

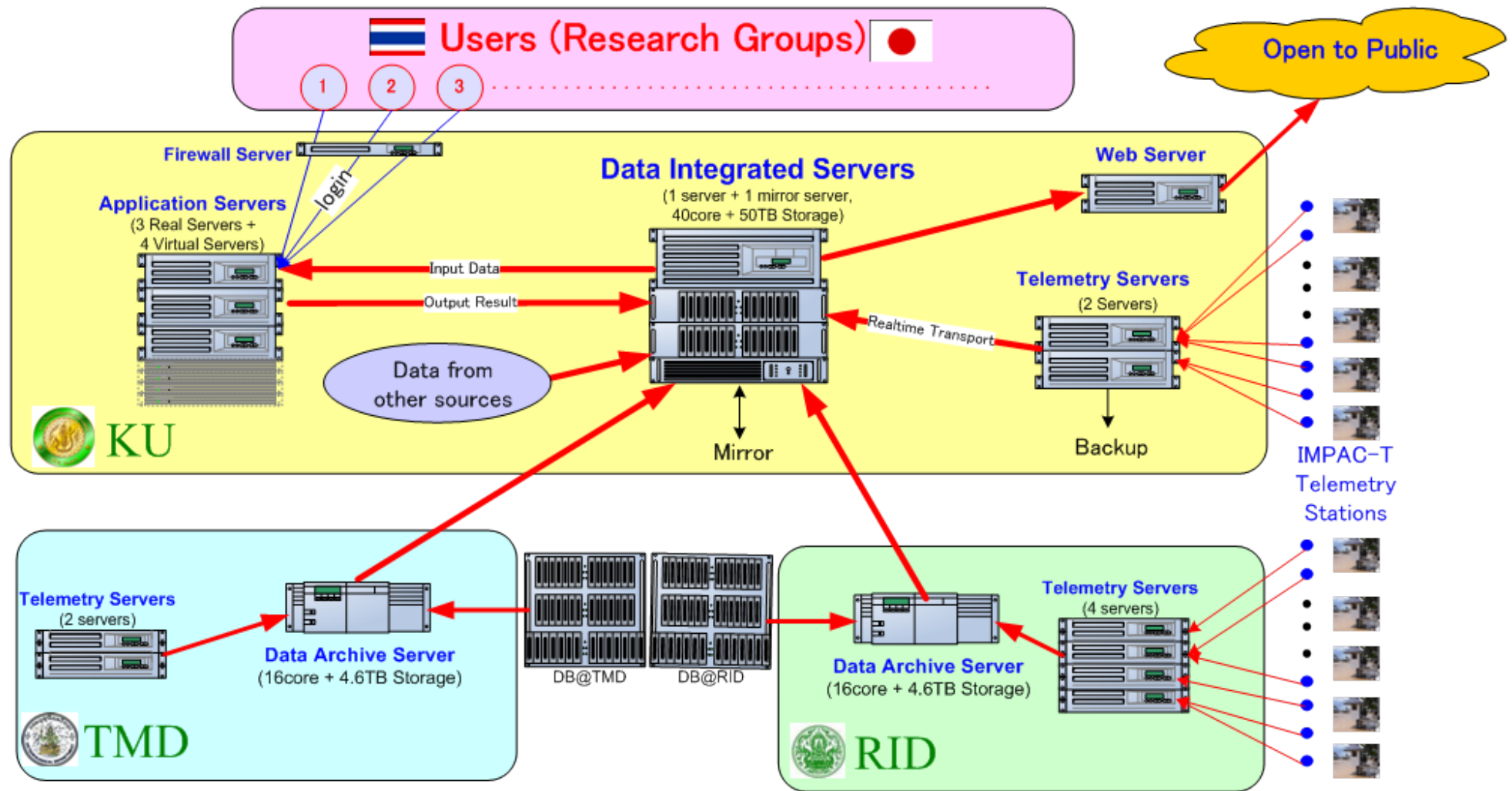
Secondary Data Sharing System

Eiji Ikoma

The Univ. of Tokyo

IMPAC-T Information Server Team

IMPAC-T Server System

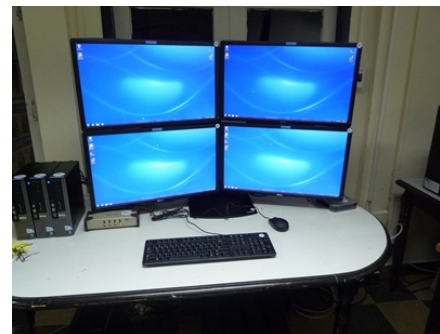


IMPAC-T Server Services

- Telemetry Data Crawling and Archive
 - Impact-t1,impact-t2@KU
 - Impact-r-t1,t2,t3,t4@RID
 - Impact-t-t1,t2@TMD
- Data Integration System(DIS)
 - Impact-di@KU
 - Impact-r-di@RID
 - Impact-t-di@TMD
- Web Server
 - Impact-www@KU
- Application Server (for models)
 - Impact-a1,ap,mr@KU
 - Impact-v1,v2,v3,v4@KU(virtual)
- **Secondary Data Sharing System**



@KU



t1-t4@RID



t1-t2,di@TMD

What is “Secondary Data” ?

⇔ Primary Data

≠ Telemetry Data

≠ Observation Data

→ Secondary Data

= Data for Secondary Use ?





[Main page](#)

Contents

Featured content

Current events

[Random article](#)

[Donate to Wikipedia](#)

▼ Interaction

[Help](#)

[About Wikipedia](#)

Community portal

Recent changes

[Contact Wikipedia](#)

► Toolbox

► [Print/export](#)

Article

Talk

Read

[Edit](#)[View history](#)

Secondary data

From Wikipedia, the free encyclopedia

Secondary data, is data collected by someone other than the user. Common sources of secondary data for **social science** include **censuses**, organisational records and data collected through qualitative

Data collected by someone other than the user

Secondary data analysis saves time that would otherwise be spent collecting data and, particularly in the case of [quantitative data](#), provides larger and higher-quality [databases](#) that would be unfeasible for any individual researcher to collect on their own. In addition, analysts of social and economic change consider secondary data essential, since it is impossible to conduct a new survey that can adequately capture past change and/or developments.

Contents [hide]

- 1 Sources of secondary data
- 2 Secondary analysis or re-use of qualitative data
- 3 References
- 4 Further reading
- 5 External links

What is “Secondary Data” ?

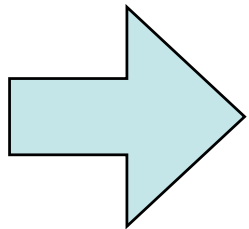
Anyway, Secondary data is **important data which can be used for research.**

→ necessary to develop
easy-to-use environment



Secondary Data Sharing System

- System for **Sharing “Data”**
- Somebody upload to Server
 - Store and Manage on Server
 - Users can use/download data



Upload, Manage, Use

Upload

- By http
 - Using Web-based Interface
- By ftp
 - Using FTP client software
- By direct-connection
 - Using USB HDD for connecting to Server

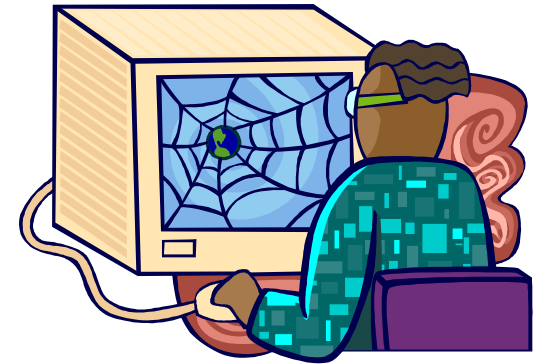
Management

- All Secondary data will be stored at IMPAC-T Data Integrated Server
- 1 Dataset = 1 Directory
- Dataset should include meta-information
 - Free style document file
 - Well-formatted text (XML etc..)
- Security/Permission Information
- Is Pre-processing necessary?

IMPAC-T
Data Integrated Server
@ KU



For Use

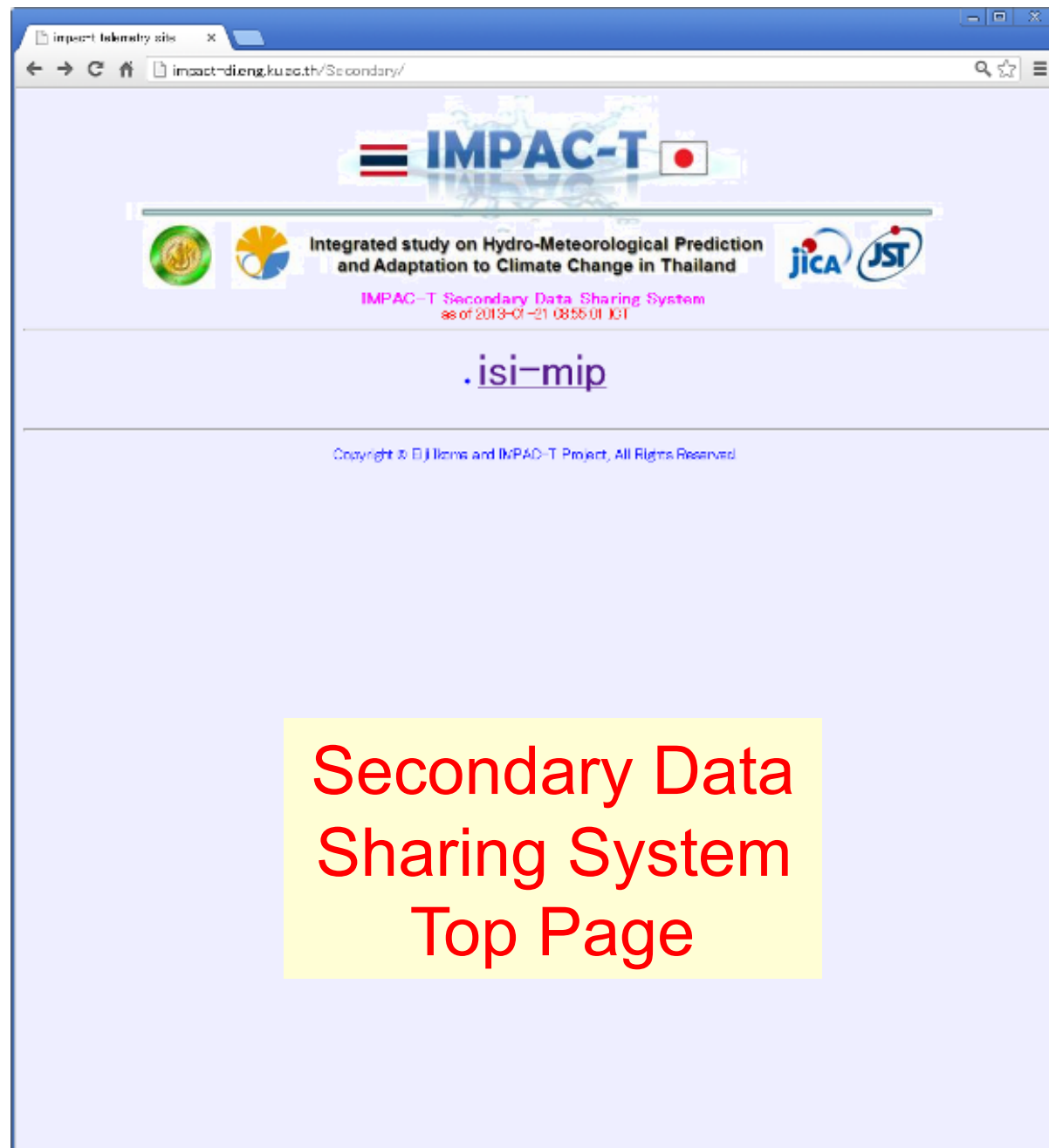


- Download
 - Web-based download Interface
 - Data Integrated Servers prepare “download data set” for each user
- Direct-use on server
 - Each user has their own environment/disk space on application server
 - Data Integrated Servers prepare data and copy to user’s home directory
- Direct-connected HDD and “carry”

Case 1



- ISI-MIP data (by Dr.Kiguchi)
 - Upload: Direct-connection (USB HDD)
 - Manage: original data and Thai-area data(pre-processing)
 - Use/Download: Prepare download interface
- All data are stored at impact-di server
= 22TB (compressed, including pre-processed data)



Secondary Data isi-mip

impact-dieng.ku.ac.th/Secondary/cgi-bin/step1-2.sh?isi-mip

Select Area ,Model and Element(s)

STEP 1(Area,Model,Element) → STEP 2(Scenario,Term)

- Dataset: **isi-mip**
- Area: ☒ global ☐ thai
- Model: ☒ cgfdl-esm2m ☐ hadgem2-es ☐ cipsi-cm5a-lr ☐ miroc-esm-chem ☐ noresm1-m
- Element: ☒ LWdown ☒ PSurf ☒ Qair ☒ Rainf ☒ SWdown ☒ Tair ☒ Wind

Choose

- Dataset
- Area(global/thai)
- Model
- Element

Secondary Data isi-mip

impact-dieng.ku.ac.th/Secondary/cgi-bin/step1-2.sh?isi-mip

Select Scenario, and Term

STEP 1(Area,Model,Element) → STEP 2(Scenario,Term)

Dataset: **isi-mip** Area: **thai** Model: **hadgem2-es** Element: **LWdown,PSurf,Qair,Rainf,SWdown,Tair,Wind**

Scenario

- ☒ historical 1951 --- 2004 (84 MB/year)
- ☒ rcp2p6 2005 --- 2099 (84 MB/year)
- ☒ rcp4p5 2005 --- 2099 (84 MB/year)
- ☐ rcp6p0 2005 --- 2099 (84 MB/year)
- ☐ rcp8p5 2005 --- 2099 (84 MB/year)

☒ Download ☐ Copy to Your HomeDirectory @ AppServer

Choose

- Scenario
- Term

Interface for Data Selection

Secondary Data isi-mip

impact-di.eng.ku.ac.th/Secondary/cgi-bin/step1-2.sh?isi-mip

Select Scenario, and Term

STEP 1(Area,Model,Element) → STEP 2(Scenario,Term)

Dataset: **isi-mip** Area: **thai** Model: **hadgem2-es** Element: **LWdown,PSurf,Qair,Rainf,SWdown,Tair,Wind**

Scenario

- ☒ historical [1951] --- [2004] (84 MB/year)
- ☒ rcp2p6 [2005] --- [2099] (84 MB/year)
- ☒ rcp4p5 [2005] --- [2099] (84 MB/year)
- ☐ rcp6p0 [2005] --- [2099] (84 MB/year)
- ☐ rcp8p5 [2005] --- [2099] (84 MB/year)

☒ Download ☐ Copy to Your HomeDirectory @ AppServer

You request following files:

• historical

YEAR	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975
LWdown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PSurf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qair	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rainf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SWdown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tair	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Wind	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

• rcp2p6

YEAR	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029
LWdown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PSurf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qair	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rainf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SWdown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tair	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Wind	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

• rcp4p5

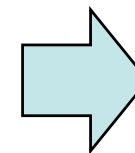
YEAR	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029
LWdown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PSurf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Qair	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rainf	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SWdown	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tair	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Wind	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

20496 MB (Total), Download from IMPAC-T Server

Is this correct?

Copyright © Eiji Ikoma and IMPAC-T Project, All Rights Reserved.

Confirmation
For Data Download



Download

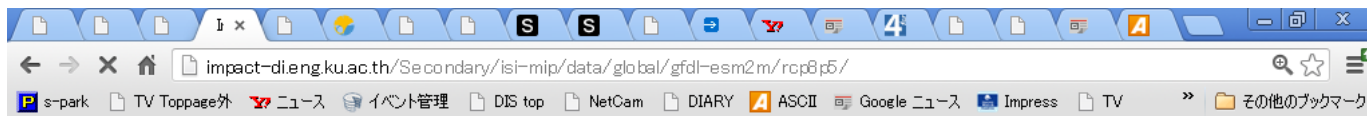
Case 2

- KAKUSHIN-MRI data (by Dr.Kiguchi)
 - Upload: Direct-connection (USB HDD)
 - Manage: original data
 - Use/Download: (under consideration)
- 70% of data are stored at impact-di server
Finally, 13TB (original data) will be stored.

→ depend on the discussion after my talk

Case 3

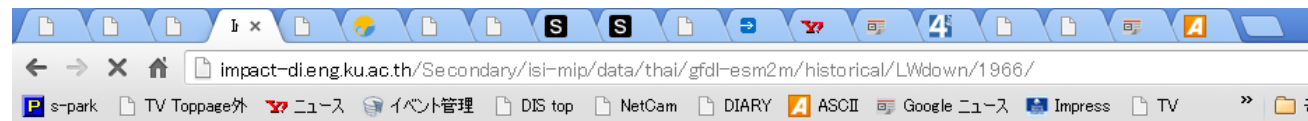
- Any data
 - Upload: Direct-connection (USB HDD)
 - Manage: original data
 - Use/Download: Web(Directory View only)
- Simplest Interface for download


















Index of /Secondary/isi-mip/data/global/gfdl-esm2m/rcp8p5

Name	Last modified	Size	Description
Parent Directory		-	
LWdown/	28-Dec-2012 20:51	-	
PSurf/	29-Dec-2012 00:52	-	
Qair/	29-Dec-2012 04:20	-	
Rainf/	29-Dec-2012 07:44	-	
SWdown/	29-Dec-2012 11:26	-	
Snowf/	29-Dec-2012 14:53	-	
Tair/	16-Jan-2013 10:48	-	
Wind/	16-Jan-2013 14:08	-	

Apache/2.2.3 (CentOS) Server at impact-di.eng.ku.ac.th Port 80



Index of /Secondary/isi-mip/data/thai/gfdl-esm2m/historical/LWdown/1966

Name	Last modified	Size	Description
Parent Directory		-	
 LWdown_1966_0101_00.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0101_03.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0101_06.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0101_09.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0101_12.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0101_15.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0101_18.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0101_21.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0102_00.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0102_03.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0102_06.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0102_09.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0102_12.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0102_15.bin	28-Dec-2012 12:50	640	
 LWdown_1966_0102_18.bin	28-Dec-2012 12:50	640	

Discussion

- What kind of data can you prepare?

Depend on the data format, size,
structure, permission...

I'll prepare appropriate system
for upload, direct-use/download Interface
-software, hardware-

Please give rough information
about your data!

Thank you for your attention

