



# KI-unterstützte Suche nach Geodaten

Dr. Simon Jirka, 52°North GmbH, [jirka@52north.org](mailto:jirka@52north.org)

14.11.2025 – Thementag „KI in der Geoinformation“

# Überblick

- Vorstellung
- KI:STE
- I-CISK
- NFDI4Earth - OneStop4All
- Ausblick

# 52°North - in a nutshell

- Angewandte Forschung im Bereich der Geoinformatik
- Ca. 25 Mitarbeitende
- Projekte in den Bereichen
  - Forschung
  - Beratung
  - Software-Entwicklung
- Themenfelder, u.a. Echtzeitdaten, Geo-AI, Cloud-Infrastrukturen, Forschungsdatenmanagement



UNIVERSITY  
OF TWENTE.



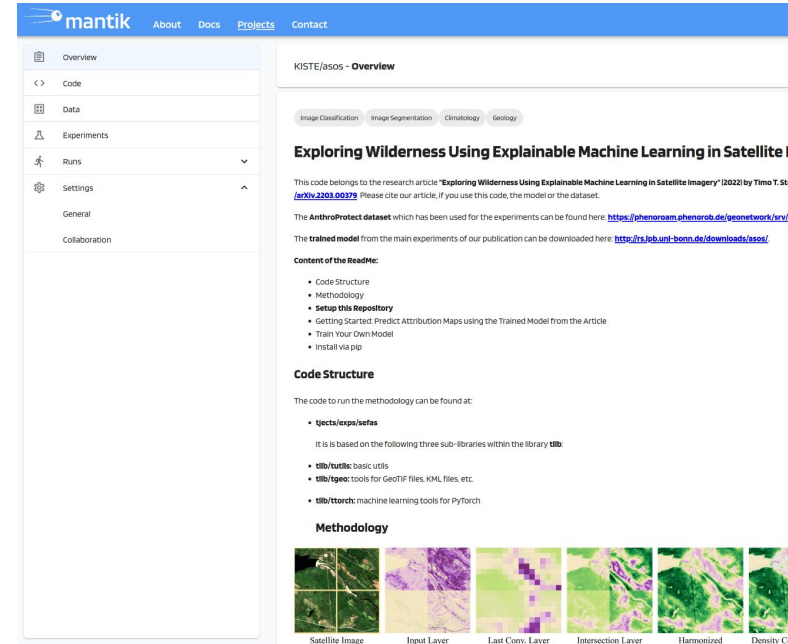
con•terra



# KI:STE - FDI für EO-Daten und KI-Analysen

Bessere Verknüpfung von Domänenwissen und (KI-)Infrastruktur

- Zugriff über verschiedene Datenanbieter hinweg
- Unterstützung verschiedener analytischer Infrastrukturen (lokal, Cloud, HPC)
- E-Learning-Ressourcen für Domänenexperten
- Aufzeichnung der Ausführung und Leistung von KI-Modellen als Metadaten



Screenshot der Mantik KISTE-Plattform der Ambrosys GmbH (MLflow).



gefördert vom



Federal Ministry  
for the Environment, Nature Conservation  
and Nuclear Safety



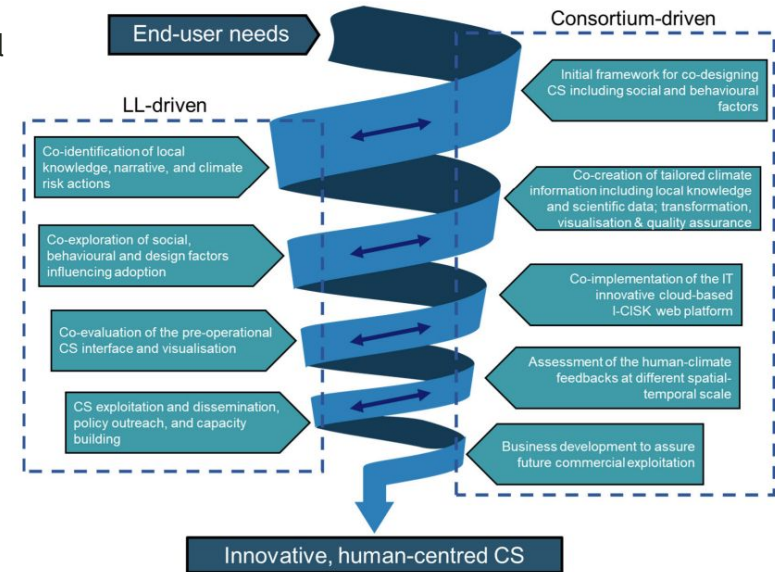
# Co-Design to Reveal the Value of Climate Services

## Ziele:

- Entwicklung von Ansätzen zur Anpassung an den Klimawandel und extreme Wetterereignisse
- Einbeziehung sozialer und verhaltensbezogener Faktoren, lokaler Bedürfnisse und lokalen Wissens
- Unterstützung von Bürgern, Interessengruppen und Entscheidungsträgern bei der Entwicklung von Maßnahmen

## Ansatz:

- Co-Design mit 7 Living Labs
- Wissenschaftliche Expertise
- Entwicklung einer flexibel anpassbaren und modularen Webplattform
- Cloud-Technologien





# Herausforderungen der Reallabore in I-CISK

📍 Upper and Middle Guadalquivir Basin, Spain  
**Agriculture, Livestock,  
Forestry and Water Scarcity**

[Read more](#)



📍 Rijnland Basin, the Netherlands  
**Water recreation, shipping  
nature conservancy**

[Read more](#)



📍 Budapest, Hungary  
**Tourism, health, and urban  
planning**

[Read more](#)



📍 Lesotho  
**Food security, drought &  
humanitarian aid**

[Read more](#)



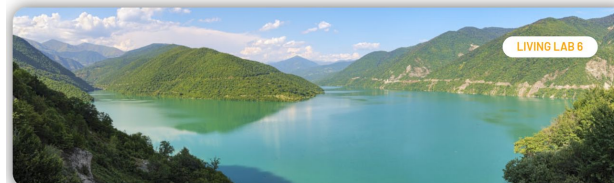
📍 Emilia Romagna, Italy  
**Agriculture and Domestic  
water resources**

[Read more](#)



📍 Crete, Greece  
**Multi hazard Sectoral  
climate change assessment  
for Tourism**

[Read more](#)



📍 Alazani River Basin, Georgia  
**Water supply, hydropower  
energy and forestry**

[Read more](#)

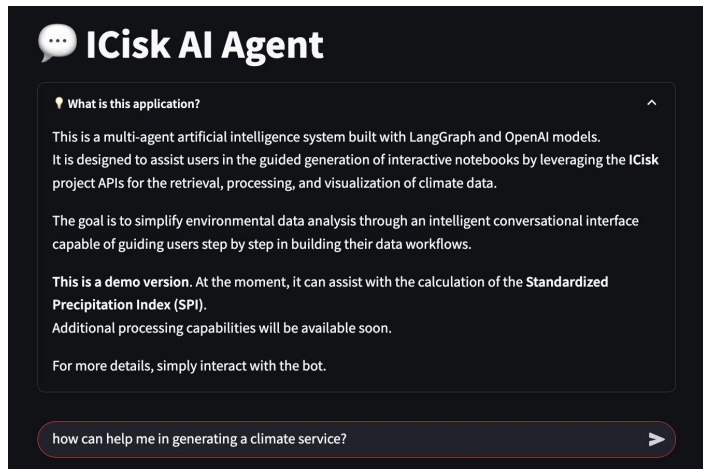
Online at: <https://i-cisk.dev.52north.org/>



**Funded by  
the European Union**



# KI-AGENT für universelle Klimainformationsdienste



1

## AI AGENT CS AutoComposer

Entwicklung eines Climate Service AutoComposer basierend auf

- LLM OpenAI API
- I-CISK Data Ingestor/Bias Correction API
- I-CISK Visualization Tools

2

## Jupyter Notebook

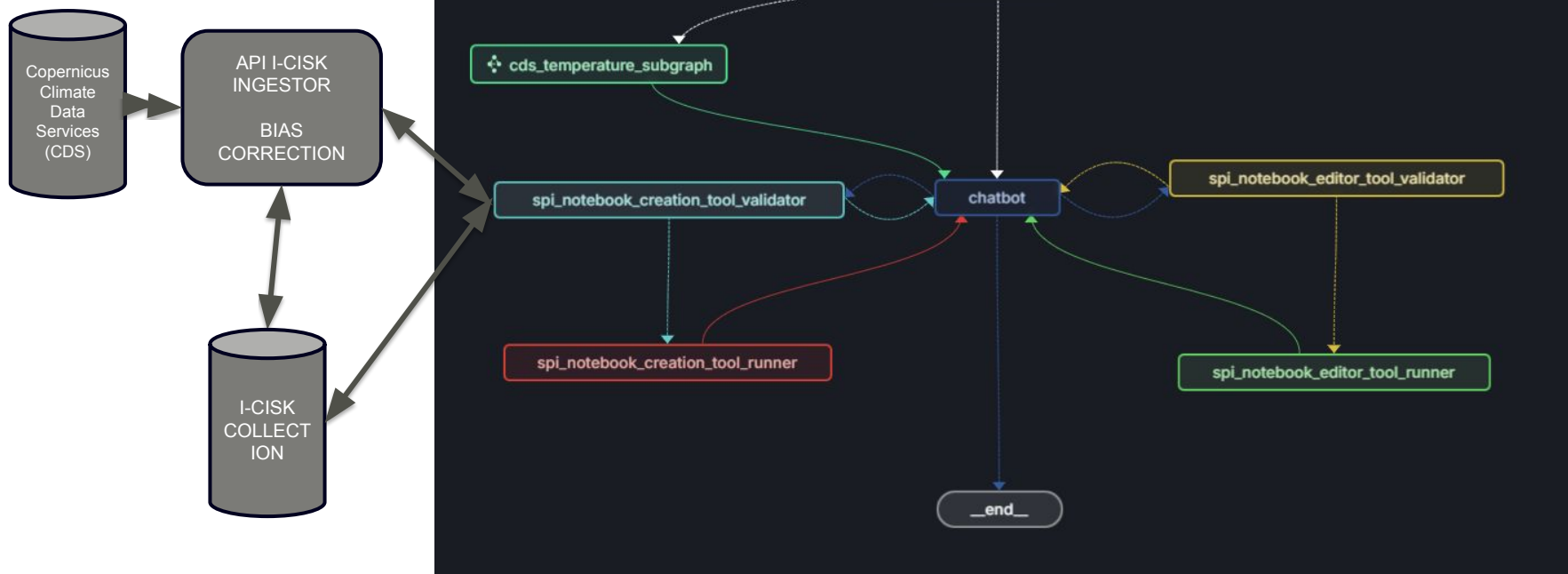
Automatische Erzeugung eines Jupyter Notebooks

3

## Visualisierung

Anbindung an Visualisierungswerkzeuge, die im Projekt entwickelt wurden

# KI-AGENT für universelle Klimainformationsdienste





# KI-AGENT für universelle Klimainformationsdienste



the bounding box is ok and the period of interest from Aug 2024 to Feb 2025



Using tool: *spi\_notebook\_creation*

Parameter	Value
region	[-9.6, 35.9, 4.3, 43.8]
reference_period	[1981, 2010]
period_of_interest	['2024-08', '2025-02']



The Jupyter notebook for calculating the Standardized Precipitation Index (SPI) for Spain has been successfully created. You can download it using the link below:

[Download SPI Notebook](C:\Users\tommaso\AppData\Local\Temp\icisk-chat\spi\_-9.6, 35.9, 4.3, 43.8]\_notebook.ipynb)

If you need further assistance or modifications to the notebook, feel free to ask!

Scrivi un messaggio



14.11.2025 - Dr. Simon Jirka - KI-unterstützte Suche nach Geodaten

```
1 import os
2 import math
3 import datetime
4 from datetime import relativedelta, relativedelta
5 import getpass
6
7 import numpy as np
8 import pandas as pd
9 import xarray as xr
10
11 import scipy.stats as stats
12 from scipy.special import gammaln, gamma
13
14 import matplotlib.pyplot as plt
15
16 !pip install "cdsapi>=0.7.4"
17 import cdsapi
18
19 !pip install cartopy
20 import cartopy.crs as ccrs
21 import cartopy.feature as cfeature
22
23
24 ... Run to see results ...
25
26
27 1 spi_ts = 1
28 2
29 3 region = [-9.6, 35.9, 4.3, 43.8] # min_lon, min_lat, max_lon, max_lat
30 4
31 5 reference_period = (1981, 2010) # start_year, end_year
32 6
33 7 period_of_interest = ('2024-08', '2025-02') # start_month, end_month
34 8
35 9 cds_client = cdsapi.Client(url='https://cds.climate.copernicus.eu/api', key=getpass.getpass)
36 10
37
38 ... Run to see results ...
39
40
41 1 filename = f'era5_land_total_precipitation_{"".join([str(c) for c in region])}_monthly_
42 2
43 3 out_dir = 'tmpdir'
44 4 os.makedirs(out_dir, exist_ok=True)
45 5
46 6 cds_out_filename = os.path.join(out_dir, filename)
47 7
48 8 if not os.path.exists(cds_out_filename):
49 9     cds_dataset = 'reanalysis-era5-land-monthly-means'
50 10     cds_query = {
51 11         'product_type': 'monthly_averaged_reanalysis',
52 12         'variable': 'total_precipitation',
53 13         'year': [str(year) for year in range(*reference_period)],
54 14         'month': [f'{month:02d}' for month in range(1, 13)],
55 15         'time': '00:00',
56 16         'area': [
57 17             region[3], # W
58 18             region[0], # W
59 19             region[1], # S
60 20             region[2] # E
61 21         ],
62 22         'data_format': 'netcdf',
63 23         'download_format': 'unarchived'
64 24     }
65 25
```

# NFDI4Earth - OneStop4All



- NFDI4Earth → Nationale Forschungsdateninfrastruktur für die Erdsystemwissenschaften
- Koordiniert durch die Technische Universität Dresden
- 52°North unterstützt die Umsetzung des OneStop4All
  - Zentrale Anlaufstelle für die Suche nach u.a.
    - Daten
    - Diensten
    - Software
    - Lernmaterial
  - Basiert auf den FAIR-Prinzipien (Findability, Accessibility, Interoperability, Reusability)



# NFDI4Earth - OneStop4All



 User Support

 About NFDI4Earth

I commit to FAIR data practices and sign the [NFDI4Earth Commitment](#).

ALL RESOURCES

Search for data repositories, software tools, information or research data materials

 SEARCH

FIND INFORMATION, GUIDANCE AND RESOURCES THAT FIT YOUR RESEARCH NEEDS

Discover datasets and repositories

Find resources, search data or take a guided tour on repositories

Explore tools and services

Learn more about existing services and software, and how to use them

Improve your skills and start learning

Improve your skills with interactive self-learning materials and events

ASK OUR BOT

# NFDI4Earth - OneStop4All

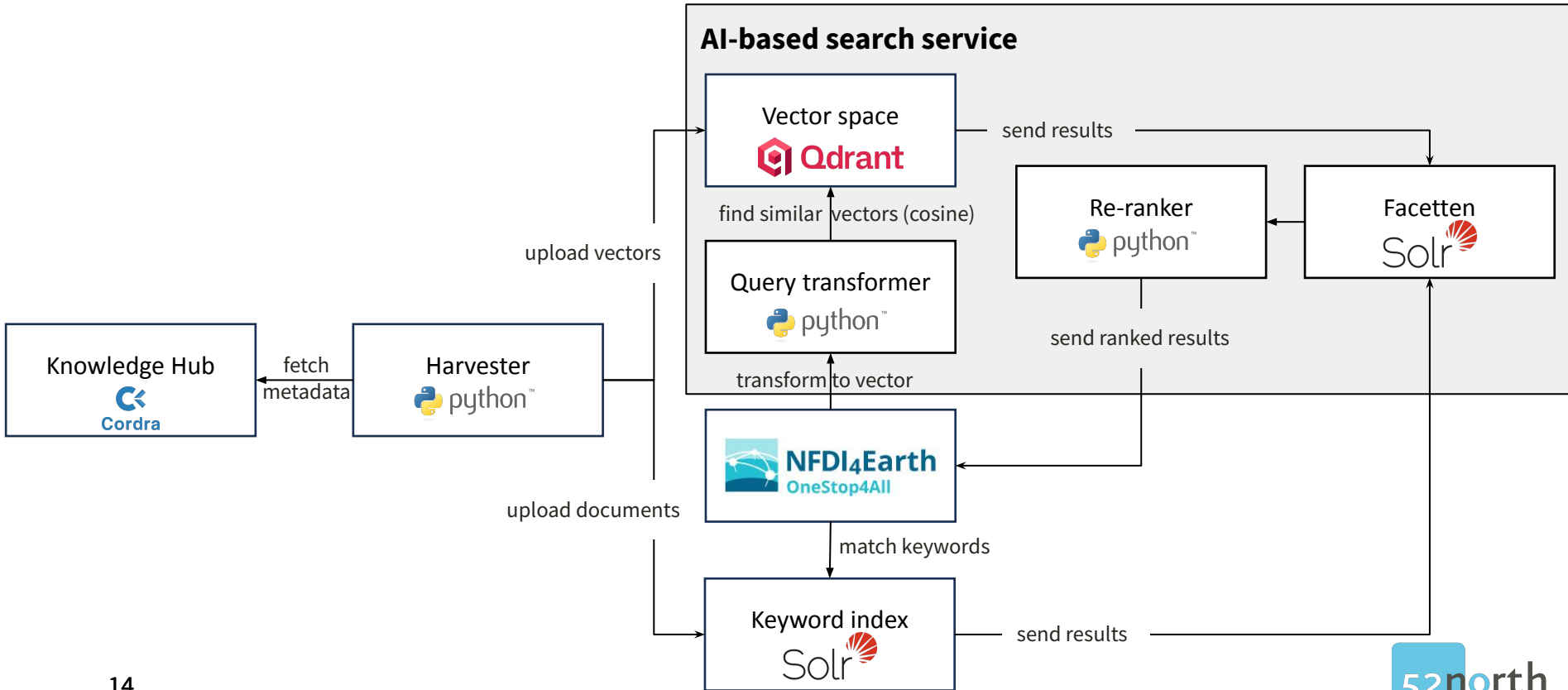
- **Ziel:** Implementierung einer KI-basierten Suche im OneStop4All
- **Gegenwärtiger Zustand**
  - Auf Basis der Metadaten im Knowledge Hub wird mit Hilfe eines Harvesters ein Solr-Index erstellt, um die Anfrage-Performance zu verbessern.
  - Such-Anfragen werden an diesen Solr-Index geschickt (keyword-based matching).
- **Idee für eine KI-basierte Suche:**
  - Transformation der Metadaten im Knowledge Hub in Vektordaten (“Embeddings”) zur Erstellung eines Vektor-Speichers
  - Transformation der Query in einen Vektor  
→ KI findet die Suchergebnisse nach semantischer Ähnlichkeit der Vektoren
- **UI:** KI-Suche ist voll integriert und läuft im Hintergrund

# Konzept: Herausforderungen

- Nicht alle Ressourcen-Typen haben eine textuelle Beschreibung und Keywords
  - Semantische Suche (Vektorsuche) funktioniert am besten mit (viel) Volltext
- Viele Datensätze haben schlechte Metadaten, z.B.
  - Beschreibung ist nicht vorhanden oder sehr kurz
  - komplette JSON-Dokumente als Beschreibung
- Schlechte Metadaten führen eventuell zu Rauschen bei der Vektorsuche
- Möglicher Ansatz: nur Ressourcen-Typen und Dokumenten mit Volltext für Vektorsuche verwenden
  - Die restlichen Dokumente wären dann nur über Keyword-Suche auffindbar



# Übersicht der Komponenten



# Ausblick

- KI bietet wichtige Ansätze um die Suche nach Geodaten zu vereinfachen und zu verbessern
- Metadaten
  - Unterstützung bei der Erstellung von Metadaten
  - Überprüfung von Metadaten
- Suchfunktionen
  - Dialogbasierte Suche nach Geodaten
  - Berücksichtigung von Semantik
  - Zusammenhänge zwischen Datensätzen
  - Umgang mit alternativen Bezeichnungen, Schreibweisen und Namen

# Vielen Dank für die Aufmerksamkeit

- Simon Jirka
- [jirka@52north.org](mailto:jirka@52north.org)
- <https://52north.org>