**MEETING MINUTES**

**Professional Computing Project Team 03**

**MEETING OVERVIEW**

**Date:** August 18, 2025  
**Time:** 12:00PM  
**Meeting Type:** Project planning and work stream organization  
**Purpose:** Organize project structure, define work streams, and establish technical implementation strategy for an AI-powered document research tool with PDF extraction and library accessibility features

**Attendees:**

* Rania Khan
* Hazel Wang
* James Felstead
* Chunyu Zheng
* Peter Fang

**KEY DISCUSSION POINTS**

**Project Structure and Work Stream Organization**

The team finalized their **4-workstream architecture** for the **8-week implementation timeline** (Aug 18 – Oct 13) with clear individual ownership and deliverables:

**Workstream 1 - Backend + Metadata (Peter & James):**

* Core pipeline connecting to **OpenAI Deep Research (GPT-5)**
* Metadata extraction and normalization engine
* Multilingual query generation and error handling
* Access classification system (Green/Yellow/Red coding)

**Workstream 2 - Frontend & UX (Rania):**

* **Google-style search interface** with advanced filtering
* **Interactive results table** with sortable columns and expandable rows
* **Color-coded access badges** and user action buttons
* Search history functionality

**Workstream 3 - Batch Search & Export (Hazel):**

* **CSV/JSON bulk processing** capabilities
* **Structured export system** organized by Country → Year → Type
* Progress tracking for large-scale document processing
* Performance optimization for "low thousands" of documents

**Workstream 4 - Database Design & Data Management (Chunyu):**

* **Robust schema** supporting documents, sources, searches, and batches
* **Data integrity features** including deduplication and geo normalization
* **Admin utilities** for data inspection and cleaning
* **ISO standards compliance** for country/region naming

**Technical Implementation Strategy**

**Development Architecture:**

* **Backend API** with /search and /batch-search endpoints
* **Database-driven approach** with normalized metadata storage
* **Scalable batch processing** supporting thousands of documents

**Document Access Classification System**

**Three-tier accessibility framework:**

* **Green (PDF):** Direct downloadable documents
* **Yellow (Remote request):** Library contact required for digital access
* **Red (Physical only):** In-person library visit required

**Implementation approach:**

* **Preprocessing strategy** for library accessibility ratings
* **Contact information integration** for Yellow/Red classified documents
* **Automated classification** based on source analysis and availability patterns

**DECISIONS MADE**

1. **Work Stream Consolidation:** Merged original work streams 1 and 2 into a single backend/metadata stream, resulting in 4 total work streams
2. **Work stream Assignments:** all 4
3. **Development Prioritization:** Focus on basic PDF extraction functionality before implementing library accessibility features
4. **Project Management System:** Implement **ClickUp** for ticket management and progress tracking
5. **Phased Development Approach:** Establish core end-to-end functionality first, then implement additional features like multi-language support
6. **Individual Component Ownership:** Team members to take ownership of specific workflow components rather than working collectively on all tasks

**ACTION ITEMS**

|  |  |  |
| --- | --- | --- |
| **Responsible Party** | **Task** | **Deadline** |
| Rania Khan | Create ClickUp tickets for all work streams and team members | This week |
| All team members | Write detailed descriptions in assigned tickets outlining approach and expected outputs | This week |
| Team | Prepare questions about query requirements for Wednesday supervisor meeting | Before Wednesday |

**NEXT STEPS**

* **Immediate:** Complete ClickUp project structure setup and technical documentation in assigned tickets while finalizing tool selection and implementation methodology
* **Short-term (1-2 weeks):** Focus shifts to active development with James delivering PDF extraction functionality, supervisor consultation for preprocessing approval, and each team member producing tangible deliverables
* **Medium-term (2+ weeks):** Complete end-to-end working prototype, implement library accessibility assessment and cloud storage integration, and evaluate multi-language support capabilities
* **Ongoing:** ClickUp progress tracking and regular team synchronization meetings for coordination

**TECHNICAL ARCHITECTURE SUMMARY**

**Core Components:**

1. **PDF Extraction Engine** - URL-to-PDF processing capability
2. **AI/Metadata Processing** - Intelligent document analysis and categorization
3. **Library Accessibility System** - Preprocessed accessibility ratings and location data
4. **User Interface** - Frontend for search, discovery, and document access
5. **Data Management** - Database and cloud storage integration

**Technology Stack Considerations:**

* **Web scraping tools** for document discovery
* **AI integration** (OpenAI and similar platforms)
* **Cloud storage solutions** for document management
* **Multi-language support** capabilities for future expansion

**CONCLUSION**

The team has successfully established a clear technical direction and organizational structure for the AI-powered document research tool project. The consolidation of work streams and individual component ownership provides a foundation for efficient development. The emphasis on core functionality delivery before feature expansion demonstrates mature project management thinking and realistic timeline assessment.

**Meeting Minutes Prepared by:** Rania Khan  
**Date Prepared:** August 18, 2025  
**Next Meeting:** Wednesday (Weekly client recurring)