



Acervo Digital: Procesamiento de Textos

¿Qué se hizo?	Implementar algoritmos para la lectura, procesamiento y análisis documentos PDF dentro del marco de acervo documental digital del gobierno.	
¿Para qué?	Para facilitar la indexación, búsqueda y generación de estadísticas de documentos digitales. Con el propósito de contar con una biblioteca de documentos que apoyen la elaboración de informes de gestión y faciliten el empalme con el gobierno entrante.	
¿Cómo se hizo?	<ol style="list-style-type: none">1. Implementación de scripts para la descarga, lectura, limpieza y procesamiento de documentos PDF.2. Implementación de API que integró todos los algoritmos de procesamiento y que se conectó con el Datálogo Colombia (https://datalogo.dnp.gov.co/)	