

Item 5.7.3, ET-WDC Tokyo meeting, 21-23 Jan 2014

WHOLE NEW WDCGG

PART 1

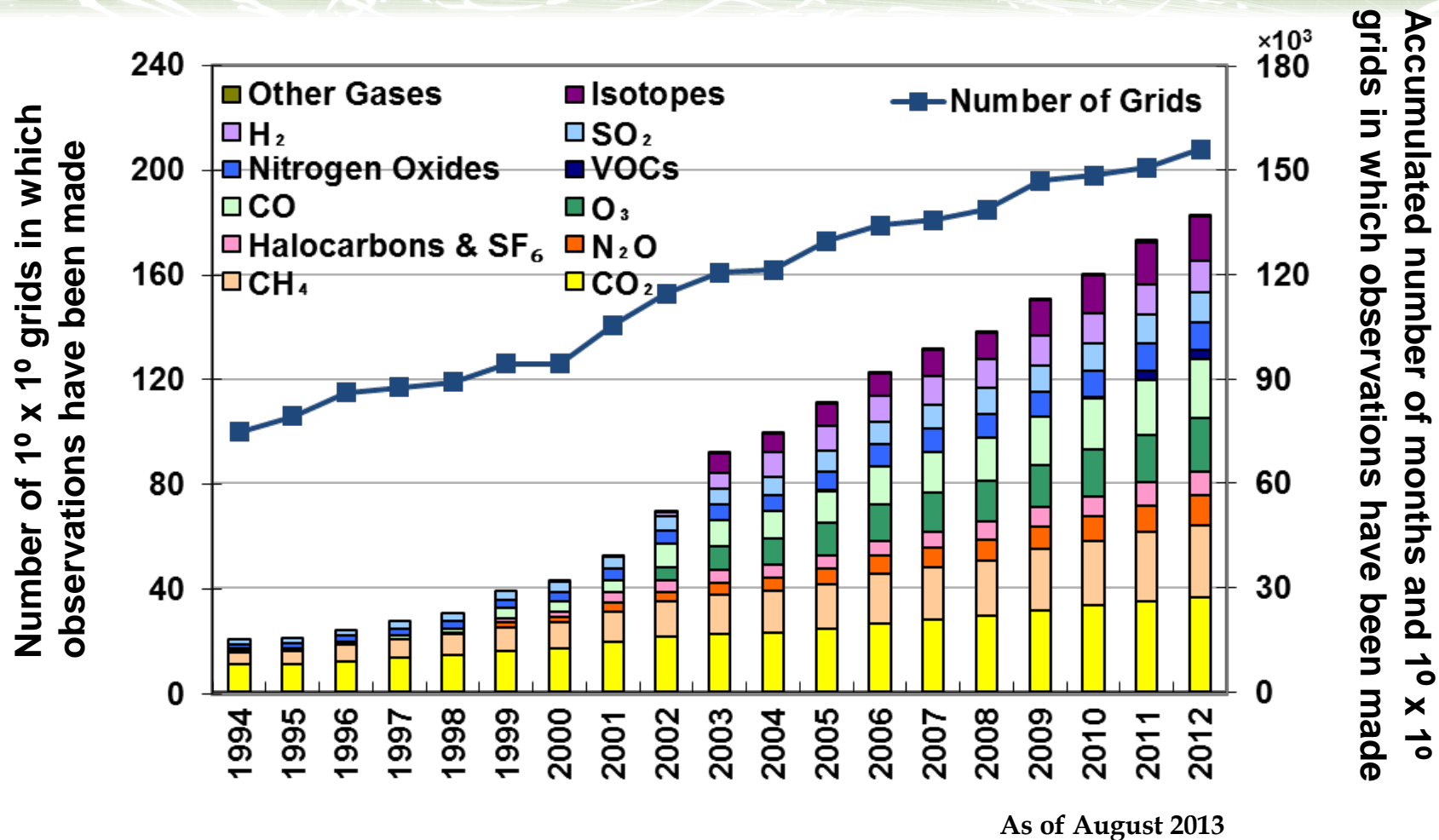
— DATA-FLOW AND DATABASE STRUCTURE —

Expert Team on World Data Centres, Tokyo meeting
21-23 January 2014

Teruo Kawasaki
World Data Centre for Greenhouse Gases
Japan Meteorological Agency
wdcgg@met.kishou.go.jp

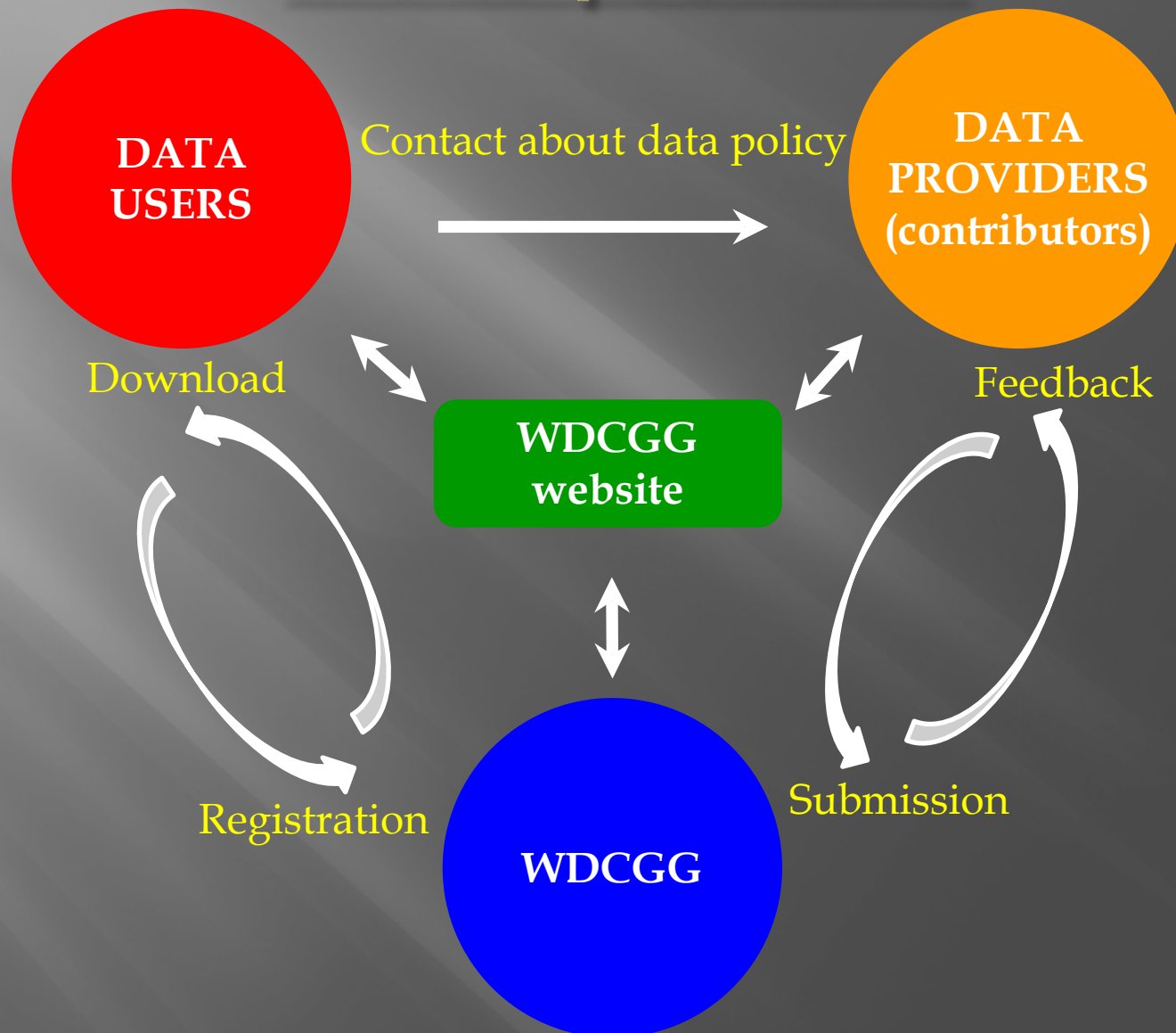
Data reported to the WDCGG

— Development of reports for different elements —



Data from fixed stations have increased in terms of locations and amounts.

The basic roles of the Data Centre are to provide :



The basic roles of the Data Centre are to provide :

1. stable and valuable services to data users
 - easy-to-use web interface, formatted files, consolidated flags, ...
2. feedback to contributors
 - statistical information, download notice, ...
3. data archive
 - reliability, update history, consistency with the metadata, ...

ISSUES FOR “DATA USERS”

1. Non-user-friendly “Data Search”

WMO Global Atmosphere Watch
World Data Centre
for Greenhouse Gases

Category Parameter Country/Territory Contributor

Updated in the last 365 days

!!Attention!!
Carbon monoxide daily and monthly data for 2012 from Giordan Lighthouse (GLH, UMLT) was provided incorrectly on this website from 19 September 2013 to 25 December 2013. If you use the data downloaded during this period, please update the data. (December 25 2013)

Note: From 26 April 2012, the number of header lines in each data file provided by NOAA/ESRL is not fixed 32 as assigned in "WDCGG Data Submission and Dissemination Guide" (GAW Report No. 188). The header parts have been modified to contain original COLLABORATORS information according to the request of NOAA/ESRL and its contributors. The length of the header lines is available at the line number 5 (C05) in each data file. See [list.txt](#) for more information. (April 26 2012)

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Sample programs **Total = 377** [Download archives of the data listed below](#) [Download metadata csv file](#) Sorted by Station Name

Station Name	Country/Territory	SITE-ID	Contributor	Parameter
Adriqole	Ireland	ADR651N00	AGAGE	CCl ₄ , CFCs, CH ₃ CCl ₃ , N ₂ O ⁺⁺
Aircraft (over Bass Strait and Cape Grim)	Australia	AIA999900	CSIRO	¹³ CO ₂ , CH ₄ , CO, CO ₂ , H ₂ , N ₂ O ⁺
Aircraft Observation of Atmospheric trace gases by JMA	Japan	AOA999900	JMA	CH ₄ , CO, CO ₂ , N ₂ O
Aircraft Orleans	France	ORL999900	LSCE	CH ₄ , CO ₂
Akademik Korolev, RV	United States of America	AKD999900	NOAA/ESRL	CH ₄
Alert	Canada	ALT482N00	CSIRO	¹³ CO ₂ , CH ₄ , CO, CO ₂ , H ₂ , N ₂ O ⁺⁺
Alert	Canada	ALT482N00	EC	¹³ CO ₂ , C ¹⁸ O ₂ , CH ₄ , CO, CO ₂ , N ₂ O, SF ₆ , MET
Alert	Canada	ALT482N00	NOAA/ESRL	¹³ CH ₄ , ¹³ CO ₂ , C ¹⁸ O ₂ , C ₂ Cl ₄ , CBrClF ₂ , CBrF ₃ , CCl ₄ , CFCs, CH ₂ Cl ₂ , CH ₃ Br, CH ₃ CCl ₃

Station Name

ISSUES FOR “DATA USERS”

2. Machine unreadable “Metadata”

C01 TITLE: CO2 monthly mean data
C02 FILE NAME: mnm224n00.jma.as.cn.co2.nl.mo.dat
C03 DATA FORMAT: Version 1.0
C04 TOTAL LINES: 280
C05 HEADER LINES: 32
C06 DATA VERSION: 201307
C07 STATION NAME: Minamitorishima
C08 STATION CATEGORY: Global
C09 OBSERVATION CATEGORY: Air sampling observation at a stationary platform
C10 COUNTRY/TERRITORY: Japan
C11 CONTRIBUTOR: JMA
C12 LATITUDE: 24.28
C13 LONGITUDE: 153.98
C14 ALTITUDE: 8
C15 NUMBER OF SAMPLING HEIGHTS: 1
C16 SAMPLING HEIGHTS: 20
C17 CONTACT POINT: y-fukuyama@met.kishou.go.jp
C18 PARAMETER: CO2
C19 COVERING PERIOD: 1993-01-01 2013-08-01
C20 TIME INTERVAL: monthly
C21 MEASUREMENT UNIT: ppm
C22 MEASUREMENT METHOD: NDIR
C23 SAMPLING TYPE: continuous
C24 TIME ZONE: Local time UTC+9
C25 MEASUREMENT SCALE: WMO X2007 scale
C26 CREDIT FOR USE: This is a formal notification for data users. "For scientific purposes, access to these data is unlimited
C27 and provided without charge. By their use you accept that an offer of co-authorship will be made through personal contact
C28 with the data providers or owners whenever substantial use is made of their data. In all cases, an acknowledgement
C29 must be made to the data providers or owners and the data centre when these data are used within a publication."
C30 COMMENT:

C31
C32 DATE TIME DATE TIME CO2 ND SD F CS REM
1993-01-01 00:00 9999-99-99 99:99 -999999.99 0 -99.999 2 0 -99999999
1993-02-01 00:00 9999-99-99 99:99 361.16 85 0.683 3 0 -99999999
1993-03-01 00:00 9999-99-99 99:99 360.74 573 0.941 3 0 -99999999

Header
(metadata)

Data

ISSUES FOR “DATA USERS”

3. Too many kinds of “Data Format”

Format

More than **10**

Example:

stationary platform, mobile platform, ice core, vertical profile, ... , meteorological data

File name

mnm224n00.jma.as.cn.co2.**nl**.mo.dat

Example:

mnm224n00.jma.as.cn.co2.**ndir**.mo.dat

mnm224n00.jma.as.cn.co2.**wmo2007**.mo.dat

auxiliary item

Data File

C32	DATE	TIME	DATE	TIME	CO2	ND	SD	F	CS	REM
	1993-01-01	00:00	9999-99-99	99:99	-999999.99	0	-99.999	2	0	-99999999
	1993-02-01	00:00	9999-99-99	99:99	361.16	85	0.683	3	0	-99999999

ISSUES FOR “DATA PROVIDERS”

1. A heavy burden of submitting metadata

➤ Non-user-friendly web interface

Measurement Information Edit Form

The measurement information should be submitted for each parameter. Please save this form by clicking the "Save" button at the bottom of this form after you complete each item here. Please refer the Appendix 4 in the WDCGG Data Submission and Dissemination Guide. If you have any questions, please don't hesitate to contact the WDCGG.

** denotes required fields*

Parameter (Station) CH3CCI3 (Ryori)

Contact Person(s) for measurement The Contact Person for the measurement is the person who is contacted regarding the submission of data and the measurement information. You should confirm or select Contact person in the below list box. The list box is referred from the "Contact Persons" information.

*Contact Person Yukio Fukuyama ▾
If you could not find your contact person in the list above, please enter the name and e-mail address of the contact person.

*name
*e-mail

Contact Person
name
e-mail

Contact Person
name
e-mail

Supporting Contributor(s) (optional) The institute or organization that supports the observation should be described here.

Responsible Investigator(s) (optional) The responsible investigator is the person who is officially or scientifically responsible for the observation. The investigator's name, office address, phone/FAX number and e-mail address should be included.

Metadata Edit Form

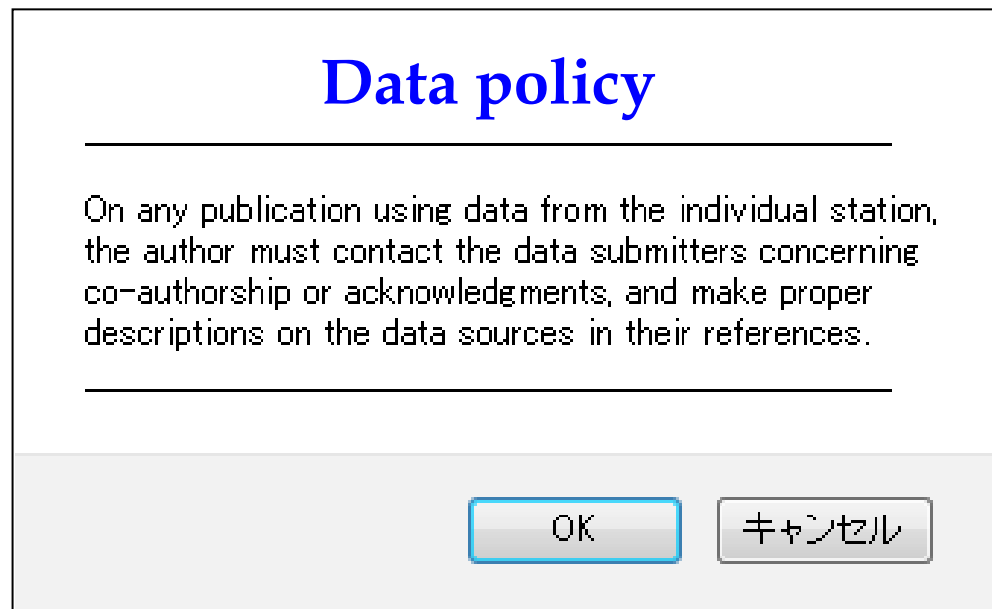
ISSUES FOR “DATA PROVIDERS”

2. No feedback to data providers

- No statistical information, No download notice
 - Who downloaded the data?
 - When/Which data was downloaded?
 - What was the data used for?

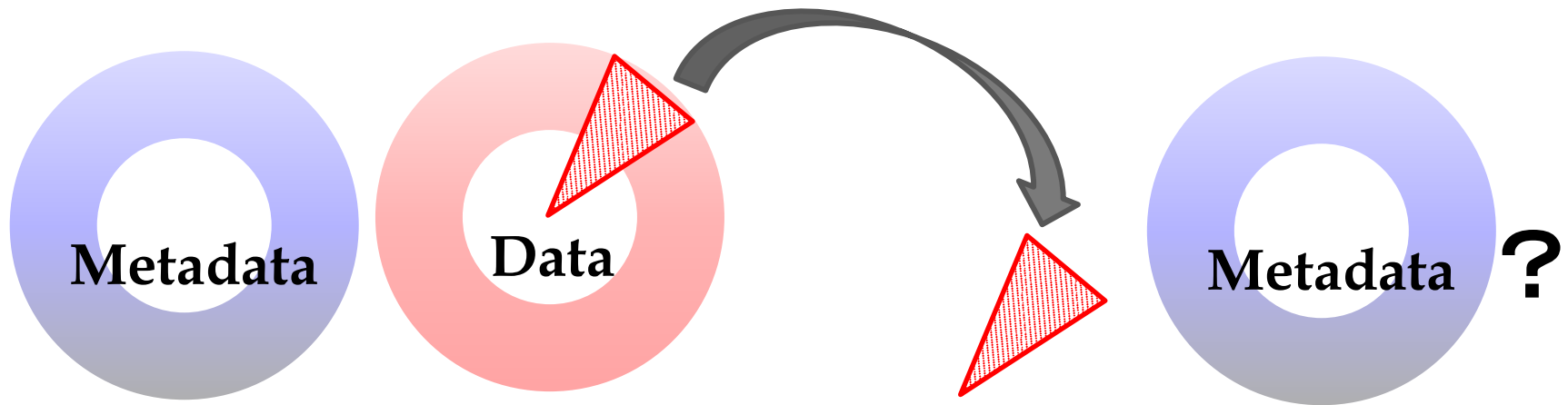
3. Unobserved Data policy

- Data users cannot always contact data providers



BOTTLENECK IN THE OPERATION

1. Most contributors submit the data in fragments



2. Main points are not easy to understand or are not submitted

- scale?, unit?, time zone?,
instrument/measurement method?
- Sometimes they change the above
in the middle of a data file without any warning
- At times, we cannot contact the contributors
to check with them

BOTTLENECK IN THE OPERATION

3. The original data is submitted in many random data formats

– Sometimes they change formats without notice

DATE		TIME	DATA (ppm)	FLAG	SD
...	
1990-12-31		22:00	399.0	0	-9999
1990-12-31		23:00	399.1	0	-9999

data1.tsv

DATETIME	DATA (ppb)	FLAG	ND	SD
..., ... , ... , ... , ...				
199101010100	, 0.3992 , 0 , 30 , -9999			
199101010200	, 0.3993 , 0 , 30 , -9999			

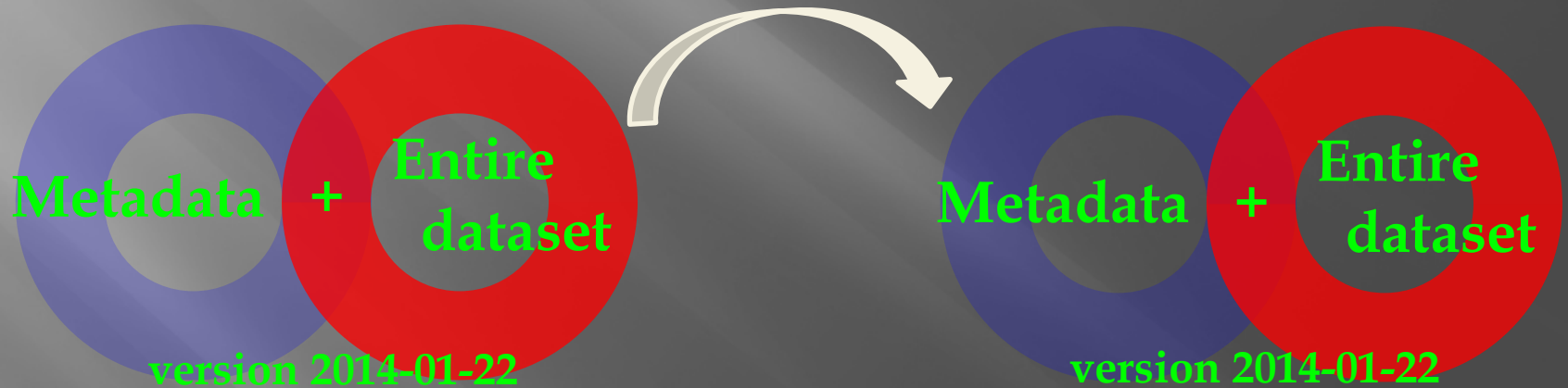
data2.csv

4. Currently, older versions of the metadata cannot be stored in our relational database system

– We cannot re-create the files backward

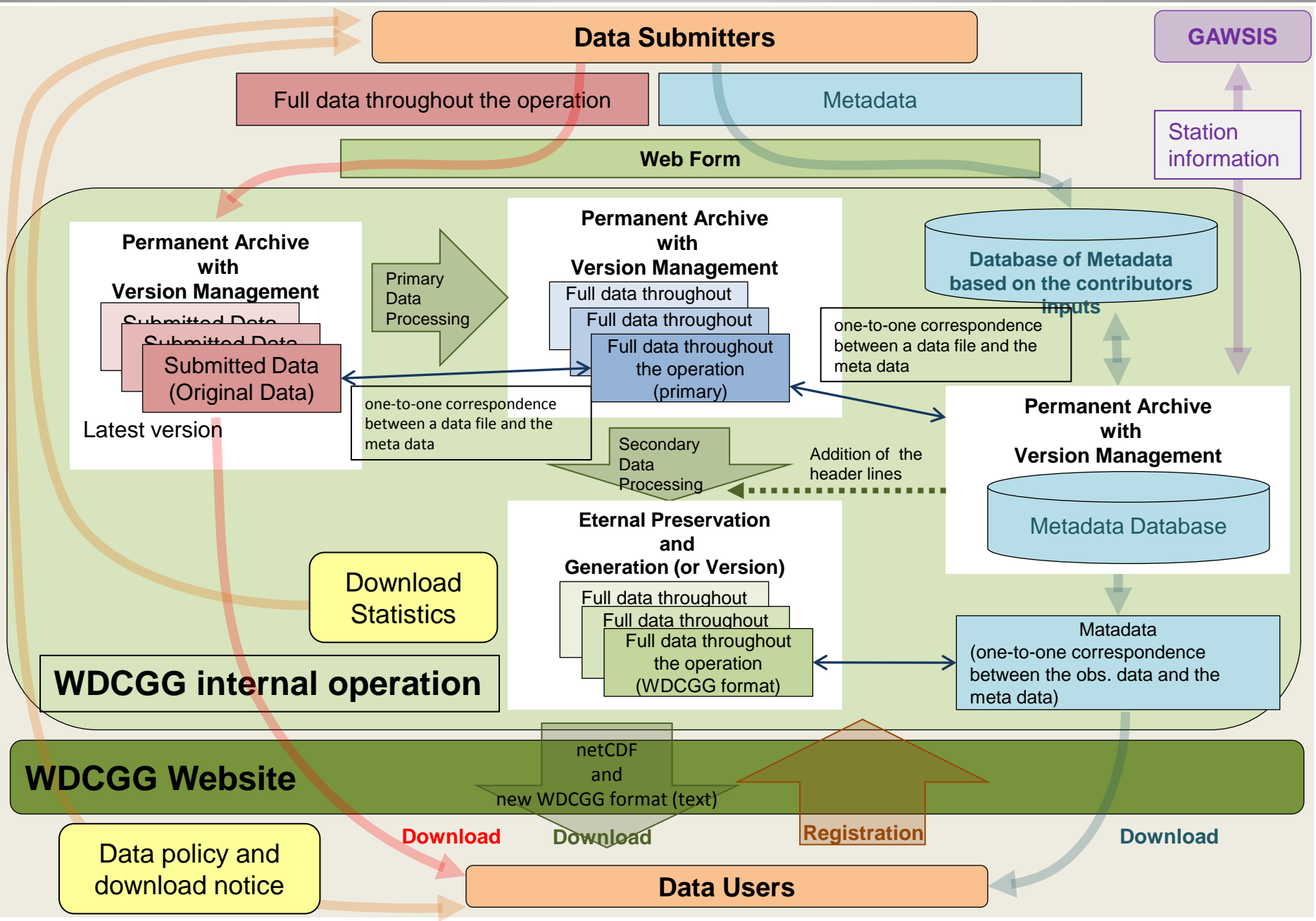
Change of Data Submission Methods

1. Submission of the entire dataset (at one time)
 - To avoid discrepancy between Laboratories and Data Centre due to WMO scale migration/Internal re-calibration



2. Submission of the metadata together with the data
 - One-to-one correspondence between data and metadata
 - Metadata has the same version as data

The data flow of the new WDCGG system (Draft)



Enhancement of the database system

1. Design to store all versions
2. Design to manage in an integrated fashion
 - Original historical data + metadata + wdcgg formatted file
3. Flexible design for contributors to demand
 - At times, they demand to divide a file by scale, measurement method, instrument method, ...
4. Clear terms and classification
 - Free-form is not permitted as much as possible
5. Harmonization with GAWSIS (GAW Station Information System) on station metadata

The database structure of the new WDCGG system (Draft)

