



SLAIF – Slovenian AI Factory & AI for Science

Artificial Intelligence for Science

Ljubljana, 20 May 2026

Prof. Saso Dzeroski – Technical Lead

WHY EVERYONE IS TALKING ABOUT AI NOW?

Because AI can solve many everyday problems now

- Not only scientific ones
- AI is now embedded in many products and services(search, translation, customer support, ...)
- The key to this is the use of general-purpose **foundation models**

What are Foundation Models?

- Foundation models (FMs) are large AI models trained on large general data
- Can be later adapted (fine-tuned) by using small data to solve specific downstream tasks

Large Language Models (LLMs)

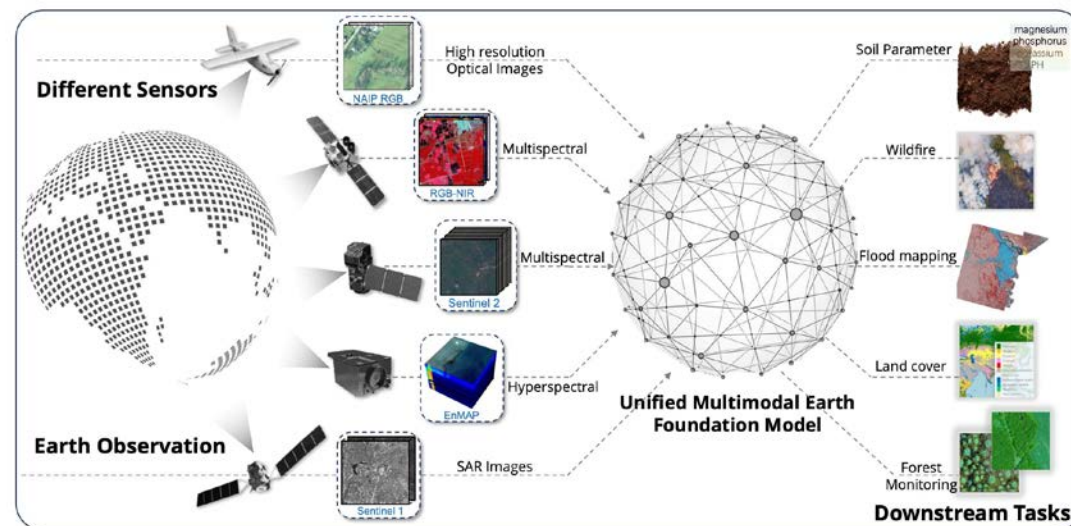
- LLMs are foundation models learned from text
- Typically by learning to predict the next word
- Can be used to generate text (generative AI)
- Downstream tasks include question answering, summarization, ...



TRAINING AND FINE-TUNING FM_s REQUIRES LARGE COMPUTE

Besides text, foundation models can also process other data, e.g. different kinds of images

Foundation Models are large neural networks
With billions of parameters (strength of connections between neurons)



Training and fine-tuning LLMs/FMs requires large compute

- LLMs/FMs are trained from massive data, requiring very large compute (weeks/months on GPUs)
- Even fine-tuning, adapting a FM to a specific domain task still requires significant compute

THE EU APPROACH TO AI:

Accelerate adoption & innovation

Regulation: The AI Act

- *Safety*
- *Trust*

Infrastructure

- *AI Factories*
- *Facilitate the development of (generative) AI models*

Innovation

- *Apply AI: Boost AI adoption in industry and public services*
- *But also, support research (e.g., AI in Science/**RAISE**)*

WHAT ARE AI FACTORIES?

Short answer: **AI factories are infrastructure for AI**

- In the 20th century, countries built roads and power grids
- In the 21st century, they must also build AI infrastructure
- This is a matter of technological sovereignty and global competitiveness

SLAIF** will ensure that Slovenia is not merely a user of artificial intelligence, but **an active contributor to Europe's technological sovereignty and global competitiveness

WHAT ARE AI FACTORIES?

AI factories lower the barriers to adoption of AI

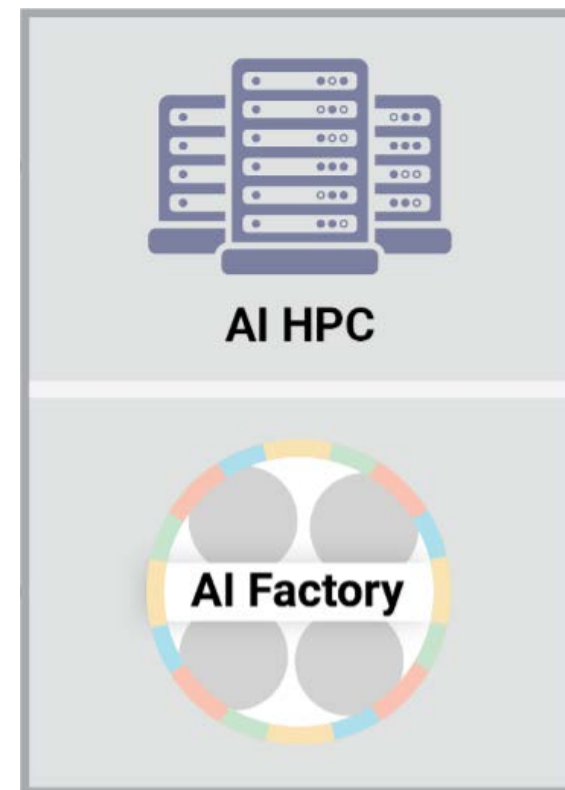
- They offer a wide and exhaustive range of services to AI startups and researchers needed for the development of European generative AI models and applications

AI factories have two main components:

- An AI-optimized supercomputer
- AI factory services and activities

They bring together:

- Compute
- Algorithms and workflows
- Data and models
- Skills, expertise, and talent



EUROPEAN AI FACTORY ECOSYSTEM



- EuroHPC and EU Member States co-finance deployment of **AI factories** (with supercomputers), and **AIF antennas** (which use the computing infrastructure of AI factories)
- Current ecosystem:
 - 19 AI factories
 - 13 antennas
- However, only 9 factories are procuring and deploying new supercomputers
- Some countries hold shares in AI factories located in other countries (e.g., Portugal and Türkiye in the BSC AI Factory).
- Slovenia will host a new supercomputer and also holds a stake in IT4LIA (€10 million)



Switzerland's participation is contingent upon the ratification of its accession to Horizon Europe.
The boundaries and names shown and the designations used on this map do not imply official endorsement or acceptance by the European Union.
This designation shall not be construed as recognition of a State of Palestine and is without prejudice to the individual positions of the Member States on this issue.
Administrative boundaries: © EuroGeographics © OpenStreetMap
Cartography: Eurostat – IMAGE, 05/2025

WHY CAN SLOVENIA HAVE AN AI FACTORY?

Although Europe will have many AI factories, not every EU country will host one. This requires vision, political will, financial support, scientific and technical excellence and human resources.

- **Slovenia has:**
 - **A strong AI community**
 - **A strong high-performance computing community**
 - VEGA was the first operational EuroHPC system
 - **Excellent collaboration between both communities** (which should not be taken for granted)

A broader understanding of infrastructure is needed

- Infrastructure is not only hardware (computers)
- Data is infrastructure (Slovenia leads CLARIN, language data)
- AI is infrastructure (especially for science; e.g., the JSI AI infrastructure center)

AI IN SLOVENIA – RESEARCH BASE

- Long-standing tradition of AI research (since the 1970s)
- Strong research community
 - **Jožef Stefan Institute, Ljubljana** (multiple research departments, UNESCO center)
 - **University of Ljubljana**
 - **Faculty of Computer and Information Science** (multiple research laboratories)
 - Faculty of Electrical Engineering
 - Faculty of Mathematics and Physics
 - University of Maribor, Faculty of Electrical Engineering and Computer Science (FERI)
 - University of Nova Gorica
 - University of Primorska
 - Faculty of Information Studies, Novo mesto
- International activity and visibility (events, projects, policy co-creation)
- ELLIS Slovenia Hub

AI IN SLOVENIA – WHERE ARE WE STRONG?

Knowledge-Based AI

- Knowledge representation (semantic technologies)
- Decision support systems

Data-Driven AI

- Machine learning of explainable models
- Language technologies (language resources, language models)
- Computer vision

AI for Science (Combining knowledge and data)



PoVeJMo
SLOVENSKI
VELIKI JEZIKOVNI MODEL

AI IN SCIENCE – PROJECTS

- Projects from the ARIS “Gravitations” Programme
 - GC-001: **AI for Science** (Development of AI methods and their application in science), JSI
 - GC-002: **LLMs for Digital Humanities**, UL FRI
 - GC-004: **AI and Materials Science** (Solutions for chemical energy storage for climate-resilient renewable energy systems – CHRONOSTORE), National Institute of Chemistry
 - GC-005: **AI and Life Sciences** (Translation-enhancing oligonucleotide therapeutics – BOOST), National Institute of Chemistry
 - GC-006: **AI and Environmental Sciences** (Geospatial information technologies for a resilient and sustainable society – GeoAI), UL FGG
- Other Initiatives
 - MSCA COFUND SMASH – *Machine Learning for Science and Humanities*
 - AI for Science vertical within SLAIF
 - RAZUM – AI for Science infrastructure center at the JSI

AI IN SCIENCE – EVENTS

- **ESSAI 2023, First European Summer School on Artificial Intelligence**, Ljubljana, 24–28 July 2023
 - Six parallel tracks, 24 courses
 - ACAI-2023: Advanced Course on Artificial Intelligence for Science
 - ESSAI has become a tradition; ESSAI 2026 will take place in Vienna
- **Artificial Intelligence for Science International Conference**, Ljubljana, 22–26 September 2025
 - 28th Discovery Science Conference
 - Additional tracks - AI & Environmental Science; AI & Life Science; AI & Materials Science; AI & Physics; AI & Space (+ Summer School); AI & Digital Humanities; AI Factories
 - The 2nd International Conference on AI for Science will be held in Mainz, 5–9 October 2026
- **AI in Science Summit**, 3–4 November 2025, Copenhagen
 - AI for Planet and Climate
 - Launch of RAISE, the EU initiative for AI in science
 - The 2026 AI in Science Summit will take place in Dublin on 15 October 2026



SLAIF – SLOVENIAN AI FACTORY

Two projects with a combined value of €135 million, co-financed by the EC and the RS MVZI (Ministry of Higher Education, Science and Innovation)

- AI-Optimized Supercomputer, €123.32 million
- Co-financed under the Digital Europe Programme
- The project consortium consists of:



- AI Factory, services and activities, €11.68 million
- Co-financed under Horizon Europe
- The project consortium consists of:



SLAIF – SUPERCOMPUTER



This part of SLAIF will be co-financed by Slovenia and EC Digital Europe Program

*A total of **123 Mio EUR** of HW procurement and operational costs*

The SLAIF AI-HPC Consortium consists of

- **IZUM:** The hosting entity
- **The Jožef Stefan Institute**
- **ARNES:** The HW will reside in the (newly built) ARNES data center in Maribor



SLAIF – SUPERCOMPUTER

Combination of CPU and GPU resources, supporting both traditional HPC and AI computing:

- CPU: approx. 10 petaFLOPS
- GPU: 30 petaFLOPS (Phase 1) +70 petaFLOPS (Phase 2) = 100 petaFLOPS in total

Two-phase procurement to ensure a longer operational lifespan:

- Supercomputer operational by the end of 2027
- Upgrade planned for 2029
- Compared with VEGA: 3× more CPU power and 32× more GPU power

At present, the following are trained on the LEONARDO supercomputer in Bologna:

- Large language model for Slovenian – GaMS (UL, FRI)
- First Slovenian vision-language model – SViLa (JSI)

Such tasks will be supported by the new SLAIF AI-optimized supercomputer

Generativni model
za slovenščino
Generative Model
for Slovene

GaMS: “



SLAIF: AI Factory Part (SLAIF HE Project)

This part of SLAIF will be co-financed by Slovenia and EC Horizon Europe Program

*With a budget of cca **12 mio EUR***

The SLAIF AI-HPC Consortium consists of ten partners of which

- **IZUM:** Administrative coordinator
- **The Jožef Stefan Institute:** Technical lead
- **Universities of Ljubljana and Maribor:** Workpackage leaders



SLAIF – TARGET USER GROUPS

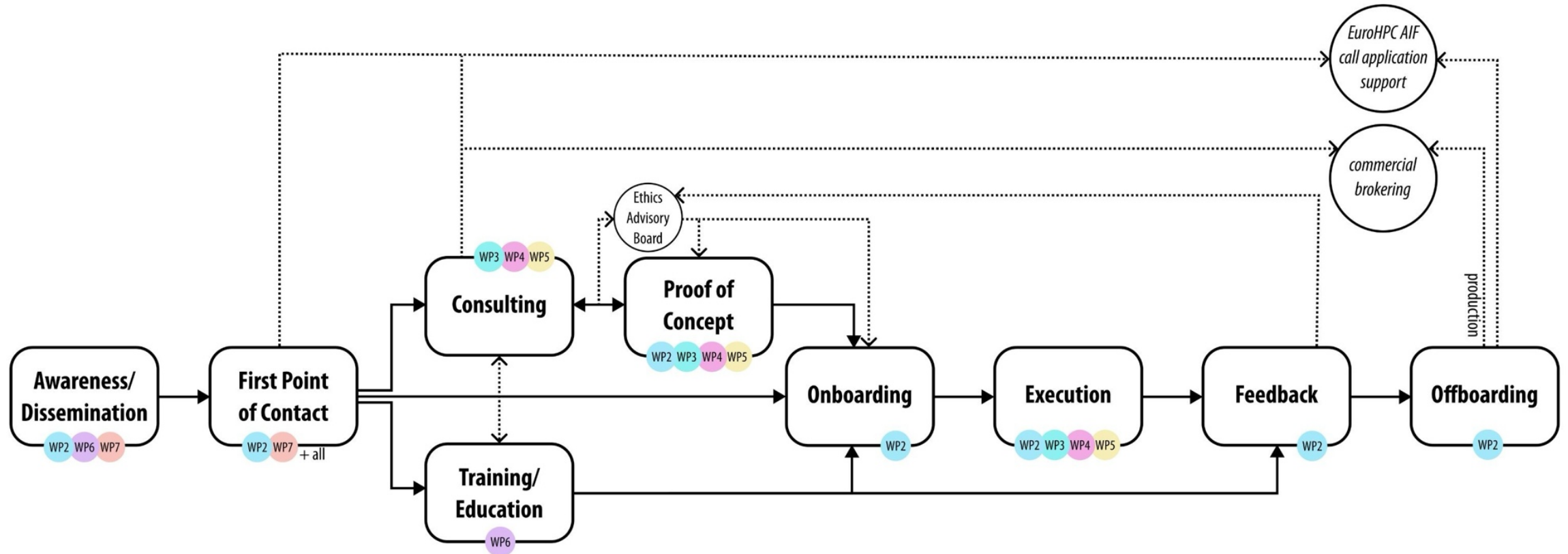
Different user groups

- Economy / Industry
 - Startups
 - Small and medium-sized enterprises
 - Large enterprises
- Public Sector
 - Healthcare
 - Public administration
 - Public companies, institutes, and agencies
- Research Community (Universities, Research institutes)

Different levels of AI readiness

The SLAIF User Journey

A structured pathway from first contact to production deployment



SLAIF SERVICES & ACTIVITIES

Access to Computing and Data Infrastructure

General Workflows for Training and Fine-Tuning Models

- Large language models
- Large speech models
- Vision-language models
- Multimodal models

Domain-Specific Workflows and Models for Different Verticals

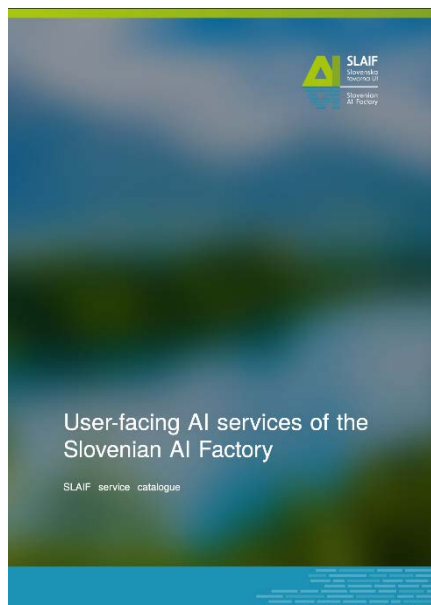
- Green transition
- Health and biotechnology
- Digital society
- AI for science

AI Education and Skills Development

SLAIF SERVICES

Different kinds of services

- User support
- Infrastructure and data
- AI development & deployment
- AI products



Contents

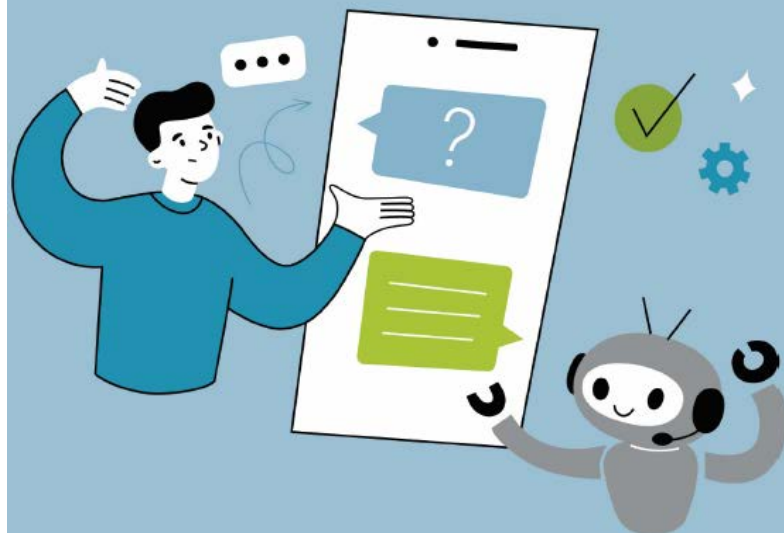
User support	9
Business Processes Review and Automation Using AI	11
Support at Distributed Training of Large-Scale AI Models	12
Support in Model Optimisation: Pruning, Quantisation, and Distillation	14
Support in Proof-of-Concept (PoC) Application Development	16
Technical User Support	18
Infrastructure and data-related services	19
Cloud Platform Service (OpenStack PaaS)	21
Container Image Registry	22
Container Platform Service (Kubernetes)	23
Data Access Services	24
Data Analytics Services	26
Semantic Data Discovery over the SLAIF Catalogue	28
Data Governance Services	30
Data Management / Sharing Services – Connectors for Trusted Sharing	31
Data preparation services	33
Data Processing Services	35
Exploration and Investigation of Disaster Events through the Slovenian EO Data Cube	37
Facebook PE-Video with Slovenian Annotations	39
FAIR-Compliant Metadata Management and Discovery Service	40
Fast Parallel Storage	42
Foundation-Model-Assisted Dataset Bootstrapping from Uncurated Visual Data (Auto-Annotation)	43
HPC Accelerated Compute Service	45
HPC Compute Service	47
Mid- and Long-Term Storage	48
Mid- and Long-Term Remote Access Storage	49
Monitoring Services	50
Orchestration Services	51
Synthetic Data Services	52
Synthetic Genetic Material Data Generation	54
Transformation and Processing Pipeline for Genetic Material Data	55
User Software Repository	57
AI development and deployment	58
Foundation Models for Human Sensing with Support for Fine-Tuning	60
Adaptation of a Large Language Model for Domain-Specific Question Answering	62
Adaptation of a Large Language Model for Domain-Specific Summarization	63
Adaptation of a Large Language Model for a Specific Task using Parameter Efficient Fine-Tuning	64
Adaptation of a Large Language Model for Sentiment and Emotion Classification using Parameter Efficient Fine-Tuning	65
Adaptation of a Large Language Model to a Specific Domain using Instruction Following	66
AI-Based Image Analysis for Industrial Anomaly Detection and Quality Control	67
AI-Based Microscopy Image Analysis	69
Automated Construction of Domain-Specific Knowledge Graphs	71
Change Detection and Description from Remotely Sensed Data	73
Embedding Selection – Intelligent Embedding Selection for Scalable AI Systems	75
ML Pipeline Exploiting Known Attribute Value Uncertainties	77
Model Evaluation for the Local Context Using Linguistic and Cultural Competence Measures	78
Per-Prompt LLM Selection – Automatically Choosing the Best Open Access LLM	80
Preference Optimization Pipeline and LLM-Based Machine Translator for English and Slovene Pairs	82
Satellite and Aerial Image Analysis	83
Temporal, Multi-Modal Foundation Models from Remotely Sensed Information	85
Visual and Interactive AI Workflows (OrangeWeb)	87
Visual and Interactive AI Workflows (OrangeWeb) for Green Transition	88
Visual and Interactive AI Workflows (OrangeWeb) for Health and Biotechnology	89
AI products	90
Access to a Multimodal Language/Sensor Model for Data Analysis	92
Agentic RAG for Administration	94
AI-Assisted Workflow for Grading Scanned Handwritten STEM Exams	96
AI-Based Human Pose Estimation and Kinematic Analysis	98
AI-Based Predictive Analytics for Strategic Resource and Trend Forecasting	100
AI-Based Visual Defect Detection and Quality Assessment for Industrial Production	102
AI-Driven Predictive Maintenance and Performance Optimization for Energy Systems	104
AI-Enhanced Weather Forecasting for Power Grid Operations	106
AI for Health and Biotechnology – AI-Assisted Lab Blood Test Interpretation	108
AI Framework for ECG Analysis	110
AI-Powered Intelligent Document Analysis and Semantic Retrieval	112
AIDA – Agricultural Intelligence and Decision Assistant	114
Automated Classification and Feature Extraction for Experimental Materials Data	116
Crop Growth Monitoring and Analytics Services	118
Domain-Agnostic Object Counting, Detection and Tracking in Images	120
Energy Distribution Optimisation	122
FoodyLLM – AI Copilot for Food Nutrition Intelligence	123
Irradiance Prediction Service for Energy Modelling	125
Multi-Label Classification for Microbial Data	127
Multilingual Idiom Semantic Search and Translation System	128
Observable-Cue Mental Health Risk Prediction & Screening Toolbox	130
Open Model for Slovenian Acoustic Prosody Annotation	132
Pre- and Post-Surgery Primary Brain Tumour Volume Estimation	133



SLAIF SERVICES

AI development and deployment

Services for adapting, fine-tuning, integrating, scaling, and operationalising AI workflows and models for real use cases.



SLAIF service catalogue



SLAIF service catalogue



AI development and deployment

Adaptation of a Large Language Model for Domain-Specific Question Answering

Responsible institution
University of Ljubljana

Vertical Cross-domain Availability Q3/2027 Type Pipeline

Target users: Industry, Large Enterprises, Scientific Institutions, Public Institutions

Overview

This pipeline enables the adaptation of a large language model to answer domain-specific questions accurately and efficiently. It is designed to support expert systems, knowledge bases, or FAQ systems in specialized fields. The prerequisites are a sufficient amount of domain-specific data and domain know-how. The pipeline first helps users in preparing data for question answering based on the domain data and provides code and know-how for synthetically expanding the training data based on the provided domain data. Next, it provides the know-how and code for fine-tuning the model with the created instruction-tuning dataset. Finally, it provides code for efficient inference using the adapted model. The pipeline leverages the Slovene LLM GaMS and similar open-source models, tools for generating domain-specific question-answer pairs using NeMo Curator in combination with the GaMS-Instruct approach, LoRA or QLoRA-based supervised fine-tuning using Hugging Face and NeMo, and vLLM for fast domain-specific question answering inference.

Expected outcome

A step-by-step guide for domain-specific question-answering training, along with scripts for data generation, model training, and inference.

Requirements

Requires several NVIDIA GPUs.

Service output

Pipeline for question answering adaptation, including data generation scripts, model training scripts, and inference code.

Specific limitations

Licensing

MIT

AI development and deployment

Adaptation of a Large Language Model for Domain-Specific Summarization

Responsible institution
University of Ljubljana

Vertical Cross-domain Availability Q3/2027 Type Pipeline

Target users: Industry, Large Enterprises, Scientific Institutions, Public Institutions

Overview

This pipeline enables the adaptation of a large language model (LLM) to generate high-quality, domain-specific summaries for technical or scientific documents, reports, research papers, or industry-specific content. It is designed to enhance the accuracy and relevance of summaries within a specific domain. The pipeline first helps users in preparing data for summarization (OCR of PDF documents) and converting it into a training dataset. Next, it provides the know-how and code for fine-tuning the model with the created dataset. Finally, it provides code for efficient inference using the adapted model. The pipeline leverages the Slovene LLM GaMS (for Slovene-language summarization) and other open VLMs (e.g., Gemma), open-source OCR libraries such as Marker or small open-source VLMs such as Nanonets, supervised fine-tuning or preference tuning using GRPO with LoRA or QLoRA-based fine-tuning, and vLLM for efficient, low-latency inference of domain-specific summarization tasks.

Expected outcome

A step-by-step guide for domain-specific summarization adaptation, along with scripts for data preparation, model training, and inference.

Requirements

Requires several NVIDIA GPUs.

Service output

Pipeline for summarization adaptation, including data preparation scripts, model training scripts, and inference code.

Specific limitations

Licensing

MIT

SLAIF – EDUCATION AND TRAINING



Different Types of Training

- Workshops
- Webinars
- Micro-credential courses

Different Training Topics

- AI methods
- High-performance computing
- AI application domains

Different Target Groups



SLAIF Training Catalogue



Large language models and domain-specific AI for the Slovenian agricultural sector

Course provider:

Jožef Stefan Institute (JSI)

Instructors:

Marko Debeljak (JSI) and colleagues

Content: The event aims to provide agricultural stakeholders with a clear, realistic, and application-oriented understanding of large language models (LLMs) and their potential role in Slovenian agriculture. Participants will be introduced to the rationale for focusing on high-quality, domain-specific data and benchmarks rather than individual model versions, and to the role of the national AI infrastructure (SLAIF) in supporting domain- and language-specific AI solutions. A key objective is to foster informed expectations, trust, and early engagement of future users by openly addressing both the capabilities and limitations of current AI technologies in regulated and practice-oriented agricultural contexts.

The event presents current activities related to the development and application of large language models within the Slovenian agricultural domain in the context of SLAIF. It introduces the fundamentals of LLMs in an accessible manner, with a focus on their applicability, risks, and constraints in agriculture. Participants will be presented with a proposed taxonomy of Slovenian agriculture, including key sectors and subdomains, along with example benchmark tasks designed to evaluate model performance in Slovenian. Preliminary benchmark results will be demonstrated to illustrate how existing models perform on selected agricultural use cases. The event further provides a structured forum for validating assumptions about users, use cases, and sectoral coverage, and introduces a dedicated web application for the collection of high-quality question-answer pairs and other expert inputs to support model adaptation and evaluation.

After the event, participants will understand the importance of domain-specific data, benchmarks, and expert validation for the sustainable development of AI solutions in agriculture. They will have a realistic understanding of what current large language models can and cannot do in the Slovenian agricultural context, and how national AI infrastructure supports their adaptation. Participants will be able to critically assess proposed use cases and benchmarks from a practitioner's perspective and will be equipped to contribute expert knowledge, feedback, and structured data inputs for the continued development and evaluation of AI tools for Slovenian agriculture.

Prerequisites: No technical prerequisites are required. The event is intended for practitioners and experts with domain knowledge in agriculture; prior experience with AI technologies is not required.

Target audience: Agricultural advisers and extension services; Farmers and agricultural practitioners; Agri-food stakeholders and industry representatives; Researchers and developers working on AI for agriculture; Public-sector stakeholders and policy makers

Language: Slovenian; English

Tentative date: Autumn 2026

Venue: ISI Ljubljana/Murska Sobota

Format: Lectures; Hands-on sessions

Duration: 6h

No. of ECTS credits: N/A

SLAIF – EDUCATION AND TRAINING

88

Unique training activities planned

59

Workshops & hands-on sessions

16

Webinars

13

EQF-aligned micro-credentials



Safe, responsible and compliant AI for SMEs: From generative AI to the EU AI Act

Course provider:

Jožef Stefan Institute (JSI)

Instructors:

Maja Škrjanc (JSI), Mitja Jermol (IRCAI)

Content: This 2-hour training is designed for SMEs, small and micro companies that are introducing—or considering the introduction of—artificial intelligence, with a particular focus on generative AI. Participants receive a structured overview of core AI concepts (including the distinction between data-driven methods and GenAI, and the role of context, memory and reasoning), alongside a business-oriented discussion of typical risks in practice (e.g., confidentiality and data leakage, GenAI misuse, deepfakes and brand/identity fraud).

FAIR data management for artificial intelligence

Course provider:

Jožef Stefan Institute (JSI)

Instructors:

Panče Panov (JSI)

Learning objectives: 1) Understand AI data assets across the lifecycle: how datasets, labels, dataset splits, features, and evaluation artifacts evolve from collection to reuse; 2) Apply the FAIR principles to AI work: make data and outputs easier to find, access, combine, and reuse (for teams and future projects); 3. Create an actionable DMP for AI projects: a lightweight plan that supports reproducibility, handover, and compliance; and 4) Handle constraints responsibly: recognize sensitive data, ethical considerations, access limitations, and industry vs research expectations.

Efficient LLMops, hosting, and model quantization

Course provider:

Jožef Stefan Institute (JSI)

Instructors:

Boshko Koloski (JSI), Matej Martinc (JSI), Usama Derebashi (JSI), Nikola Marić (JSI), Sašo Džeroski (JSI)

Content: This course covers the operational aspects of large language models (LLMOps), efficient hosting, and model quantization. LLMOps represents an extension of MLOps for generative models; it includes practices for managing the entire lifecycle of LLMs—from selection and customization to deployment and continuous monitoring. Participants learn the differences between MLOps and LLMOps and why generative models require additional procedures, such as robust data pipelines, version control, and flexible resource orchestration.

Large language models and domain-specific AI for the Slovenian agricultural sector

Course provider:

Jožef Stefan Institute (JSI)

Instructors:

Marko Debeljak (JSI) and colleagues

Content: The event aims to provide agricultural stakeholders with a clear, realistic, and application-oriented understanding of large language models (LLMs) and their potential role in Slovenian agriculture. Participants will be introduced to the rationale for focusing on high-quality, domain-specific data and benchmarks rather than individual model versions, and to the role of the national AI infrastructure (SLAIF) in supporting domain- and language-specific AI solutions. A key objective is to foster informed expectations, trust, and early engagement of future users by openly addressing both the capabilities and limitations of current AI technologies in regulated and practice-oriented agricultural contexts.

SLAIF: AI Factory Part – The Structure of the SLAIF HE Project



WP2

**Workflow and Data Orchestration
Infrastructure**

WP3

Core AI Platform - AI Workflows for training, fine-tuning, evaluation, and for LLMs/VLMs/multi-modal foundation Models; Explainable & AutoML; Semantic catalogue

WP4

User-centric AI Services - Data Storage Connectors (e.g. COPENICUS data space); User Interface and Experience; User Engagement and Feedback

WP5

Vertical Services for Sectorial Applications (incl. customized workflows)

- Green Transition
- Health & Biotech
- Digital Society
- AI for Science

WP6

AI Skills and Training

GREEN TRANSITION: Climate & Environment

Climate change is causing sea level rise and more extreme weather

- By 2100, +30 cm to +1 m (IPCC)
- With +1 m, parts of Piran could be flooded twice per day

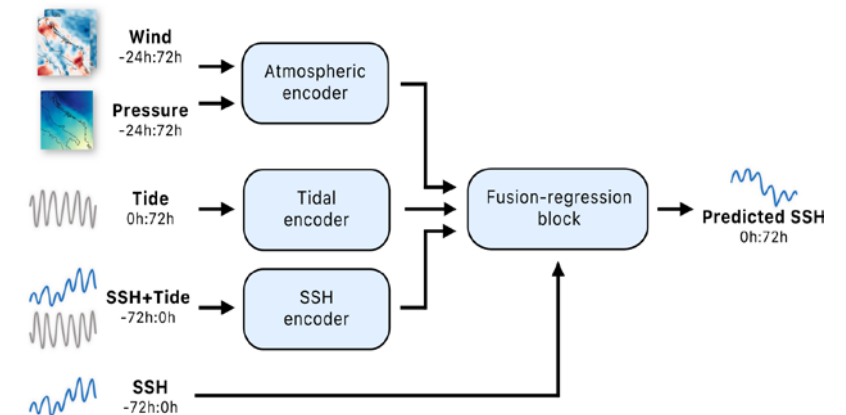
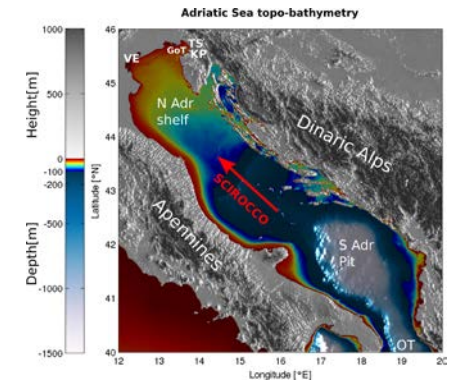
HIDRA (developed by University of Ljubljana (FRI) and ARSO) uses AI to forecast sea levels and coastal flooding risk

HIDRA can produce forecasts hundreds of thousands of times faster than classical simulation models

It also improves prediction quality vs traditional approaches

Practical importance: Enables early warning and protection of people, infrastructure, ecosystems

Training and scaling forecasting models requires GPU compute and data workflows





GREEN TRANSITION: Land Use/Cover from Satellite Images

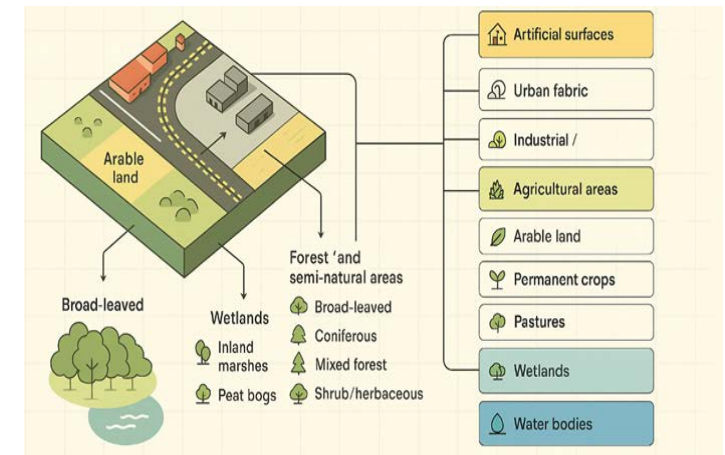
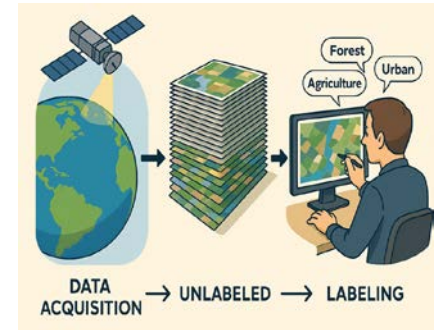
At JSI, we develop methods for semi-supervised classification of satellite images

- Copernicus satellite imagery is **massive**, but labeled data is **scarce and expensive**
- Land cover is naturally **multi-label** and **hierarchical** (e.g., Corine Land Cover)

***Pre-training (Self-supervised learning):** Learn general visual representations from unlabeled remote-sensing images*

***Fine-tuning (Semi-supervised learning):** Use a small labeled subset large unlabeled subset to predict multiple labels within the hierarchy*

**SLAIF + Copernicus Data Space =
Compute & Workflows + High-volume EO data**
Enables scalable AI services for agriculture and environment



HEALTH & BIOTECH:

Fine-tuning an LLM with nutrition data

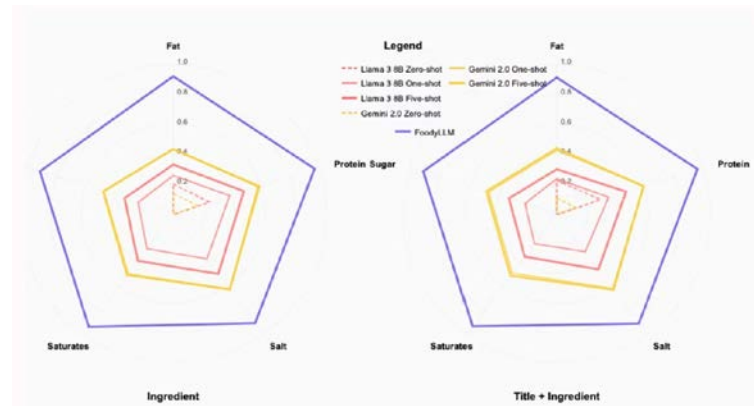


The tasks:

- *Traffic Light Classification of Recipes: healthy (green), moderate (orange), limit consumption (red)*
- *Assessing Recipe Nutrient Values (Salt, Sugar, Fat, ...)*

The approach: Fine-tune Llama3 with curated data

The results: Fine-tuned model much better (assessing recipe nutrient values results shown below)





DIGITAL SOCIETY: An LLM for Slovenian

GaMS – Slovenia's own generative language model

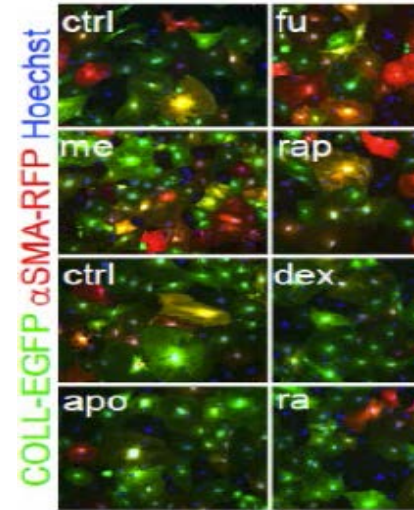
- *GaMS = Generative Model for Slovene – LLM tailored to Slovenian language & culture*
- *Designed to fill gaps left by global models that under-represent Slovene*
- *Open-access and developed by PoVeJMo (UL CJVT, UL FRI, INZ, ZRC SAZU)*
- *Multiple model sizes (1B to 27B+ parameters) are planned or available*

GaMS – A strategic national infrastructure

- *Language sovereignty: ensures Slovene digital presence and quality AI services*
- *Cultural understanding: better handling of Slovene syntax, idioms, cultural context*
- *Open for innovation: supports research, startups, industry, public services*
- *Ecosystem build-out: national data collection campaigns and instruction datasets.*

Generativni model
za slovenščino
Generative Model
for Slovene

GaMS: “



- **Drug Discovery**

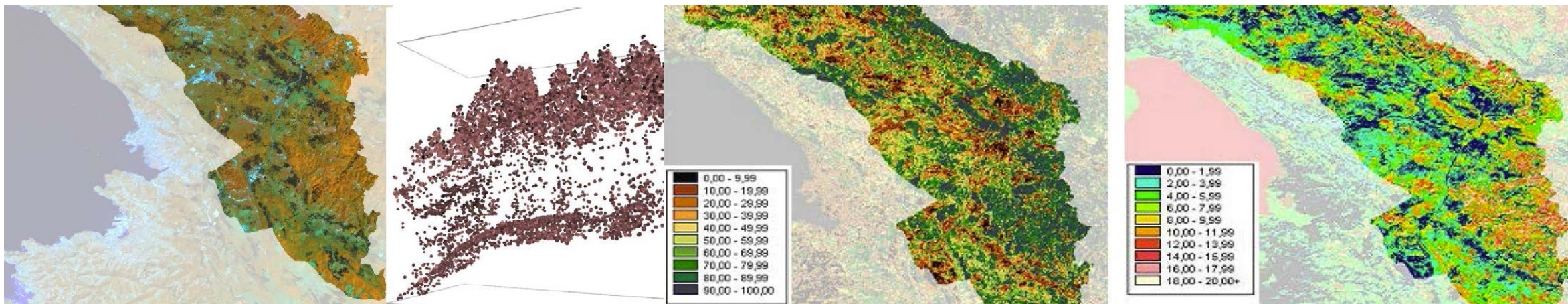
- Predicting the effects of different compounds
- Tuberculosis, salmonella, pulmonary fibrosis

- **Personalized Medicine**

- Predicting the effect of peptide insertion on protein function
- Designing gene therapies

AI IN ENVIRONMENTAL SCIENCES

- Environmental Monitoring Using Remote Sensing
 - Machine learning from combined 2D (satellite) and 3D (LiDAR) imagery
 - Estimation of forest height and density (including biomass)



- Wildfire Risk Assessment
- Assessing the impact of climate change (e.g., altered precipitation patterns)
- Planning appropriate mitigation measures

AI IN MATERIALS SCIENCE

- **Materials are critically important for industry**
 - They are a bottleneck of the green transition
 - They determine the performance, cost, and sustainability of industrial products
 - Energy storage/conversion; Manufacturing/mobility; Electronics/sensors
- **AI can significantly shorten the materials design cycle**
 - Guiding experiments and simulations; Self-driving laboratories
- Example from the Jožef Stefan Institute: **Design of foamed glass using a combination of machine learning and optimization**
 - Low-cost, durable, non-toxic, fire-resistant and insulating material
 - Large-scale production – often from recycled waste glass
 - Optimization targets: density, porosity
 - New materials synthesized based on AI-generated proposals demonstrated desired properties in laboratory testing



Thank you for your attention!



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA VISOKO ŠOLSTVO,
ZNANOST IN INOVACIJE

SLAIF: Slovenska tovarna umetne inteligence je prejela financiranje Skupnega podjetja za evropsko visokozmogljivo računalništvo v okviru sporazuma o nepovratnih sredstvih št. 101254461. Projekt financirata Ministrstvo za visoko šolstvo, znanost in inovacije Republike Slovenije ter program Horizon Europe Evropske unije.

SLAIF: Slovenian AI Factory has received funding by the European High Performance Computing Joint Undertaking under the grant agreement no. 101254461. The project is funded by the Ministry of Higher Education, Science and Innovation of Republic of Slovenia and by Horizon Europe Programme of the European Union.