focused

## 19. Final Notes

### You’re Ready! 🚀

You have everything you need: - ✅ Infrastructure built - ✅ Guidelines written - ✅ Tools ready - ✅ Training plan - ✅ Quality control system - ✅ Team management strategy

### This Week’s Action Items:

**Day 1-2:** Recruit 3-4 annotators  
**Day 3:** Prepare training materials  
**Day 4:** Training Session 1  
**Day 7:** Training Session 2 & Qualification  
**Day 8-9:** Evaluate & provide feedback  
**Week 2 Monday:** LAUNCH PRODUCTION! 🎉

### Expected Outcome (8 Weeks Later):

* ✅ 2,500-3,500 annotated events
* ✅ High-quality training dataset
* ✅ Validation metrics met
* ✅ Ready for machine learning phase
* ✅ Thesis milestone achieved!

## 20. Next Phase Preview

**After Annotation (Week 9+):**

We’ll build the NLP pipeline and ML models: 1. Text preprocessing 2. Stanford CoreNLP integration 3. Feature extraction 4. Event detection models 5. Entity extraction models 6. Hierarchical classification 7. Knowledge base design 8. Q&A system development

**I’ll help you build all of this!** But first, let’s get this training data collected.

**Good luck with your annotation launch!** 🚀

You’ve got this! Remember: Every great ML system starts with great data. Take the time to do this right, and everything else will follow.

**Document Status:** Ready to Execute  
**Last Updated:** October 2025  
**Version:** 1.0

**Questions? I’m here to help every step of the way!** 💪