System Development
Architecture

Sergi Blanco-Cuaresma and the ADS Team

ADS Users Group Meeting, 9-10 Nov. 2022
Back-Office Overview

Classic Back-office

Import pipeline

Data pipeline

Master pipeline

Apache Solr

Bibliographic database

PostgreSQL
Python-based pipelines overview

- Import pipeline: Bibliographic data
- Data pipeline: Metrics
- Full-text pipeline
- Augment pipeline: Affiliations
- Master pipeline
- ORCID pipeline
- myADS pipeline
- Citation Capture: Software records

Language Model:
- Félix
- Tom

NER Efforts:
- Golnaz
- Tom

Apache Solr
Bibliographic database
PostgreSQL
Microservices Overview

- ADSWS
- Harbour
- Resolver gateway
- Author affiliation
- Metrics
- Solr service
- Biblib
- Object
- Tugboat
- Citation helper
- Oracle
- Vault
- Export
- Orcid
- Core
- Graphics
- Resolver service
- Reference

User interface (Browser)

Tim, Jennifer

kubernetes
Future Event-driven Architecture
Summary

- **Work completed**
  - Designed an incremental plan to migrate to a new architecture
  - Progress in multiple fronts to replace old components such as:
    - Refactored UI / Scan Explorer
    - Document matching / Reference extraction / …

- **Work ahead**
  - Experiment with new technologies for the new architecture
    - gRPC
    - Kafka
  - Continue replacing classic pieces and plug them into the new architecture
  - Replace main obsolete API component (adsws)