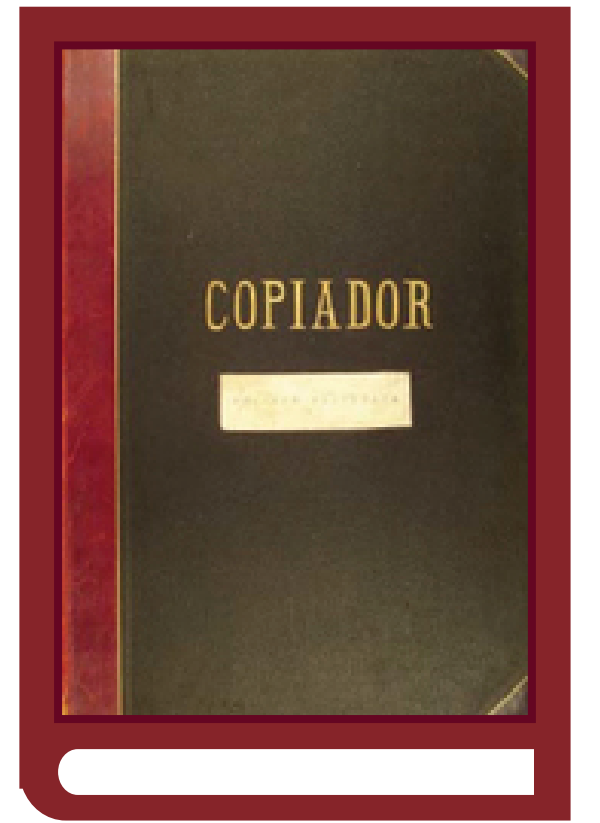




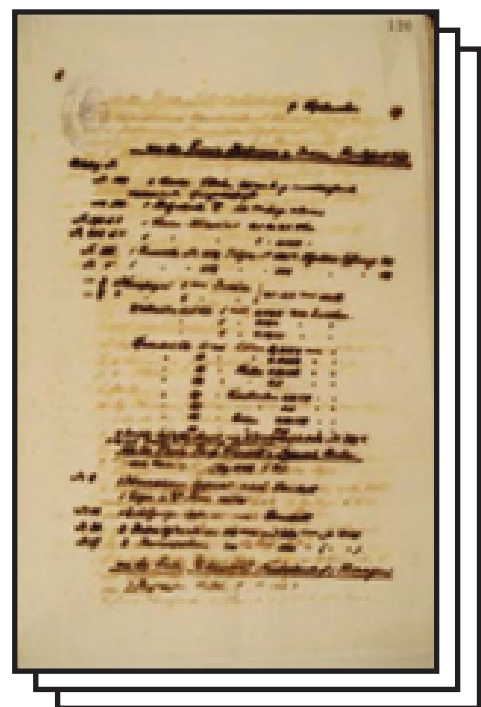
PARTICULARIDADES Y DESAFÍOS EN LA DIGITALIZACIÓN

LIBRO COPIADOR - FACULTAD DE CS. FÍSICAS, MATEMÁTICAS Y ASTRONÓMICAS (1918-1925)

Un caso de uso de un repositorio institucional para la preservación de Memoria Institucional



SEDICI y el Museo de Física de la Facultad de Ciencias Exactas de la UNLP destinaron personal para la digitalización de un documento archivístico: el libro *Copiador - Facultad de Ciencias Físicas, Matemáticas y Astronómicas (1918-1925)*. Se siguieron los estándares internacionales para la digitalización (IFLA, NARA, FADGI, etc.), pero **muchas de las dificultades que presentó el material no estaban contempladas en la bibliografía.**



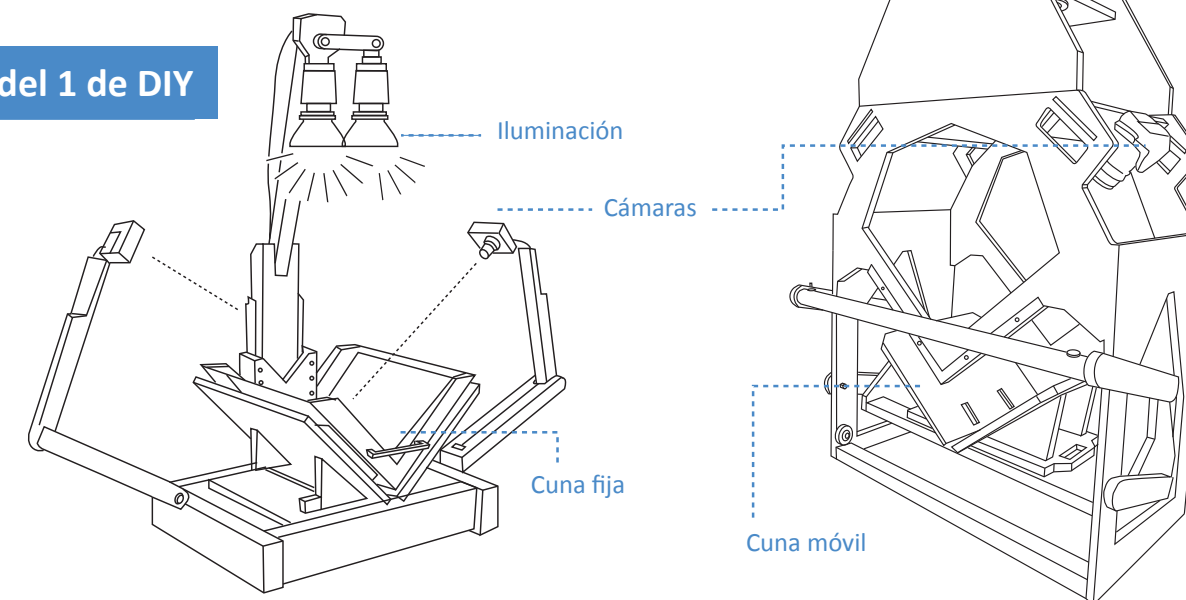
ESTADO DE CONSERVACIÓN

- Papel amarronado, débil, friable con roturas y desprendimientos
- Escritura manuscrita con tinta difundida en el papel y transferida a los siguientes, con pérdida de nitidez e imagen doble
- Escritura mecanográfica poco legible.
- Encuadernación con especiales requerimientos de manipulación

ESCÁNERES Y SOFTWARES DESCARTADOS

Los escáneres **DAL** utilizados para el escaneo de libros en buen estado no pudieron utilizarse porque el giro de las páginas hacía que el papel pudiera quebrarse.

Model 1 de DIY



Derecho a leer
<http://derechoaleer.org/blog/>

Archivista 2014

Presentaba muchas mejoras pero no pudo emplearse por el tamaño y estado de deterioro del Libro Copiador.

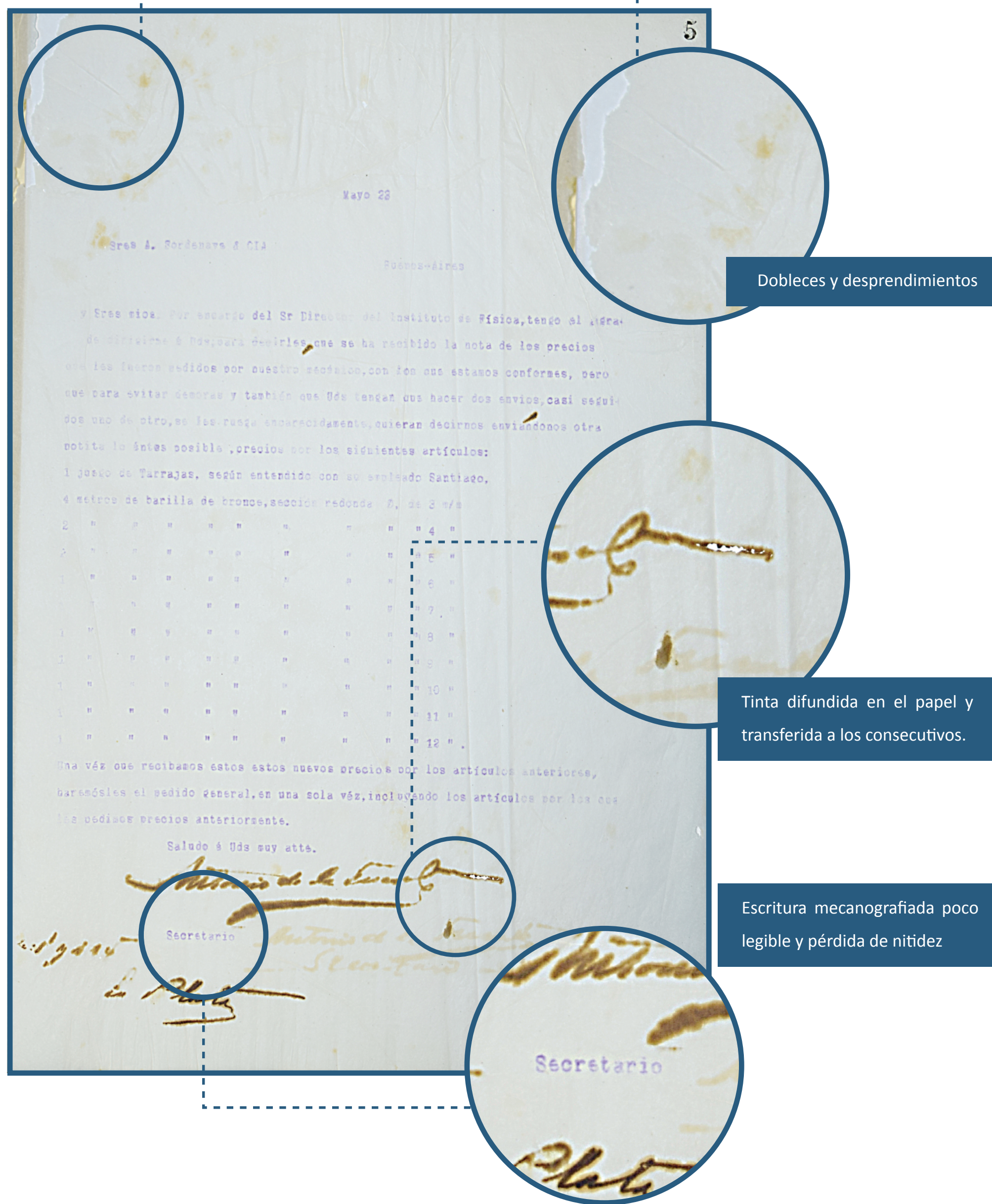
Los softwares utilizados para la captura de imágenes y posprocesamiento debieron adaptarse y transformarse dejando de lado el firmware modificado (CHDK) y el Scan Tailor.



El Scan Tailor no tenía funciones avanzadas de ajuste de colores y nitidez.



Firmware modificado CHDK: requería la configuración in situ de cada cámara por separado y el uso de una tarjeta de memoria para mover las imágenes a la computadora.



ESCÁNER CON CÁMARA CENTRAL CONSTRUIDO PARA EL TRABAJO CON SOFTWARE DE CAPTURA



Se optó por un sistema de escaneo **rediseñado a partir del Model 1** de DIY, con una cámara cenital apuntando hacia el libro, junto con dos luces LED dicroicas de tonalidad cálida cuya temperatura no daña el material.

El **digiCamControl 2.0.72.0** fue un programa rápido, confiable y versátil para la captura de imágenes y permitió el manejo de las cámaras directamente desde la computadora.



Photoshop CS6 permitió mayor libertad en la manipulación de las imágenes destinadas al reconocimiento de texto. Se aplicaron filtros y se automatizó el procedimiento estándar para todas las imágenes.

PRUEBAS REALIZADAS CON SEPARADORES DE COLOR PARA PAPEL TRANSLÚCIDO



Color elegido para resaltar y eliminar más fácilmente las manchas amarillentas. Acabó resaltando demasiadas manchas e hizo más difícil la limpieza del texto.



Se buscaba destacar y resaltar más los textos y despegarlos del fondo. Se obtuvo un fondo muy oscuro con contraste insuficiente y pliegues de las hojas demasiado definidos que dificultaba la legibilidad.

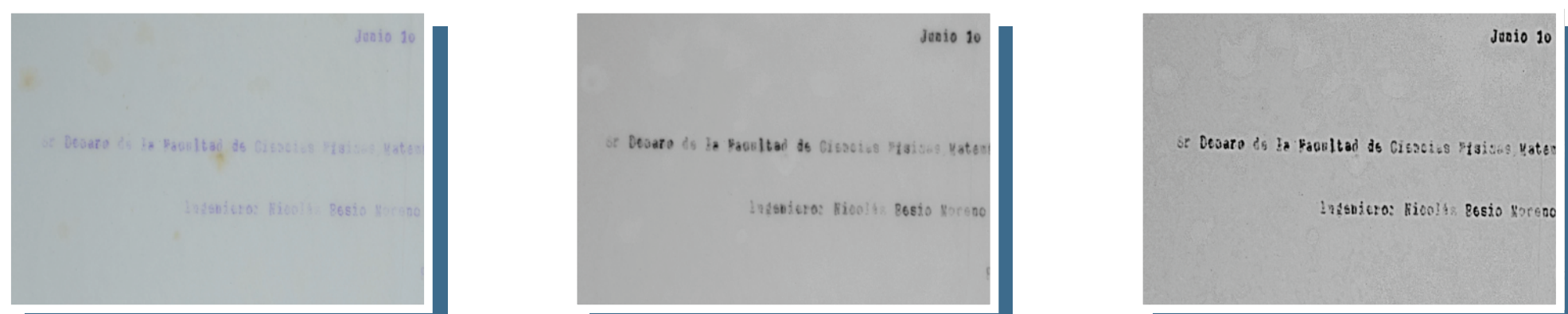


El separador de color blanco que mejoró el contraste del papel con la tipografía solucionó el inconveniente de las transparencias del papel.

POSTPROCESOS DE AJUSTE DE IMAGEN Y ENFOQUE



Se utilizaron dos filtros para mejorar la imagen capturada con el fin de hacer el OCR del documento



- **Desaturación por color (black and white filter)**: este filtro desatura los colores por separado. Esto permite seleccionar las tonalidades que representan manchas, suciedades y atenuarlos hasta que la superficie se vea homogénea.
- **Enfocar (smart sharpen)** para acentuar el borde de la tipografía en la imagen y mejorar el contraste con el fondo.

El proceso completo se automatizó por medio de las funciones *Actions* y *Droplet* de Photoshop

IMAGEN ORIGINAL E IMAGEN MEJORADA LISTA PARA EL RECONOCIMIENTO DE TEXTO

