

3TU.DATACENTRUM SERVES DATA-CURATION NEEDS OF GEODESY, HYDROLOGY, AND CLIMATE SCIENCE IN NL

by **Alenka Prinčič, Jeroen Rombouts, Madeleine de Smaele.**
 TU Delft Library, Prometheusplein
 1, 2628 ZC Delft, NL.



What is 3TU.Datacentrum?

A scientific data archive & data lab, a platform for science- and technology research data, built in collaboration of three universities of technology in the Netherlands. It serves the data-curation needs of scientists in different disciplines of engineering varying from water management, hydraulic engineering, and remote sensing to nanoscience. This versatility in disciplines and thus versatility in needs is a challenge of the data center.

Why use 3TU.Datacentrum?

The data repository provides research communities with:

- ❑ advice on data management
- ❑ a long-term archive for scientific data
- ❑ permanent access to research data
- ❑ tools for reuse of research data
- ❑ support in dynamic sharing of data

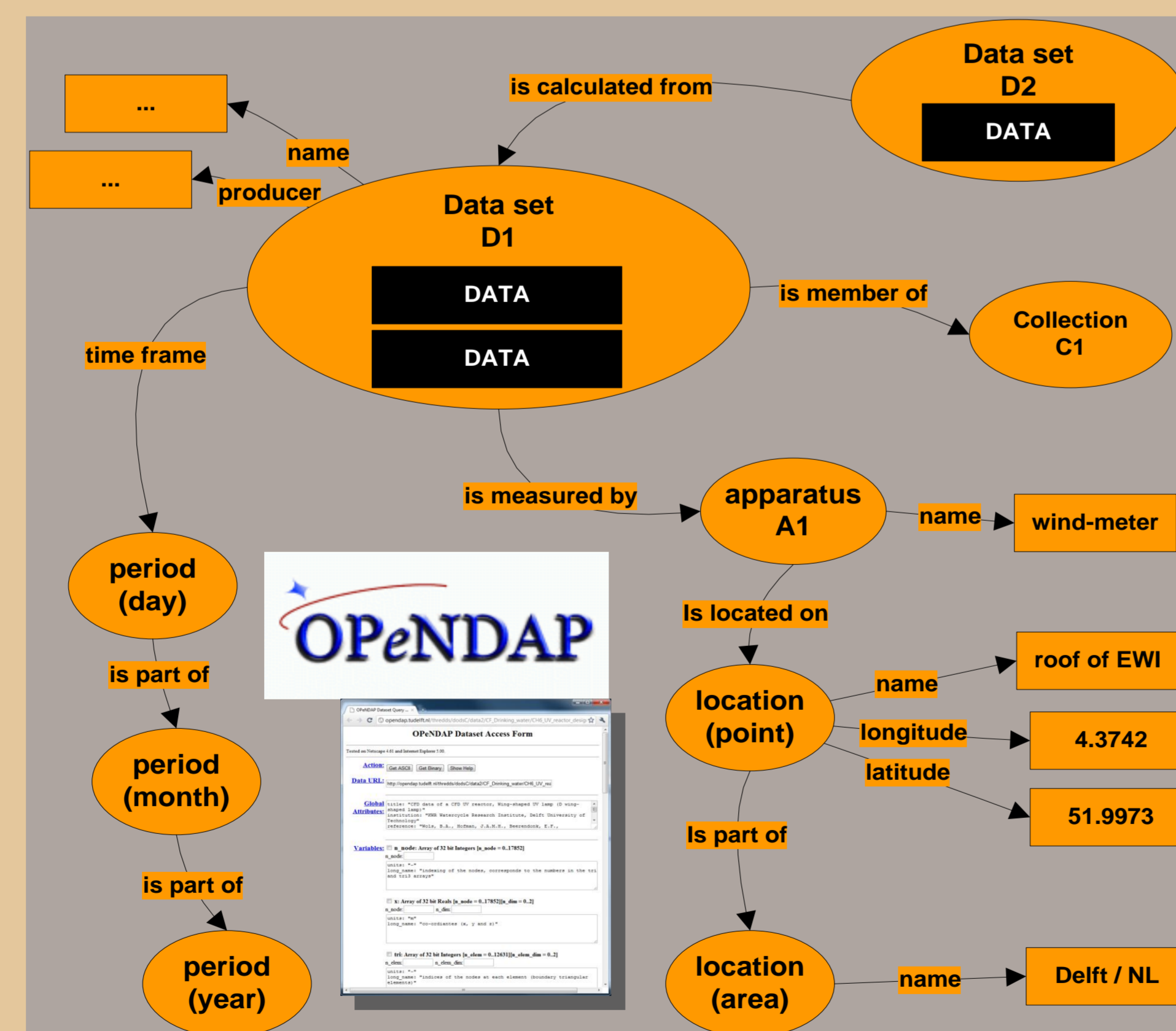
Ever worried that your data were lost due to a CD-ROM that is no longer readable? A data center can help you avoid such risks.

What does it offer?

3TU.Datacentrum offers solutions for different needs of data management in geosciences and other disciplines. We show two diverse cases – one (A) with simple data sets within a complex collection and one (B) with complex data sets within a simple collection.

(A) Complex collection of simple data sets.

Consists of hydrological measurements in a river basin in Luxemburg with several project partners collecting data with several instruments on several locations.



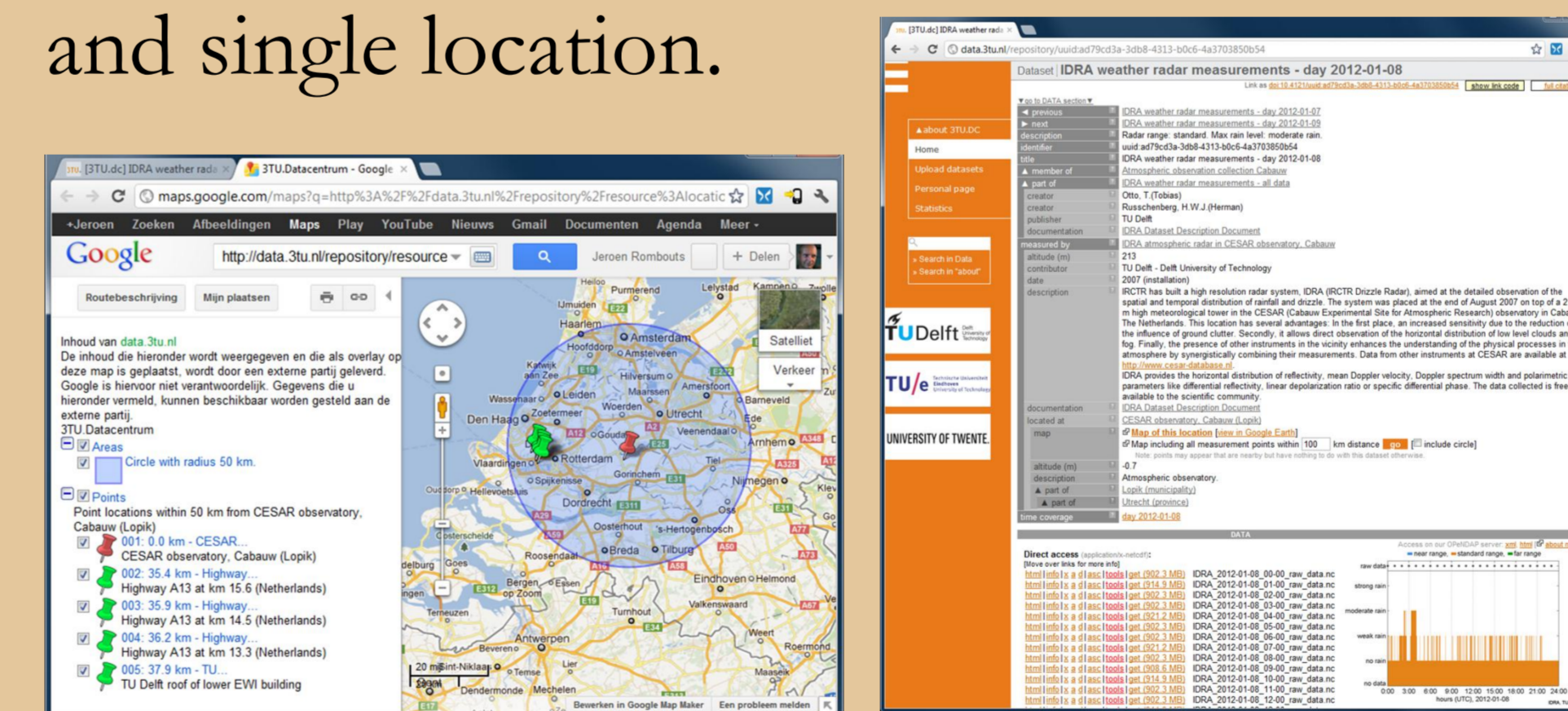
The data model in 3TU.Datacentrum: www.data.3tu.nl.

Due to multiple users there was a need for central storage to be able to share the data,

and for normalization to enable easy combination of data sets. Because the measurements are not repeatable, a proper preservation of data is essential. This collection is stored in a model where the datasets, instruments, locations and even periods of time are all objects in their own right, with own metadata and interconnected through rdf relationships, the cornerstone of linked data and the semantic web. Alongside NetCDF also other formats such as xml or xls can be used, which offers extra functionality to the user. The authorship and the citability of this data collection are secured by a persistent identifier (DOI).

(B) Simple collection of complex data sets.

Consists of data from a single instrument and single location.



The measurements of the drizzle radar produce huge amounts of high-resolution temporal and spatial data. Due to the volume of the data, a proper preservation that enables easy sharing of the data is essential.

Both raw and processed data are stored including a quick-look file and a user-friendly diagram. The collection has a DOI to secure the ownership and citability when the data sets are reused. Using OPeNDAP protocol queries within and across datasets are made possible.

	A (Darelux)	B (IDRA)
size	small	large
storage location	Fedora	OPeNDAP
use (dip)	NetCDF, xml, xls, csv	NetCDF
OPeNDAP value	Glue together	slice apart
metadata	data preview	quick look
data	processed	raw & processed
persistent identifier	DOI	DOI

3TU.Datacentrum offered solutions - as standardized as possible yet different where required - for both cases. Both data collections are now safely stored and available to scientific community for future research.

Future perspective.

In a similar way 3TU.Datacentrum can serve the project RainGain in which data will be collected using advanced radar technologies and covering different geographical areas (see poster nr. 2400, session HS7.1).

