

## EOai4BIO

EO-based hybrid artificial intelligence methods for monitoring habitat distribution and diversity

<b>Programm / Ausschreibung</b>	ASAP, ASAP, ASAP 18. Ausschreibung (2021, KP)	<b>Status</b>	abgeschlossen
<b>Projektstart</b>	01.06.2022	<b>Projektende</b>	31.07.2023
<b>Zeitraum</b>	2022 - 2023	<b>Projektlaufzeit</b>	14 Monate
<b>Keywords</b>	earth observation; machine learning; expert-based; deep-learning; biodiversity		

### Projektbeschreibung

Die Grundlage einer intakten Biodiversität in Österreich ist die ausreichende Vielfalt, Qualität und Verbreitung natürlicher Lebensräume. Dennoch gibt es zum aktuellen Zeitpunkt in Österreich kein einheitliches, operationelles System zur regelmäßigen, nationalen Erhebung und Evaluierung der Verbreitung und Entwicklung von Lebensräumen.

Ein wesentliches europaweites und nationales Instrument zum Monitoring der Habitat- und Pflanzendiversität ist die Erhebung der gemäß FFH-Richtlinie, Anhang 1 geschützten Lebensräume für die Berichtspflicht nach Artikeln 11 & 17 FFH Richtlinie.

Bezüglich der flächendeckenden Verbreitung der gemäß FFH-Richtlinie geschützten Lebensräume ist die Datenlage in Österreich jedoch teilweise lückig, bzw. ist eine regelmäßige Aktualisierung von Verbreitungsdaten aktuell nicht machbar (EU-Berichtspflicht gemäß Artikel 11 FFH-Richtlinie).

Biotopkartierungen der österreichischen Bundesländer sind uneinheitlich gestaltet, nicht flächendeckend und unregelmäßig verfügbar.

Satellitengestützte Fernerkundung kann zur großflächigen Habitat- / Biotoptyperhebung beitragen, in dem diese durch eine hohe Datendichte (big earth data) die spektralen, strukturellen und phänologischen Eigenschaften erfasst. Allerdings fehlt zum aktuellen Zeitpunkt die systematische qualitative Charakterisierung von Habitaten / Biotopen mittels Fernerkundung in einer Sample-Datenbank, welche zu einer Erhöhung der Nachvollziehbarkeit und Übertragbarkeit von selbstlernenden Verfahren der künstlichen Intelligenz beiträgt.

Das Projekt untersucht daher die Anwendung von hybriden, wissensbasierten und selbstlernenden, somit optimal übertragbaren und versatil einsetzbaren Verfahren im Bereich der satellitengestützten Erdbeobachtung (Earth observation, EO) zur Erhebung der Verbreitung von Habitaten, gemäß der einheitlichen Standards der FFH-Richtlinie (Anhang I).

Ziel ist es, auf Basis von hybriden Ansätzen der künstlichen Intelligenz (artificial intelligence, AI) in bildgebenden Verfahren eine Strategie für übertragbare Ansätze zu entwickeln, um eine AI-gestützte flächendeckende, automatisierte Erhebung der Verbreitung von (klassifizierten) Habitaten zu sondieren. Die im Projekt entwickelten methodischen Ansätze zielen darauf ab, eine ausreichend hohe Ergebnissicherheit durch die Einbindung von objektivierbarem Expertenwissen zu erreichen.

Langfristig gesehen trägt das geplante Projekt somit dazu bei, durch die Verbesserung der Wissensgrundlagen, die Basis für eine wissensbasierte, wissenschaftlich begleitete und fundierte Biodiversitätspolitik in Österreich zu schaffen.

Wesentliche Innovationen eines hybriden AI- Ansatzes zur satellitengestützten Identifikation von Habitaten sind: 1) Kombination von regelbasierten Heuristiken und Ansätzen des maschinellen Lernens (machine / deep learning) zur Optimierung von Verfahren des maschinellen Sehens (computer vision) und des Bildverstehens (image understanding), 2) Großräumliche Übertragbarkeit der entwickelten Ansätze, 3) Kosteneffizientere Erhebungsmöglichkeiten für zukünftige Kartierungen durch vollautomatisierte Vorklassifikation, 4) Möglichkeiten zur flächendeckenden und lückenlosen Erfassung der Verbreitung von Habitaten

Die angestrebten Erkenntnisse und Ergebnisse beinhalten: 1) Darstellung der Machbarkeiten für eine satellitengestützte Identifikation ausgewählter Habitate, auf Basis des aktuellsten Standes der Technik im Bereich von hybriden wissenskonditioniert selbstlernenden Systemen, 2) Auslotung der flächendeckenden, österreichweiten Übertragbarkeit der entwickelten methodischen Ansätze, 3) Entwicklung und Bereitstellung eine sog. benchmark samples Datensatzes für das Testen weiterer hybrider AI Ansätze der Forschungsgemeinschaft, 4) Darstellung von Möglichkeiten und Grenzen der erarbeiteten Ansätze, 5) Darstellung der Projektendergebnisse in einem Projektendbericht inklusive einer Demoversion für Anwendungsmöglichkeiten

## **Abstract**

The basis of a functioning biodiversity in Austria is a sufficient diversity, quality and distribution of natural habitats. However, currently no consistent, operational system for regular, national surveys of the distribution, development and evaluation of habitats, exists in Austria. An essential tool for monitoring habitat and plant diversity in Europe, is the mapping of habitats protected under Annex 1 of the Habitats Directive for the reporting obligation under Article 11 & 17 of the Habitats Directive. Regarding data on the distribution of FFH-habitats, annex 1, the data situation in Austria for reporting duties to the EU according to article 11, is partially incomplete, and regular updating of distribution data is currently not possible. Biotope mapping in the Austrian provinces is inconsistent, does not cover the whole country, and is irregularly available. Satellite-based earth observation (EO) is able to contribute to area-wide habitat / biotope type assessment by characterising habitats with big earth data based spectral, structural and phenological signatures. However, a sample database including systematically the specific habitat characteristics is still missing, which is needed to increase the transparency and transferability of yet less transparent self-learning artificial intelligence methodologies.

This project investigates the application of hybrid, knowledge-based and self-learning, thus optimally transferable and versatile applicable methods in the field of EO mapping of the distribution of habitats, according to the European-wide, uniform standards of the FFH-habitats directive (Annex I). The project`s objective is to develop transferable approaches based on hybrid artificial intelligence (AI) with the aim to achieve an AI-supported, automated mapping of the distribution of classified habitats with a sufficient degree of certainty. In the long term, the planned project therefore contributes to establishing the basis for a knowledge-based, scientifically accompanied and profound biodiversity policy in Austria, by improving the knowledge base for habitat distribution.

Essential innovations of a hybrid AI approach for satellite-based identification of habitat types are: 1) combination of rule-based heuristics and machine / deep learning approaches to optimise computer vision and image understanding, 2) Large-scale transferability of the developed approaches, 3) More cost-efficient survey options for future mapping through fully automated pre-classification, 4) Possibilities for a comprehensive and complete survey of the distribution of habitat types

The intended results and findings include: 1) Demonstration of the feasibility of satellite-based identification of selected

habitat types, based on the latest state of the art in the field of hybrid, knowledge-conditioned, self-learning systems, 2) Exploration of the Austria-wide transferability of the developed methodological approaches, 3) Development and provision of a so-called benchmark samples data set for testing further hybrid AI approaches of the research community, 4) Presentation of possibilities and limitations of the approaches developed, 5) Presentation of the project results in a final project report including a demo version for possible applications.

## **Endberichtkurzfassung**

Feasibility assessment for the integration of Fauna-Flora (FFH) habitat types into a remote sensing-based approach for identification and delineation

An expert-based screening of the European Union's Habitats Directive Annex I habitat types, shows that Fauna-Flora-Habitats (FFH) prospectively have varying chances of success for delineation and identification through a remote-sensing modelling approach.

As the respective expert-based screening was conducted via a Multiple Criteria Decision Analysis Matrix (MCDA), the Matrix itself includes descriptive attributes per assessment criteria for each FFH habitat type. The developed MCDA Matrix therefore results in a profound knowledge database for habitat type characteristics and their suitability to be identified by remote-sensing based modelling approaches.

Furthermore, the MCDA Matrix evaluates and displays the suitability of habitat type characteristics for integration into remote-sensing based modelling approach.

The results of the MCDA feasibility assessment show that habitat types associated with the habitat-group forests prospectively have the highest chance for successful identification and delineation in large-scale, country-wide applications using a remote-sensing based modelling approach.

In contrast, habitat types that are heavily dependent on the identification of very small-scale entities for their distinct assessment (e.g., differentiation of grass species) seem unsuitable for country-wide RS-based assessment. This affects mainly habitat types associated with the habitat-group of natural and semi-natural grassland and raised bogs, mires and fens.

In steep terrain the delineation of habitat types becomes even more difficult. Depending on the aspect of surfaces, topographic shadows may cover and influence such areas for analysis. Topographical errors of spectral data may also increase with the steepness of the terrain (e.g., habitat types associated with the habitat-group rocky habitats and caves).

Unsurprisingly, the habitat "caves not open to the public" show the lowest score, because what cannot be seen in remote sensing data is therefore not detectable. This is also valid for all other habitat types with key characteristics that are not included in remote sensing data (e.g., species and species communities covered by a forest canopy).

Due to the results of this feasibility assessment, alluvial and riparian forest habitats (FFH habitat type 91E0\* "Alluvial forests with *Alnus glutinosa* and *Fraxinus excelsior* " and FFH habitat type 91F0 "Riparian mixed forests of *Quercus robur* , *Ulmus*

laevis and Ulmus minor , Fraxinus excelsior or Fraxinus angustifolia along the great rivers”) were chosen as suitable for the development of hybrid, knowledge-based and self-learning methodological approaches for satellite-based earth observation.

Establishment of a workflow & development of methodological approaches

For the development of a remote-sensing based and AI-supported methodology to automatically delineate the alluvial and riparian forest habitat types 91E0\* and 91F0, pilot regions needed to fulfil the following criteria:

For the pilot regions, already existing and readily available reference data needed to be available.

The pilot regions needed to cover the total spectrum of eligible biotope-types and tree species for alluvial forest habitats in Austria.

As a result, the “Nationalpark Donauauen” (Danube-floodplains) and the “Natura 2000 site Salzachauen” (Salzach-floodplains) were selected as study areas.

Research and review of already existing possible reference- and evaluation data for these study areas showed that not all available data were applicable as input training data or evaluation data. Thus, reference- and evaluation field mapping additionally needed to cross check the quality of the readily available data and complement the data with own field mapping data.

To fulfil the standards for input training-data, the corresponding data needed to be spatially explicit and clearly verifiable. Therefore, the main task of the reference mapping was to create a dataset with a sufficient number of high quality, spatially explicit, and up-to-date indicator tree species data. For each habitat-type, indicator tree species (dominating and the sub-dominant tree species) needed to be mapped.

The focus of in field mapping for evaluation purposes was to establish biotope- and habitat maps with full coverage of the study sites. For this purpose, the data quality of already existing data on habitats and biotopes was cross checked and verified in the field. If found necessary, the data was replaced and missing data was complemented.

For the digitization of both the reference- and the evaluation mapping data, it was reasonable to digitize the data based upon LiDAR-based normalized digital surface models (nDSM) (point cloud data visualizing a top-down view on the earth’s surface) in combination with Pléiades satellite data and high resolution and up-to date orthophotos. This digitization procedure created spatially explicit datasets and reduced possible distortions.

The results of the field mapping were the production of high quality, spatially explicit, and up-to-date tree species data which served as input training data for modelling purposes, and comprehensive evaluation data with full coverage of the study sites, which served as input for the accuracy assessment and the evaluation of the certainty of results for modelling outputs.

For the development of an Austrian-wide remote sensing approach to derive the alluvial and riparian forest habitat types

91E0\* and 91F0, cost-free available high-resolution Sentinel-2 and mono-temporal very high-resolution Pléiades satellite data were tested to identify advantages and limitation for discriminating tree species and tree species composition. Tree species composition is the basis for modelling of 91E0\* and 91F0 habitat types in respect to their definition and mapping guidelines.

Multitemporal Sentinel-2 data with a dense temporal frequency provide information on phenological differences, that can be used for discriminating tree species. However, Sentinel-2 data lacks in spatial resolution to identify target biotic entities (tree species and tree species composition). Thus, Pléiades satellite data with a spatial resolution of 0.5 m (pansharped) was tested complementary following an approach of initial single tree species classification and modelling of habitat types based on tree species composition of forest patches.

#### Achieved certainty of results for modelling outputs

The outcomes of the accuracy assessments show that the certainty of results for the modelling approach is higher for indicator tree species than for FFH habitat-types 91E0\* and 91F0. Also, the certainty of results is higher for the modelling outcomes in reference test sites (= the same test sites as on which the methodological approach had been developed) than in the transfer test site (= spatially distinct study site for testing the transferability and applicability of methodology that was developed on the reference test sites).

On the reference test sites in the study region of the "Nationalpark Donauauen" the achieved certainty of results for the automatized and remote-sensing based detection of most indicator tree species was very high, nearly achieving scores for overall accuracies that might guarantee for legal certainty. The certainty of results for identification of FFH habitat types 91E0\* and 91F0 on the reference test sites in the "Nationalpark Donauauen" was still quite high (achieving overall accuracies of approx. 60 % - 70 %), though not able to guarantee legal certainty.

In contrast, the certainty of results for modelling outcomes on the FFH habitat types 91E0\* and 91F0 in the transfer test site in the Natura 2000 site Salzachauen, was quite low. The overall accuracies achieved for modelling results in the transfer site of the Natura 2000 site Salzachauen, show that the outcomes should be used for future research on establishing a workflow, modelling approaches and the transfer of developed methodologies into spatially distinct areas.

#### Networking and cooperation between Stakeholders

Apart from the major objective of creating a workflow and developing methodological approaches, the project "EOai4BIO" also aimed at connecting the major stakeholders within Austria, which either cover expertise with respect to remote-sensing based habitat mapping or have needs for the practical application of such a methodology.

This networking and capacity building was conducted partly within the regular project's team meetings, including discussions with external scientists, external experts, possible facilitators, and possible future users and clients in the team meeting. In those extended team meetings, for example, the project "EOai4BIO" was presented, further ongoing, and planned projects and possibilities for data- and knowledge exchange were discussed.

Additionally, at the end of the project, a stakeholder uptake workshop was organised by the project team. The uptake and interest in the workshop were high: In total the workshop was participated by 29 stakeholders ranging from scientist of universities, technical colleges, to experts from private companies or privatized / semi-privatized research companies and managements of nature conservation areas. Also, representatives from the federal provinces of Austria attended the workshop. The project's stakeholder workshop, successfully achieved to connect experts and to initiate an exchange of current scientific developments and expert knowledge with respect to methodological approaches on an automatised, remote-sensing based habitat detection, for the first time in Austria. The workshop also included possible future clients and policy decision makers in the discussion, to ensure that research and development activities directly correspond to the needs of society, possible clients, and policies. Representatives of the Austrian federal provinces stated that the needs for practical application of automatized habitat detection are high, and a timeline for practical application of methodological approaches with a high certainty of results, is needed. Also, private companies and privatized research companies stated their needs for technological progress and exchange with respect to this topic. In the final discussion round, as a conclusion, it was found that the key to practical application of existing methodological approaches is the increased collaboration especially of all technical experts, but generally also of all involved stakeholders in near future.

## Conclusions

The European Union's Habitat's Directive Annex 1 habitats have varying chances of success for delineation and identification through a remote-sensing modelling approach.

It was found that the Fauna-Flora-Habitats of alluvial and riparian forests (91E0\* and 91F0) have rather high chances of success for modelling through remote-sensing based approaches, as the remote-sensing based detection of indicator tree species can be used for methodological approaches. Therefore, in the project "EOai4BIO" alluvial and riparian forests (91E0\* and 91F0) were selected as habitats, to develop remote-sensing based modelling approaches.

The following aspects crucially influenced the modelling results for alluvial and riparian forest habitats (91E0\* & 91F0) in the project "EOai4BIO":

A high quality, a very high quantity and sufficient spatial distribution of data on indicator tree species, applied as input training data, are essential preconditions for achieving satisfying modelling results.

Some of the respective indicator tree species are easier to detect through remotes sensing based modelling approaches than others.

Also, the structure of habitats, in terms of homogeneity and heterogeneity have a high influence on the modelling outcomes. The transfer of the developed methodological approaches for application in other, spatially distinct study areas, remains challenging due to possible differing area characteristics.

The certainty of results for the remote-sensing based modelling outcomes for alluvial and riparian forest habitats (91E0\* and 91F0) was varying depending on the study area and its respective characteristics. Possible practical implementation of the developed model requires targeted application, depending on the respective application needs (e.g.: research purpose,

landscape evaluation, biodiversity monitoring, legal certainty for environmental impact assessment).

The certainty of results for practical application of existing methodological approaches can be improved through increased national collaboration especially of all technical experts, but generally also of all involved stakeholders in near future.

### **Projektkoordinator**

- Universität Salzburg

### **Projektpartner**

- DI Nina Maria Weber