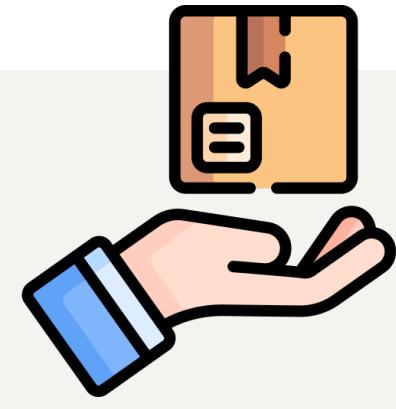


Advancing FAIRness for global air quality data analyses

Designing the last mile of FAIRness in the multi-service data infrastructure of the Tropospheric Ozone Assessment Report (TOAR-II) and Artificial Intelligence for Air Quality (IntelliAQ)

A. Mozaffari^{ID}, S. Schröder^{ID}, M. Romberg^{ID}, E. Epp^{ID}, N. Selke^{ID}, J. Ahring^{ID}, C. Betancourt^{ID}, L. H. Leufen^{ID}, F. Kleinert^{ID}, M.G. Schultz^{ID}
Contact: a.mozaffari@fz-juelich.de Jülich Supercomputing Centre (JSC), Forschungszentrum Jülich GmbH, Germany

Products & Services



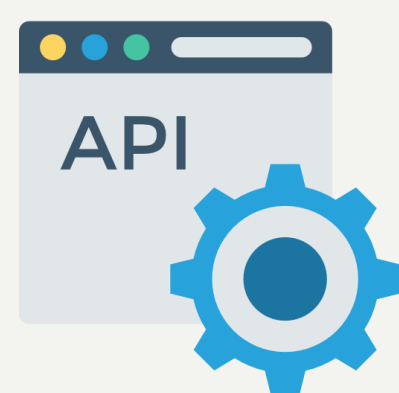
The website provides linkage to other ozone data repositories

TOAR:DATA PORTAL



GEOspatial Point Extraction and Aggregation Service with REST API interface

A Representational State Transfer (API) allows querying all meta-data and data in TOAR. Statistical aggregations on-the-fly; comprehensive analysis tools in preparation.

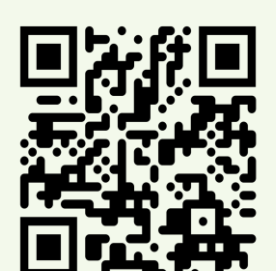


Data publishing services for datasets accompanying TOAR, in collaboration with EUDAT (B2SHARE)

About the project



TOAR
tropospheric ozone assessment report



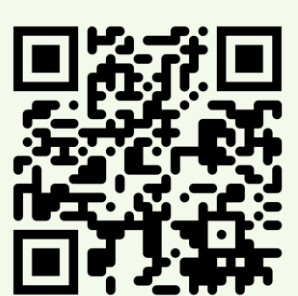
TOAR₂ is a joint effort to provide an up-to-date scientific assessment of tropospheric ozone's global distribution and trends from the surface to the tropopause.

TOAR Phase I : (2014 – 2019)
TOAR Phase II : (2020 – 2024)

TOAR community consists of >150 members from 35 countries who are organized into 14 working groups to produce new analyses of the tropospheric ozone distribution and trends.



IntelliAQ

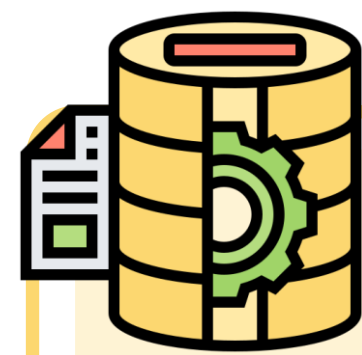


IntelliAQ₃ is a European project developing novel deep learning approaches for the analysis and synthesis of global air quality data based on deep neural networks by building a linkage of several different types of data.



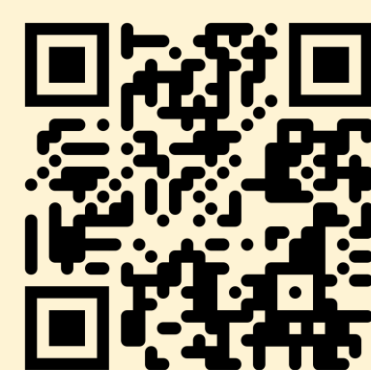
Certification

We have applied for Core Trust Seal to improve data and metadata FAIRness, data license transparency, and long-term documentation, storage, and backup strategies.



TOAR Data & Database₁

toar-data.org



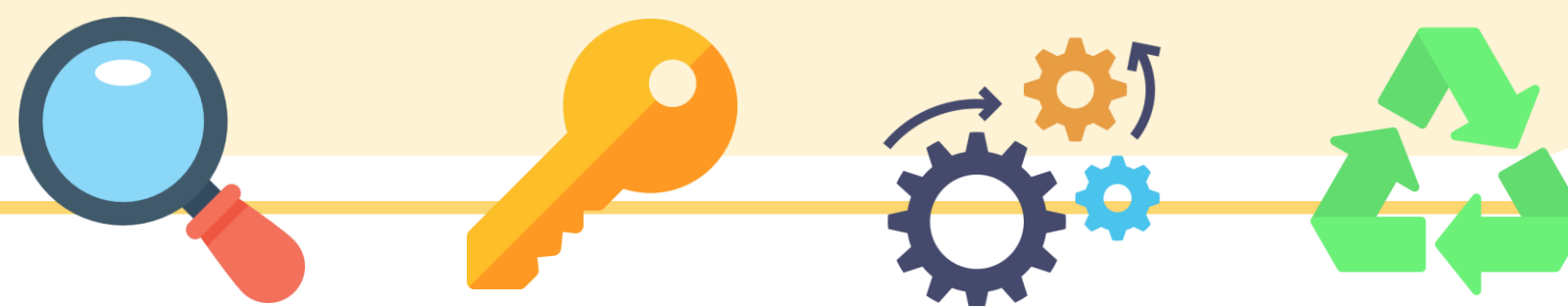
TOAR-II data portal for global measurements of ozone and its precursors

toar-data.fz-juelich.de



quality-controlled surface ozone measurements and its precursors from various environmental agencies and programs, and also from universities and individual researchers.

Via this harmonized data collection, it is possible to provide globally consistent metrics for analyses of health, vegetation, and climate impacts from ozone air pollution and analyses of air pollution trends globally. FAIR principles and the desire to build a trustworthy repository guided the development of our metadata schemas and workflows.



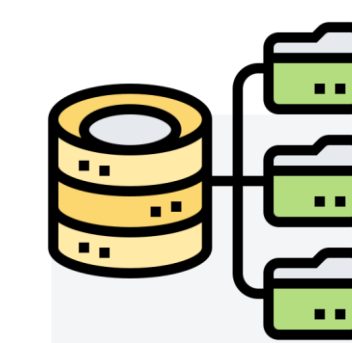
What is next?

IntelliAQ and TOAR aim to produce datasets that can be reused for more than a decade. Besides its main role as a community data repository, the TOAR data centre acts as a platform to test novel, high-performance workflows for heterogeneous data sets, primarily in the context of machine learning applications for air quality forecasting, interpolation and quality control.

For busy people



TOAR and IntelliAQ are building a multi-level data service infrastructure for air quality and weather data. FAIR principles and modern data science guide the design at every level.



ML workflow based on TOAR

Layered Analysis Module



We envision a workflow where user requests for global trend analyses will be efficiently calculated and saved in a caching system. Each query will be saved as a Fair Digital Object (FDO) to make them traceable and reusable. Also, scientists can reference their analyses by a Persistent Identifier (PID) without the need to duplicate data at any step during the workflow.

ML AIR



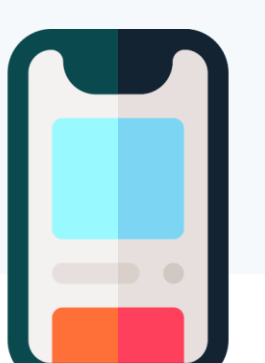
Data from the TOAR data portal are used for different machine learning (ML) applications. For example, the standard workflow of the tool MLAir₄ uses the API interface of the TOAR data portal to automatically load data and create forecasts based on ML approaches such as deep neural networks.

TOAR:APP



STIFTERVERBAND

in conjunction with the TOAR data and AI application, we have proposed the TOAR: APP that provides an easily accessible near-real-time assessment of ozone damage on vegetation for the agricultural and local communities. The proposal won the Open Data Award by Stifterverband - Germany in 2021 and it is currently in the development phase.



1. Schröder, S., Schultz, M.G., Selke, N., ... et al (2021), TOAR Data Infrastructure , <http://doi.org/10.34730/4d9a287dec0b42f1aa6d244de8f19eb3>
2. Schultz, M.G., Schröder, S., Lyapina, O., ... et al., (2017), Tropospheric Ozone Assessment Report: Database and Metrics Data of Global Surface Ozone Observations, Elem Sci Anth, 5, p.58
3. Schultz, M.G., Kleinert, F., Leufen, L.H., ... et al. (2022) Artificial intelligence for air quality , <https://doi.org/10.54050/PRJ1218384>
4. Leufen, L.H., Kleinert, F., Schultz, M.G., (2021) MLAir (v1.0) – a tool to enable fast and flexible machine learning on air data time series, <https://doi.org/10.5194/gmd-14-1-2021>



IntelliAQ is funded by the EU's ERC programme, Grant Agreement 78576.