

## Deliverable D4.1

Notebooks and algorithms repository (accessible from the AI4Life website)

Project Title	Artificial Intelligence for Image Data Analysis In The Life Sciences
Project Acronym	AI4Life
Project Number	101057970
Project Start Date	01.09.2022
Project Duration	36 Months

WP N° & Title	WP4: Contributor Services
WP Leaders	IGC
Deliverable Lead Beneficiary	Instituto Gulbenkian de Ciência
Dissemination Level	PU
DOI	10.5281/zenodo.7987775
Contractual Delivery Date	31.05.2023 (M9)
Actual Delivery Date	30.01.2023
Authors	Ricardo Henriques
Contributors	Iván Hidalgo, Estibaliz Gómez de Mariscal
Reviewers	Arrate Muñoz Barrutia, Florian Jug, Rachel Robinson-Lehtinen, Wei Ouyang



## Change Log

Version	Date	Description of changes
v0.1	20.04.2023	Initial draft
v0.2	25.05.2023	Draft submitted for internal review
v0.3	30.05.2023	Approved for submission

## Acronyms and Abbreviations

AI	Artificial Intelligence
D	Deliverable
ML	Machine Learning
PU	Public
RI	Research Infrastructure
v	Version
WP	Work Package

## Table of contents

<b>Deliverables</b>	<b>1</b>
<b>Acronyms and Abbreviations</b>	<b>2</b>
1. Introduction	3
2. Description of work	4
2.1 Online distribution and accessibility to notebooks and algorithms.	4
2.2 Software tools for the interaction with deep learning models	4
These libraries are synchronised with the model standards defined in the BioImage Model Zoo. Therefore, they provide an easy path for new software and method developers to quickly adopt such standards and disseminate their work among the life sciences community.	4
2.3. Open source and user-friendly notebook platform to support the interaction with AI4Life.	4
3. Conclusion	5

## 1. Introduction

The AI4Life project is part of the European Union's Horizon Europe research and innovation programme, led by the project coordinator Euro-BioImaging and participated by ten partners, four of them being European Research Infrastructures themselves. The project started in September 2022 and will continue until September 2025.

AI4Life aims at bringing state-of-the-art AI-based image analysis to life scientists by establishing and supporting innovative services that target both researchers in the life sciences and computational methods developers in the AI and computer vision fields.

Work Package 4 (WP4) focuses on 'Contributor Services' and contributes to Objectives 1–3 and 5 of AI4Life:

- Objective 1: Democratised availability of AI-based image analysis methods as a FAIR service accessible through the AI4Life service landscape and computationally powered by the European Open Science Cloud (EOSC) infrastructure.
- Objective 2: Establish standards for the submission, storage and FAIR access of reference data, reference annotations (ground-truth), trained AI models, and trainable AI methods.
- Objective 3: Simple model deployment, sharing, and dissemination of AI-based methods as a new developer-facing service of the [BioImage Model Zoo \(BMZ\)](#).
- Objective 5: Empower common image analysis platforms with AI tools.

To do so, four different tasks were defined:

- Task 4.1 Establish containerization structure and services.
- Task 4.2 Easy development and deployment of new methods
- Task 4.3 Enable Human-in-loop BioImage Model Zoo workflow.
- Task 4.4 Continuous integration services (Automated testing).
- Task 4.5 Implement data sharing and protection guidelines.

This first deliverable (D4.1) focuses on Task 4.2 and 4.3 by providing a repository of notebooks and algorithms that enable the interaction (programmatic and in a user-friendly manner) with the BioImage Model Zoo content (the central repository to interconnect the tools related to the AI models and methods).

In this document, we describe this repository and how users can browse and exploit it.

## 2. Description of work

### 2.1 Online distribution and accessibility to notebooks and algorithms.

The Biolmage Model Zoo is the central point for interaction with AI-trained models and methods. It establishes a user-friendly space to search and launch models, methods, notebooks and even find datasets to train new models. Thus, the new tools developed within WP4 are directly distributed through the Biolmage Model Zoo, and the algorithms live in the Biolmage.IO GitHub community.

The Biolmage Model Zoo is directly accessible through the AI4Life webpage. Also, the ZeroCostDL4Mic platform is accessible through both the AI4Life website and the Biolmage Model Zoo.

### 2.2 Software tools for the interaction with deep learning models

To allow developers to deploy, consume, export and import [Biolmage Model Zoo](#) models directly or within their tools, two different packages are available:

- [Biolmage.IO Core library](#) (Python package).
- [Java deep learning library \(JDLL\)](#) (Java package).

These libraries provide a common backend for the current (i.e., ZeroCostDL4Mic, ilastik, deeplmagej, ImJoy and deeplcy) and future community partners to easily adapt the standards defined by the Biolmage Model Zoo. Moreover, they establish the assessment and testing of the current continuous integrating system (Task 4.4) of new models, which ensures the correct functionality of models and methods in the user-friendly software tools compatible with the Biolmage Model Zoo.

These libraries are synchronised with the model standards defined in the Biolmage Model Zoo. Therefore, they provide an easy path for new software and method developers to quickly adopt such standards and disseminate their work among the life sciences community.

### 2.3. Open source and user-friendly notebook platform to support the interaction with AI4Life.

The Biolmage.IO Core library provides a series of [example notebooks](#) to guide newcomers into the deployment, consumption, export and import of the models in the Biolmage Model Zoo.

We also took advantage of the infrastructure and collection of Jupyter notebooks already within ZeroCostDL4Mic to stimulate the contribution of new methods and models that can be connected with the Biolmage Model Zoo (Task 4.2). ZeroCostDL4Mic is an entry-level platform (located on a GitHub [repository](#)) with a collection of Jupyter Notebooks that follow an easy-to-use graphical user interface. Its main goal is to develop these Jupyter Notebooks in a self-explanatory manner so that users with little or no coding expertise are able to use deep learning networks with microscopy data quickly.

The most popular notebooks within the platform (i.e., 2D, 2D multilabel and 3D U-Nets, and StarDist 2D) enable the consumption of pre-trained models in the Biolmage Model Zoo and the export of new models trained with different datasets.; So the users can quickly populate the model repository and become compatible with existing user-friendly tools (e.g., ilastik, deeplmagej) (Task 4.3). This connection is enabled by the BiolmageIO.Core Python library, described in the next section.

These exemplar notebooks constitute a template and standards for developers to create new user-friendly notebooks (Task 4.2) that can be directly connected with the Biolmage Model Zoo. These notebooks can also be distributed directly through the Biolmage Model Zoo or the ZeroCostDL4Mic platform, following the [guideline](#) online.

### 3. Conclusion

The collection of notebooks within D4.1 and the ready-to-use libraries constitute a bridge between computer scientists developing AI methods and the FAIR ecosystem that we are building in AI4Life. These resources give developers and new contributors a chance to easily participate in this ecosystem, which is necessary to boost the integration of AI technologies into life-sciences research.



*AI4Life has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement number 101057970.*

