



DKRZ

DEUTSCHES
KLIMARECHENZENTRUM

Publication of climate data

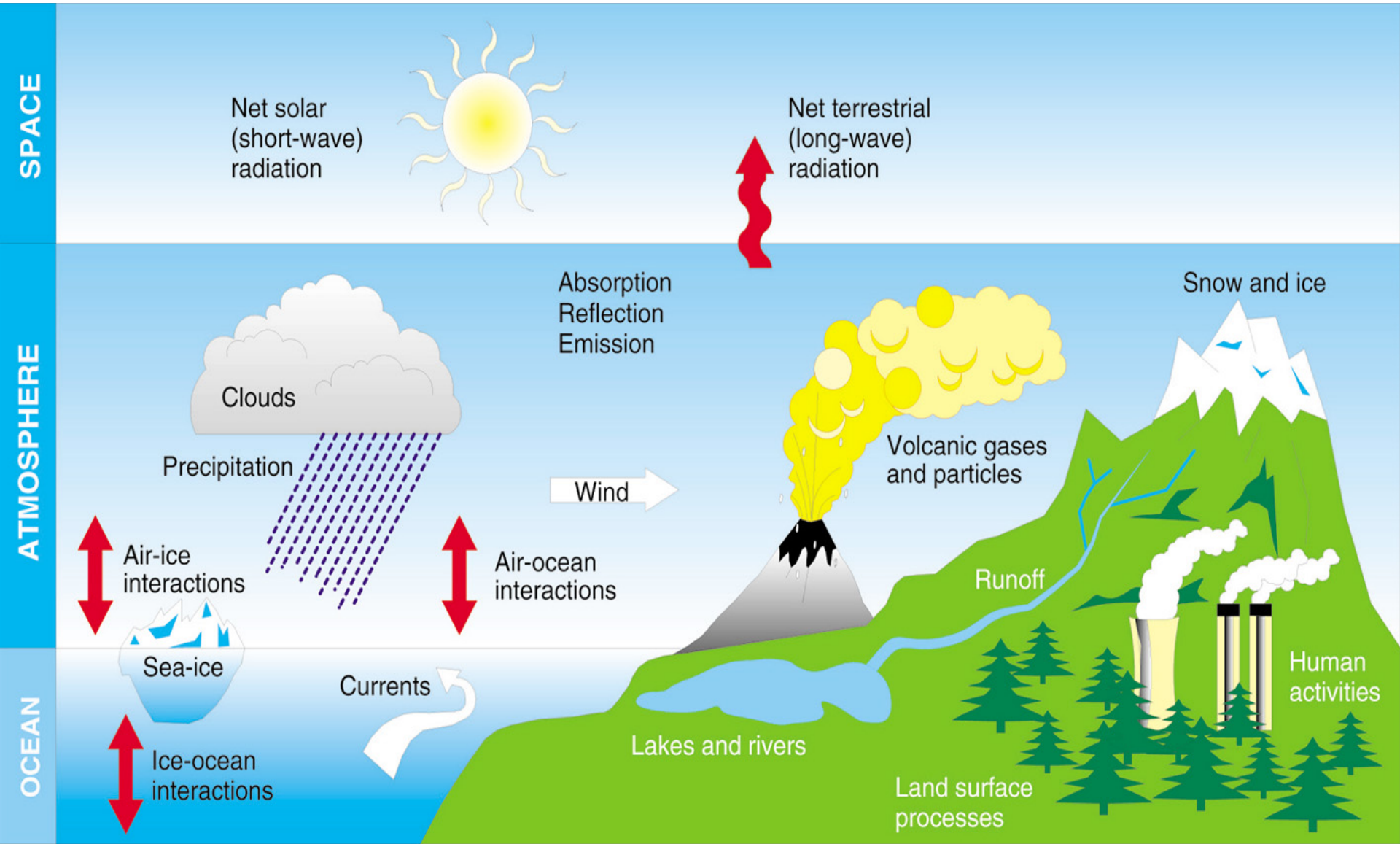
std-doi Publication of Climate Data at WDCC
DataCite Summer Meeting 7./8. June 2010

*Heinke Höck
World Data Center for Climate
(WDCC)*

Content

- Climate Data and Metadata at WDCC
- Preconditions
- Workflow
 1. Permission
 2. SQA - Scientific Quality Assurance
 3. TQA - Technical Quality Assurance
 4. Publication
- Future

Climate System



Climate Data at WDCC

- Climate model results from global and regional climate models from different climate modelling centres

CCCma, CCSR/NIES, CSIRO, GFDL, HADLEY, MPIfM , NCAR
based on IPCC-emission scenarios

- Data from scientific projects

HOAPS (satellite data), CARIBIC (civil aircraft data), GOP, COPS

- Model like Observations

Reanalyses data

General Statistics and Structure of Data

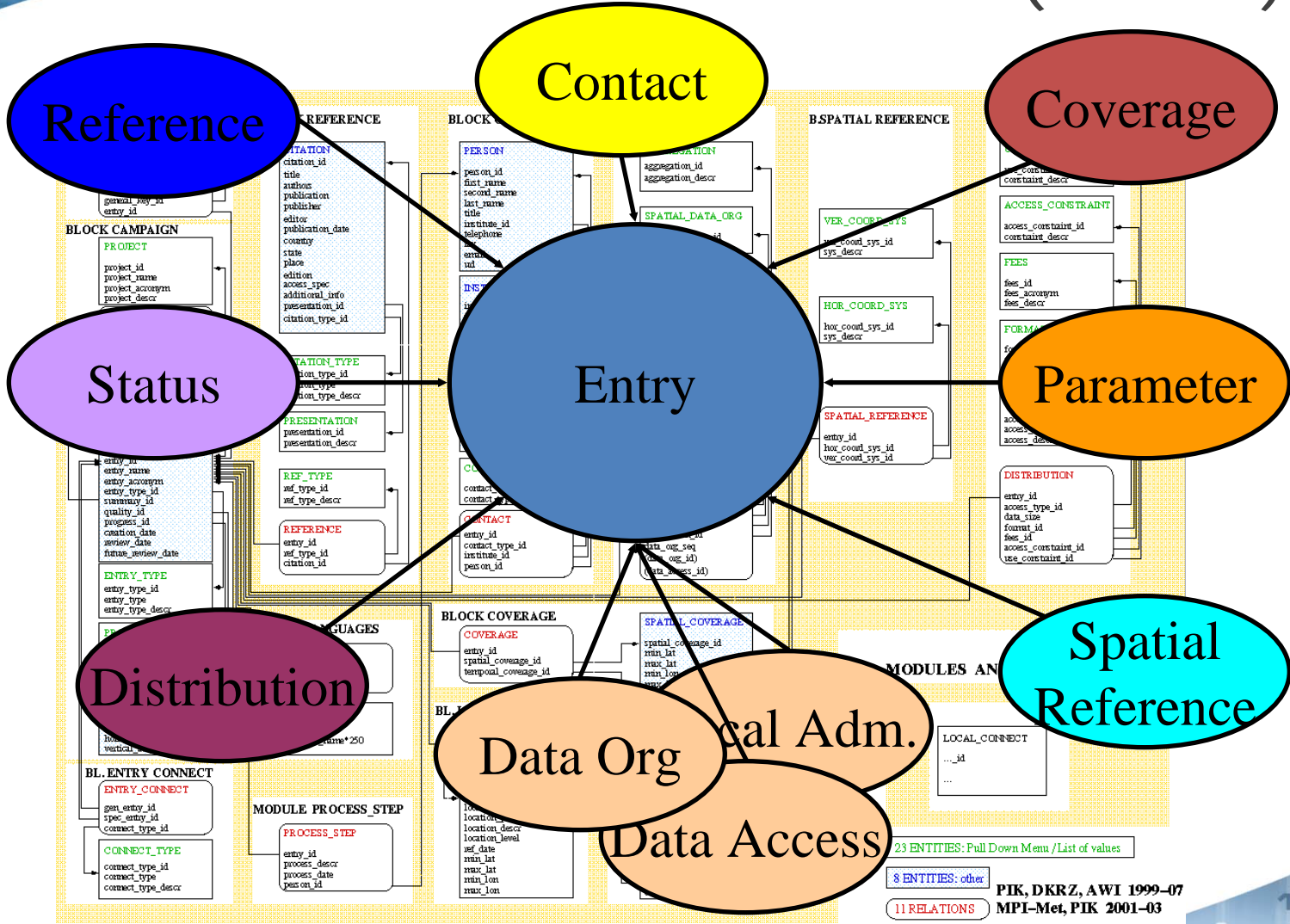
WDCC Database Size: 428 Tbyte

EXPERIMENTS: 1400
std-doi publication

collection of datasets

DATASETS: 170 000
formats: GRIB (WMO), NetCDF ...

Metadata at WDCC (CERA2)



std-doi publications at WDCC

- **Climate model results**

std-doi publication scheme is developed

Data publications can be obtained from TIB library catalogue

Future: implementation for IPCC-AR5

- **Data from scientific projects**

Project funded by DFG together with University of Bonn and the Bonn-Rhine-Sieg University of Applied Sciences

Development of an **automated** standard procedure for std-doi publication process

Preconditions

long term availability of data at WDCC
long term availability of metadata at WDCC
open access to data and metadata

Workflow Processes

	Permission	SQA	TQA	Publication
Scientist				
WDCC				
TIB				

Scientific Quality Assurance – SQA

TIME

Technical Quality Assurance – TQA

Who is allowed to initiate the std-doi
publication process ?

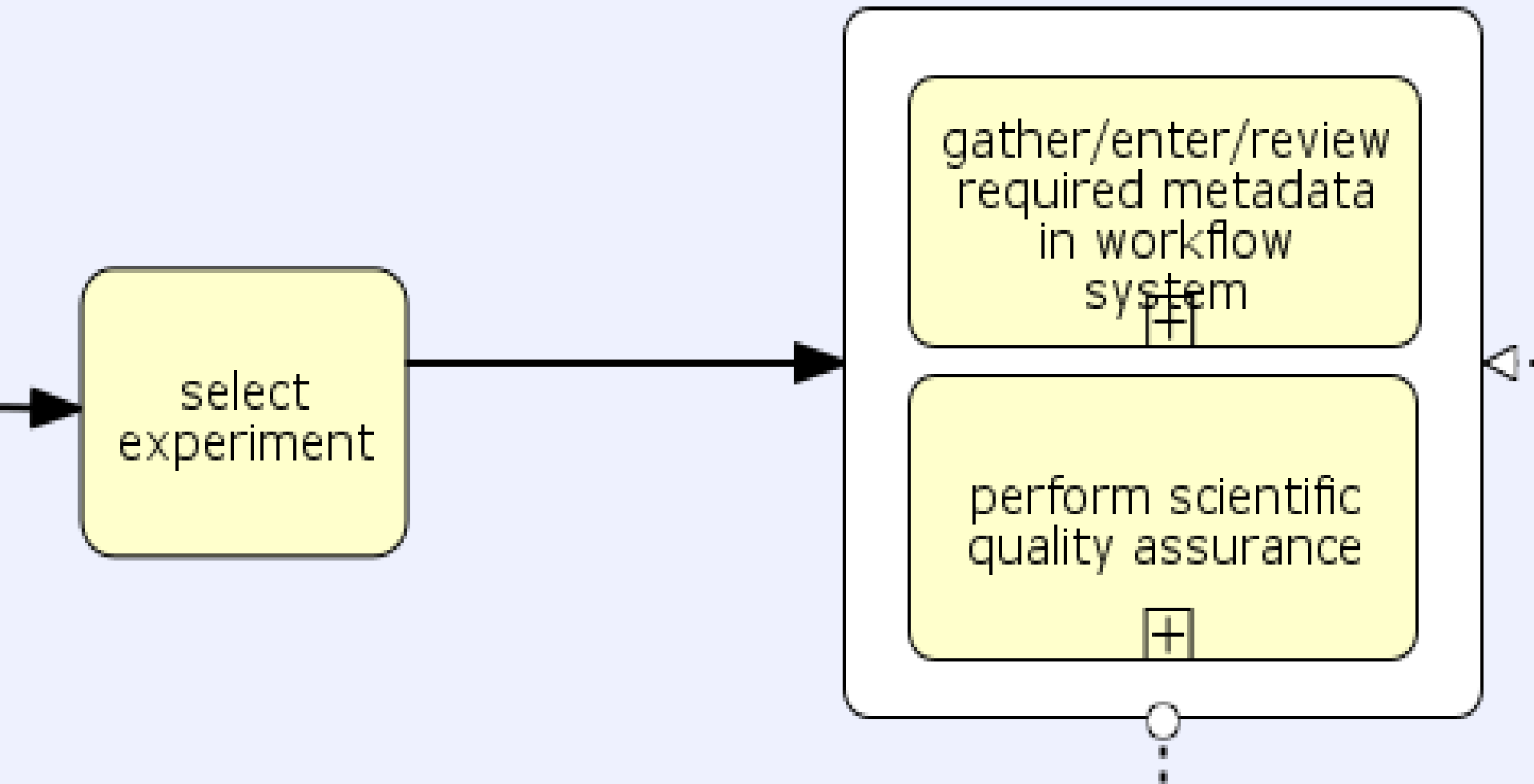
no standard process

WDCC: Investigator of the Experiment

- Today: e-mail
- Future: browser interface with account and
according experiment list

SQA - Scientific Quality Assurance Scientist

Business Process Modeling Notation
(www.signavio.com/en.html)



SQA Review Required Metadata

- std-doi profile sent to TIB

Title

Publication Date

Author(s)

Description

Datasize

Data Format(s)

- Metadata of DOI/URN resolved compact site

Location(s)

Spatial and Temporal Coverage

Contact

List of Datasets (Topics)



World Data Center for Climate, Hamburg



Not logged in ([Login](#)) | [Process List](#) [CERA UI Home](#) | [WDCC Home](#)

Always quote citation when using data!

DOI for Scientific and Technical Data
10.1594/WDCC/EH5-T63L31_OM-GR1.5L40_A1B_1_6H

Title
IPCC-AR4 MPI-ECHAM5_T63L31 MPI-OM_GR1.5L40 SRESA1B run no.1: atmosphere 6 HOUR values MPImet/MaD Germany

Citation
Roeckner, Erich; Lautenschlager, Michael; Schneider, Heiko 2006; IPCC-AR4 MPI-ECHAM5_T63L31 MPI-OM_GR1.5L40 SRESA1B run no.1: atmosphere 6 HOUR values MPImet/MaD Germany. [doi: 10.1594/WDCC/EH5-T63L31_OM-GR1.5L40_A1B_1_6H]




Publication Date
2006-09-25

Author(s)
Roeckner, Erich; Lautenschlager, Michael; Schneider, Heiko

Summary
The data represent 6 hourly values of selected variables for the Data Distribution Centre (DDC) of the Intergovernmental Panel on Climate Change (IPCC). (see also: <http://ipcc-ddc.cru.uea.ac.uk>)
The A1B scenario is the part of the A1 family which describes a balance across all energy sources. The experiment has been initialized in year 2000 of the 20C_1 run and continues until year 2100 with anthropogenic forcings (CO2, CH4, N2O, CFCs, O3 and sulfate) according to A1B. The experiment is extended until year 2200 with all concentrations fixed at their levels of year 2100 (stabilization experiment). Datasets with monthly average values are also available.
Technical data to this experiment:
The experiment is using ECHAM5.2.02a coupled to MPI-OM Vers. 1.0 GR1.5L40 and was run on a NEC-SX (hurrikan). The output from the model run: hurrikan.dkrz.de:/ut/k/k204098/EXP000/run012

SQA Data Scientist

- Approved by author(s)
- Short description of quality checks done by author(s)
- Quality check protocol files done by author(s)

Quality Checks	Virtual window
1.	<input type="text" value="approved by author"/>
2. Description of Quality Checks	<div style="border: 1px solid black; padding: 5px;">description of quality checks (1900 characters)</div>
3. Data Level ⓘ	<input type="text" value="2a - description"/>
4. Quality Check Protocols	Help text explaining which files to upload... Download demo protocol for EXP2
<input type="text" value="Path"/>	<input type="button" value="Browse"/> <input type="button" value="Upload"/>
Uploaded files:	
1	logfile.pdf 324 kB 
2	protocol.pdf 523 kB 
3	explanation.txt 32 kB 

SQA Example Short Description

Experiment: 10.1594/WDCC/CLM_A1B_2_D3

1) Quality documentation see

- 'README, Plots and Reports for CLM regional climate model runs' in CERA2
- http://cera-www.dkrz.de/WDCC/ui/Entry.jsp?acronym=CLM_PLOTS_2008 and 'CLM Technical Report'
- http://www.mad.zmaw.de/fileadmin/extern/documents/reports/MaD_TechRep3_CLM.pdf Chapter 4 and 6.

2) Control of timeseries:

- creation of minimum, maximum, mean, average timeseries of every record
- control of timeseries with statistical analysis

TQA - Technical Quality Assurance

WDCC

- Number of data sets is correct and not equal 0
- Size of every data set is not equal 0
- The data sets and corresponding metadata are all accessible via internet
- The data size is controlled and correct
- The time description (metadata) and existence of data are consistent.

complete, start- stop date consistent, continuous time steps are correct

- Format is correct
- Variable description and data are consistent

Fixing Data and Metadata

persistent identifiers (DOI/URN) need persistent objects (DATA)

Data

- **no change** is possible after std-doi publication

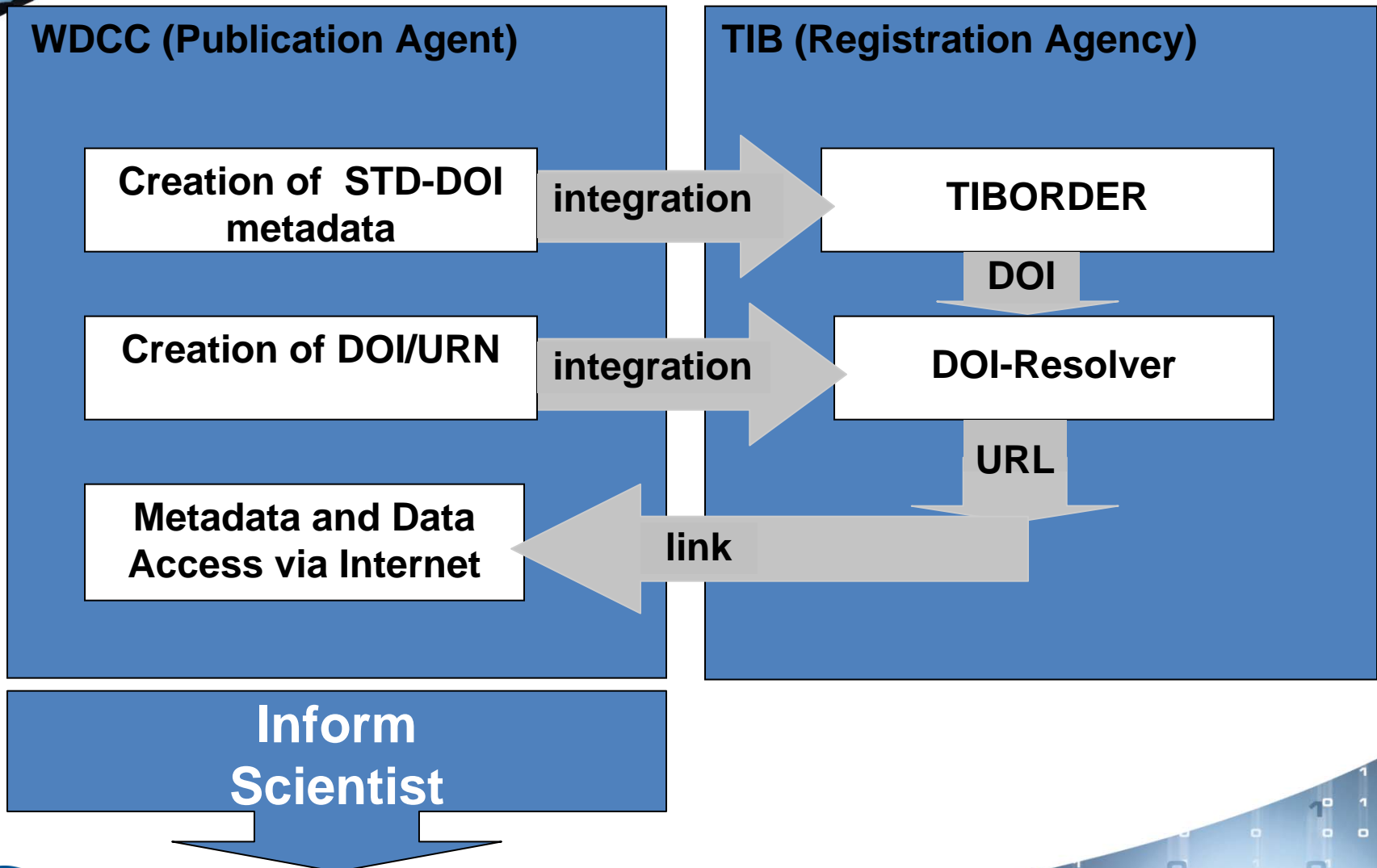
Metadata of distribution is fixed

- Datasize
- Data Format(s)

Metadata of citation is fixed

- Author(s), Title and Publication Date

Publication



Future

- SQA and TQA toolbox (examples) for standardization of quality control
- Two browser interfaces for workflow processes
 1. Scientist (virtual windows, deployment tests)
 2. Publication agent



Thank you for your Attention!

<http://www.dkrz.de>

<http://www.wdc-climate.de>

<http://umwelt.wikidora.com>