



OODOCUMENT

Integración de LibreOffice a IA²



¿ Quiénes somos y qué hacemos?

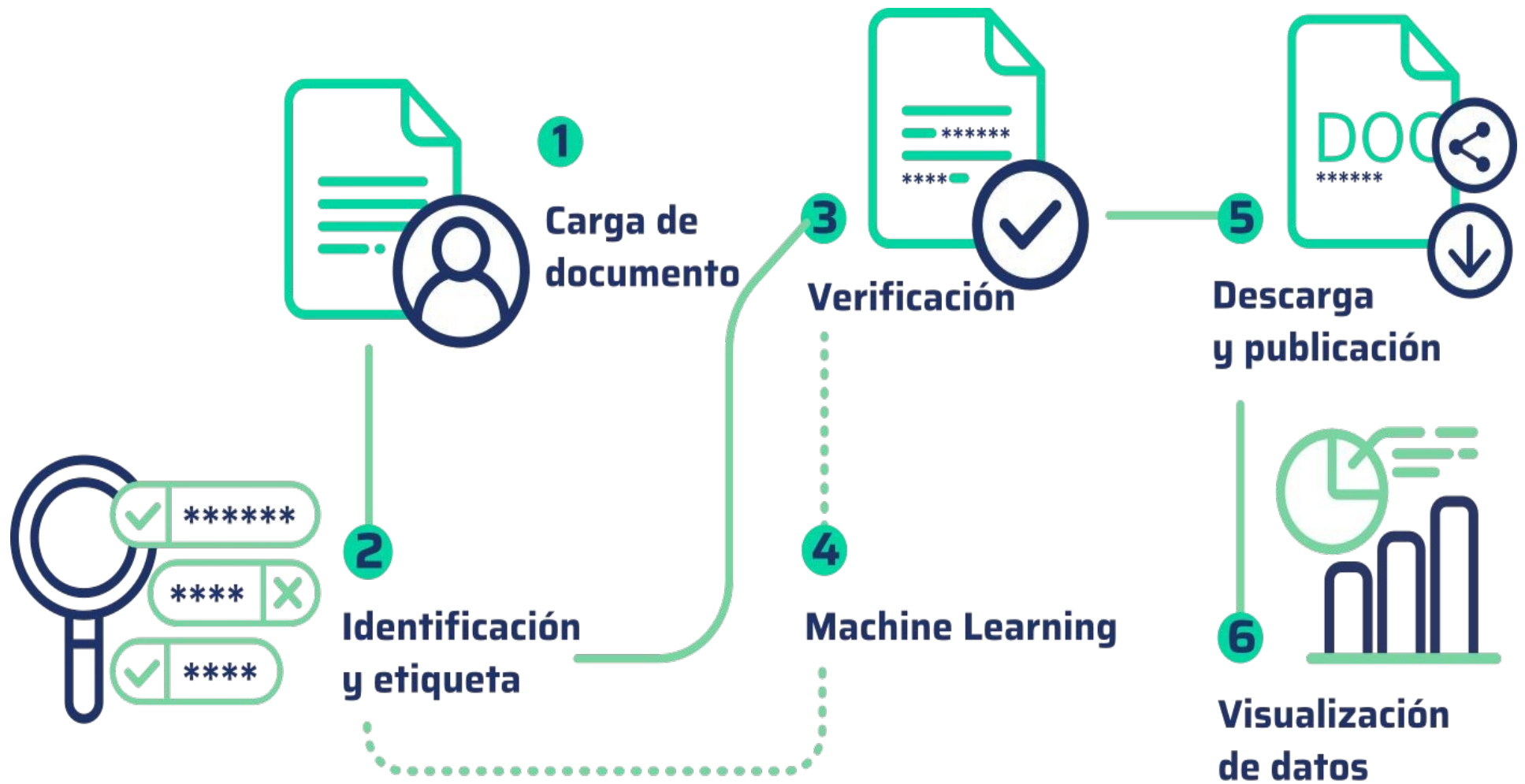
● **Cooperativa de Trabajo Cambá**

● **FACTTIC**

● **IA²**



#Funcionamiento



¿ Qué necesidades teníamos ?

- **Conversión multiformato (odt, doc, docx, pdf, txt)**

- **Manipulación y creación de estilos**

- **Reemplazo de palabras en un texto**

- **Infraestructura del proyecto en Python**

- **Servicio externo**



Libreoffice al rescate

- Python-UNO bridge
- LibreOffice headless service
- Integración con Docker
- Paquete PyPi: Oodocument



Integración: Docker y Python-UNO

Dockerfile

```

7 FROM python:3.7
8
9 RUN apt-get update && apt-get install -y --no-install-recommends \
10     python3-uno \
11     libreoffice-writer \
12     default-jre \
13     && rm -rf /var/lib/apt/lists/*
14
15 RUN cp /usr/lib/python3/dist-packages/uno.py /usr/local/lib/python3.7/site-packages/
16 RUN cp /usr/lib/python3/dist-packages/unohelper.py /usr/local/lib/python3.7/site-packages/
17
18 CMD ["python3"]
  
```

Ejemplo de uso

```

docker build -t my-python-uno .
docker run -it --rm --name my-running-script
    -v "$PWD":/usr/src/myapp
    -w /usr/src/myapp my-python-uno python tu_script.py
  
```



Inicio de servicio de Libreoffice

```
/usr/bin/libreoffice --headless --accept=  
"socket,host=0.0.0.0,port=8001;urp"
```

Inicializa en “headless mode” que permite usar la aplicación sin interfaz gráfica (GUI).

Este modo especial puede ser usado cuando la aplicación es controlada por un cliente externo via una API.



Ejemplos de uso de Oodocument

Conversion de .docx a .txt

```
from oodocument import oodocument
oo = oodocument('./input.docx', host='0.0.0.0', port=8001)
oo.convert_to('./output.txt', 'txt')
oo.dispose()
```

Codigó fuente ejemplo

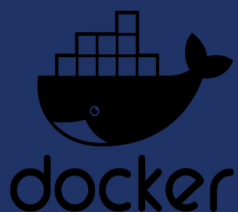
```
def convert_to(self, dest, format):
    # Filters https://git.libreoffice.org/core/+refs/heads/master/filter/
    formats = {
        "pdf": "writer_pdf_Export",
        "txt": "Text",
        "docx": "MS Word 2007 XML",
        "odt": "writer8",
        "doc": "MS Word 97",
    }
    filter = PropertyValue()
    filter.Name = "FilterName"
    filter.Value = formats[format]
    self.save(dest, filter)
```

Reemplazo con indices y estilo

```
from oodocument import oodocument
data = {}
data['holamundo'] = 'xxx'
oo = oodocument('./input.docx', host='0.0.0.0', port=8001)
oo.set_font_color(255, 255, 0)
oo.set_font_back_color(255, 0, 0)
oo.set_clear_hyperlinks(False) # default value is True
oo.replace_with(data)
oo.dispose()
```



Enlaces



<https://github.com/instituciones-abiertas/python3-uno-dockerfile>

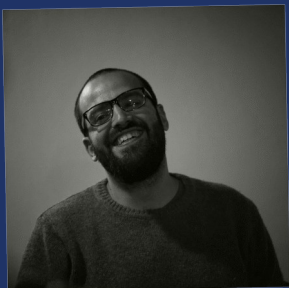


<https://github.com/instituciones-abiertas/oodocument>



<https://github.com/instituciones-abiertas>





charlie@camba.coop



josx@camba.coop



Gracias por escucharnos
¿Preguntas?

