✅ Annotation Strategy Checklist

🛠️ Project Setup

- [ ] Define annotation objectives (e.g., classification, NER, bounding box)
- [ ] Select and configure annotation tool
- [ ] Develop initial ontology / label schema
- [ ] Create comprehensive annotation guidelines with examples
- [ ] Prepare golden dataset for QA and training
- [ ] Set up versioning system for datasets

👩‍🏫 Train Annotators

- [ ] Onboard annotators with kickoff session + tool walkthrough
- [ ] Distribute documentation on edge cases and escalation rules
- [ ] Assign hands-on practice tasks using golden dataset
- [ ] Provide automated and manual feedback on practice batches
- [ ] Pair new annotators with experienced mentors
- [ ] Confirm annotators meet quality threshold before production
- [ ] Encourage questions and record new edge cases for future training

---

🔍 Quality Control & Error Detection

- [ ] Measure inter-annotator agreement (e.g., Cohen's Kappa)
- [ ] Perform statistical anomaly checks (timing, label ratios, IoU/spans)
- [ ] Monitor annotator performance over time
- [ ] Run automated validation scripts (e.g., missing tags, overlapping boxes)
- [ ] Check cross-annotation consistency for similar inputs
- [ ] Set up tiered review workflows for low-confidence annotations
- [ ] Use active learning to prioritize high-uncertainty samples

---

🤝 Disagreement Resolution

- [ ] Analyze root causes of disagreements (e.g., guideline ambiguity)
- [ ] Resolve via additional voting, expert review, or updated labels
- [ ] Escalate recurring issues to stakeholders and refine ontology
- [ ] Update living documentation with resolved cases
- [ ] Share disagreements as examples in training reviews

📦 Data Preparation for ML

- [ ] Standardize annotation format (e.g., IOB for text, COCO for images)
- [ ] Apply preprocessing (tokenization, resizing, normalization)
- [ ] Create edge-case test set for robustness evaluation
- [ ] Use stratified sampling for train/val/test split
- [ ] Address class imbalance via oversampling, undersampling, or SMOTE
- [ ] Apply data augmentation (e.g., back-translation, image flips)
- [ ] Validate schema and clean metadata
- [ ] Document dataset characteristics and preprocessing steps

---

🤖 Support ML Team

- [ ] Track dataset versions and link to model runs
- [ ] Monitor bias and fairness across demographic groups
- [ ] Conduct regular syncs to update priorities and share feedback
- [ ] Identify gaps using model predictions (active learning loop)
- [ ] Maintain centralized docs and onboarding guides
- [ ] Share helper scripts or dashboards for model and data inspection