DA là hệ thống quản lý văn bản.  
- web

- yêu cầu: + đọc các văn bản mẫu ( pdf, excel, doc, docx) do admin upload convert sang text.

                + chuyển đổi sang dạng text để có khả năng search keyword ( tên văn bản, các keyword có trong văn bản....).

+ Sử dụng thuật toán N-Gram để phân tích ra keyword (viết thành dạng lib càng tốt. đầu vào là text đâu ra là 1 list<String> keywords) . lưu ý cần có 1 bước tiền xử lý văn bản text trước khi đưa vào thuật toán N-Gram. (<https://spark.apache.org/docs/latest/ml-features.html>).

+ cho phép người dung search theo keywords hoặc nội dung trong text.

+ Giao diện người dung gần giống như 3 ảnh ở ngoài (optional có thể khác nếu hay hơn).

                + cung cấp link down văn bản gốc do admin upload cho user.

+ Có 2 cấp quản lý người dung là admin và trưởng phòng . admin có thể tạo phòng và trưởng phòng . trưởng phòng tạo nhân viên cho phòng(users) . users chỉ search được các văn bản do trưởng phòng upload. ( có thể nhận được mail thông báo có văn bản mới hoặc thay thế nếu có … opitional)