Challenge

Over 70 years of legacy project data - including:
- PDFs (e.g., technical reports, images)
- Scanned documents (e.g., handwritten notes, drawings)
- Spreadsheets (e.g., statistical data)
- Videos (very few)

Problems:
- Data is siloed, unstructured, and hard to access
- Scattered across teams and formats
- Retrieval relies on human memory or personal networks
- Limited visibility into historical decisions and outcomes

Current Situation

- Data is stored in project-based folders, organised by roles
- Deltek ERP system only contains metadata (e.g., client name, project name)
- Actual documents are not integrated or easily searchable
- Maybe many PDFs are scanned versions, some with handwritten content

Project Opportunity

Generally, use AI to transform legacy data into a structured, searchable, and intelligent knowledge system.
- Natural language processing (NLP) for extracting meaning from reports and text documents
- Computer vision (CV) for interpreting images and drawings
- Optical character recognition (OCR) for digitising scanned files and handwritten notes
- Knowledge graphs and tagging to build connections between project contents and decisions

Potential Efforts

- Enhancing project scoping with precedent knowledge
- Benchmarking past project outcomes across time and locations
- Improving internal knowledge transfer
- Reusing prior design solutions to reduce duplication