Unidad 9 · Capturar, clasificar y recuperar documentos

Resumen · Libro del alumno, págs. 137–152

Un sistema de gestión documental reúne **repositorio, metadatos, equipos, red y roles**. Escanear no autentica: el **reconocimiento de texto** solo hace buscable el documento y puede equivocarse (1.200 leído como 12.000, diez veces mayor), así que las cifras importantes se **cotejan con el original**. Los **metadatos** (identificador, tipo, fecha, estado, responsable y acceso) permiten encontrar y controlar el documento, con valores de una lista cerrada. La prueba de recuperación termina cuando se recupera **el documento correcto en su versión vigente**. Una copia en la misma carpeta sincronizada **no es independiente**, y el Código de Comercio fija seis años de conservación para la documentación mercantil.

9.1 · Elementos del sistema

- **Metadatos.** Son los datos que describen un documento: identificador, tipo, fecha, versión, responsable y nivel de acceso. Permiten encontrarlo y controlar quién lo ve. Sin metadatos suficientes, un documento es difícil de clasificar y controlar.
- **Más que carpetas.** Un **sistema de gestión documental** (SGD) es el conjunto de elementos que permite capturar, ordenar, encontrar, proteger y conservar documentos.
- **Documento 1 · Componentes del sistema de Vértice.** (ficticio)
- **Roles y permisos.** Cada persona hace una cosa: quien consulta, quien registra y quien administra.
- **Qué no es un SGD.** Una carpeta sincronizada, un correo con adjuntos o un disco compartido no son por sí solos un sistema de gestión documental, aunque guarden los mismos archivos.
- **Quién usa el sistema.** Toda la utilidad de un sistema documental está en que muchas personas lo usen sin pisarse.

9.2 · Captura y OCR

- **Capturar con calidad.** La **captura** convierte un papel en imagen digital.
- **El reconocimiento de texto.** El **OCR** lee la imagen y la convierte en texto que se puede buscar.
- **Documento 2 · Registro de captura de E01, E02 y E03.** (ficticio, martes 23/03/2027)
- **Escanear no es autenticar.** Una copia digital sirve para trabajar, pero no demuestra por sí sola que el documento sea el original ni que no se haya modificado.
- **Qué cotejar siempre.** No hace falta comparar cada palabra, pero sí lo que, si falla, cambia una decisión: importes, cantidades, fechas, identificadores y nombres de personas o de empresas.
- **Registrar el control.** Quien coteja lo anota: qué documento, qué se comparó, qué se corrigió, quién lo hizo y cuándo.

9.3 · Metadatos y clasificación

- **Campos obligatorios.** Un registro útil exige un **identificador** único, el **tipo** de documento, la **fecha**, el **estado** y el **acceso**.
- **Documento 3 · Campos y valores permitidos.** (ficticio)
- **Nombres con regla.** El nombre del archivo sigue un patrón: «UD09_P09_albaran_v01» dice unidad, paso, tema y versión.
- **Documento 4 · Registro de partida B01, seis documentos.** (ficticio; vista parcial, sin la columna Acceso)
- **Clasificar es decidir quién ve qué.** El campo de acceso no es un detalle: un contrato en borrador no debe ser público, y una solicitud con datos personales solo la ve quien la tramita.
- **Pocos campos, bien elegidos.** La tentación es pedir muchos campos, pero cada campo obligatorio es trabajo para quien registra y una oportunidad de error.

9.4 · Búsqueda y recuperación

- **Buscar por campos.** Con metadatos se puede consultar por **identificador**, que devuelve un único documento, o por **estado**, que devuelve un grupo.
- **Documento 5 · Prueba de recuperación PR-09 sobre B01.** (ficticia, martes 23/03/2027)
- **Recuperar el correcto.** La prueba no acaba cuando aparece un archivo: acaba cuando se abre, se comprueba que es el documento esperado y que es la versión vigente.
- **Probar lo que no existe.** Una buena prueba incluye también consultas que no deberían devolver nada, como un identificador que no existe, y comprueba que, en efecto, no devuelven nada.
- **Cuando falla.** Si una consulta devuelve un resultado que no se esperaba, se anota el fallo, se corrige el metadato o el nombre que lo causó y se repite la prueba completa.

9.5 · Custodia y sostenibilidad

- **Custodiar es proteger y poder devolver.** La **custodia** reúne tres cosas: quién accede, cuántas copias hay y cuánto tiempo se conserva.
- **Documento 6 · Tipos de copia.** (ficticio)
- **Conservar el tiempo necesario.** Algunos documentos mercantiles se conservan seis años desde el último asiento, pero seis años no son un plazo universal para cualquier dato.
- **Papel evitado, almacenamiento duplicado.** Tramitar en digital 100 expedientes de 4 hojas evita imprimir 400 hojas, menos de una resma de 500.
- **Acceso, copia y plazo.** Estas tres decisiones se toman juntas.
- **Un ejemplo con fecha.** Si el último asiento de un libro mercantil es del 15/03/2027, seis años después llegan el 15/03/2033.

CASO RESUELTO · TRES DOCUMENTOS ESCANEADOS

Cifra corregida cotejando el original, archivo renombrado con identificador y metadatos, y recuperación de E02 comprobada por su identificador.

NO CONFUNDAS

- Escanear y autenticar: el OCR solo hace buscable.
- Archivo con nombre («nuevo.pdf») y archivo con identificador y metadatos.
- Copia en la misma carpeta y copia independiente.

DÓNDE SE EXPLICA EN EL LIBRO

- **9.1** Elementos · 139–140
- **9.2** Captura y OCR · 140–142
- **9.3** Metadatos · 142–143
- **9.4** Búsqueda · 144–145
- **9.5** Custodia · 145–146

Material complementario del libro del alumno. Las personas, empresas y situaciones de los casos son ficticias, salvo que se indique otra cosa.