LifeWatch ERIC e-Science for NIS research workshop | 20-21 May 2021

An event organised within the framework of the LifeWatch ERIC Internal Joint Initiative and with the support of the ENVRI-FAIR project

Session 1: Long-term monitoring of hard-bottom marine communities
ARMS Workflow - development, current status, functionalities

Antonio José Sáenz-Albanés | LW-ERIC ICT-Core Operations Coordinator

1. Forewords about ICT presentations

Two sections in each ICT presentation

Associated Workflow

Explain technical aspects related to the current deployment

Tesseract Architecture

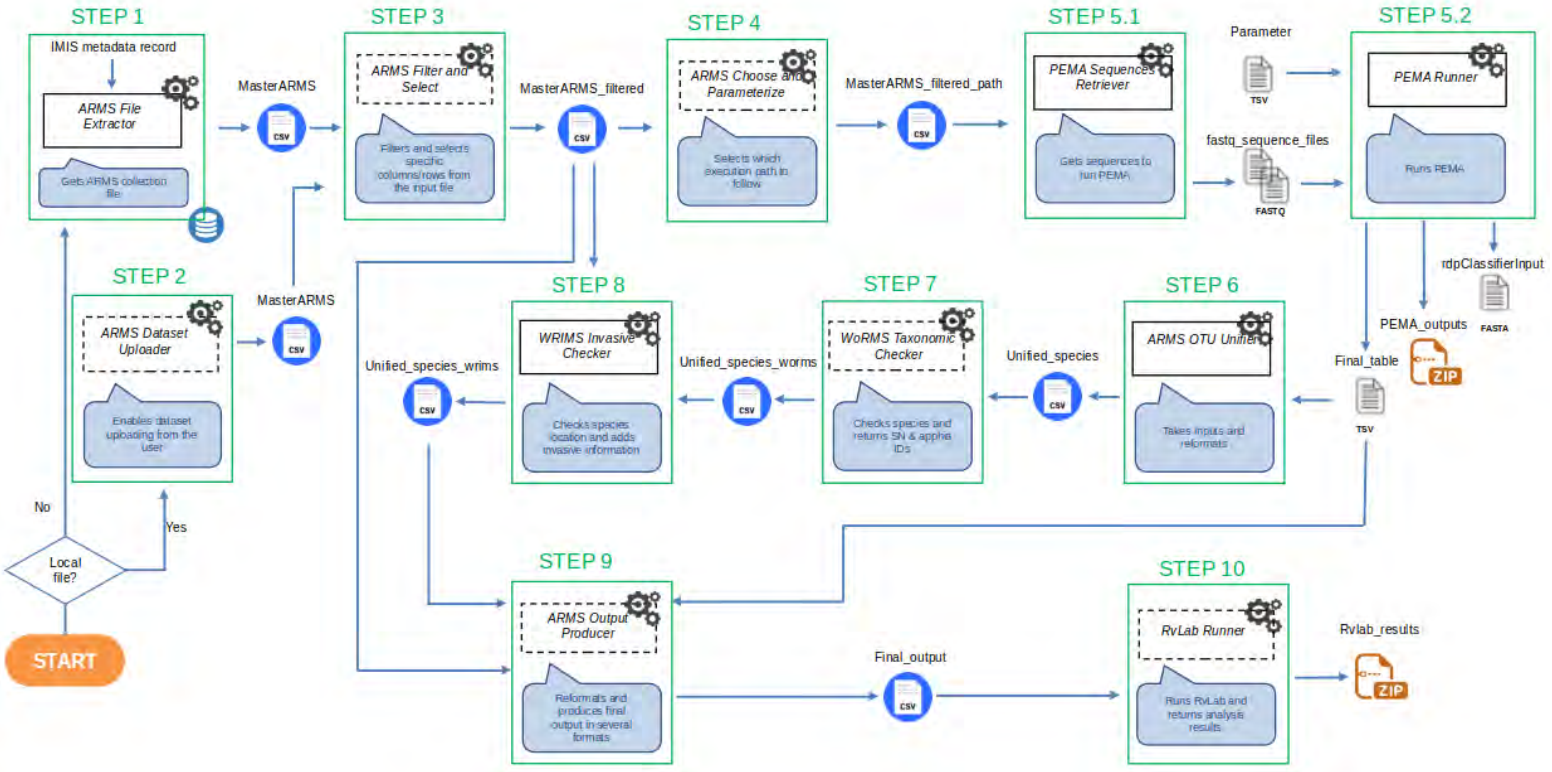
Explain technical aspects of the architecture **tested** in the development of the workflows

2. Current deployment of ARMS workflow

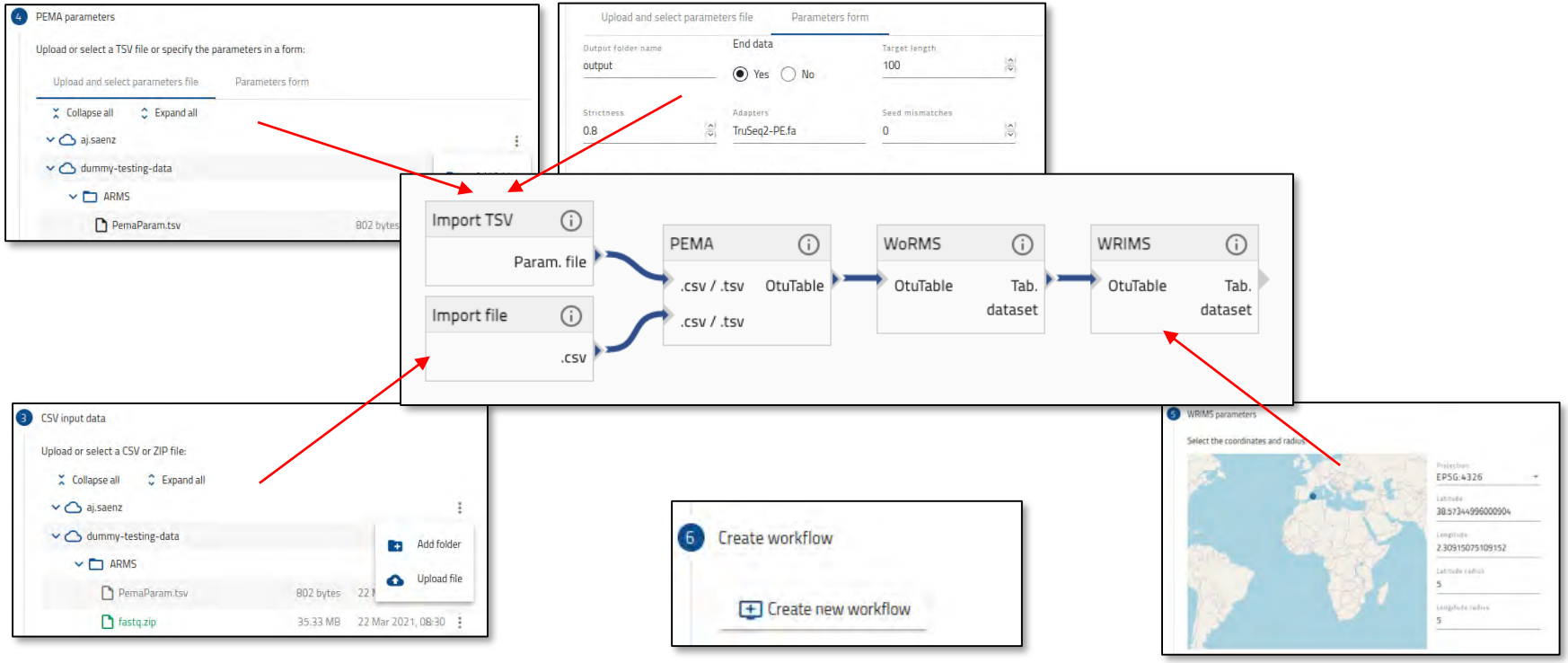
Associated Workflow

Explain technical aspects related to the current deployment

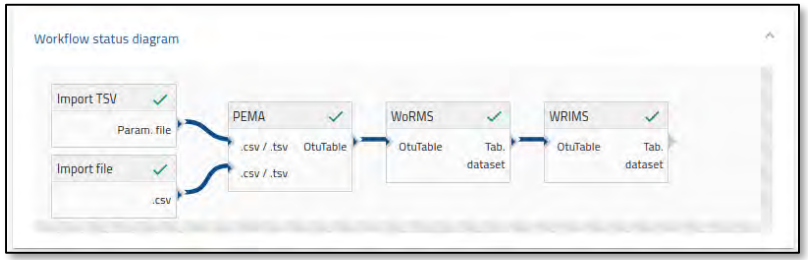
ARMS Validation Case

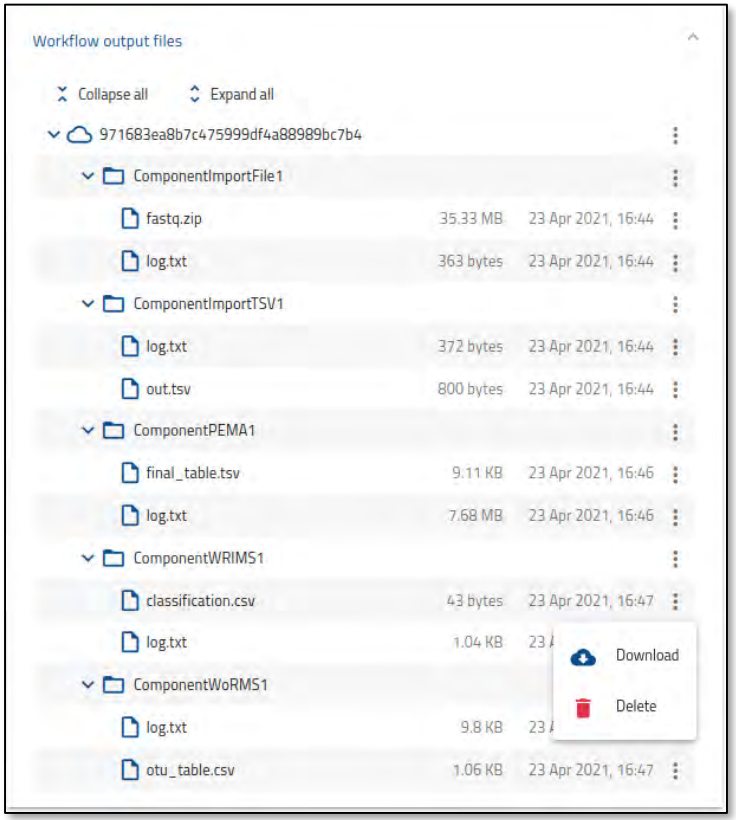


ARMS UI Current Implementation

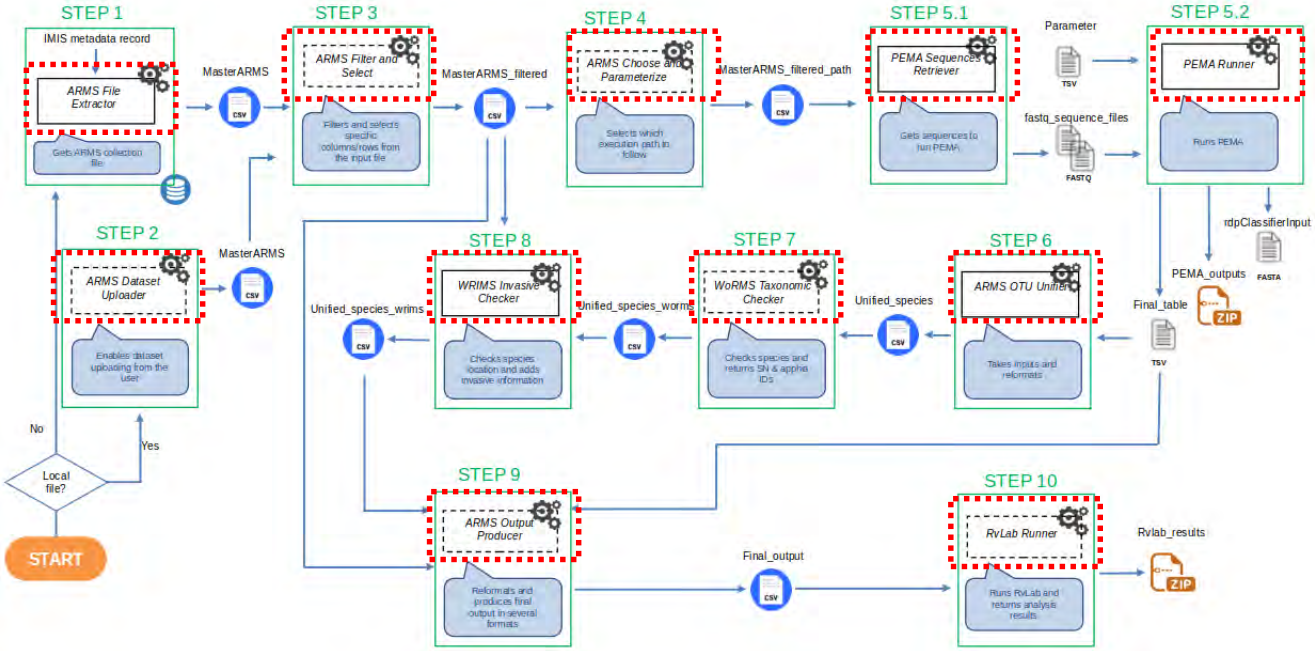


ARMS UI Current Implementation





ARMS Wrappers Implementation and Generalization



A “wrapper” provides two aspects:

- An interface to the workflow engine
- An execution environment and implementation

One of our goals is to generalize them so that we can reuse

ARMS File Extractor

Label: ARMS File Extractor

Description: Retrieves MasterARMS.csv

Input: ARMS URL

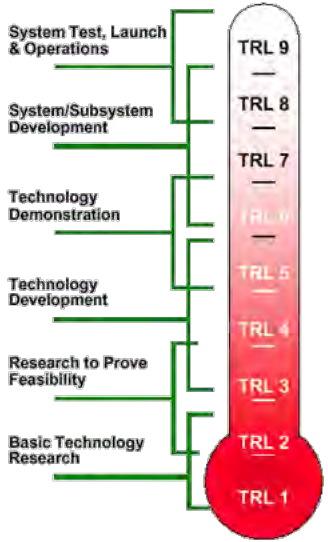
Output: MasterARMS (csv)

Metadata Harmonized: YES

TRL: 5

GUI Integration: YES

Technology: R / Python



Technology Readiness Level

PEMA Sequences Retriever

Label: PEMA Sequences Retriever

Description: Retrieves Sequences fasta files for PEMA

Input: MasterARMS (csv)

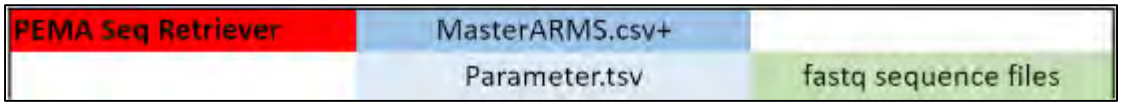
Output: fastq_sequence_files (fastq)

Metadata Harmonized: YES

TRL: 4

GUI Integration: YES

Technology: R / Python



PEMA Runner

Label: PEMA Runner

Description: Executes PEMA code and produces results

Input: fastq_sequence_files (fastq)

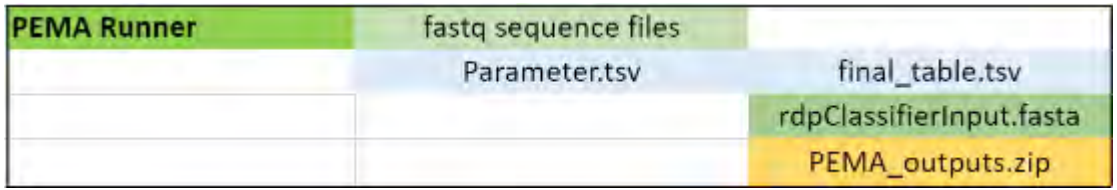
Output: Final_table (tsv), rdpClassifierInput (fasta), PEMA_outputs (zip)

Metadata Harmonized: YES

TRL: 5

GUI Integration: YES

Technology: Docker / R / Python



ARMS OTU Unifier

Label: ARMS File Extractor

Description: Takes input CSV, TSV and XLS and unifies them

Input: Final_table (tsv)

Output: Unified_species (csv)

Metadata Harmonized: YES

TRL: 2

GUI Integration: NO

Technology: R / Python

WoRMS Taxon Check

Label: WoRMS Taxon Check

Description: Species names checked by WoRMS

Input: Unified_species (csv)

Output: Unified_species_worms (csv)

Metadata Harmonized: YES

TRL: 3

GUI Integration: YES

Technology: R / Python

WRIMS Invasive Check

Label: WRIMS Invasive Check

Description: Checks Species location and returns invasive

Input: Unified_species_worms (csv), MasterARMS (csv)

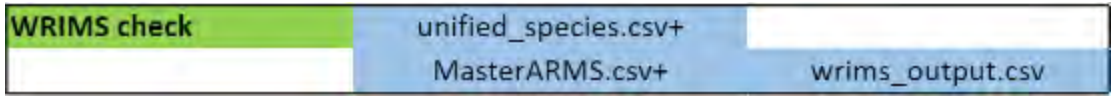
Output: MasterARMS (csv)

Metadata Harmonized: YES

TRL: 2

GUI Integration: NO

Technology: R / Python



Produce output

Label: Produce output

Description: Produces final output in various formats

Input: MasterARMS_filtered (csv),
Unified_species_wrims (csv),
Final_table (tsv)

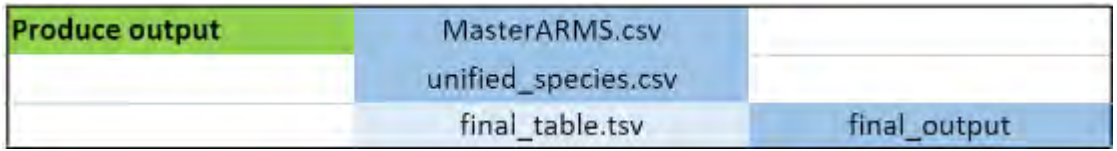
Output: Final_output (csv)

Metadata Harmonized: YES

TRL: 3

GUI Integration: YES

Technology: R / Python



Run RvLab

Label: Run RvLab

Description: Runs RvLab and returns analysis results

Input: Final_output (csv)

Output: Rvlab_results.zip

Metadata Harmonized: YES

TRL: 3

GUI Integration: NO

Technology: Docker / R / Python

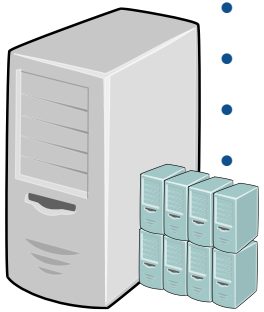
3. Security and Distributed Heterogeneous Executions

Tesseract Architecture

Explain technical aspects of the architecture **tested** in the development of the workflows

User Browser

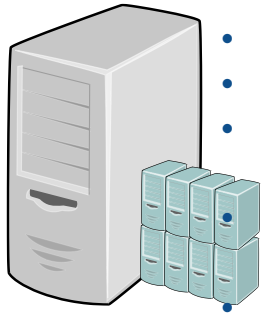
Partial Restricted Access



Services:
- Single Sign On
- Workflow UI
- Authorization
- Monitorization

Advanced Perimeter Security

Secured Channel



Cloud Services:
- Workflow Engine
- Distributed Storage
- Native Execution
- Distributed Execution
- Remote Service Invocation
- Monitorization

Secured Channel

Cloud-HPC Gateway

HPC

Thanks!

aj.saenz@lifewatch.eu