PPT Outline: eScriptorium for OCR & Digitalization Workflow

Slide 1: Title Slide

- *Title: Enhancing Our OCR Workflow with eScriptorium*
- *Subtitle:* How eScriptorium Can Improve Our Digitalization Process

1. Understanding eScriptorium

Slide 2: What is eScriptorium?

- **Open-source & fully offline** tool for **transcribing & annotating historical texts**
- Supports **OCR (printed text) & HTR (handwritten text)**
- Developed for **challenging, old documents**
- Can be **hosted locally**—no reliance on cloud services

Slide 3: Key Features

✅ **AI-Powered OCR & HTR**: Works with printed & handwritten texts
✅ **Automated & Manual Annotation**: Train custom models
✅ **Supports Complex Layouts**: Multi-column, marginalia, mixed printed & handwritten
✅ **Offline & Self-Hosted**: No cloud dependency; full control over data

2. Why Should We Use eScriptorium?

Slide 4: Problems in Our Current Workflow

- **LabelMe segmentation is slow & manual**
- **Standard OCR tools fail on historical texts**
- **No handwriting recognition (HTR) in our current workflow**
- **MSSQL stores data, but no built-in transcription & model training**
- **Dependence on cloud-based OCR tools for difficult texts**

Slide 5: How eScriptorium Solves These Problems

✅ **Offline & Secure:** No internet needed, full data control
✅ **Train Custom OCR & HTR Models:** Improves accuracy over time
✅ **Combines Segmentation + OCR in One Tool:** Faster workflow
✅ **Works with MSSQL:** Export recognized text easily

3. Integration into Our Workflow

Slide 6: Current vs. New Workflow

🟡 **Current Process:**

1️⃣ **LabelMe** (Manual segmentation)

2️⃣ **OCR tool** (Standard recognition, struggles with old texts)

3️⃣ **Manual corrections** (Slow & inefficient)

4️⃣ **Store in MSSQL**

🟢 **New Process with eScriptorium:**

1️⃣ **eScriptorium** (Segmentation + OCR/HTR + Model Training)

2️⃣ **Export structured text** (ALTO, XML, CSV)

3️⃣ **Store in MSSQL**

Slide 7: How We Can Use It

4. Next Steps & Implementation

Slide 8: Plan for Adoption

📌 **Step 1:** Install & test eScriptorium locally

📌 **Step 2:** Train custom models on our historical documents

📌 **Step 3:** Compare accuracy vs. existing OCR tools

📌 **Step 4:** Integrate with MSSQL for storage & retrieval

📌 **Step 5:** Transition from LabelMe segmentation to eScriptorium

Slide 9: Conclusion

💡 **eScriptorium improves OCR accuracy, speeds up workflow, and gives us full offline control**

🚀 **Next step: Start testing it with our historical documents!**