

Control de Versiones SharePoint - CUN

Sistema de monitoreo y control de versiones que sincroniza documentos desde SharePoint a Excel y expone la información mediante una API REST.

Descripción

El sistema consta de dos componentes principales:

1. **Conector SharePoint-Excel:** Sincroniza periódicamente los documentos desde SharePoint a un archivo Excel local
2. **API REST:** Lee el archivo Excel y expone los datos para consulta y monitoreo detallado

Características Principales

Conector (main.py)

- Sincronización automática cada 2 minutos con SharePoint
- Extracción de metadatos de documentos
- Generación de archivo Excel con información actualizada
- Sistema de logs para seguimiento de sincronización
- Detección automática de cambios y duplicados
- Validacion de Data

API (api.py)

- Lectura y procesamiento del archivo Excel
- Análisis temporal de cambios (última hora, día, semana, mes)
- Detección y seguimiento de duplicados
- Estadísticas por tipo de archivo
- Endpoints mejorados para consulta de:
 - Lista actual de archivos con metadatos
 - Historial detallado de cambios
 - Análisis de duplicados y anomalías
 - Estadísticas temporales y por categoría

Tecnologías Utilizadas

- **Backend:** Python 3.9+
- **Framework API:** FastAPI
- **Origen de Datos:** SharePoint REST API
- **Almacenamiento Intermedio:** Excel
- **Procesamiento de Datos:** Pandas
- **Documentación:** Swagger/OpenAPI

API Endpoints

Documentación

- Swagger UI: <http://localhost:8000/docs>
- ReDoc: <http://localhost:8000/redoc>

Endpoints Principales

| Endpoint | Método | Descripción |
|--------------------------------------|--------|--|
| /archivos | GET | Lista actual de archivos con metadatos completos |
| /historico | GET | Historial completo organizado por fechas |
| /cambios | GET | Análisis detallado de cambios y estadísticas |
| /actualizar-archivos | POST | Actualiza la lista de archivos desde Excel |

Detalles de Endpoints

GET /cambios

- Resumen general de archivos
- Cambios por períodos temporales
- Análisis de duplicados
- Estadísticas por tipo de archivo
- Ordenamiento por fecha más reciente

GET /historico

- Organización por fechas
- Resumen diario de cambios
- Lista detallada de modificaciones
- Enlaces a documentos

Monitoreo y Análisis

El sistema proporciona:

- **Análisis Temporal:**
 - Cambios en la última hora
 - Cambios en las últimas 24 horas
 - Cambios en la última semana
 - Cambios en el último mes
- **Análisis de Duplicados:**
 - Detección automática
 - Ubicaciones múltiples

- Fechas de modificación
- Estado de cada copia

- **Estadísticas por Tipo:**

- Cantidad de archivos
- Tamaño total
- Últimas modificaciones
- Top 5 archivos recientes

Interfaz de Pruebas

- Interfaz web para pruebas de API
- Visualización de resultados en formato JSON
- Acceso directo a documentos
- Filtros y ordenamiento

Seguridad

- Autenticación mediante credenciales de SharePoint
- Almacenamiento seguro en AWS S3
- Variables sensibles en .env

Autores

- **Kevin Claro** - *Desarrollo Inicial* - [kevin_claro@cun.edu.co]

Licencia

Este proyecto está bajo la Licencia MIT - ver el archivo [LICENSE](#) para más detalles.

Estructura del Proyecto

SegundoConector/

```

├── api/ # Componentes de la API
│   ├── models.py # Modelos Pydantic
│   ├── routes.py # Rutas de la API
│   └── services.py # Servicios y lógica de negocio
├── core/ # Núcleo de la aplicación
│   ├── config.py # Configuraciones
│   └── logging.py # Configuración de logs
├── data/ # Datos y archivos temporales
│   ├── historial_archivos.json # Historial de cambios
│   └── novedades.json # Registro de novedades
├── models/ # Modelos de datos
├── repositories/ # Capa de acceso a datos
├── services/ # Servicios adicionales
└── utils/ # Utilidades y helpers

```

- |— .env # Variables de entorno
- |— api.py # Servidor API REST
- |— main.py # Conector SharePoint-Excel
- |— Documentos_SharePoint.xlsx # Archivo de sincronización
- |— requirements.txt # Dependencias del proyecto