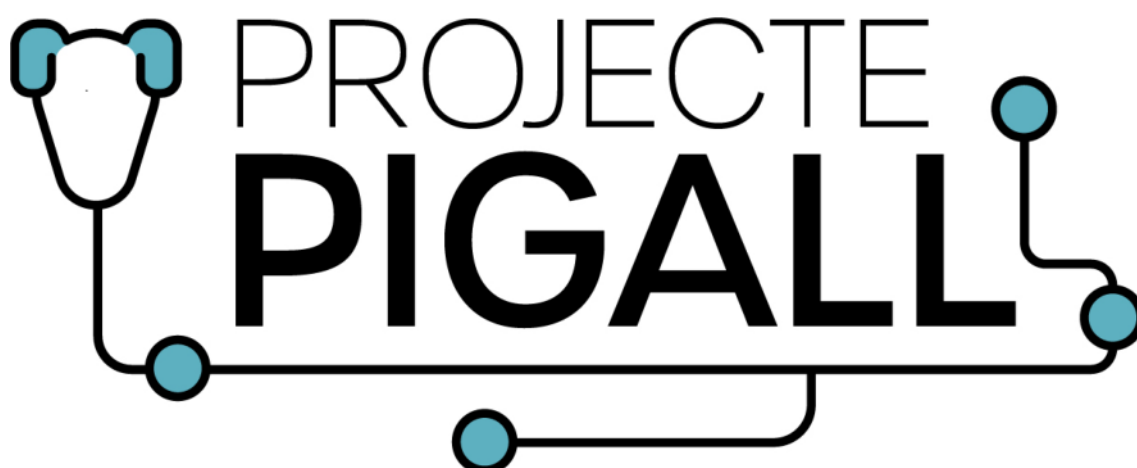


MANUAL D'INFRAESTRUCTURA TÈCNICA



1. Introducció

1.1 Sobre el projecte Pigall

El projecte Pigall té com a objectiu la construcció d'una eina basada en Intel·ligència Artificial que permeti entendre les dinàmiques socials de la ciutat de Girona i anticipar possibles escenaris de risc futurs. L'eina Pigall ha de facilitar la presa de decisions per a una millora de la planificació i la intervenció social, basant-se en les evidències que proporcionen les dades. També ha de permetre una optimització dels recursos disponibles.

L'objectiu és, doncs, la construcció d'una eina amb les següents finalitats:

1. Analitzar la realitat social de la ciutat, proporcionant tota la informació necessària que permeti identificar problemàtiques.
2. Ajudar en la planificació, el disseny d'estratègies i l'assignació dels recursos disponibles.
3. Permetre fer prediccions i anticipar-se a escenaris de risc futurs.
4. Automatitzar algunes tasques mecàniques realitzades pels tècnics de Serveis Socials que no requereixin prendre decisions.

1.2 Funcionalitats de l'eina Pigall

Aquestes finalitats que ha de tenir el projecte, a la pràctica, s'han traduït en la construcció d'una eina que ofereix les següents funcionalitats:

1. **Quadre de comandament**, que permet visualitzar dades a través de gràfics, mapes de calor, etc, així com fer consultes directament a la base de dades i generar anàlisis estadístics automatitzats.
2. **Quadre d'avisos**, que notifica automàticament d'anomalies detectades a la ciutat, o de l'aparició de possibles escenaris de risc en el futur proper.
3. **Escenaris de futur**, una eina de simulació que permet fer prediccions a mitja i llarg termini sobre com poden evolucionar diversos aspectes de la ciutat.
4. **Eines de consulta**, una miscel·lània d'eines per facilitar algunes tasques de gestió i administratives.

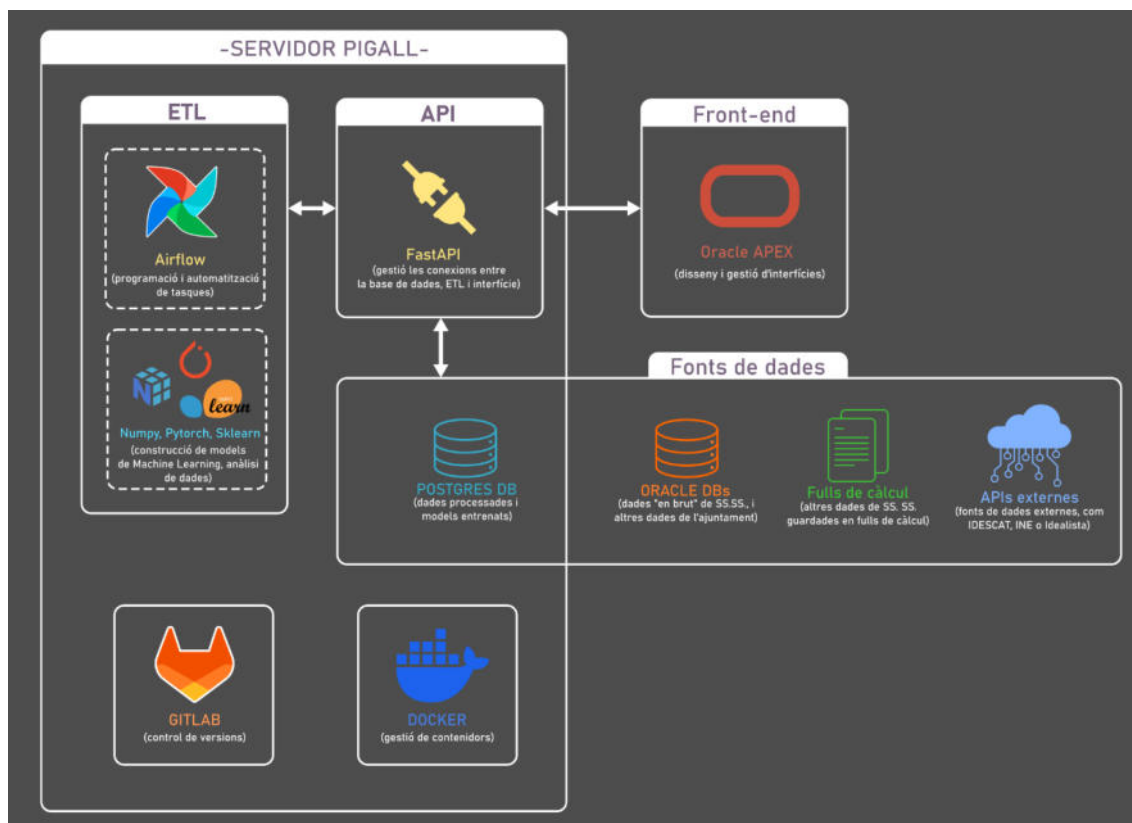
1.3 Sobre la replicabilitat

La metodologia seguida pel Pigall i que s'explica al llarg del dossier metodològic és replicable a altres àrees i administracions. No obstant, a la pràctica, si es vol replicar el Pigall s'haurà d'adaptar la infraestructura a la disponible a l'entitat que vol replicar.

També caldrà fer un tractament de dades adaptat a les que hi hagi disponibles, atès que cada àrea, àmbit i institució pot registrar-les de manera diferent. Aquest fet pot fer variar tant la qualitat de les dades com la seva completesa.

2. Infraestructura

En aquesta secció s'explica quina infraestructura s'ha utilitzat per desenvolupar el projecte Pigall. La següent figura mostra un diagrama del funcionament de la infraestructura de l'eina Pigall:



1. Diagrama de funcionament del projecte Pigall

2.1 Servidor Pigall

L'Eina Pigall s'executa dins d'un servidor Linux que està funcional les 24h del dia. Dins d'aquest servidor hi ha instal·lats:

- **Gitlab**: sistema de control de versions per tal de poder facilitar als tecnòlegs desenvolupar el projecte de forma col·laborativa.
- **Airflow**: sistema per agendar i automatitzar tasques. Permet, per exemple, establir cada quan cal re-entrenar els models d'aprenentatge automàtic, o anar poblant les bases de dades del Pigall amb dades agregades.
- **FastAPI**: proporciona un conjunt d'eines per construir APIs i gestionar crides.
- **Docker**: permet encapsular i executar per separat els diferents programes i scripts de l'aplicació, permetent-los funcionar de manera independent i modular.
- **PostgreSQL**: gestor de dades relacionals. Part de les dades processades, així com els models d'aprenentatge automàtic entrenats, es guarden aquí per al seu posterior ús.

Els scripts que permeten el funcionament de l'eina Pigall estan desenvolupats en llenguatge Python, i es divideixen en dos grups, cada un amb el seu propi repositori.

- **Scripts ETL:** gestionen les tasques computacionalment més intensives de l'eina, com ara el preprocessament de dades, nodreixen la base de dades, entrenen els models, etc.
- **Scripts API:** fan prediccions amb els models ja entrenats o presenten dades ja prèviament processades.

2.2 Arquitectura tecnològica

– Quadre de comandament

El quadre de comandament es nodreix de les dades generades a les bases de dades **PostgreSQL** i **Oracle**. L'usuari disposa de formularis per fer sol·licituds de dades específiques segons criteris especificats. Aquestes peticions no accedeixen directament a la base de dades, sinó que es canalitzen mitjançant una capa intermèdia de serveis.

Aquesta comunicació es gestiona mitjançant codi **PL/SQL** o **REST Data Sources** d'**Oracle APEX**, consumint l'API desenvolupada per l'equip Pigall. Aquesta API permet estructurar, filtrar i transformar la informació abans de mostrar-la al frontend.

– Avisos

Els avisos es computen amb diferents granularitats temporals mitjançant **processos ETL** específics orquestrats per **Apache Airflow**. Aquests scripts ETL s'encarreguen de consolidar l'informació i aplicar transformacions necessàries a les dades, generant resultats estructurats segons les necessitats de l'eina. Tota aquesta informació processada s'emmagatzema dins de taules específiques de **PostgreSQL** dissenyades per optimitzar la consulta i el rendiment en entorns d'exploració analítica.

Finalment, el front-end es construeix amb **Oracle Apex** i consumeix aquestes dades a través de consultes i serveis intermedis.

– Escenaris de futur

Els escenaris de futur es construeixen a partir de formularis web interactius que permeten ajustar els paràmetres dels diferents simuladors. Un cop l'usuari ha completat un formulari, la informació introduïda es processa i s'envia a backend o bé mitjançant **codi PL/SQL** o bé a través de **REST Data Sources** propis d'**Oracle Apex**. En ambdós casos, la comunicació es realitza usant una API desenvolupada especialment per a aquest propòsit.

– Eines de consulta

Les eines de consulta tenen una arquitectura equivalent a la del quadre de comandament, amb la diferència que les consultes que fa l'eina serveixen per poblar informes que després poden ser descarregats pels usuaris.

2.3 Scripts ETL

Com hem dit, els scripts ETL (acrònim per Extract, Transform, Load en anglès) són aquells programes i algorismes que s'ocupen de realitzar els càlculs més costosos del Pigall. Aquests scripts es divideixen en tres grups:

- **Actualització de taules:** són aquells scripts que nodreixen la base de dades PostgreSQL amb noves dades processades, per exemple preu de l'IPC actualitzat, altes i baixes del

padró, etc, perquè puguin ser consultades pel quadre de comandament i les eines de consulta.

- **Entrenament de models:** són aquells scripts que s'ocupen de crear, entrenar (i re-entrenar) els models d'aprenentatge automàtic que s'utilitzen després per generar avisos i predir escenaris de futur.
- **Creació d'avisos:** analitzen les dades, en ocasions fent ús dels models d'aprenentatge automàtic prèviament entrenats, i detecten situacions de risc. També poblen la taula d'avisos.

2.4 Scripts API

Els scripts de la API s'ocupen de:

- **Servir dades en temps real** per mostrar a la interfície d'usuari, per exemple en el quadre de comandament.
- **Fer crides a models d'aprenentatge automàtic** per tal de fer les simulacions dels escenaris de futur.
- **Generar informes automàticament** que poden ser descarregats a través del quadre de comandament, els avisos, o les eines de consulta (depenent de la tipologia d'informe).