

# TOAR-II Data

Martin G. Schultz<sup>id</sup>, Sabine Schröder<sup>id</sup>, Niklas Selke<sup>id</sup>, Eleonora Epp<sup>id</sup>, Mathilde Romberg<sup>id</sup>,  
Jianing Sun<sup>id</sup>, Jessica Ahring, Amirpasha Mozaffari<sup>id</sup>, Max Lensing, Clara Betancourt<sup>id</sup>,  
Lukas Leufen<sup>id</sup>, Björn Hagemeyer<sup>id</sup>, Rajveer Saini<sup>id</sup>, Sander Apweiler<sup>id</sup>

*Jülich Supercomputing Centre, Forschungszentrum Jülich, Germany*



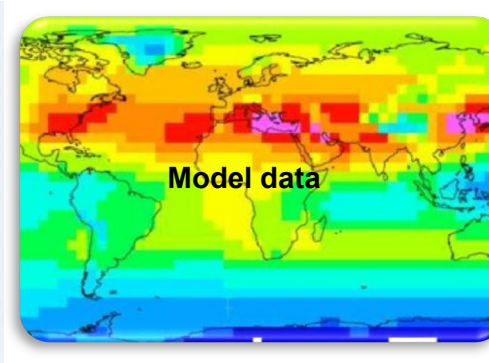
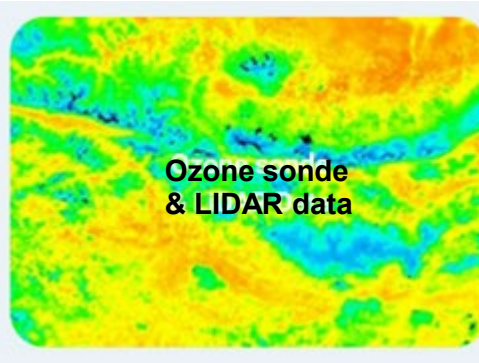
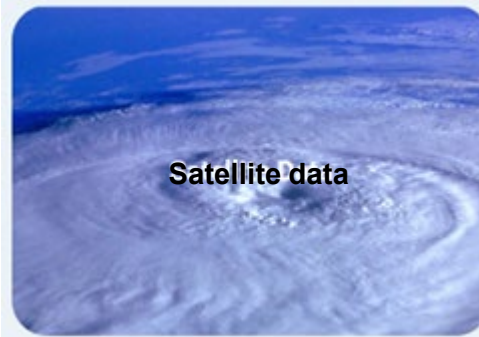
TOAR-II Workshop  
virtual, 16-18 Nov 2021



# What kind of data do we expect in TOAR-II?



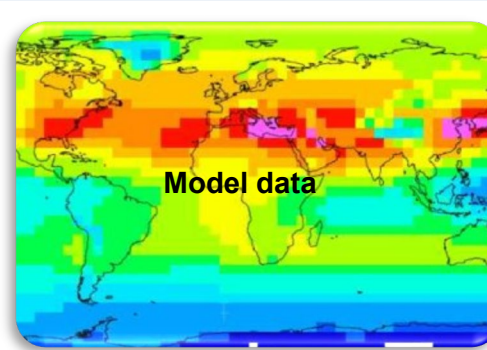
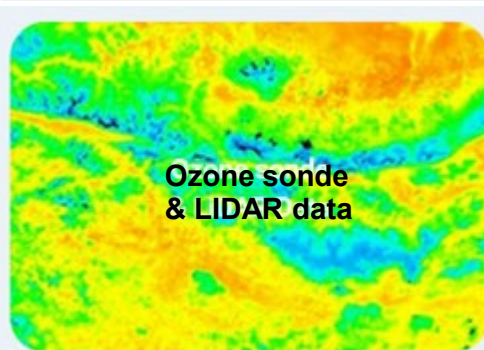
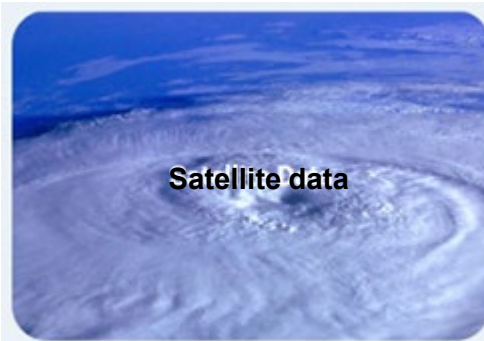
~15,000 stations  
10 variables  
~ 20 bio data points



# How will TOAR-II data be managed?



TOAR database



TOAR data portal



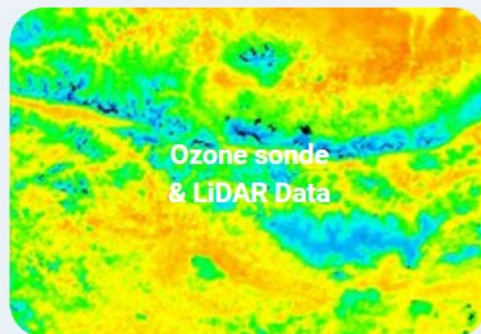
A one-stop-shop to locate and access tropospheric ozone data from a large variety of measurement platforms

# TOAR Data Portal

The Home of Tropospheric Ozone Data



<https://toar-data.org>



Latest News



**TOAR**  
tropospheric  
ozone  
assessment  
report  
*Phase II*





## IAGOS DATA PORTAL

In-service Aircraft for a Global Observing System (IAGOS) is a European Research Infrastructure for global observations of atmospheric composition from commercial aircraft. ... [More](#)

[REST:API](#)[Contact](#)[License](#)

## ATOM TRACKING CHEMISTRY,

ATOM provides information on greenhouse gases and human-produced air pollution, including atmospheric carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), tropospheric ozone (O<sub>3</sub>), and black carbon (BC) from airborne campaigns conducted by NASA, ... [More](#)

[REST:API](#)[Contact](#)[License](#)

links to original web site

<https://toar-data.org>



# Contribute descriptions and links to data: This is a task for the TOAR-II WGs!



**TOAR:DATA PORTAL** Home News Data Access ▾ Get In Touch ▾ About ▾

## Links and descriptions of other data centres or repositories for inclusion in the TOAR data portal

Add an entry to the TOAR data portal via the following online form, or download the **registration-form**, with explanations, fill it in and send it back to us. You can also use these forms to request an update to an existing entry (please indicate this) or inform us about an error in one of the listed entries.

**Repository registration form**

Name  Email

Data Category

Satellite

[https://toar-data.org/contribute/#contribute\\_toar](https://toar-data.org/contribute/#contribute_toar)

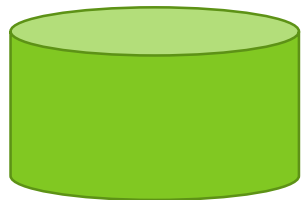
Web URL of data set description

Logo for the dataset ( PNG,or SVG ) , Maximum file size 2 MB.  
Durchsuchen... Keine Datei ausgewählt.

## 2

## The TOAR Database Infrastructure

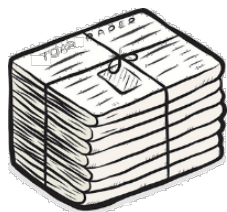
... comprises of



The TOAR database of global ground-level air quality observations



Web services for accessing and processing of TOAR data and metadata



Publication services for original TOAR datasets

011001100  
100010010  
011101010  
010110011  
010010110

Tools, source code, and documentation



# 2

## The TOAR Database in a nutshell

### Contents:

- Harmonized and quality-controlled surface ozone measurements and related data from all over the world (one of the largest collections of global air quality data)
- Globally consistent metadata to describe station characteristics and aid the interpretation of ozone data

### Purpose:

- To provide globally consistent metrics for analyses of health, vegetation, and climate impacts from ozone air pollution

### Data sources:

- Various environmental agencies and programs
- Universities and individual researchers
- OpenAQ (in version 2)
- COSMO (version 1) and ERA-5 (version 2) reanalysis data



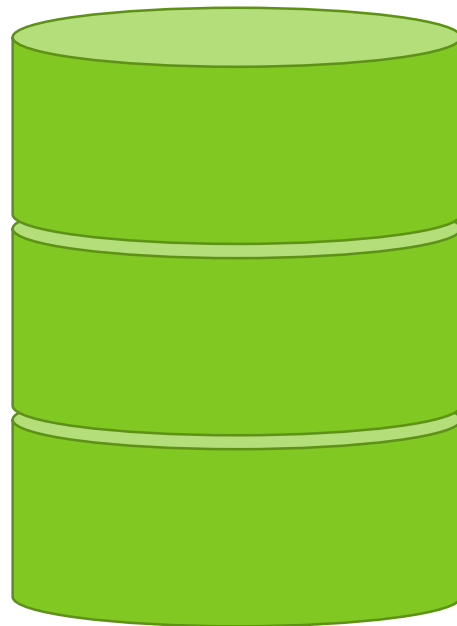


## Global ozone-related air quality and weather data

(as of August 2021)

**13,257 stations****103,735 time series****~12,000,000,000 data points**

Ozone, NO, NO<sub>2</sub>, CO, PM<sub>2.5</sub>,  
PM<sub>10</sub>, ethane, propane,  
benzene, toluene, irradiance,  
cloud cover, T, RH, Q, u, v,  
wdir, wspeed, Rn



~35 contributing networks and  
research groups

harmonized data and metadata  
enhanced station metadata

1970 – 2013  
some datasets extended to  
2017



The TOAR database version 2 is under construction. We expect that most data will be inserted by Q3/2022.



## TOAR Data Infrastructure

### Mission Statement

The Tropospheric Ozone Assessment Report (TOAR) data centre is the central hub for data access in support of research assessing the impacts of ozone air pollution on human health, vegetation, and climate. Besides maintaining a [data portal](#) with links to ozone data sets from research organisations all over the world, we operate a database of harmonised surface ozone measurements and related data. This is one of the largest collections of quality controlled air pollution measurements in the world. All data in the database are easily accessible through open, freely available and well documented web services. The TOAR data centre team is committed to the FAIR principles and aims to achieve the highest standards with respect to data curation, archival, and re-use.

## TOAR V2

TOAR V2 supports the [second phase](#) of the TOAR activity

### Services

- REST API to access TOAR V2 (will soon be available)
- REST API service for flux-based vegetation damage assessment (DO3SE) (will soon be available)

### Documentation

#### User Guides

- [Data Portal](#) (DRAFT-pdf) - Brief Introduction
- [Database](#) (DRAFT-pdf) - Content, access, variables, metadata and FAIR Data self-assessment
- [Data Submission](#) (DRAFT-pdf) - Data contributions' HowTo

#### Technical Guides

- [Infrastructure](#) (in preparation) - setup of systems and services
- [Data Processing](#) (in preparation) - data harmonisation and curation
- [OAI Mapping](#) (DRAFT-pdf) - TOAR Data Infrastructure and the Open Archival Information System reference model

#### General Guides

- [Glossary](#) (DRAFT-pdf)

<https://toar-data.fz-juelich.de>

## TOAR V1

Martin G. Schultz

TOAR-II workshop, 16-18 Nov 2021



TOAR  
tropospheric  
ozone  
assessment  
report  
Phase II

JÜLICH  
Forschungszentrum | JÜLICH  
SUPERCOMPUTING  
CENTRE

TOAR Data User Guide #3

## The TOAR Database User Guide

Version 1.0 | 02 August 2021



Forschungszentrum Jülich GmbH  
ESPE LUIS - 600  
524

## User guides

TOAR  
tropospheric  
ozone  
assessment  
report  
Phase II

JÜLICH  
Forschungszentrum | JÜLICH  
SUPERCOMPUTING  
CENTRE

TOAR Data Technical Guide #2

## TOAR Data Input and Processing

Version 0.7 | 13 July 2021



Forschungszentrum Jülich GmbH  
FSDF LUIS - FSD

## Technical guides

## Detailed documentation (gitlab pages)

toardb\_fast

TOAR II

### Models

#### Contact

Name	Type	Description	Required
id	integer	for internal use only	Yes
person	Person	A contact is either a person or an organisation	No
organisation	Organisation	A contact is either a person or an organisation	No

#### Coordinates

Name	Type	Description	Required
lat	number	longitude coordinate of station (decimal degrees_north). This is our best estimate of the station location which is not always identical to the official station coordinates (see potential changelog entry).	No
lng	number	latitude coordinate of station (decimal degrees_east). This is our best estimate of the station location which is not always identical to the official station coordinates (see potential changelog entry).	No
alt	number	altitude of station (in m above sea level). This is our best estimate of the station altitude, which is not always identical to the reported station altitude, but frequently uses the elevation from google earth instead (see potential changelog entry).	No

#### Data

Name	Type	Description	Required
------	------	-------------	----------

#### TOAR II Database

- Models
  - Contact
  - Coordinates
  - Data
  - Organisation
  - Person
  - Stationmeta
  - StationmetaAnnotation
  - StationmetaAuxDoc
  - StationmetaAuxImage
  - StationmetaAuxUrl
  - StationmetaChangelog
  - StationmetaCore
  - StationmetaCoreBase
  - StationmetaGlobal
  - StationmetaRole
  - Timeseries
  - TimeseriesAnnotation
  - TimeseriesChangelog
  - TimeseriesProgramme
  - TimeseriesRole
  - Variable
  - Geolocation URLs
- Controlled Vocabulary
  - Role Code
  - Role Status
  - Kind Of Annotation

- PostGIS database (supports geographic queries)
- Improved metadata schema:
  - Data versioning & documented QC
  - Better attribution to PIs, providers, etc (roles)
  - Use of controlled vocabulary
  - More flexibility to handle provider-specific metadata
  - Automated workflows including statistical QC
  - Station characterisation from higher resolution geospatial data
- Clear data license and data use policy (CC-BY 4)
- Provision of hourly data





## Attribution 4.0 International (CC BY 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).

### You are free to:

- Share** — copy and redistribute the material in any medium or format
- Adapt** — remix, transform, and build upon the material for any purpose, even commercially.

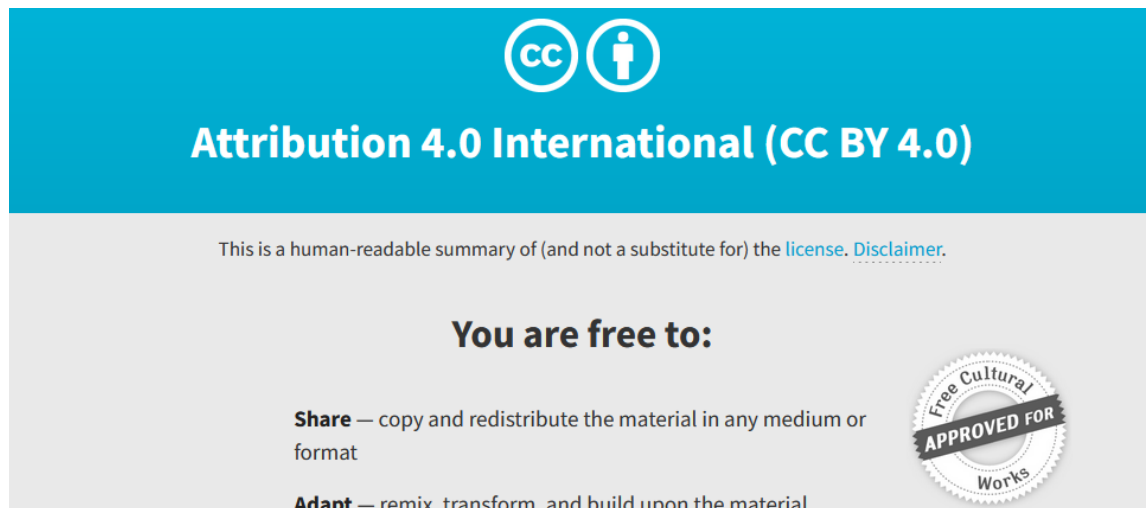
The licensor cannot revoke these freedoms as long as you follow the license terms.

---

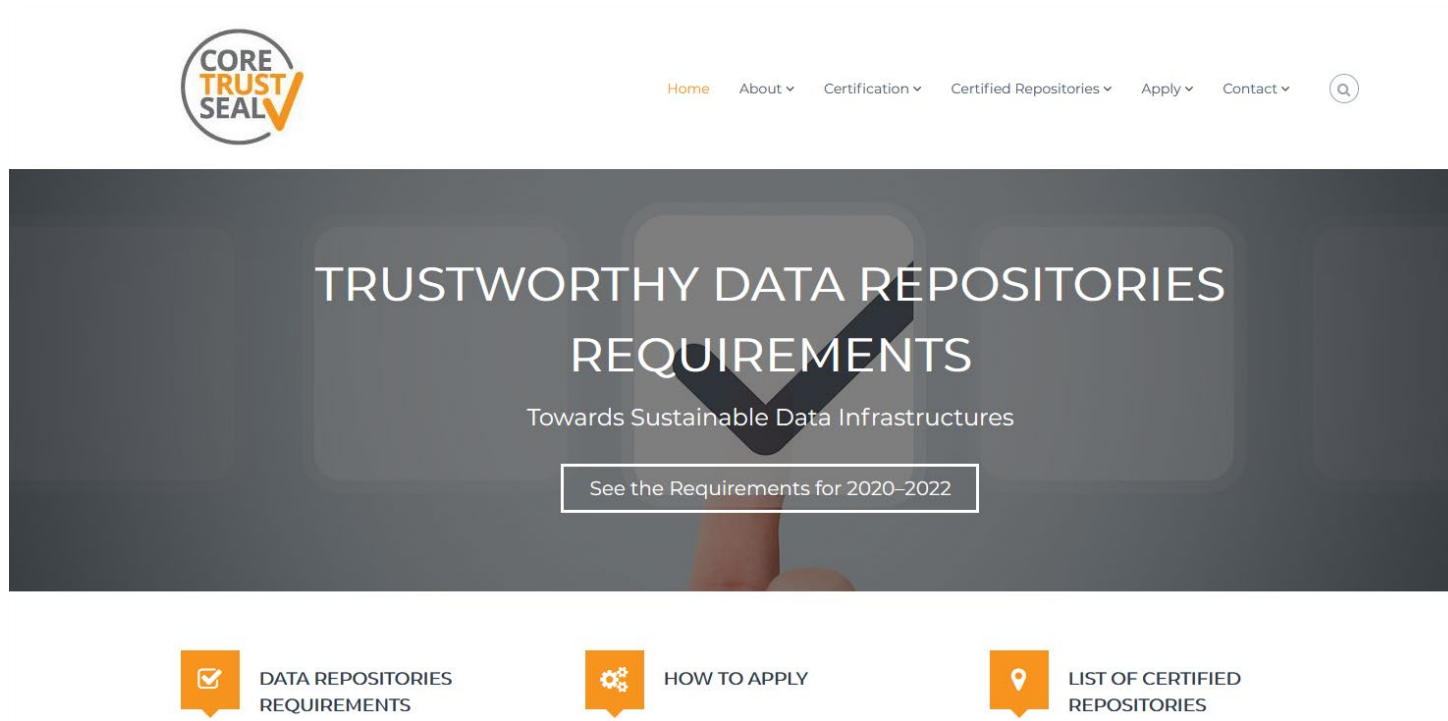
### Under the following terms:

**Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.





**Important consequence:**  
We must ask you (as data provider) for  
consent that we may publish your data  
under this license!  
No consent = no data!



Application submitted in September 2021

# TOAR-II Data and Services Timeline



2022 2023  
I D | J | F | M | A | M | J | J | A | S | O | N | D



Collect requests for API features

Update time series (Q2/2023)



FAIR workflows  
Data screening  
Documented processing

Statistics  
Data aggregation  
Metadata, formatting, ...



## TOAR-II Data training workshops

### 3 TOAR-II data training workshops planned:

- May – July 2022
- One onsite event + 2 virtual events
- Hackathon style (lectures + hands-on training)

### Program:

- Guided tour through database
- Use of REST API
- Develop Jupyter notebooks for data analysis
- Data aggregation and statistical analyses





<https://www.shadestation.co.uk/designer-watches/Ice-Watch/Ice-Clock/Ice-Clock--Wall-Clock-/IWF-RD>

... if there is time,  
let us try out the REST API:

<https://toar-data.fz-juelich.de/api/v2/>

Example 1

stationmeta, list

Example 2

stationmeta, single station

Example 3

time series metadata

Example 4

ozone data, single station



# Acknowledgements

This work is funded through the programme Engineering Digital Futures of the Helmholtz Association's research field Information and the European Research Council, H2020 Research Infrastructures, grant IntelliAQ (787576).

