

# THE “GDSCLIENT” COLLECTING TOOL FOR NETWORKED SOLID EARTH SCIENCE DATA

H Nagao<sup>1,2\*</sup>, S Tsuboi<sup>1</sup>, Y Ishihara<sup>1</sup>, and H Yanaka<sup>3</sup>

<sup>1</sup>Institute for Research on Earth Evolution, Japan Agency for Marine-Earth Science and Technology, 3173-25, Showa-machi, Kanazawa-ku, Yokohama 236-0001, Japan

\*Email: [nagao@jamstec.go.jp](mailto:nagao@jamstec.go.jp), [tsuboi@jamstec.go.jp](mailto:tsuboi@jamstec.go.jp), [y-ishihara@jamstec.go.jp](mailto:y-ishihara@jamstec.go.jp)

<sup>2</sup>The Institute of Statistical Mathematics, 10-3, Midori-cho, Tachikawa, Tokyo 190-8562, Japan

Email: [hnagao@ism.ac.jp](mailto:hnagao@ism.ac.jp)

<sup>3</sup>Fujitsu Limited, 1-9-3, Nakase, Mihama-ku, Chiba 261-8588, Japan

Email: [yanaka@jp.fujitsu.com](mailto:yanaka@jp.fujitsu.com)

## ABSTRACT

The data center of our institute distributes solid earth science data obtained by the Ocean Hemisphere Project (OHP) network through the website of Pacific 21. We have developed Java-based software “GDSClient”, which enables us to collect not only the data of the OHP network but also those distributed from other data centers by means of the web service technology. It is possible to request the data controlling parameters such as data centers, observatories, a data period, and other auxiliary detailed parameters. It is unnecessary to know differences between data centers with preparing a WSDL (Web Services Description Language) file, in which information of user interface is described in XML format. The latest GDSClients are released from the website of Pacific 21.

**Keywords:** GDSClient, NINJA, Seismic data, Geomagnetic data, GPS data, WSDL, XML, XSLT, HTTP

## 1 INTRODUCTION

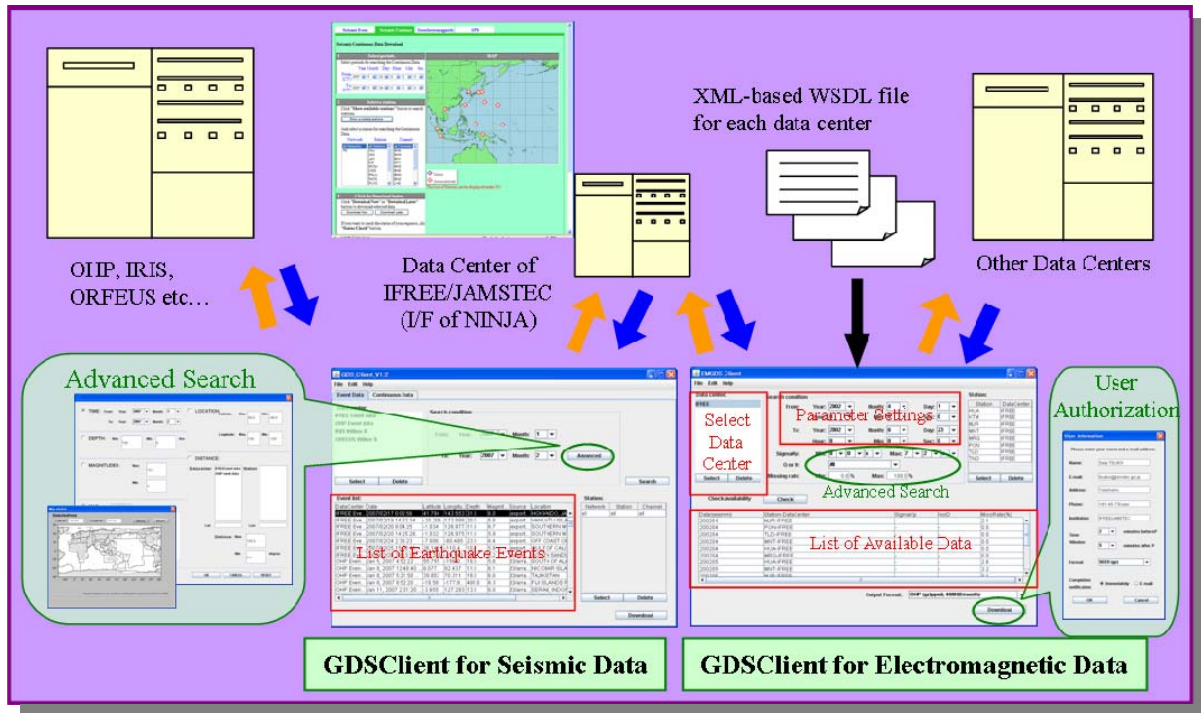
The Institute for Research on Earth Evolution, Japan Agency for Marine-Earth Science and Technology (IFREE/JAMSTEC), together with the Earthquake Research Institute of the University of Tokyo, has carried out long-term geophysical observations in the western Pacific Ocean called “the Ocean Hemisphere Project (OHP) network” for more than ten years. The main purpose of the OHP network is to fill in the spatial gap of geophysical observatories and to obtain more precise models of the structure and the activities of the Earth’s interior (e.g., Kawakatsu et al., 1998). In this OHP network, we are operating 25 seismic observatories, 7 geomagnetic observatories, and 8 GPS observatories. The observational data are distributed via the web site of the data center of IFREE/JAMSTEC called “Pacific 21” (<http://www.jamstec.go.jp/pacific21/>). We developed a geophysical data distribution system “NINJA”, which has a function to collect not only the OHP data but also geophysical data distributed from other data centers (Takeuchi et al., 2002). This networked data collection is possible through a single user interface with inputting parameters such as observatories and a data period to retrieve. However, this function of NINJA requires communications with each data server, and this becomes to be difficult recently from a point of view of the network security. In spite of this, requests from geophysicists to develop useful data collection software that releases from access to each data center are increasing.

We have been developing the NINJA portal and the GDSClient (Geophysical Data Service Client) in order to realize the collection of networked data overcoming the tight network security. The NINJA portal enables us to collect the seismic event and continuous data distributed in multiple data centers, and a prototype version is now available via the website of Pacific21 (Tsuboi et al., 2008). The GDSClient is PC-based software using the web service technology, i.e., the HTTP is used for both communications with data servers and data transfers. Currently available data are not only seismic event and continuous data but also geomagnetic data (e.g., Nagao et al., 2008). We mention the overview of the GDSClient in the next section.

## 2 OVERVIEW OF “GDSCLIENT”

### 2.1 The Role of “GDSClient” in the Flow of Geophysical Data

Figure 1 schematically illustrates each role of the NINJA and the GDSClient in the flows of geophysical data



**Figure 1.** Schematic illustration showing flows of networked distributed geophysical data, each role of the NINJA and the GDSClient in the data flows, and functions of the GDSClient.

distribution in the network. It is necessary to take into consideration not only user-friendliness but also recent strict network security in developments of data servers or data collection tools. The data center of IFREE/JAMSTEC entrusts the distribution of our own observational data to the NINJA, and the collection of the networked data to the GDSClient.

The GDSClient of seismic version can download two types of seismic data, i.e., continuous data and event data. The continuous data mean sequential observational data themselves while the event data mean parts of the continuous data filed for each earthquake event. Currently available data centers are the IFREE data center, the OHP data center of the University of Tokyo, IRIS (Incorporated Research Institutions for Seismology, USA), and ORFEUS (Observatories and Research Facilities for European Seismology), etc. The GDSClient of electromagnetic version can download the continuous data. Currently available data centers are the IFREE and OHP data centers, and another data centers can be added if necessary.

Even in the case of geophysical data available for anyone, it is usually prohibited to re-distribute collected data to a third party without a permission of the relevant data center. Therefore software that archives collected geophysical data once into a server may violate this rule. The GDSClient is free from this problem of data re-distribution because users access and download the geophysical data directly from data centers, although it would be better, of course, to negotiate with the relevant data center in advance.

## 2.2 Development of the “GDSClient”

Various operating systems as typified by Windows, MacOS, and UNIX are used in geophysics, therefore it would be ideal to develop data collecting tools to be executable also on any operating systems. The GDSClient was developed with the Java language by taking this point into consideration, and is executable on the Java platform of a version greater than 1.4.

Although input items needed in requesting geophysical data depend on the design of user interface of each data center, some of the input items are usually similar, e.g., a data period, observatories, and channels, in the case of time series of the solid earth science data. Therefore collecting tools should absorb the differences between user interfaces in order to achieve user-friendliness. The GDSClient refers to a WSDL (Web Services Description Language) file for each data center if distributed, in which information of user interface and data schema are described in the XML format. Figure 2 shows the basic structure of a WSDL file in the case of the geomagnetic database of the OHP network in NINJA.

```

<?xml version="1.0"?>
<definitions xmlns:gds="http://www.jamstec.go.jp/pacific21/GDSSchema/">
  <message name="stationlistWF_input">
    <part name="IFsearchPeriod" />
    <part name="IFsearchStations" />
    <part name="IFsearchAdvancedCondition" />
  </message>
  ...
  <portType name="GDSPortType">
    <operation name="stationlistWF">
      <input name="stationlistWF_input" ... />
      <output name="stationlistWF_output" ... />
    </operation>
  </portType>
  ...
  <binding name="GDSGetBinding" type="tns:GDSPortType">
    <operation name="startEM">
      <http:operation location="topEM.do"/>
      <input>
        <gds:translate type="text/xsl" href="http:// ... " />
      </input>
      <output>
        <gds:translate type="text/xsl" href="http:// ... " />
      </output>
    </operation>
    ...
  </binding>
  <service name="IFREEemDataService">
    <documentation>IFREE geomagnetic data service</documentation>
    <port name="GDSGetPort" binding="tns:GDSGetBinding">
      <http:address location="http://www.jamstec.go.jp/pacific21/NINJA"/>
    </port>
    <port name="GDSPostPort" binding="tns:GDSPostBinding">
      <http:address location="http://www.jamstec.go.jp/pacific21/NINJA"/>
    </port>
  </service>
</definitions>

```

**Figure 2.** The basic structure of a WSDL file served from a data center. This example is the case of the DSClient of electromagnetic version. Explanations of some significant XML tags are shown in Table 1.

**Table 1.** Some significant XML tags used in WSDL files that are referred by the GDSClients

Tags	Explanations
<code>&lt;definitions xmlns:gds="http:// ..."&gt;</code>	Declaration of the WSDL description. URL in which the definitions of XML schema used in the GDSClients can be specified in this <code>&lt;definitions&gt;</code> tag.
<code>&lt;message name="stationlistWF_input"&gt;</code> <code>&lt;message name="stationlistWF_output"&gt;</code>	<code>&lt;message&gt;</code> tag defines a name space concerning input and output workflows.
<code>&lt;part name="IFsearchPeriod"&gt;</code> <code>&lt;part name="IFuserInfo"&gt;</code>	<code>&lt;part&gt;</code> tag defines detailed name spaces used in each name space described in the <code>&lt;message&gt;</code> tag.
<code>&lt;operation name="IFsearchPeriod"&gt;</code> <code>&lt;operation name="stationList"&gt;</code>	<code>&lt;operation&gt;</code> tag indicates functions to be called in remote servers.
<code>&lt;gds:translate type="text/xsl" href="http:// ... " /&gt;</code>	<code>&lt;gds&gt;</code> tag indicates URL of a XSLT file in which translation rules between input text commands and internal commands are described.

The WSDL, roughly speaking, defines what kind of parameters can or should be specified, how to translate input text commands into internal commands used in the data server and vice versa, i.e., the XSLT (Extensible Stylesheet Language Transformations), and which URL to send the users' requests. Users need not to know the differences between user interfaces of each data center, and the GDSClient can reduce drastically its system size owing to the WSDL files, although a WSDL file is not always necessary.

### 2.3 Detailed Functions of "GDSClient"

In both GDSClients of seismic and electromagnetic versions, the mandatory parameters to be input through the user interface are data centers, observatories, channels or components, and a data period. In order to avoid a data download of large amount, a limitation is set for each parameter not to overload the data server. When one of the input parameters or the file size of requested data exceeds these limitations, the GDSClient interrupts the downloading process showing a warning message. The GDSClient has an "advanced search" mode, which can specify auxiliary parameters in order to control the users' requests in more detail. In the case of the seismic version, it is possible to specify magnitudes of earthquakes and depths of hypocenters, for example. Adding to this, an interested area is also selectable from the world map in the case of seismic event data. In the case of electromagnetic version, the range of the Sigma  $Kp$  index, the quiet or disturbed days, or the range of missing rate can be specified in the advanced search. Then candidate data matched to the input search parameters are displayed. Most of data centers require inputting user's information such as name, affiliation, and e-mail address before the data download, and the GDSClient carry out this user information process instead of the data center (Figure 1). Finally it is possible to retrieve the geophysical data requested to the local PC by the HTTP.

## 3 FUTURE WORKS AND CONCLUSIONS

It is possible to download the GDSClient of the latest version from the website of Pacific 21 (Figure 3). The GDSClient is still now updating, e.g., increase the number of accessible data centers, more rapid communications with HTTP servers in sending requests and data retrieval, and more friendly user interface. We believe that the best solution for the collection of networked geophysical data is to utilize a web service technology like the GDSClient.

**PACIFIC 21**

**What's Pacific 21 ?**

Pacific21 is a generic name for the geophysical observation network in Western Pacific region operated by research institutions in Japan. Operation of the network is coordinated by the committee for global geophysical observation under the Science Council of Japan. Observed data are distributed through each research institution and Global Network Data Center of IFREE/JAMSTEC.

[Map of OHP Stations on Google Earth](#)      [Map of OHP Seismic Stations](#)

**Download OHP Data via NINJA**

<b>Seismic Continuous Data</b> <a href="#">Check Data Availability</a>	<b>Seismic Event Data</b> <a href="#">List of Seismic Stations</a>
<b>Geoelectromagnetic Data</b> <a href="#">Check Data Availability</a>	<b>Geodetic Data</b> <a href="#">Check Data Availability</a>

[NINJA Portal](#)

**Powerful Tools Developed by the Data Center of IFREE/JAMSTEC**

<b>GDSClient</b> software for collecting geophysical data	<b>Integration Tool for Geochemical and Geophysical Data</b>
--	--

**PACIFIC 21 Geophysical Data Service**

Geophysical data service is a distribution system of the geophysical data. It is based on the Web Services technologies.

GDS Client software, ver. 1.2

- [Read me](#)
- [Download software](#)

**Figure 3.** The left panel shows the top page of “Pacific 21” (<http://www.jamstec.go.jp/pacific21/>) operated by the data center of IFREE/JAMSTEC. It is possible to access to the NINJA interfaces and the NINJA portal from this page, and users can download each kind of geophysical data, i.e., seismic continuous, seismic event, electromagnetic, and GPS data obtained in the OHP network. The website of the GDSClient shown in the right panel can be accessed from the link (red circle) on the top page, and the GDSClient packages of both seismic and electromagnetic versions and their operation manuals are available from this page.

## 4 ACKNOWLEDGMENTS

We are grateful to the members of the Earthquake Research Institute of the University of Tokyo for their cooperation on the distribution of the OHP data.

## 5 REFERENCES

- Kawakatsu, H., Bina, C., & Forte, A., (Eds.) (1998) Ocean Hemisphere network Project (OHP) special issue, *Earth Planets Space*, 50, 893-1065.
- Nagao, H., Tsuboi, S., Ishihara, Y., & Yanaka, H., GDSClient: Collecting tool for geophysical data using web service technology, (2008) *Abstracts of Japan Geoscience Union Meeting*, J161-P003.
- Takeuchi, N., Tsuboi, S., Watada, S., Fukao, Y., Kobayashi, M., Matsuzaki, Y. & Nakashima, T., Application of distributed object technology to seismic waveform distribution, (2002) *Seism. Res. Lett.*, 73, 166-172.
- Tsuboi, S., Takeuchi, N., Ishihara, Y., Nagao, H., Yamagishi, Y., Watanabe, T., Yanaka, H., & Yamaji, H., Development of XML schema for broadband digital seismograms and data center portal, (2008) *Abstracts of Japan Geoscience Union Meeting*, S144-005.