Imaging data and services for aquatic science

Reporting

Project Information

iMagine

Grant agreement ID: 101058625

DOI
10.3030/101058625 ↗

EC signature date
26 May 2022

Start date	**End date**
1 September 2022	31 August 2025
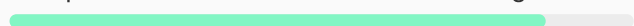
Funded under
Research infrastructures

Total cost
€ 4 500 000,00

EU contribution
€ 4 500 000,00

Investment in EU policy priorities

Digital agenda	○	Clean air	○
Artificial Intelligence	○	Climate action	◑
Biodiversity	○		

Coordinated by
STICHTING EGI
Netherlands

Periodic Reporting for period 2 - iMagine (Imaging data and services for aquatic science)

Reporting period: 2023-09-01 to 2024-08-31

Summary of the context and overall objectives of the project ⌄

iMagine offers AI-powered image analysis tools and marine and freshwater research datasets, promoting more efficient data processing. This framework, integrated into the European Open Science Cloud, supports AI model development, benefiting various applications such as water pollution mitigation, biodiversity studies and others.

The project leverages the EGI federation infrastructure, providing neural networks and distributed data analysis capabilities. Over 9 million images and 8 AI applications from 13 Research Infrastructures are shared via this framework, fostering Best Practices development. Synergies among aquatic use cases facilitate shared solutions in data management, quality control, and FAIRness, contributing to harmonization across RIs and Best Practice guidelines.

Work performed from the beginning of the project to the end of the period covered by the report and main results achieved so far ⌄

In the second year of iMagine, significant advancements were made in AI-powered image analytics for marine and freshwater research. Key achievements and activities are summarised:

Objective 1: Deliver a scalable, shared IT platform for image analysis in marine and freshwater research:
The iMagine AI Platform, initially built on the DEEP Platform, transitioned to AI4OS in its second year and saw adoption by all project and three external use cases. With four integrated cloud providers, platform usage surged, totaling 183,262 GPU hours, 4.2 million CPU hours, and 495 TB-months of storage, supporting over 60 users across 12 countries. Although GPU usage exceeded initial estimates, providers will continue support as focus shifts from model training to service delivery in year three.

Objective 2: Advance existing image analytical services to increase research performance in aquatic sciences:
WP3's five image analytics use cases advanced toward production-ready services, with efforts focused on model training, accuracy validation, and integration with the OSCAR inference platform. Competence Centre meetings and adoption of third-party models like YoloV8 facilitated this progress. The first public services (UC1 and UC5) launched in September 2024, now accessible through iMagine, have served 17 users in the initial two months.

Objective 3: Develop prototype new image analytical services and datasets that can accelerate progress towards healthy oceans, seas, coastal and inland waters:
WP3 introduced three innovative AI-based image processing use cases, focusing in the second year on image collection and annotation for model training. Each use case faced unique challenges due to varied data sources: audio clips (UC6), CCTV streams (UC7), and microscopic images (UC8). A total of 8 datasets, including UC7's two Zenodo datasets, are now available on the iMagine website, with over 6 million images published and plans for enhanced metadata for marine science accessibility.

Objective 4: Capture and disseminate development and operational best practices to imaging data

and image analysis service providers:
The project gathered good practices from platform providers and use case developers, publishing a comprehensive guide in D3.4 which has seen over 60 downloads and was shared at the EGI 2024 Conference. Collaboration with AI4Life led to a demonstrator for data interoperability, expanding multidisciplinary image analysis potential. Bilateral meetings with ANERIS, Blue-Cloud-2026, and ENVRI focused on technology transfer, sustainability, and service integration.

Objective 5: Deliver a portfolio of scientific image and image analytics services targeting researchers in marine and aquatic sciences:
Five mature use cases continued development in the project's second year, participating in bi-weekly Competence Centre sessions with AI experts for guidance and platform optimization. Two use cases —Marine Litter Assessment and Flowcam Phytoplankton Identification—reached production, with the remaining three expected to follow soon. Each service will offer trained models for image classification and FAIR images for model training, stored on Zenodo. Collaboration with the Zenodo team aims to establish a specialized metadata template to better document marine images according to FAIR principles.

Progress beyond the state of the art and expected potential impact (including the socio-economic impact and the wider societal implications of the project so far) ⌄

The iMagine AI platform saw a significant boost over the second year, with eight use cases from within the consortium and four external use cases relying on it. These use cases primarily focused on porting existing training images to the platform, training various models (or model configurations) with those images, and validating prediction accuracy. An iterative approach was used to identify the most accurate models and the most relevant image sets. In some cases, new images were collected to enhance model accuracy and minimize prediction bias. WP3 actively supported platform adoption, notably through biweekly "Competence Centre" video conferences that connected use case teams with AI and platform experts.

The project successfully onboarded four external use cases during the year, with three selected through the iMagine Open Call and one proposed by MARIS, the project's scientific coordinator. These external use cases complement the eight existing project use cases, enhancing the project's impact on aquatic sciences. Together, these use cases contributed 14 models to the iMagine AI Platform and published 11 datasets on Zenodo.

Among the five mature project use cases aiming for production delivery by the end of Year 2, two— UC1 (Marine litter assessment) and UC5 (Flowcam phytoplankton identification)—achieved this goal. They have since provided Virtual Access to 17 users across seven countries, including researchers affiliated with the LifeWatch and Jerico Research Infrastructures. The remaining three mature use cases (UC2 "Zooscan", UC3-OBSEA, UC3-Azores, UC3-Smartbay, UC4 "Oil spill detection") are preparing for service delivery in the coming months, pending final model configurations.

In the second year, the project focused on building partnerships with key initiatives: engaging with Blue-Cloud to explore EOSC Node involvement, collaborating with AI4Life for technology adoption from iMagine, working with ANERIS for model integration from the iMagine AI Platform, and partnering with Zenodo to create a customized metadata schema for aquatic images. These collaborations are set to deepen in the third year.

In the third year, the project will intensify its efforts in service promotion and user engagement, building on its presence at four conferences in PY2. Targeted outreach will focus on key stakeholders aligned with each of the five domain-specific use case services, as outlined in the "D2.5 Innovation Management and Exploitation Updated Plan." This strategy aims to drive traffic to the central access form, which facilitates access to the five thematic services. Alongside service delivery, the capabilities of these services will continue to evolve, with particular emphasis on customizing user interfaces to better meet the needs of domain scientists and reduce barriers for new users.

Last update: 7 February 2024

Permalink: https://cordis.europa.eu/project/id/101058625/reporting

European Union, 2025