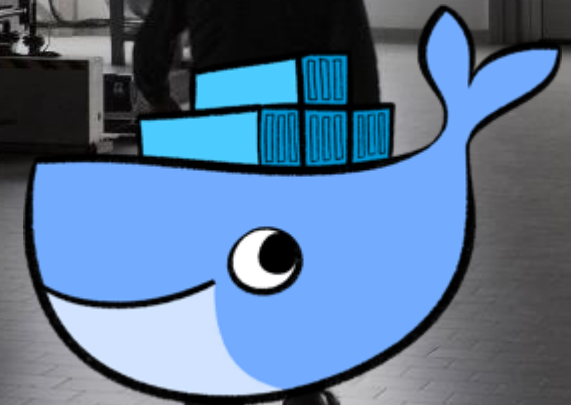


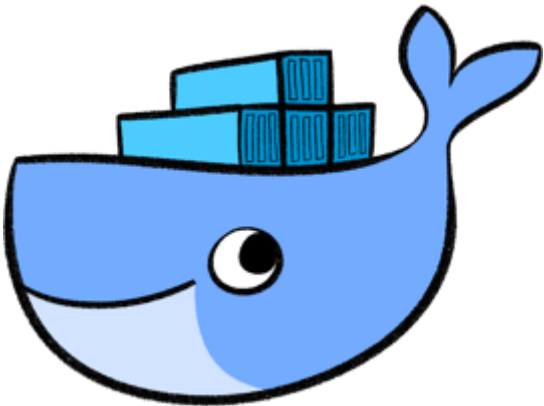


# Construye tu plataforma de Data Science con Docker



# Pre-requisitos

- Instalar Docker
  - <https://www.docker.com/get-started>
- Descargar repositorio
  - <https://github.com/alehpineda/TL-Docker-Data-Science>

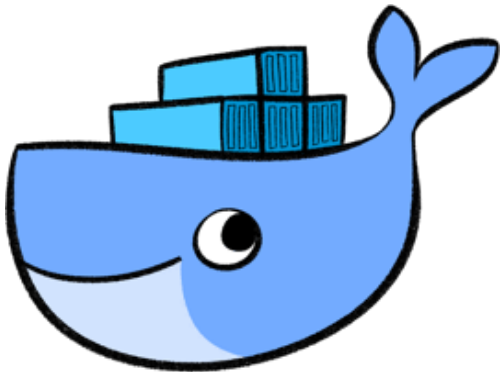


**ABInBev**



# Pre-requisitos

- Crear una cuenta en Docker Hub
  - <https://hub.docker.com/>
- Descargar las siguientes imágenes de Docker en la terminal:
  - `docker pull jupyter/scipy-notebook`
  - `Docker pull tylerfowler/superset`
  - `Docker pull postgres`
  - `Docker pull minio/minio`
  - `Docker pull python:3`
  - `Docker pull puckel/docker-airflow`



**ABInBev**





# Quien es este we?

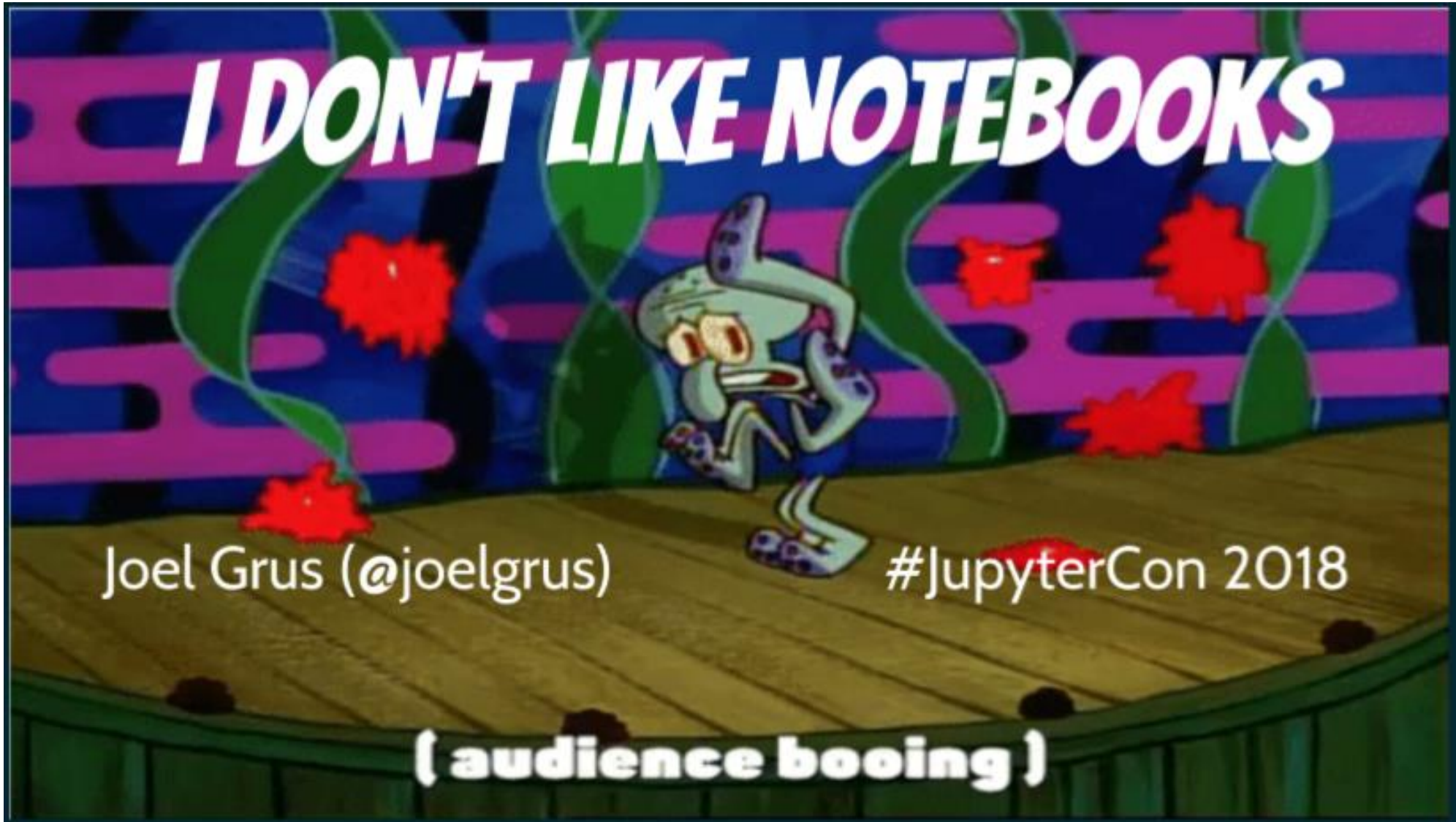
**ABInBev**

**But why?**









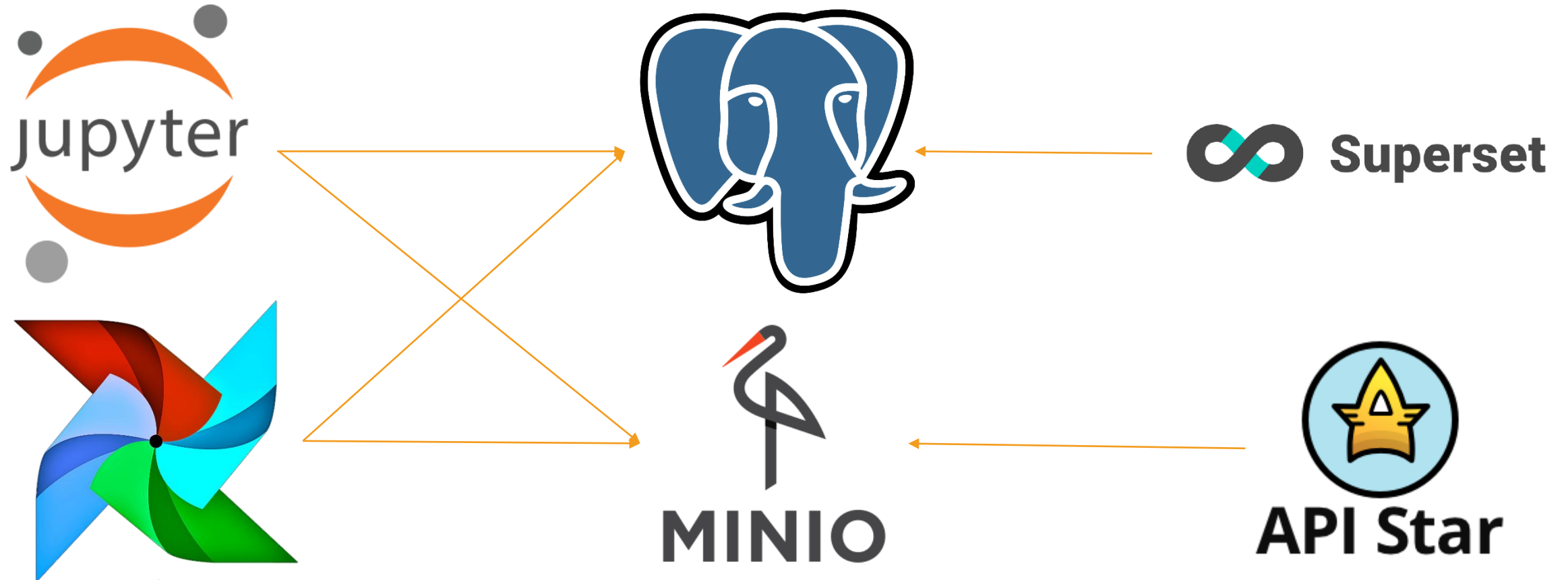
# Retos actuales en el área

- Visualizaciones Estáticas
- Archivos aislados
- Dificultad de Reentrenamiento
- Gran reto para compartir modelos



# Metas

Plataforma de Data Science 100% Open Source



# Arquitectura



# Jupyter

version: "3"

services:

jupyter:

container\_name: "tl\_jupyter"

image: jupyter/scipy-notebook

ports:

- 8888:8888

Live Demo: <http://18.210.180.148:8888>

## Retos actuales en el área

### • ~~Visualizaciones Estáticas~~

- Archivos aislados
- Dificultad de Reentrenamiento
- Gran reto para compartir modelos



# Arquitetura



# Superset

superset:

container\_name: "tl\_superset"

image: tylerfowler/superset

ports:

- 8088:8088

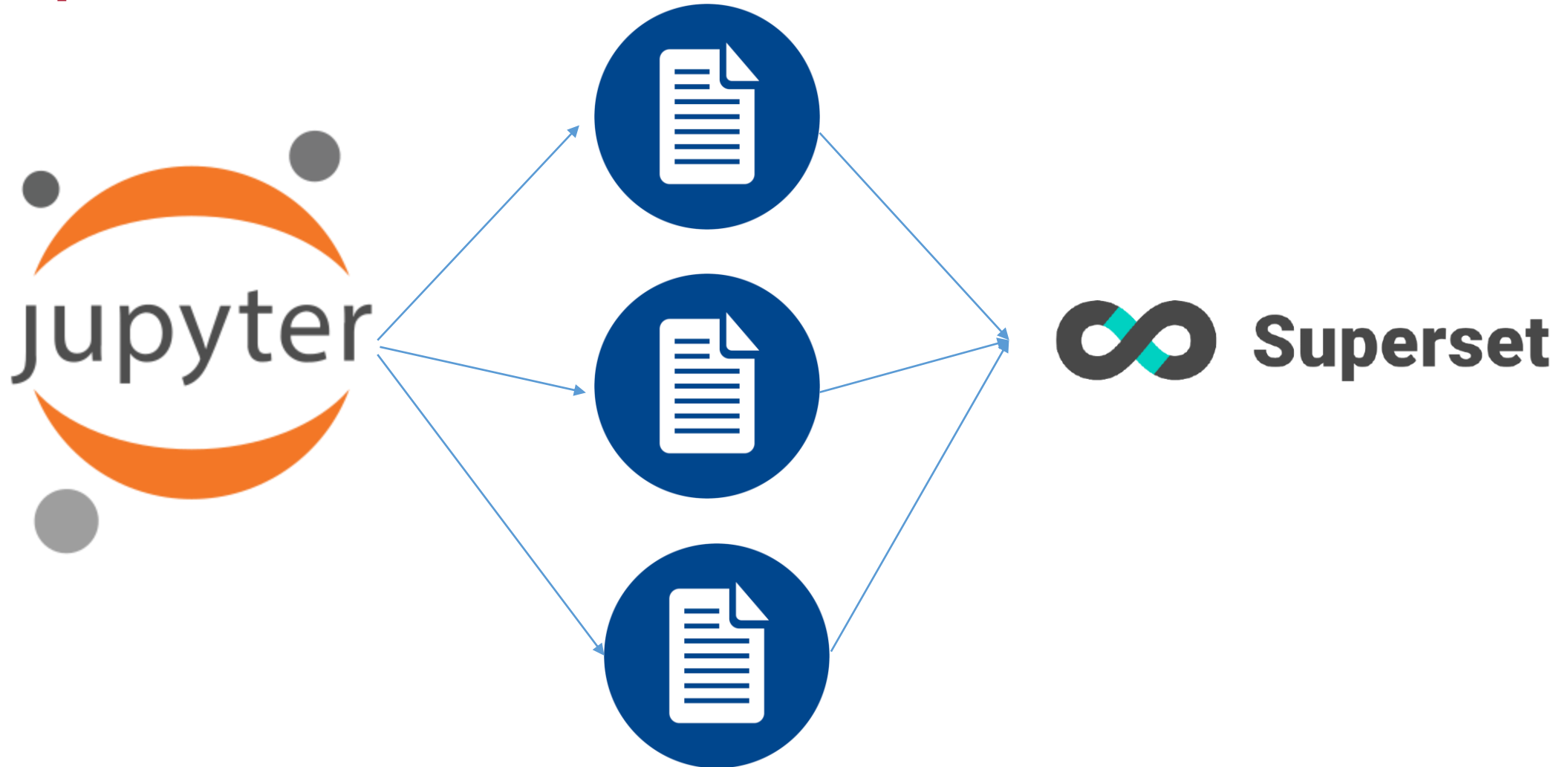
# Superset

Live Demo: <http://18.210.180.148:8088>

## Retos actuales en el área

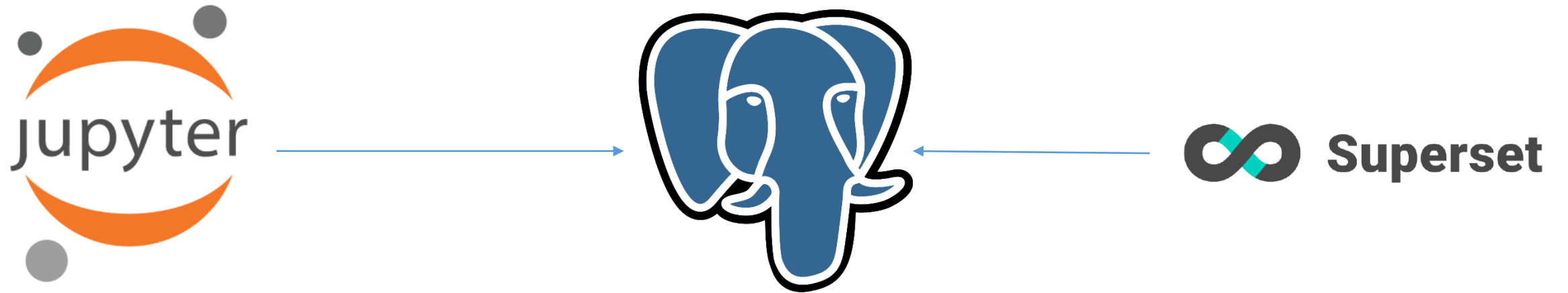
- ~~Visualizaciones Estáticas~~
- ~~Archivos aislados~~
- Dificultad de Reentrenamiento
- Gran reto para compartir modelos

# Arquitetura





# Arquitetura



# PostgreSQL

postgres:

container\_name: "tl\_postgres"

image: postgres

volumes:

- pg\_volume:/var/lib/postgresql/

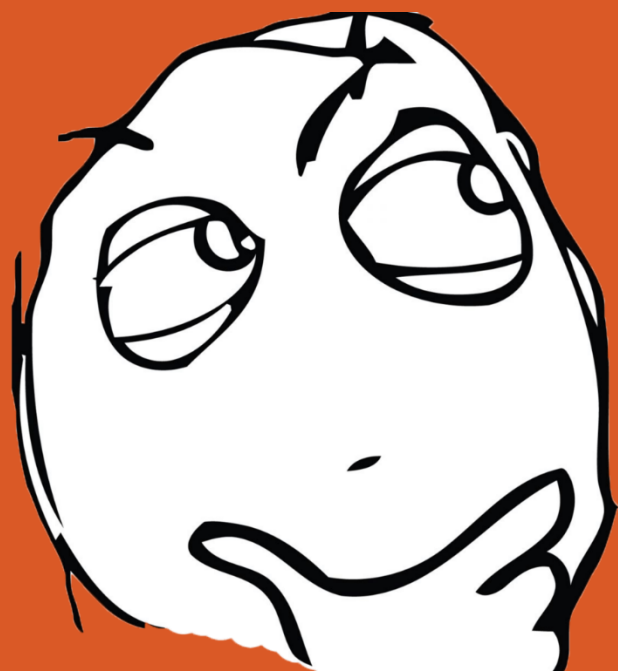
ports:

- 5432:5432

# PostgreSQL

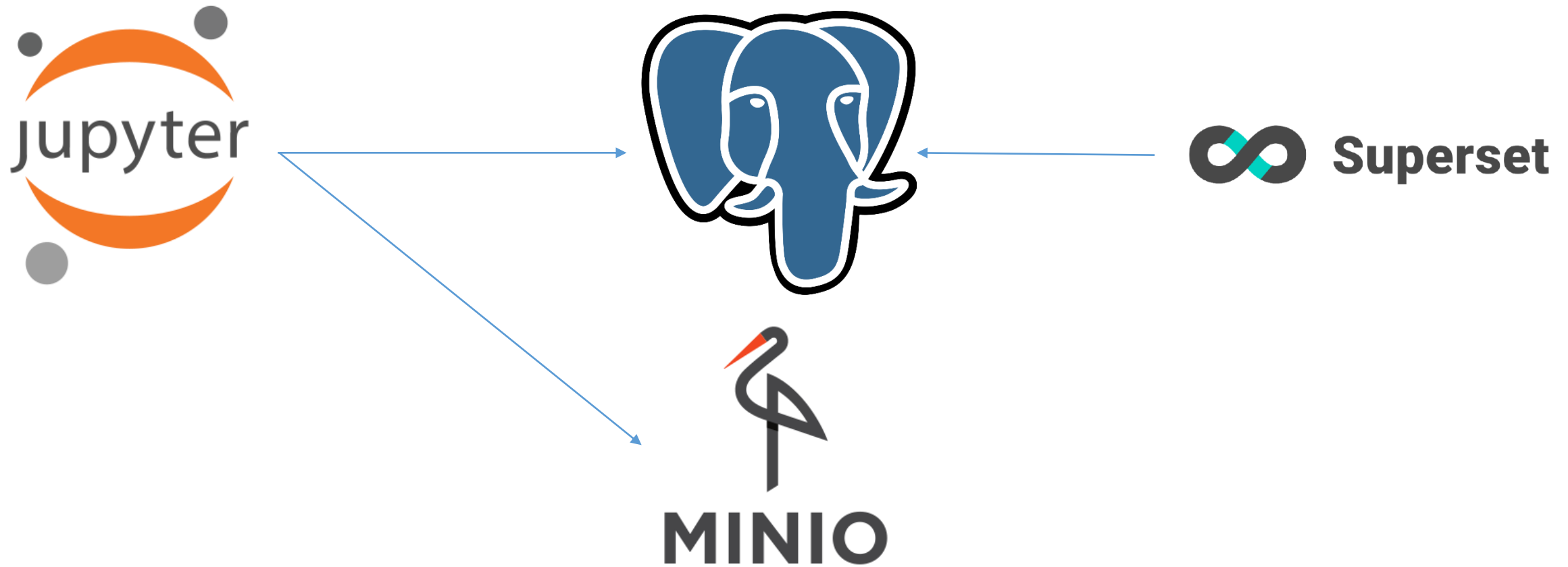
- postgresql://postgres:postgres@postgres

Live Demo: <http://18.210.180.148:8088>



# Y el modelo?

# Arquitetura





# Minio

minio:

container\_name: "tl\_minio"

image: minio/minio

environment:

MINIO\_ACCESS\_KEY: esto es una clave de acceso 123.

MINIO\_SECRET\_KEY: esto es una clave secreta 123.

volumes:

- minio\_volume:/data
- minio\_config:/root/.minio

ports:

- 9000:9000

command: server /data

Live Demo: <http://18.210.180.148:9000>

## Retos actuales en el área

- ~~Visualizaciones Estáticas~~
- ~~Archivos aislados~~
- ~~Gran reto para compartir modelos~~
- Dificultad de Reentrenamiento

# Predict-as-a-Service?



# Arquitectura



# API Star



# Apistar

apistar:

container\_name: "tl\_apistar"

build: services/apistar

volumes:

- ./services/apistar/api:/usr/src/app

ports:

- 8000:8000

command: gunicorn app:app -b 0.0.0.0:8000

# Apistar

Dockerfile:

FROM python:3

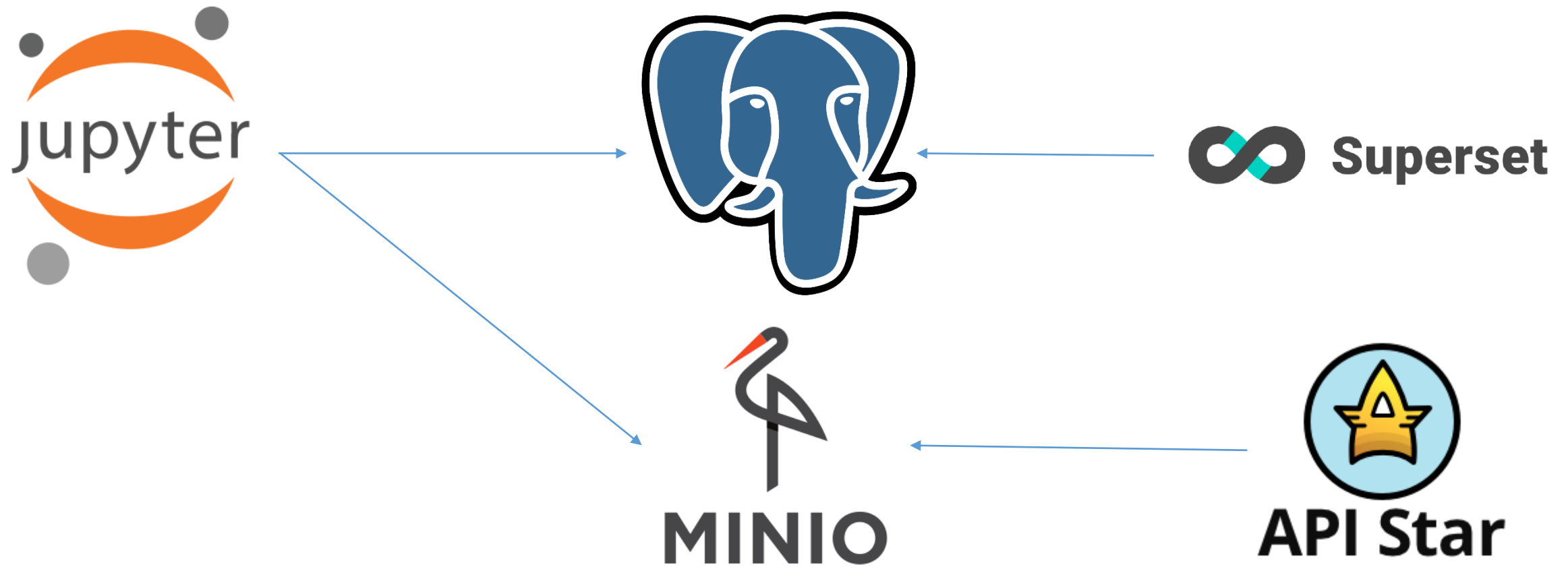
WORKDIR /usr/src/app

RUN pip install --no-cache-dir apistar==0.5.41 gunicorn

EXPOSE 8000

Live Demo: <http://18.210.180.148:8000>

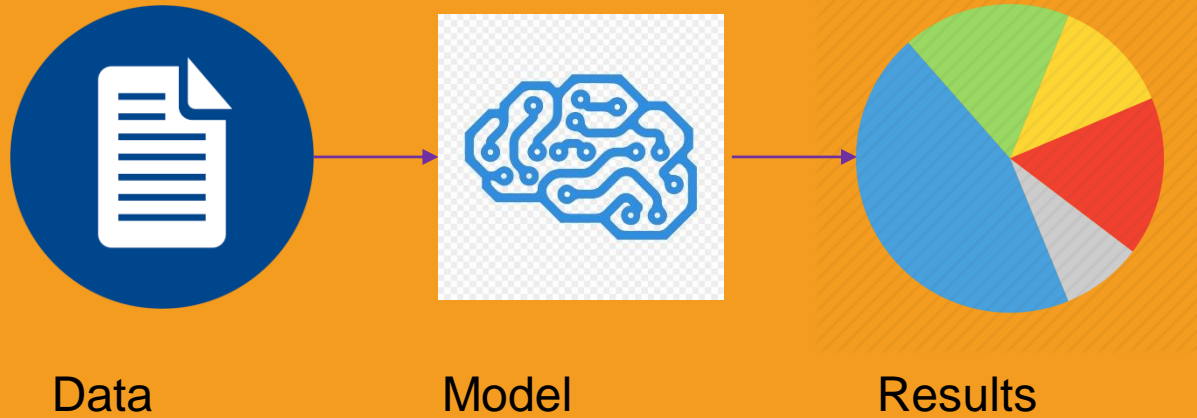
# Arquitetura



## Retos actuales en el área

- ~~Visualizaciones Estáticas~~
- ~~Archivos aislados~~
- ~~Gran reto para compartir modelos~~
- ~~Dificultad de Reentrenamiento~~

# Modelo Ideal



# Modelo Inicial



# Modelo Maduro

# Airflow





# Airflow

airflow:

container\_name: "tl\_airflow"

image: puckel/docker-airflow

depends\_on:

- postgres

ports:

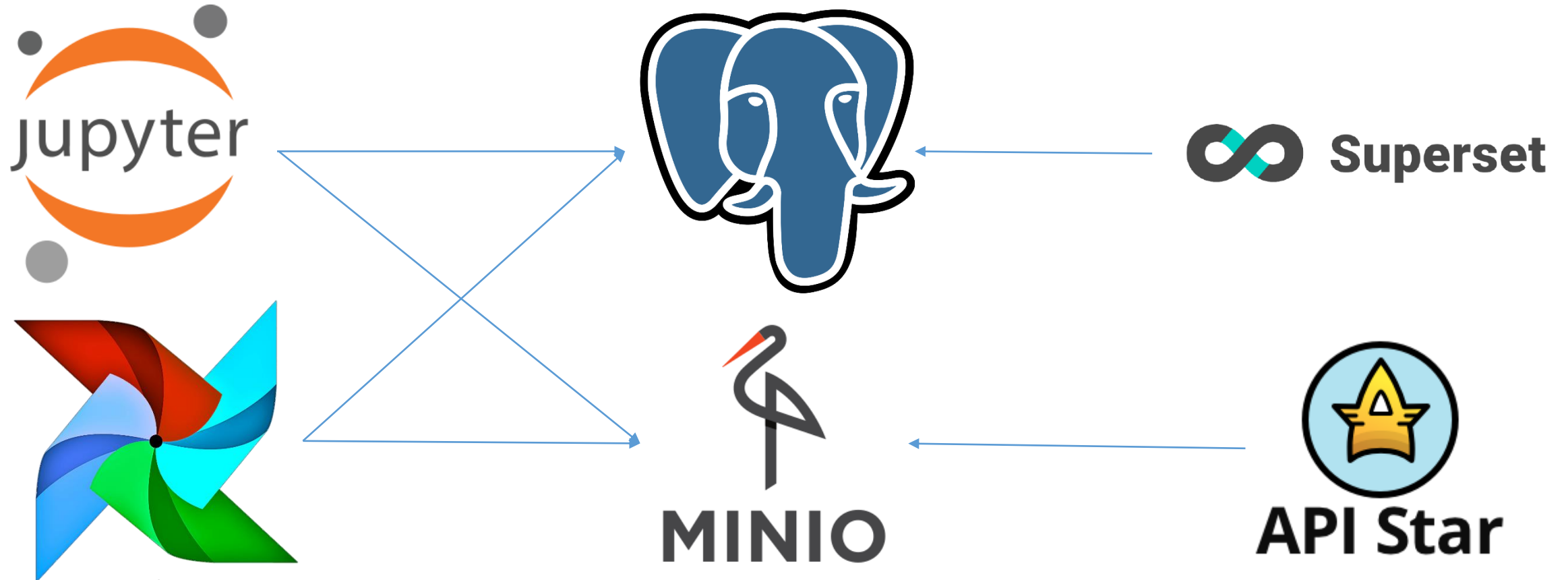
- 8080:8080

command: webserver

Live Demo: <http://18.210.180.148:8080>

# Arquitetura

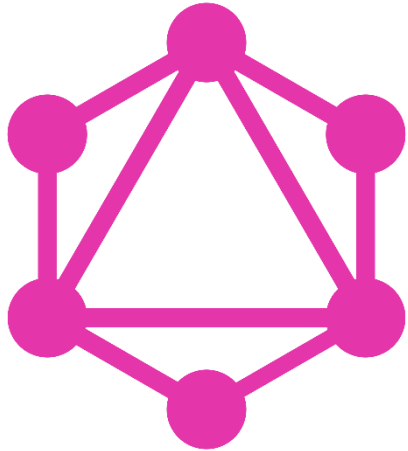
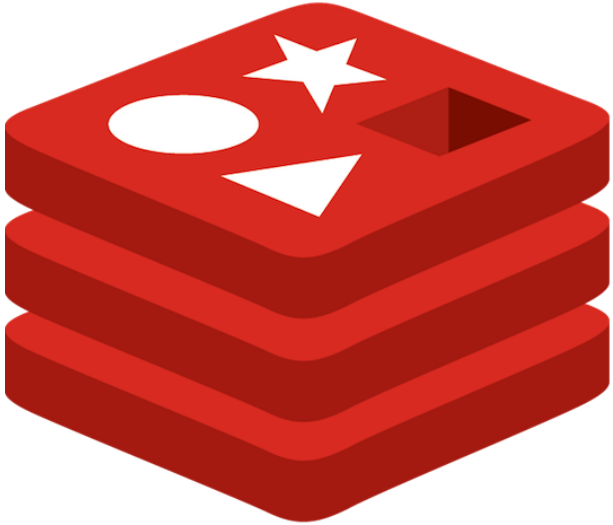
Plataforma de Data Science 100% Open Source



## Metas en el área

- Portabilidad
- Mantenibilidad
- Flexibilidad

Extiéndelo a tu gusto!



# Demasiada info!!!





# Muchas gracias!

Fuga por unas frias!







# Q&A





**Alejandro H. Pineda**

**mail: [ale.hpineda@gmail.com](mailto:ale.hpineda@gmail.com)**

**web: <https://alehpineda.github.io/cv>**