

# minib 11

marketing of scientific  
and research organizations  
no. 1(11)/2014



***DIGITAL REPOSITORY OF SCIENTIFIC  
INSTITUTES — RCIN***

## DIGITAL REPOSITORY OF SCIENTIFIC INSTITUTES — RCIN

**Kamila Kaczyńska. M.Sc.**  
Institute of Aviation, Poland  
e-mail: Kamila.Kaczynska@ilot.edu.pl  
DOI: 10.14611/minib.11.01.2014.13



### Abstract

The paper describes the Project of Digital Repository of Scientific Institutes RCIN and presents opportunities for promoting science by digitization and sharing them on the Internet. The Repository has been created by the 16 Scientific Institutes in Warsaw, Krakow and Bialowieza to modernize the science-research and IT infrastructure, to increase digital resources of mathematical, technical, natural and medical sciences, and to popularize and promote of Polish science. That dissemination and popularization of science affects its development and competitiveness in the international arena and it allows transfer of research results to the economy. In addition, Institutes of RCIN providing contemporary and archival materials of science, support the intellectual capital of Polish science and raise awareness of professional literature of search on the Internet. Project RCIN is implemented in the years 2010–2014 and financing is provided by the funds of the European Fund of Regional Development.

**Keywords:** digitalization, RCIN, infrastructure, human resources, research institute

The age of the “new renaissance”<sup>1</sup> in Polish libraries started at the beginning of the 21<sup>st</sup> century. Digitalization<sup>2</sup> of collections became one of the main goals for libraries. On the one hand digitalization of collections on the one hand secures the originals, on the other hand the process allows disseminating collections and promoting institutions. The whole process of building digital libraries is very complicated and requires the involvement of many specialists. In many cases libraries abandon plans for digitalization of their resources for fear that they may encounter difficult obstacles. There are also libraries which take up the challenge by forming consortia with other libraries.

This way over 100 digital libraries<sup>3</sup>, including Digital Repository of Scientific Institutes<sup>4</sup> have been formed on the Polish library arena. This lecture describes the process of establishment of RCIN, its infrastructure, goals and the benefits from RCIN.

## **RCIN — the beginning**

Digital Repository of Scientific Institutes was formed in order to satisfy the needs of scientists associated with search for archival and latest results of research, information and bibliography. Such an integration, dissemination and commercialization of Polish science boosts its development, competitive edge on the international arena and makes it possible for the economy to take advantage of research results. The goals of RCIN are noble and fill a gap in the world of digitalization. RCIN opens Polish science to the world and promotes it by giving an opportunity to learn about its achievements. Apart from scientific publications and archives, yet not published versions of works and the results of research from the area of mathematical, technical, life, medical, social and humanistic sciences will be published. Resources of the Repository contribute to the national Digital Libraries Federation, which strengthens the awareness of searching for scientific information in digital form, has an impact on dissemination of research results and growth of the number of citations.

## **Raising funds**

As has already been mentioned, the process of digitalization is very complicated and this means that it is also expensive. Scientific institutions themselves, despite

strong determination, are unable to cover the costs, that's why raising external funding is such an important element. This is one of challenges that the members of RCIN consortium had to face. By joining forces, 16 scientific institutes from Warsaw, Kraków and Białowieża gained financial support worth more than PLN 36m from the European Regional Development Fund. The four-year project financed with the assets of the Operational Programme Innovative Economy has been carried out since 2010 by:

- Museum and Institute of Zoology, Polish Academy of Sciences (MiIZ PAN) — leading institution
- Institute of Geography and Spatial Organization of the Polish Academy of Sciences (IGiPZ PAN)
- Institute of Mathematics of the Polish Academy of Sciences (IM PAN)
- Institute of Organic Chemistry of the Polish Academy of Sciences (IChO PAN)
- Institute of Physical Chemistry of the Polish Academy of Sciences (IChF PAN)
- Institute of Literary Research of the Polish Academy of Sciences (IBL PAN)
- Institute of the Polish Language of the Polish Academy of Sciences (IJP PAN)
- Institute of Archeology and Ethnology of the Polish Academy of Sciences (IAiE PAN)
- Institute of Slavic Studies of the Polish Academy of Sciences (IS PAN)
- Nencki Institute of Experimental Biology (IBD PAN)
- Mirosław Mossakowski Medical Research Center of the Polish Academy of Sciences (IMDiK PAN)
- Institute of Philosophy and Sociology of the Polish Academy of Sciences (IFiS PAN)
- Institute of Fundamental Technological Research of the Polish Academy of Sciences (IPPT PAN)
- Tadeusz Manteuffel Institute of History of the Polish Academy of Sciences (IH PAN)
- Mammal Research Institute of the Polish Academy of Sciences (IBS PAN)
- Institute of Electronic Materials Technology (ITME)

The raised funds have been allocated to covering the costs of hiring and training employees, purchasing high-class equipment (computers with peripherals, scanners), software, among others, dLibra 5.0 [2] and servers (Libra, OCR DjVu and Fine Reader), on adaptation and modernization of rooms, maintenance of infrastructure, as well as on external services, eg. legal consulting etc.

## Human Resource Management

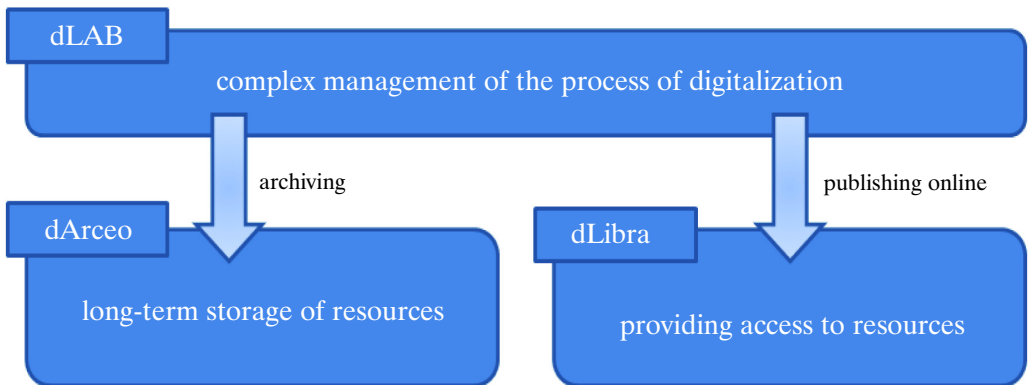
Due to the fact that the process of digitalization is the latest trend in Polish libraries, there are not too many specialists on the market. In such case it is necessary to pay much attention to training employees and raising their qualifications. This is also what happened in case of establishment of the Repository. Particular institutes raised employment by gaining additional librarians and IT specialists. Both groups of employees were trained in the area of digitalization. Librarians were taught both about the rules of creating metadata<sup>5</sup> and their Dublin Core<sup>6</sup> format, as well as about cooperation with NUKAT Centre<sup>7</sup>. Employees scanning library resources learned about the highest quality equipment and software. The efficiency of trainings was raised by training bases dLibra and Horizon<sup>8</sup>. Next, the scope of duties was delegated to particular institutions forming the Repository.

## Infrastructure

What reflects the complexity of the process of digitalization is the need for modernization and expansion of infrastructure, especially IT infrastructure. Close cooperation of 16 institutes is the main concept of the Repository. In order to coordinate their activities and raise efficiency a system of three programmes developed by Poznań Supercomputing and Networking Centre<sup>9</sup> has been used. For the purpose of building the digital library the most popular software on Polish market — dLibra — has been used. The software makes it possible to store files in any format (PDF, DjVu, MP3, FLV, JPG), importing from external databases, managing metadata, delegating users' powers, flexibility and scalability of portal's development, compatibility with external IT systems and library standards. dArceo system is responsible for storing and archiving source text, graphic and audio-visual data. The process of digitalization is supervised by dLab system, which makes it possible to monitor the course of work together with reporting.

Such complex approach to project management was necessary to coordinate the work of not just 16 institutes, but also 5 specialist scanning workshops, which carry out their orders on a rotational basis. For all workshops high-quality scanners, including fragile material scanners were purchased. These special scanners make it possible to scan in colour old books and publications without UV rays. Moreover, scanners for microfilms, sheet-feed scanners allowing double-sided scanning of catalogues and large-format scanner for maps and charts were purchased.

Picture 1. dLAB, dLibra and dArceo systems — rule of cooperation



Source: <http://dlab.psn.c.pl/kompleksowe-wdrozenia/>, 15.10.2013.

Before the scanned materials are presented in the RCIN database, documents have to be processed using the OCR method. Also, it is necessary to create metadata and the so-called parent files have to be archived.

The implementation of the project influenced the IT infrastructure of institutes: speed of Internet connections was improved, server rooms were expanded and an internal site for coordination and administrative, accounting and substantial cooperation of employees was established.

### Selection and providing access to resources

In the process of granting access to resources to a broader group of Internet users, selection of materials is very important. This seemingly nicest part of the process of digitalization quite often brings most problems. Every work is subject to legal regulations defined by:

- Act on protection of databases from 2001. (Dz.U. 128, poz. 1402 ze zm.)
- Civil code from 1964 (Dz.U. nr 16 poz. 93 ze zm.)
- Act on copyright and associated rights from 1994 r. (Dz.U. z 2006 r. nr 90, poz. 631 ze zm.)

Works presented on the Internet have to be subject to analysis in terms of often

vague copyrights. This is the most dangerous area, that's why libraries often resort to the expertise of specialists — lawyers. It turns out that almost anything associated with a piece of work can be subject to copyright. Beginning with the text of the work, to the cover and pictures located within the text. In case when the situation is not clear, works are published on internal pages and access to them is granted to a chosen group of users. The problem disappears as soon as a particular publication becomes a part of public domain, that is, it is not legally protected by copyright, when copyright has expired or the work is not creative or individual in character and thus isn't legally protected<sup>10</sup>. This happens when 70 years have passed since the death of the author, or when we are dealing with official documents. In case of scientific institutions such restrictions limit the room for promoting latest research results. One of solutions to this situation is signing an exclusive, or non-exclusive license agreement with the holder of copyrights. Such a move resolves all doubts and allows full publication. Being aware of and knowing legal regulations is just the first step towards success. Contacting creators and getting a license often turns out to be impossible, despite the involvement of resources. The process of gaining approval for dissemination is often highly time-consuming and expensive. In case of orphaned works, that is, in cases where it is impossible to identify the authors of works, digitalization is impossible.

In the RCIN project a very responsible approach to the issue of copyright was adopted. It was assumed that everything subject to digitalization is the result of creative activity and materials which raise any doubts are published according to the rules on fair use<sup>11</sup>, on external terminals. Only 23% of publications<sup>12</sup> found in the Repository belong to the public domain, that's why the users of RCIN have to display understanding and respect for law and the interest of the authors. In case when access to certain materials is restricted, it is possible to view them in the institute holding their physical version.

The main assumption of the project is promoting digital resources and granting possibly easiest access to them. Thus, every user with a computer and Internet access can view the resources of the Repository through a browser displaying graphics (Mozilla Firefox version 2.0 and higher, Internet Explorer version 6.0 and higher, Google Chrome, Opera, Safari). Books, magazines, catalogues, maps and audiovisual resources are presented in PDF, or DjVu format, depending on the size and resolution. The user's computer must have a virtual Java machine from Sun or Microsoft corporation installed. The user should also have Adobe Reader, DjVu and Media Player applications for viewing files. Sound files are presented in mp3,



or mp4 formats and audiovisual materials are provided in avi, wmv, mp4 or flv formats, which don't require installing additional applications or plug-ins.

As has already been mentioned above, the access to resources is not equal for everyone. Depending on their rights, users have access to chosen collections. Regardless of this, every reader can set up an individual account and create a list of favourite works, publish it through RSS channel and add own tags to the publications.

## Collections

In order to systematize the resources of the Repository a catalogue containing collections and their sub-collections common for all 16 institutes was established. The collections were defined according to subject area. The following are the main collections:









- Books
- Magazines
- Didactics and popularization of science
- Maps and atlases
- Diploma theses
- Institutes' publications
- From the activity of Institutes
- Cultural heritage
- Regional issues
- Manuscripts
- Old prints
- Exhibitions

Apart from this, collections of each institute were allocated to separate categories. For example, in the main collection of the Institute of Electronic Materials Technology we can find the following sub-collections:

- Electronic materials
- ITME works
- Doctoral and habilitation works
- Articles
- Books

Every work in the Repository is ascribed to at least one collection. Moreover, publications contain formal information, that is, metadata (author, title, year of publication and key word), which also facilitate browsing through and viewing the content of collections.

Looking through metadata it is possible to obtain all formal information concerning a publication. Moreover, in the upper left corner of the page there is a menu with icons<sup>13</sup> serving the purpose of managing the viewed document:

-  Description (contains metadata)
-  Information (concerning allocated collections, date of last modification)
-  Structure (list of contents)
-  Content (with the option of choosing the display of the document)
-  Content (new window)
-  Download
-  Similar publications (link to similar publications)
-  Display options

Another function facilitating the use of the Repository is grouping works. This function is especially useful in case of magazines and series. Document, as an individual publication is ascribed to an integrating collection shown under one, common name such as magazine title. Such objects are only “paperclips” which contain no content.

You can search for publications using a traditional search bar placed on the website. It is possible to use the line of contextual search by entering a piece of searched title. It is also possible to use the option of advanced search and expand questions with logical operators. Moreover, the search website presents ten recently updates and ten most often viewed publications.

## Promotion

The fact that a scientific institution has a digital repository contributes to the institution's prestige and allows the institution to promote itself by publishing its works. Such activities are associated with professionalism, openness to the world and following the latest trends. Digitalization is currently the main topic in libraries. It is a goal which is supposed to transfer libraries to the world of new technologies and prove their devotion to progress. An institution which invests in employees and IT infrastructure is

regarded as modern and friendly to the external environment. This is very important in the process of building a company's image.

RCIN project is aimed mainly at the Polish and foreign scientific sector, including scientific and research units, higher education institutions, industrial sector e.g. branch centres, economic sector, cultural institutions, as well as private individuals, hobbyists. Supraregional character of digitalization makes it possible to lift all geographical and social limitations. Thanks to this the recipients = clients can be both high-school graduates from small Polish agglomerations, as well as scientific employees of the biggest research institutes on the other side of the Atlantic. This diversified group of recipients has one thing in common, that is, the need to gain knowledge. In this case the main tool for promotion is a well-positioned Internet website.

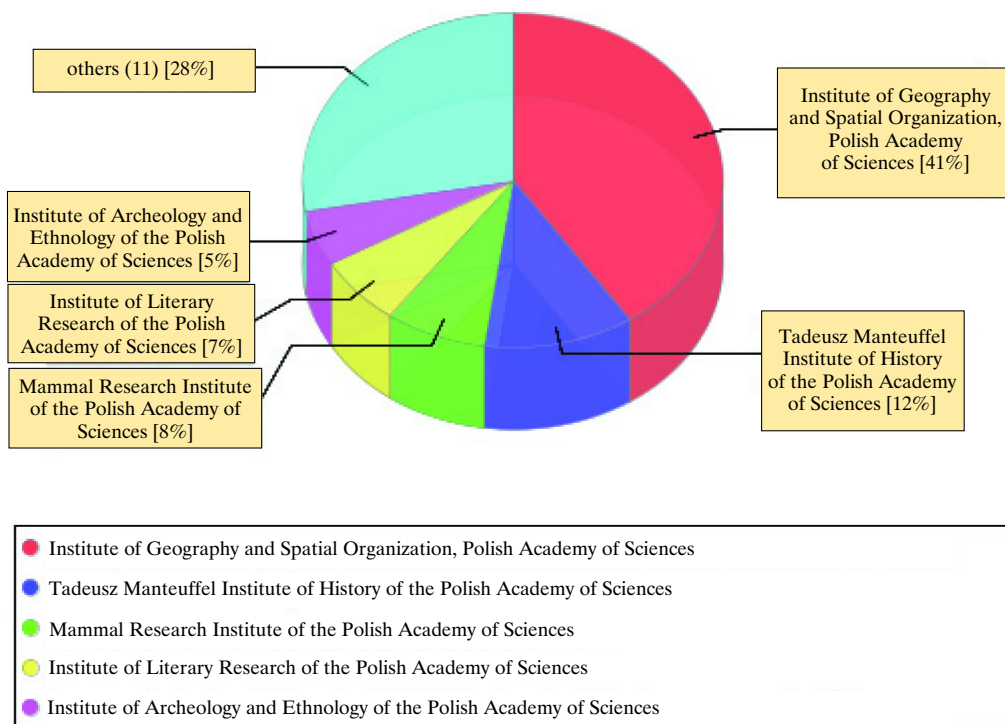
This is also the case with Digital Repository of Scientific Institutes. An Internet website <http://rcin.org.pl/dlibra> devoted to the project has been created. On the website you can find logos of institutions participating in the RCIN project. The website serves the purpose of searching for and browsing the contents of collections. Moreover, on the website of each institute there is information about the Project. Thanks to this consortium partners can promote each other. Institutes build their image by providing own publications in the Repository, by holding meetings at which recipients learn about the product. They also give lectures on the subject at conferences. A nice gesture which strengthens the awareness of the brand among the recipients is distributing gadgets with the logo **RCIN**.

The coverage of the venture is best illustrated by figures. Since July 1, 2008 a total of almost 4.5 million readers have been registered. The following chart shows the way they used RCIN's resources.

The figure in brackets is the total of views of all publications containing the provided attribute<sup>14</sup> value in the description.

1. Institute of Geography and Spatial Organization, Polish Academy of Sciences [641327]
2. Tadeusz Manteuffel Institute of History of the Polish Academy of Sciences [182071]
3. Mammal Research Institute of the Polish Academy of Sciences [122073]
4. Institute of Literary Research of the Polish Academy of Sciences [106731]
5. Institute of Archeology and Ethnology of the Polish Academy of Sciences [85902]
6. Nencki Institute of Experimental Biology [72066]
7. Institute of the Polish Language at the Polish Academy of Sciences [67822]
8. Institute of Philosophy and Sociology of the Polish Academy of Sciences [55923]

Picture 2. Chart showing the proportions of displayed publications



Source: <http://rcin.org.pl/dlibra/pubstats?statType=29>, 18.11.2013.

9. Museum and Institute of Zoology of the Polish Academy of Sciences [53395]
10. Institute of Mathematics of the Polish Academy of Sciences [45850]
11. Institute of Electronic Materials Technology [40127]
12. Institute of Physical Chemistry of the Polish Academy of Sciences [37959]
13. Mossakowski Medical Research Centre of the Polish Academy of Sciences [31931]
14. Institute of Fundamental Technological Research of the Polish Academy of Sciences [15214]
15. Institute of Slavic Studies of the Polish Academy of Sciences [11970]
16. Institute of Organic Chemistry of the Polish Academy of Sciences [8362]

A series of marketing activities leads not just to raising awareness of the Repository together with the possibilities and benefits it gives, but also boosts the already strong brand of Institutes participating in the Repository project. It is worth

pointing out here that the very choice of partners, which we are associated with, plays a huge role. 15 institutes of the Polish Academy of Sciences and one branch institute collaborate on the Repository project. It is important to choose partners with strong position and positive image, when you are building the image of your own institution.

## Prospects

RCIN will be continued at least five years after the period of financing is over — according to the conditions of the agreement. What happens after that time depends on funding. The venture has been very successful, that's why a concept for a new project — RCIN+ — has emerged. RCIN+ is supposed to be a continuation of RCIN project. A consortium consisting mainly of the institutes of the Polish Academy of Sciences has invited other prominent scientific institutions, including the Institute of Aviation, to participate in the project.

## Summary

The article outlines the possibilities that digital reality gives in the process of building a strong brand. It was pointed out that it is not only about building the image of own company, but also about the comprehensive image of Polish science on the international arena. Supraregionality and multidisciplinary of the Digital Repository of Scientific Institutes has an impact both on the world of science and the economy. Moreover, it builds the awareness of digital resources, dissemination of research results and growth of the number of citations. Resources of the Repository can reach a broader group of recipients thanks to participation in the Digital Libraries Federation and in the digital library of the European Union — Europeana. Promoting Polish science and scientific units through dissemination of research results seems most obvious, however, it is not always possible. Digitalization of library resources is a task requiring major financial assets. Investments are necessary to create adequate infrastructure and a group of specialists. Among other obstacles are copyrights, which in a particular way limit access to knowledge.

Despite certain difficulties, involvement in the promotion of Polish science brings immeasurable benefits both to the society, archiving, securing and publishing scientific works and to the economy, opening to the industry and stimulating its growth.

## References

- <sup>1</sup> “New renaissance” — bringing Europe's heritage online — report of analytical group “Comité des Sages”.
- <sup>2</sup> Digitalization — transferring library resources into digital form from paper form by scanning or photographing.
- <sup>3</sup> Data according to Digital Libraries Federation, <http://fbc.pionier.net.pl/owoc/list-lib>, 30.10.2013
- <sup>4</sup> Polish abbreviation RCIN.
- <sup>5</sup> Metadata — data about data, eg. classic library catalogues. Za pomocą metadanych opisywane są dokumenty elektroniczne, w szczególności dokumenty dostępne poprzez sieci komputerowe, np. strony WWW dostępne w sieci Internet, jak również dokumenty tworzące nowoczesne biblioteki cyfrowe.
- <sup>6</sup> Dublin Core (Dublin Core Metadata Element Set, DC, DCES) — general standard of metadata. Approved as ISO standard ISO 15836–2003.
- <sup>7</sup> NUKAT — central catalogue of Polish scientific libraries formed by means of shared cataloguing.
- <sup>8</sup> Integrated library system.
- <sup>9</sup> PCSS; <http://www.man.poznan.pl/online/pl/>, 15.10.2013
- <sup>10</sup> “Procedury prawne i organizacyjne tworzenia bibliotek cyfrowych. Obszary ryzyka w procesach budowania bibliotek cyfrowych” Barbara Szczepańska [in:] *Biblioteki Cyfrowe a Prawo Autorskie*. L edycja seminarium z cyklu Akademia Prawa Komputerowego, Warszawa 21 lipca 2011: Centrum Promocji Informatyki.
- <sup>11</sup> <http://rcin.org.pl/dlibra/text?id=copywright-low>, 14.11.2013
- <sup>12</sup> <http://rcin.org.pl/dlibra/pubstats?statType=16>, 14.11.2013
- <sup>13</sup> <http://rcin.org.pl/dlibra/docmetadata?id=35960>, 10.10.13
- <sup>14</sup> <http://rcin.org.pl/dlibra/pubstats?statType=29>, 15.11.2013

## Bibliography

1. Czarnocka-Cieciura D. [et al.], *Repozytorium Cyfrowe Instytutów Naukowych — RCIN*, Warszawa, [2011].
2. Stanisławska-Kloc S., *Cyfrowe biblioteki i archiwa — w świetle polskiego i europejskiego prawa autorskiego*, [in:] *Problemy cyfryzacji zasobów muzeów, bibliotek i archiwów*: Forum Instytucji Kultury, 19 października 2011, Warszawa, Centrum Promocji Informatyki.
3. Stanisławska-Kloc S., *Prawo autorskie a biblioteki cyfrowe. Legalność bibliotek cyfrowych — odpowiedzialność cywilno-prawna z tytułu naruszeń praw autorskich*, [in:] *Biblioteki cyfrowe a prawo autorskie*: L edycja seminarium z cyklu Akademia Prawa Komputerowego, 21 lipca 2011, Warszawa, Centrum Promocji Informatyki.
4. Szczepańska B., *Procedury prawne i organizacyjne tworzenia bibliotek cyfrowych. Obszary ryzyka w procesach budowania bibliotek cyfrowych*, [in:] *Biblioteki cyfrowe a prawo autorskie*: L edycja seminarium z cyklu Akademia Prawa Komputerowego, 21 lipca 2011, Warszawa, Centrum Promocji Informatyki.

## Websites

- Biblioteka Narodowa — <http://www.bn.org.pl/>
- Centrum NUKAT — <http://centrum.nukat.edu.pl/>
- Europejska Biblioteka Cyfrowa Europeana — <http://www.europeana.eu/>
- Federacja Bibliotek Cyfrowych <http://fbc.pionier.net.pl/owoc>
- dLibra <http://dlibra.psnc.pl/>
- Poznańskie Centrum Superkomputerowo-Sieciowe <http://www.man.poznan.pl/online/pl/>
- Repozytorium Cyfrowe Instytutów Naukowych <http://rcin.org.pl/dlibra>
- Wikipedia, wolna encyklopedia — <http://pl.wikipedia.org/>

**Kamila Kaczyńska, M.Sc.** — Librarian. Graduate of the University of Warsaw — Faculty of Physics, and Chodkowska University in Warsaw — Faculty of Management and Marketing. She completed a post-graduate course in library science at the University of Warsaw. She gained professional experience in the National Library. Currently, as the head of Scientific-Technical Library of the Institute of Aviation she takes care of modernization and computerization of library work.



Institute of Aviation  
Scientific Publishers  
al. Krakowska 110/114  
02-256 Warsaw, Poland  
phone: (+48 22) 846 00 11 ext. 551  
e-mail: [minib@ilot.edu.pl](mailto:minib@ilot.edu.pl)

[www.minib.pl](http://www.minib.pl)  
[www.twitter.com/EuropeanMINIB](https://www.twitter.com/EuropeanMINIB)