**🎯 OMR MASTER PLAN**

**Objective:** Build a scalable, high-accuracy OMR system that can process real-world music scores (especially jazz PDFs), handle handwritten-like inputs, and output structured formats like MusicXML or MIDI.

**🧱 PHASE 0: Setup & System Design**

**🔧 Step 0.1 – Define Goals**

* Input: scanned or digital sheet music (PDFs, images)
* Output: structured MusicXML or MIDI
* Capabilities:
  + Detect symbols (notes, rests, clefs, barlines, etc.)
  + Detect staff structure and layout
  + Infer semantic relationships (pitch, duration, rhythm, articulation)
  + Handle real jazz scores, even noisy or irregular layouts

**🔧 Step 0.2 – Environment Setup**

* Create Python virtual environment
* Install core tools:

bash

CopyEdit

pip install ultralytics opencv-python pillow pdf2image music21

* Install MuseScore and ensure CLI works
* Install Label Studio for manual annotation:

bash

CopyEdit

pip install label-studio

**🧪 PHASE 1: Synthetic Data Generation (Bootstrapping the Model)**

**🔹 Step 1.1 – Generate Random Music Snippets**

* Use music21 to create simple melodies and chords
* Export as MusicXML

**🔹 Step 1.2 – Convert MusicXML to PNG**

* Use MuseScore CLI to render

bash

CopyEdit

musescore -o output.png input.musicxml

**🔹 Step 1.3 – Generate Ground Truth Labels Automatically**

* Use known note positions + clefs from MusicXML
* Create bounding boxes for:
  + Noteheads
  + Clefs
  + Key/time signatures
  + Rests
* Save in YOLOv8 format

**🔹 Step 1.4 – Organize Dataset for Training**

bash

CopyEdit

dataset/

├── images/train/

├── images/val/

├── labels/train/

├── labels/val/

└── data.yaml

**🧠 PHASE 2: First Model – YOLOv8 Symbol Detection**

**🔹 Step 2.1 – Train YOLOv8 on Synthetic Dataset**

* Train initial detector on noteheads, clefs, barlines, etc.
* Evaluate bounding box accuracy on validation set

**🔹 Step 2.2 – Export Predictions**

* Apply model on:
  + Synthetic validation data
  + Unlabeled real jazz PDFs (converted to images)

**📖 PHASE 3: Real Jazz Book Integration (Human-in-the-Loop Learning)**

**🔹 Step 3.1 – Prepare Real PDF Data**

* Use pdf2image to convert jazz PDFs into PNGs

**🔹 Step 3.2 – Run Model on Jazz Pages**

* Generate predictions for symbols
* Export to YOLO label format

**🔹 Step 3.3 – Correct Model Mistakes**

* Import predicted images + labels into Label Studio
* Manually correct bounding boxes + labels
* Export corrected dataset

**🔹 Step 3.4 – Retrain Model**

* Combine synthetic + corrected jazz labels
* Retrain model with increased diversity
* Evaluate again on real jazz PDFs

**🧩 PHASE 4: Staff Detection and Structural Layout**

**🔹 Step 4.1 – Use OpenCV for Staff Line Detection**

* Detect staff lines using morphological filters + projection
* Group lines into staves
* Classify staves as single or grand using:
  + Spacing
  + Clef combination
  + Barlines

**🔹 Step 4.2 – Train YOLOv8 for Staff Block Detection (Optional)**

* Label full staff regions (bounding boxes)
* Label type (single\_staff, grand\_staff, etc.)

**🔁 PHASE 5: Semantic Graph Parsing (Notes to MusicXML)**

**🔹 Step 5.1 – Build Symbol Graph**

* Each detected symbol becomes a node
* Edges = proximity, beams, stems, ties

**🔹 Step 5.2 – Infer Pitches**

* Use staff position + clef to compute pitch
* Use MusicXML-compatible pitch notation (e.g., <step>C</step>)

**🔹 Step 5.3 – Infer Durations**

* Use symbol type + dots + beaming

**🔹 Step 5.4 – Output to MusicXML**

* Use music21 to build MusicXML tree
* Export .musicxml or .mid

**🌐 PHASE 6: Feedback & Deployment**

**🔹 Step 6.1 – Streamlit Web Interface**

* Upload image
* See predicted symbols overlaid
* Download MusicXML
* Optional: Correct symbols and resubmit

**🔹 Step 6.2 – User Correction Loop**

* User corrects model output in GUI or Label Studio
* Corrections fed back into training loop

**🧠 PHASE 7: Advanced Features (Post-MVP)**

* Handwriting adaptation via fine-tuning on MUSCIMA++
* Chord symbol detection (OCR + music context)
* Tuplets and complex rhythms
* Articulations, slurs, dynamics
* Multi-page score stitching
* Ensemble and orchestral layout recognition

**✅ Summary of the Plan**

| **Phase** | **Name** | **Output** |
| --- | --- | --- |
| 0 | Setup | Tools ready (YOLOv8, MuseScore, Label Studio) |
| 1 | Synthetic Dataset | PNG + YOLO labels for clean music |
| 2 | First Model | Trained YOLOv8 to detect musical symbols |
| 3 | Jazz Books | Refined model with real scanned scores |
| 4 | Staff Logic | Staff blocks, types, and structural layout |
| 5 | Parsing | MusicXML reconstruction (notes, rhythms) |
| 6 | Feedback UI | Streamlit app + Label Studio review loop |
| 7 | Advanced | Handwriting, chords, dynamics, orchestral scores |