IMPAC-T Data Archive System

- 2009.6.1 1st Joint Coordinating Committee in Bangkok, Thailand-

Eiji Ikoma, Masaru Kitsuregawa
EDITORIA/IIS, Center for Information Fusion
The University of Tokyo

Introduction

- Who am I ?/ Why I am here?

→My Research Topic = **Computer Science**

-Database, Visualization, User-Interface

- What can I do in this project?

→ To develop **Data Archive system** base on my experiences.

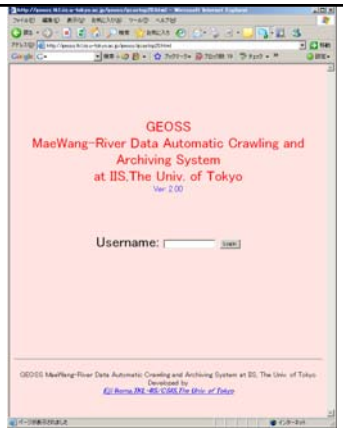
Keywords about my research

- Huge-scale database
- Web-service, User Interface
- User support, Personalization
- Data Portal
- Workbench
- Data Crawling
- Data Mining
- Data Visualization
- Quality Control
- Data Comparison

some of my works

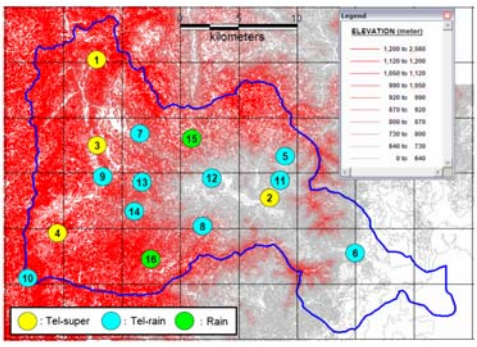
- SiB2 on Web – Workbench for model
- Real-time Data Visualization system for GMS-Satellite data
- GSWP2-ICC/DDC system
- GEOSS-Mae Wang River data Automatic crawling and Archiving System
- CEOP Data Quality Control System
- Data Mining and Visualization System
- AWCI Data Upload and Archive System

……..

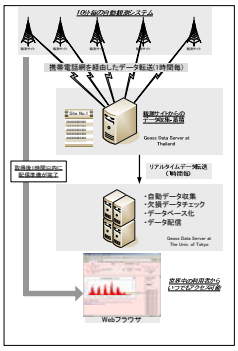


- 2006-2008

Telemetry System



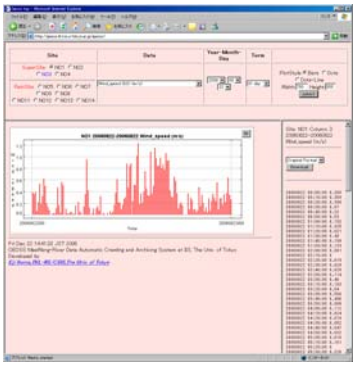
System Structure



- Collecting 16-point Data in Thailand to Thai-Server using cell-phone
- Send data to UT every 1hr
- At UT Server,
  - Automatic-reformatting
  - Error correction
  - Archive and Indexing on DB
  - Send data to other server

→ Open to public

User Interface on WEB



- You can search data from “site”, “Element”, “Date”, “Term”
- Data Download and Visualization
- Easy-operation Interface
- Automatically refreshed every 1 hour

What is important point? For Data Archiving System

Observation, recording data, archiving data…. are of course very important.

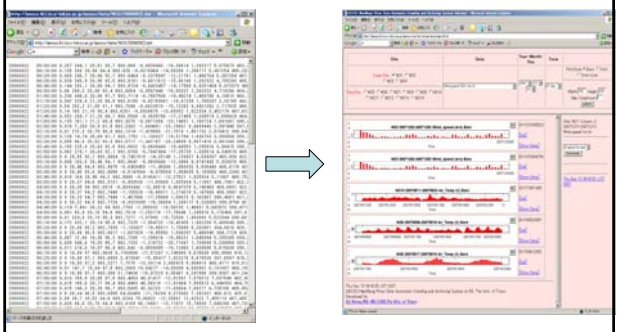
However,

Data should be changed to Information.

Any data which cannot be used is not information. Data should be used.

Archiving system should provide **Information**.

And, with easy operation and easy feeling.



Data

Information

Key Technology

1. Stable System

Network, Storage, Server, Human-Operation

2. Easy-Use System

For users (UI) and

For Administrator (system architecture)

3. “Fresh” System

“Fresh” data has much value.

In This Project,

- For Stable System
  - At First, we should consider Network structure and data flow -where is “1st archive”, where is “2nd archive”… From where each server get data?
  - Then, we should design storage system-- Capacity, RAID-level, Backup (storage,server,netowork) Policy ← cost effective.
  - And, Only Running System has a value.
- For Easy-Use System
  - For what purpose? How to use? Visualization is enough? Do you want to download data and use your own environment? Or, You want to load data to model directory?＝Should we develop workbench?
  - For each purpose, we should discuss and collaborate, important point is “feedback”. Then. “good-feeling” system will be developed.
- For “Fresh”
  - We should consider “smooth” data-flow. Data logger → User
  - This is very simple.
  - “Automatically”. Any oparater should not handle data every time.
  - Automatic system is also good for “decreasing human error”→stable

Thank you
for your attention.

No Data, No System!

DIAS Server(1PetaByte Storage + Servers)
@ IIS, Univ. of Tokyo