














SIH KMRL DOCUMENT OVERLOAD AN AUTOMATED SOLUTION

KMRL, like any large infrastructure company, generates and receives **huge volumes of documents** daily:

- Engineering drawings 
- Maintenance records 
- Vendor invoices 
- Purchase orders 
- Safety and compliance documents 
- Regulatory circulars 
- Incident reports 
- HR and employee files 

Pain Points


-  **Searchability** – Hard to quickly find the right document.
-  **Categorization/Tagging** – No consistent way of classifying documents.
-  **Access Control** – Not everyone should see every file.
-  **Time Wastage** – Employees spend too much time searching/organizing.
-  **Scalability Issues** – As KMRL grows, documents multiply fast.

Tech Stack Ideas

- **Frontend:** React, Next.js, or Angular
- **Backend:** Node.js, Django, or Spring Boot
- **Database:** MongoDB (good for storing metadata) + Cloud Storage (for actual files)
- **AI/NLP:** Python (spaCy, HuggingFace, TensorFlow) for classification & tagging
- **Search:** Elasticsearch or simple full-text search in MongoDB/Postgres
- **Authentication:** JWT or OAuth

Benefits of Access Control & Document Availability

1. Security & Confidentiality

- Not every employee should see **all documents** (e.g., financial invoices, regulatory notices).
- Access control ensures **only authorized roles** (Admin, Finance Team, HR, Engineers) can view or edit specific files.
 Prevents data leaks & misuse.

2. Compliance & Audit Readiness

- KMRL deals with **government regulations, contracts, and safety documents**.
- Proper access logs + controlled availability ensure compliance with:

- RTI (Right to Information) rules
 - Government audit checks
 - Safety and regulatory inspections
 - 👉 Makes audit trails transparent & easy to verify.
-

3. Operational Efficiency ⚡

- Employees don't waste time searching through irrelevant documents.
 - A maintenance engineer sees **only maintenance-related files**, not HR forms.
 - 👉 Saves time and increases productivity.
-

4. Reduced Risk of Errors ❌

- Imagine if an intern mistakenly edits/deletes a critical regulatory file 🤖.
 - With **role-based restrictions**, only the right people can make changes.
 - 👉 Protects the integrity of important documents.
-

5. Better Collaboration 🤝

- Teams can **share documents safely** with other departments or vendors without exposing everything.
 - Example: Vendor can see **purchase orders** but not **internal incident reports**.
-

6. Scalability for Large Organizations 🌐

- As KMRL expands, documents will **multiply across departments**.
 - Access control scales well, keeping things organized per team/role.
-

🎯 Real-World Example at KMRL

- **Finance Dept** → Sees only invoices, vendor payments.
- **Operations Dept** → Sees maintenance reports, job cards.
- **HR Dept** → Sees employee records, training documents.
- **Top Management** → Has access to everything.

The Problems (Simplified)

1. Information Latency ⌚

- Managers waste hours scanning long documents just to find the 2–3 important lines.
- This delays decisions (like when to run trains, approve payments, assign staff).

2. Siloed Awareness 🏢

- Different departments don't know what others are doing.
- Example: Procurement signs a spare parts contract while Engineering already planned a design change.
- HR schedules training but doesn't see the latest safety circular.

3. Compliance Risks ⚠️

- Important government/regulatory updates are lost in inboxes.
- This can lead to **missed deadlines** or legal troubles.

4. Knowledge Loss

- When experienced staff retire or transfer, their knowledge is stuck in static files that new people can't easily use.

5. Duplicated Work

- Teams keep creating their **own summaries** of the same document.
- Leads to wasted time, confusion, and version mismatches.