

# Participation of Roshydromet's World Data Centers in the International Polar Year 2007–2008 Programme

M. Z. Shaimardanov, A. M. Sterin, and A. A. Kuznetsov

All-Russian Research Institute of Hydrometeorological Information – World Data Center (RIHMI-WDC), Roshydromet, Obninsk, Russia

Received 16 September 2007; accepted 26 October 2007; published 15 December 2007.

[1] RIHMI-WDC, which performs the functions of the World Data Center B on Meteorology and Oceanography, is to considerable extent involved in the implementation of both national and international components of the IPY Programme. To ensure IPY data management several regulatory documents have been developed ranging from the “Concept of information exchange in the period of the International Polar Year 2007–2008” to data collection and storage formats. Basic principles of national data collection and exchange are determined in the document “Policy of national data management in the period of the International Polar Year 2007–2008.” The main working document to be used by all IPY participants is the “IPY 2007–2008 Data Management Plan.” To implement data integration and exchange the system IPY-Info has been developed. It is based on the recent advancements in the field of Web-technologies, specifically on the technologies and information structure developed under the Unified State System of Information on the State of the World Ocean (UISWO). The system will make it possible to monitor IPY data and metadata collection and allow users to perform remote search, displaying and retrieval of IPY data irrespective of their physical location. *INDEX TERMS:* 3349 Atmospheric Processes: Polar meteorology; 5462 Planetary Sciences: Solid Surface Planets: Polar regions; 5754 Planetary Sciences: Fluid Planets: Polar regions; 9315 Geographic Location: Arctic region; *KEYWORDS:* WDC System, International Polar Year, meteorology, oceanography, data archiving.

**Citation:** Shaimardanov, M. Z., A. M. Sterin, and A. A. Kuznetsov (2007), Participation of Roshydromet's World Data Centers in the International Polar Year 2007–2008 Programme, *Russ. J. Earth. Sci.*, 9, ES3001, doi:10.2205/2007ES000232.

## 1. Introduction

[2] The state organization RIHMI-WDC fulfils functions of the World Data Center B on meteorology and oceanography and is in a considerable degree involved into realization of the national and international parts of the IPY program, in particular, in solution of the following problems:

- Development of the interdepartmental standardizing base, order and standing orders of collection, accumulation, exchange and distribution of the data and information on IPY 2007–2008, taking into account the Russian laws and international standards in this field;
- Creation of a complete and high-quality database accumulating the results of scientific studies during IPY

2007–2008 for the disciplines allocated to the World Data Center B, and also the observational data in the polar regions of the Earth for the previous historical period;

- Creation of the integrated system of distribution of information resources of IPY on the basis of Web-technologies including the system of management of the distributed data bases of IPY and means of the remote access of the users;
- Creation of a reliable national system of management of the IPY data including monitoring of the data collection and exchange.

## 2. IPY Data Management

**Table 1.** Disciplinary Data Centers of IPY

Brief name of the organization	Disciplines (types of observations)	The organization address
AARI	Ice observations	38 Beringa st., St.-Petersburg, 199397
RIHMI-WDC	Hydrometeorology, oceanology, climatology	6 Koroleva st., Obninsk, Kaluzhskaya Region, 249035
GC RAS	Geophysical data	3 Mologezhnaya st., Moscow, 119296
GIN RAS	Geological data	7 Pyzhevskii lane, Moscow, 119017
ZIN RAS	Biological data	1 University emb., St.-Petersburg, 199034
IG RAS	Ecosystems, glaciology, and social-economy development	29 Staromonetny lane, Moscow, 109017
PMRI NSMU	Medical and biological data	51 Troitskii ave., Arkhangel'sk, 163000

management system. It is obvious that it should provide integration of the observational data and scientific investigations concerning the polar regions of the Earth, be dynamical concerning variety of forms of the collected data, means of their delivery and transfer, be based on the acting systems of collection and distribution of the data on the natural environment. In essence, the completeness and the level of realization of this task determines the efficiency of realization of the scientific program of IPY on the whole, providing integration of heterogeneous data and information, well-timed access to information resources for study and practical application.

[4] In order to provide the management of the IPY data, a whole series of normative-methodical documents, beginning with “Concept of Information Exchange in the Period of IPY 2007–2008” and ending by the formats of collection and storage of the data has been developed. The main principles of the collection and exchange of the national data based on the statements of the international policy of data exchange are determined in the document “Policies of the Data Management in the Scientific Program of Participation of Russian Federation in Conduction of the International Polar Year”. The main guidelines which should be used by all participants of IPY is “Plan of Management of the IPY 2007–2008 Data”.

[5] According to the “Concept”, the IPY data management is performed by the organizations-participants of the national program of IPY and the data centers of IPY nominated out of the organization-participants of the National Program of IPY on the basis of the disciplinary principle. The right and obligation to fulfill the functions of the IPY data centers are given to the international data centers (WDC) in the ICSU system [Webster, 1996] and to some other organizations (research institutes) having the corresponding experience and infrastructure abilities (Table 1).

[6] The methodology of the data exchange in the period of IPY conduction is based on the end-to-end (from observation platform to the final user) management of the data on the natural environment state as a mutually related set of information, technological and organization decisions realizing integrated processes of collection, accumulation, process-

ing and distribution of the data and information production during IPY.

[7] The experience in conduction of the previous observational programs and experiments (such as ALPEX, [http://findarticles.com/p/articles/mi\\_m1310/is\\_1987\\_Feb/ai\\_4752956](http://findarticles.com/p/articles/mi_m1310/is_1987_Feb/ai_4752956); GARP, [http://gcmd.nasa.gov/records/GCMD\\_ds304.0.html](http://gcmd.nasa.gov/records/GCMD_ds304.0.html); and Transects, [http://books.nap.edu/openbook.php?record\\_id=1521&page=15](http://books.nap.edu/openbook.php?record_id=1521&page=15)) showed that it is reasonable to select three levels in the collection system (see Figure 1) which in application to the IPY program could be called in the following way:

- Sea and land expeditions;
- Organizations executors of the IPY projects;
- IPY data centers (World Data Centers and specialized data centers).

[8] At each level, particular (pertinent only for this level) functions on reception, processing and transfer of the data to higher level are executed. The IPY-info database performs functions of data and metadata integration interacting with the two upper levels.

[9] At the expedition level, preparation and formatting of the data, metadata and documentation (scientific and technical reports, etc.) is performed in the electronic form according to the recommendations. The obtained observational results are compiled in the form of separate disciplinary data sets and are transferred to the organizations what are users of this or that type of the observations (often they are simultaneously the conductors of the observations).

[10] The observational data and related metadata obtained in an expedition in the scope of an IPY project should be presented into the organization conducting the project.

[11] In organizations-performers, the data should be registered, checked in the content, form of presentation, and quality. The data could also be undertaken to an additional processing in order to fit the established standards.

[12] The task of the transfer of observational data and scientific studies (the produced data) obtained in the scope of the projects and of the accompanying metadata into the

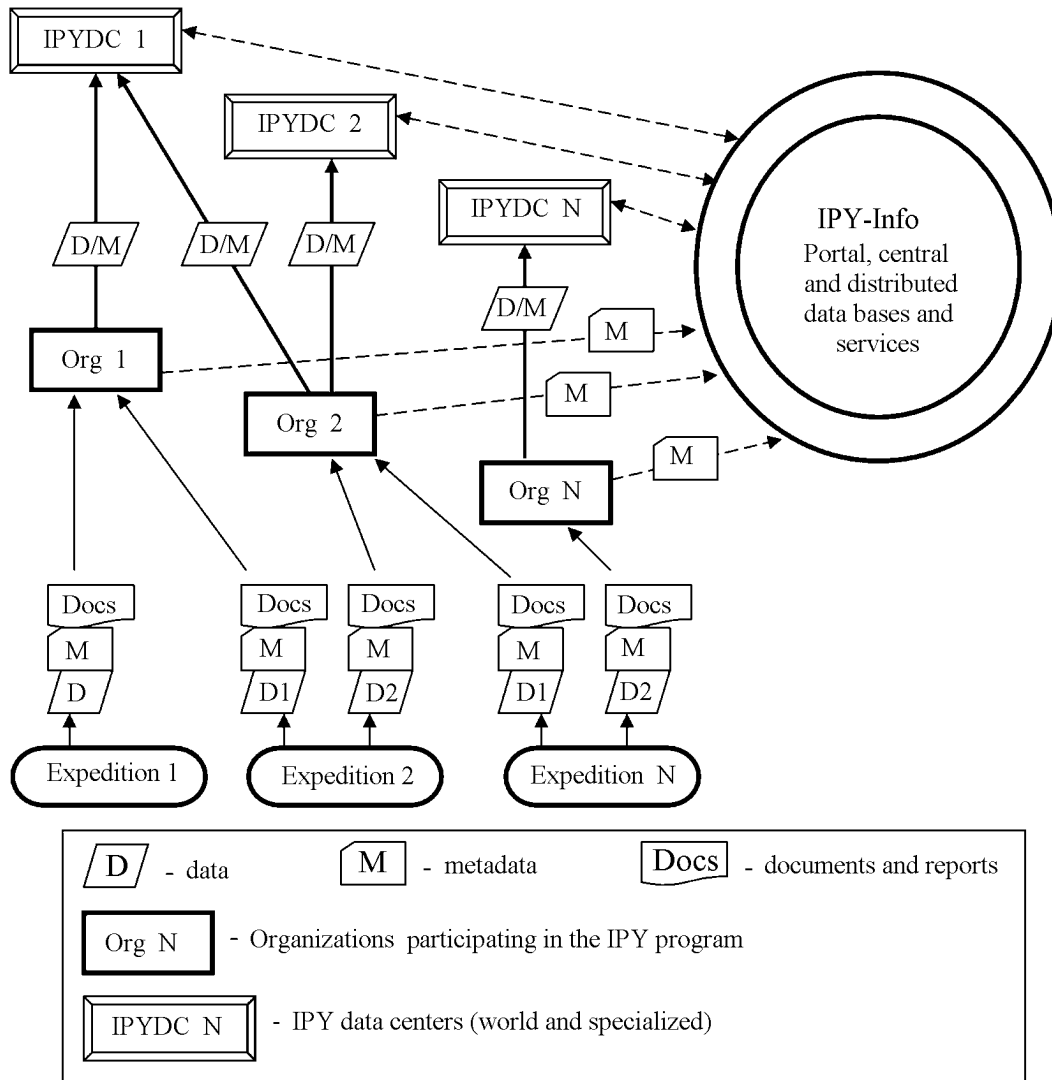


Figure 1. The general scheme of collection the IPY data and metadata.

data centers of IPY (depending on their disciplinary areas) is laid on the organizations conducting IPY projects and on the projects leaders. The task is also laid on these organizations of including of the metadata (description of data sets, models, software, and formats) into the Centralized metadata base (CBMD) by filling in of the corresponding Internet-forms of the IPY-Info system.

[13] The principal feature of the collection system due to the application of WEB-technologies is the fact that the metadata, in the first turn the descriptions of expedition observations and compiled data sets, are placed in CBMD of IPY-Info by the organizations-performers of the projects and become available to the users before the data itself arrive to IPY data centers.

[14] The IPY data centers perform registration and control of the presented data, formation of complete high-quality bases of real-time and archival data, inclusion of the data into the distributed system of the IPY databases and granted

storage of the data according to the profile of the center activity. They also take measures for widening of the information resources of IPY-Info by involving data from other sources including commercial and international organizations.

[15] Long-term storage is realized by recording of the accumulated data, metadata, and their catalogs at technical carriers of information. A robotized library on magnetic cartridges is used for this aim in RIHMI-WDC, laser disks are used in other organizations.

[16] The international exchange of IPY data is carried out in the scope of international projects registered by the joint ICSU/WMO Committee on planning and coordination of the IPY performance in the order established by normative law acts of Russian Federation in this area. Copies of the data presented by the organizations-performers of projects for the international exchange, in an obligatory way are sent to the World Data centers (RIHMI-WDC, AARI, WDC-B

**Table 2.** Formats Used for Collection of the IPY Data

Format name	Format aim	References
SEAGRID	For the data on sea ice	<a href="http://www.aari.nw.ru/gdsidb/docs/wmo/sigrid%20(2).pdf">http://www.aari.nw.ru/gdsidb/docs/wmo/sigrid%20(2).pdf</a>
AEROMETACT	For the data of marine meteorological, actinometrical, and aerological observations on board Research Vessels	<a href="http://data.oceaninfo.ru/meta/formats/index.jsp">http://data.oceaninfo.ru/meta/formats/index.jsp</a>
IMMT3	For the data of shipboard meteorological observations conducted at HMS of the 3 grade	<a href="http://data.oceaninfo.ru/meta/formats/index.jsp">http://data.oceaninfo.ru/meta/formats/index.jsp</a>
OCEANPIBM	For the data on physical oceanography	<a href="http://data.oceaninfo.ru/meta/formats/index.jsp">http://data.oceaninfo.ru/meta/formats/index.jsp</a>
PERSONA	For hydrometeorological data of the seashore stations	<a href="http://data.oceaninfo.ru/meta/formats/index.jsp">http://data.oceaninfo.ru/meta/formats/index.jsp</a>
INTERMAGNET	For the data on geomagnetism (hourly and minute)	<a href="http://www.intermagnet.org/FormatData_e.html">http://www.intermagnet.org/FormatData_e.html</a>
GRID	For the data in the mesh-point of the regular net	<a href="http://www.wmo.ch/web/www/DPS/grib-2.html">http://www.wmo.ch/web/www/DPS/grib-2.html</a>
Exchange format	For the data of scientific studies and observational data having no established format	<a href="http://data.oceaninfo.ru/meta/formats/index.jsp">http://data.oceaninfo.ru/meta/formats/index.jsp</a>

on Marine Geology and Geophysics, WDC on Physics of the Solid Earth, and WDC on Solar-Terrestrial Physics) according to the disciplinary principle.

[17] Provision of the availability of the obtained data to a wide range of users nowadays and in the future is one of the principal requirements of the IPY data politics.

[18] One should understand the availability in a wide plane not only as a physical ability for the users to obtain a copy of the data on a technical carrier or to look them through via the Internet, but also the form (format) of the data presentation making its use easier.

[19] First of all, one should mention the requirement according to which the data should be accompanied by the metadata, that is, by the complete description needed for understanding of the data and their correct application by other users.

[20] Due to the multi-disciplinary character of the IPY program, especially urgent are the problems of standardization of the data and metadata forms. Because of a large variety of the observed parameters and regions of application of the information on polar regions of the Earth, development of the standard format of the data and organization of its application in the practice of information activity present a complicated problem.

[21] To solve this problem, the Plan of the IPY Data management plans application, on one hand, of formats widely used in various disciplines, in particular, for collection and storage of the data of regular meteorological, oceanographic, geophysical and other observations (Table 2).

[22] On the other hand, for the observational data having no formats, or instead of formats having found no broad use, it is proposed to use the specification of the format of data exchange on the marina environment developed in RIHMI-WDC (RIHMI-WDC Technical Specifications, 2007).

[23] This specification combines three classes of formats: the so-called table-text format, format of a document, and format of a map.

[24] The table-text format is, as a rule, used for the results of observations expressed in numerical, or alphabetic-numerical form. Widely used program products of the MsOffice (Excel, MsAccess), or in the worst case any text editor could be used for preparation of the data in this format. However using the mentioned software does not guarantee fulfilling the standard and correspondence between the data description and their actual structure. The data in the table-text format are accompanied by a formalized description which makes it possible to perform a syntactical control of the data and their grouping over elements. So an autonomic program is supplied to check the format of the collected data as well as syntactic and minimal semantic verification.

[25] A simple text format (\*.txt), or the format of the MsWord (\*.doc) and Windows RTF (\*.rtf) editors are recommended as formats for documents. The formats \*.doc or \*.rtf are recommended to be used for text-graphic information. The exchange SHP-format used in GIS is used for cartographic information.

### 3. Integration and Giving the Access to the IPY Data

[26] In current conditions the term “data exchange” includes not only the exchange and distribution of data files on technical carriers, for example, on laser disks, but also providing a remote access to these data using the Internet. For the realization of this task, the IPY-Info System is developed

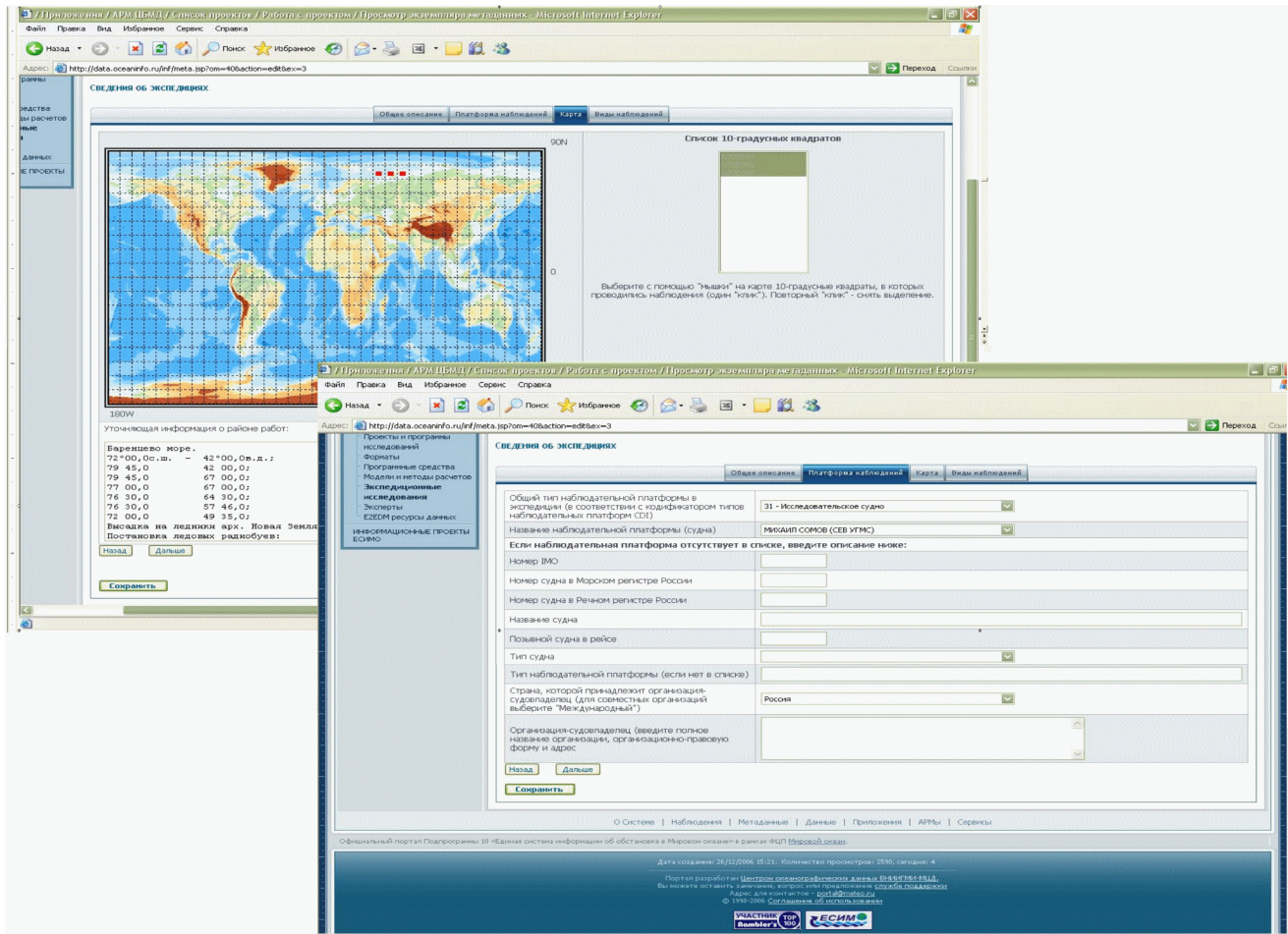


Figure 2. Interface of AWP of the remote entry of an expedition description.

based on the recent achievements in the web technologies, in particular, on the technologies and information infrastructure created in the scope of the Unified Information System on the State of the World Ocean (Government Resolution of 29 December 2005).

[27] One of the principles of the exchange politics is a compulsory accompanying of the data by the metadata, that is, by the information on measuring devices and methods of measurements or observations, on observational platforms, methods of the data processing, etc. To save the metadata in the IPY-Info System database, remote working places are developed making it possible to the users-providers introduce the metadata via corresponding forms using the Internet (Figure 2, see also <http://www.mpg-info.ru>) and to save the metadata in CBMD.

[28] Thus, the descriptions of the expeditions carried out, scientific studies, formed data sets and databases become available to a wide circle of users.

[29] The access to the data is realized on the basis of the technology of integration of distributed information resources (E2EDM) which makes it possible to perform a remote search, performance, and selection of the IPY data for scientific analysis or for specialized processing independently

of their physical location.

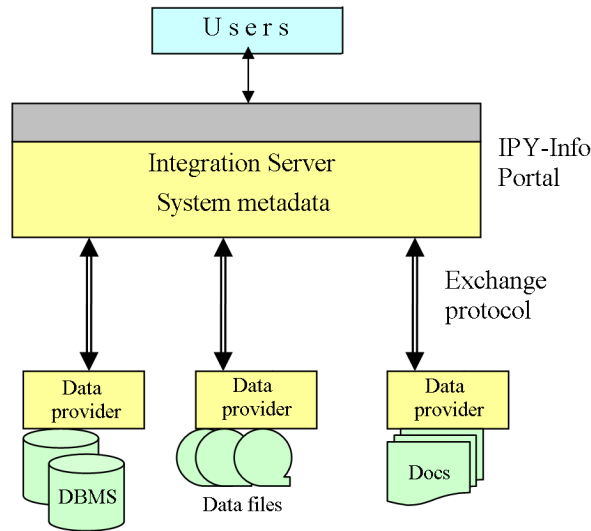
[30] This technology makes it possible to connect up to the system information resources what would be accumulated in the IPY data centers. The following type of the data and metadata are permitted to be stored in the centers:

- SQL databases;
- Structured (by a certain format) data files;
- Unstructured data files (simple text, documents, images, etc.).

[31] The program realization of the E2EDM technology is based on the customer-server model (Figure 3) and consists of program complexes (Data Provider and Integration Server), interacting according to the Protocol of the E2EDM Data Exchange.

[32] The Integration Server Program Complex is aimed to performance of the following goals:

- Identification and formalization of the distributed data sources;
- Registration of the distributed data sources after the installation of the Data Provider;



**Figure 3.** Architecture of the integration technology of E2EDM.

- Realization of the mechanism of interaction between the external users of the Integration Server and the Data Providers;
- Monitoring of the data sources and of their resources;
- Provision of the unified system of parameters coding.

[33] The Data Provider Program Complex is aimed at fulfilling of the following tasks:

- Switching in of the local system to the distributed data base system DDBS;
- Processing of the external calls for the Integration Server;
- Appealing to the local data system in the case of a request for data from the Integration Server and recoding it into the format of the transport file (object data files are transferred to the Integration Server without transformations);
- Generation of messages-responses to the Integration Server;
- Synchronization of codes in the metadata;
- Automatic support of the descriptions of the data source resource.

[34] When the Data Provider is on the side of the local data, that means that the local data system becomes a source of a system of distributed databases and this source may be involved into the automatic information data exchange. At the initial stage, it is planned to connect by the E2EDM technology the information resources of the IPY data centers, and in the future to involve other organizations.

#### 4. Formation of the IPY Data Depository on Meteorology and Oceanography

[35] According to the concept, the formation of the IPY data fund is realized in the form of a centralized base of metadata and a system of distribution of the disciplinary databases created in the IPY data centers.

[36] According to the disciplinary tasks of RIHMI-WDC, as the WDC-B on meteorology and oceanography, there will be formed a fund of IPY data. The fund will include data of the current and historical observations carried out at meteorological, aerological, and sea-shore stations and observational points located in the polar regions and also on board of research vessels and passing ships.

[37] In particular, a specialized set of historical and current meteorological data for the polar zone of Russia will be formed. The set will include the daily mean and monthly mean values of such elements as the air temperature, pressure, wind, relative and absolute humidity, visibility, number of days with mists, characteristics of the snow cover, and characteristics of precipitations. Similar set of data over the polar regions of the Earth will be prepared on the basis of observations conducted at the aerological stations.

[38] For the marine parts of the polar regions, a set of the meteorological data obtained at ships over the polar regions of the Earth will be prepared in the IMTT-3 format which will include the historical and current data on the water surface temperature, waves, wind velocity, ice conditions and other parameters.

[39] The set of oceanographic data will also include the results of historical and current observations by various instrumental means of the vertical sounding of the ocean: measurements of currents, waves, and water level.

[40] The real-time flows of meteorological and oceanological data collected via channels of the WMO Global Telecommunication System will be located in the real-time database and also will be available to users via Internet. First of all, those are the data of meteorological stations located in the sea-shore zone and at islands over the North Polar Circle, the data of meteorological observations from passing ships and drifting buoys.

[41] According to the Plan of Data management, RIHMI-WDC is also responsible for disposing at the IPY-Info Portal of the data of observations and scientific investigations what would be obtained in other organizations in the scope of IPY projects related to studies of the climate of the atmosphere and ocean and providing the access to these data.

#### 5. Monitoring of the Data Collection and Formation of the IPY Data Depository

[42] The monitoring of realization of the Plan of IPY Data management in the part of data accumulation and formation of the IPY Data Depository will be performed on the basis of the technologies laid into IPY-Info. The main documents used for estimation of the correspondence of the obtained and included into the Depository data to the expected re-

sults are Plan of Realization of the Scientific Program of Participation of Russian Federation in International Polar Year 2007–2008 and Plan of IPY Data management.

[43] First of all, the results of observations conducted in every expedition should be described in IPY-Info by the organizations heading the expeditions via the specialized remote AWP. In the same way, the projects performed in the frames of IPY and the sets of observational data and scientific analysis obtained in them should be via the remote AWP described and registered in the central base of the IPY metadata. It would fix in IPY-Info the relation between IPY projects and organizations responsible for their performing, on one hand, and the created information production, on the other. On this basis, automatic analysis and compilation of reports on the state of formation of the IPY Data Depository

at its various cuts, directions, disciplines, and organizations will be performed.

## References

Webster, F. (1996), *World Data Centre System: Geophysical, Solar, and Environmental Data, Principles, Locations, Data, Services, April 1996*, 109 pp., ICSU Panel on World Data Centres, Paris, Boulder.

---

M. Z. Shaimardanov, A. M. Sterin, and A. A. Kuznetsov, All-Russian Research Institute of Hydrometeorological Information – World Data Center (RIHMI-WDC), Roshydromet, Obninsk, Russia