

## AI-PROMETHEUS

Data-driven kilometerscale forecasting Digital Twin for the Alpine region

<b>Programm / Ausschreibung</b>	Themenübergreifend, Themenübergreifend, Common Pot : Digitaler Zwilling Österreich	<b>Status</b>	laufend
<b>Projektstart</b>	01.01.2025	<b>Projektende</b>	31.12.2026
<b>Zeitraum</b>	2025 - 2026	<b>Projektlaufzeit</b>	24 Monate
<b>Keywords</b>	data-driven forecasting; LAM AI model; DT for Alpine region; Destination Earth Digital Twins exploitation		

### Projektbeschreibung

Im Rahmen von Destination Earth werden derzeit digitale Zwillinge (Digital Twins, DTs) sowohl auf globaler als auch auf regionaler Ebene entwickelt, die u.a. darauf abzielen, extreme Wetterereignisse präzise vorherzusagen. Obwohl der regionale DT on-demand für flexibel wählbare Regionen gerechnet wird, sind nach wie vor Schwächen aufgrund der spezifischen Merkmale der Alpinregion mit ihren komplexen Bedingungen (z. B. Orographie, Physiographie, Grenzschichtprozesse) vorhanden. Datengetriebene KI Modelle haben in jüngster Vergangenheit erhebliche Fortschritte gemacht und immens an Qualität gewonnen. Obwohl das Training dieser Modelle erhebliche Rechenressourcen erfordert, sind sie während der Inferenzphase schnell und ressourcenschonend. Um den grünen und digitalen Wandel zu unterstützen, zielt das gegenständliche Projekt darauf ab, einen datengetriebenen DT für den Alpenraum zu entwickeln, implementieren und in weiterer Folge zu verfeinern, um Kenntnis über seine Prognosegüte während relevanter Wetterereignisse zu gewinnen. Diese zu untersuchenden Ereignisse stehen im Zusammenhang mit den komplexen Eigenschaften der Alpenen Grenzschicht und umfassen vor allem jene mit hoher sozioökonomischer Bedeutung. Da datengetriebene KI-Modelle stark von hochauflösenden und langen Zeitreihen von Re-Analysedaten abhängen, wird ein Schwerpunkt auf der Erweiterung und Verfeinerung des auf AROME basierenden C-LAEF-Reanalyse-Datensatzes als Hauptinput für das Training liegen. Darüber hinaus werden CERRA-Reanalysen und ein experimenteller C-LAEF-Reanalyse-Datensatz mit einer Auflösung von 1 km genutzt und der Impact des transfer learning durch Fine-tuning des Trainings untersucht. Ebenfalls ist geplant, die Möglichkeiten eines Ensemble DT auszuloten und Unsicherheitsschätzungen zusammen mit den deterministischen Vorhersagen zu erzeugen. Neben Fallstudien, die in der Projektanfangsphase definiert werden, wird die Evaluierung für längere Zeiträume ein besseres Verständnis in die Güte eines regionalen, rein datengetriebenen DT liefern. Die Ergebnisse der DT Vorhersagen werden eingehend mit den Ergebnissen unserer hauseigenen operativen Modelle (AROME, CLAEF) und insbesondere mit den Ergebnissen von Destination Earth verglichen, d. h. dem globalen und regionalen extreme DT und dem AIFS. Das Projekt wird in enger Zusammenarbeit mit einem wissenschaftlichen Beirat umgesetzt, der aus nationalen und internationalen Experten auf dem Gebiet der KI und speziell der datengetriebenen Vorhersage besteht. Besonderes Augenmerk wird auf eine enge Zusammenarbeit mit den Kollegen des ECMWF gelegt, deren Anregungen und Feedback während der Projektumsetzung berücksichtigt werden.

## Abstract

Within Destination Earth, digital twins (DTs) on both global and regional scale are currently developed, specifically tailored to accurately forecast extreme weather situations. Although the regional DT will run on flexible domains and on-demand, they still exhibit some deficiencies due to the specific characteristics of the Greater Alpine Region (GAR) with its complex conditions (e.g. orography, physiography, boundary layer processes). Data-driven models have gained tremendous skill during the last months and although the training of these models require substantial computational resources, they are fast and computationally cheap during the inference phase. In order to support the strategic goal of a green and digital transformation, the project aims at implementing and refining a data-driven DT to the Alpine region to complement the DTs already developed within the DE initiative. This development is necessary to gain knowledge on its forecast skill during weather regimes relevant in the GAR. These weather situations are related to the complex characteristics of the GAR, and involve events with high socio-economic impact. As data-driven AI models are heavily relying on high resolution and long time series of re-analyses data, the extension and refinement of the AROME-based C-LAEF re-analysis data set will serve as major input for the intended training. In addition, CERRA re-analyses and an experimental C-LAEF re-analysis data set on 1 km resolution will be exploited and the impact of transfer learning by updating the model investigated. It is also envisaged to explore the possibilities of an ensemble DT and to provide uncertainty estimates along with the deterministic forecasts of the data-driven DT. Beside case studies being defined during the kickoff phase, long-term evaluation will provide knowledge into the capability of a regional data-driven DT to accurately forecast specific situations relevant in regions with complex orography. The results of the DT forecasts will be thoroughly compared with the results from our operational model (AROME, CLAEF) and in particular with outputs from Destination Earth, i.e. the global and regional extreme DT and the AIFS. The project will be implemented in close cooperation with a scientific advisory board consisting of national and international experts in the field of AI and specifically data-driven forecasting. Special emphasize is given to a close collaboration with the colleagues at ECMWF whose advice and feedback will be taken into account during project implementation.

## Projektpartner

- GeoSphere Austria - Bundesanstalt für Geologie, Geophysik, Klimatologie und Meteorologie