

PHOTOGRAPHS IN THE DIGITAL COLLECTION
ISSUES WITH THE DESCRIPTIVE AND SUBJECT CATALOGUING
ON THE BASIS OF THE DIGITAL LIBRARY OF THE WARSAW
UNIVERSITY OF TECHNOLOGY

Maria Miller

Main Library of Warsaw University of Technology

e-mail: miller@bg.pw.edu.pl

Małgorzata Wornbard

Main Library of Warsaw University of Technology

e-mail: wornbard@bg.pw.edu.pl

Maria Miller is a graduate of the Institute of Information and Book Studies of the University of Warsaw. She has worked in a number of IT centres and for the past number of years has worked in the IT centre in the Main Library of Warsaw University of Technology, principally managing the Digital Library of the Warsaw University of Technology. Miller's main interests focus on digital libraries, digitizations, repositories, new technologies allowing for the availability of information and trends related to the co-participation of users in the processes of creating information (Web 2.0, Web 3.0). She has published among other things *Od Library 2.0 do library 3-D – drugie życie bibliotek* [From Library 2.0 to Library 3-D – the second life of a library], *Biuletyn EBIB* 2008, no. 1 (92).

Małgorzata Wornbard is a graduate of the Department of Russian Studies, Slavonic Studies and Applied Linguistics of the University of Warsaw. In 2006 she completed Post-graduate Studies in Information Science at the Institute of Information and Book Studies at the University of Warsaw. She is the custodian in the Cataloguing Section of the Main Library of Warsaw University of Technology. At present, she is serving as the system librarian for interaction with the National Universal Central Catalogue NUKAT. Wornbard is supervisor of the computer catalogue of books in The Main Library of Warsaw University of Technology. Her interests focus on issues associated with formal cataloguing in traditional and digital libraries. Wornbard has discussed these issues at the conference *Polskie biblioteki cyfrowe 2008* [Polish digital libraries 2008], which treated various aspects of the description of iconographic documents in digital libraries.

KEYWORDS: Iconographic documents. Electronic documents. Archival material. Photographs. Documentary photographs. Special collections. Metadata. Cataloguing. Computer catalogues. Subject cataloguing. Subject analysis of documents. Bibliographical description. Formats. Data structures. Digitization. Digital libraries. Internet. Semantic network. Keywords.

ABSTRACT: In terms of digitization and the growing significance of photography from the perspective of cultural

heritage, many efforts have been made in recent years to arrive at uniform standards for the description of documents of this type, which are collected by libraries, archives and museums and made available in digital libraries. Based on their experience working with the Digital Library of the Warsaw University of Technology, the authors of this article discuss issues associated with descriptive and subject cataloguing, resulting from the specificity of iconographic documents (where there is no text data for creating a description as well as the lack of Polish norms and rules). They lay emphasis on the new challenges which the creators of digital libraries are having to contend with when it comes to new technologies and ideas, and transformations in current working practices.

PHOTOGRAPHS AND CULTURAL HERITAGE

Iconographic materials, in particular photographs, have not always been regarded as valuable in terms of national legacy. However, they constitute an important means of documenting historical events. Photographs, describing the image of some place, the daily lives of people, a moment in time, allow us to not only understand the past better, but mark changes in our immediate surroundings and the wider world. They facilitate individual interpretation and contextualize history from different perspectives. The content of photographic collections is varied, illustrating such subjects as: events, architecture, means of transport, fashion, social and family associations, competition, trade, cultural objects, and the widely understood notion of lifestyles of a given age.

Until recently, photographic collections were taken for granted. Photographs were catalogued collectively, and were either left without a description or provided with a perfunctory description such as “container box with photographs showing...” (Klijn, 2005). In archives they were stored together with paper documents, and even in the 1980s these were not given the status of archival records. A change in the evaluation of photographs took place when photography became a popular field of art. Today, photographs are seen as valuable documents, worthy of examination, proper storage, and editorial commentary. What is important here is the idea of cultural legacy, popularized by the European Commission and the European Council, as well as the development of computer technologies. Conclusions published in the document “i2010: Digital Libraries” so stated: “Information technologies are creating the possibility of other «discoveries» of European cultural and scholarly heritage and making it available for diverse current and future applications. This task requires a joint effort on the part member states, whose efforts in digitization and digital conservation lack cohesion. Effective responses to the new challenges may be enhanced by the process of digitization, facilitating access to information, and ensuring the long-term preservation of digital stores” (Communique, 2005).

Notions of national or cultural legacy are constantly evolving. Photographic documents, sometimes from private, niche collections, can prove to be a unique testament to cultural heritage, on a local or regional level, and important resource. As a result, institutions of national memory (archives,

libraries, museums) have changed their approach towards the photograph – particularly visual collections, which have been made available to the wider public by way of creating digitized collections and making them accessible online on the Internet.

MODELS FOR THE DESCRIPTION OF PHOTOGRAPHS

In 1994 the European Commission on Preservation and Access (ECPA) was established, which aimed to facilitate wider access to European archive or library collections. In 2004-2008 the ECPA initiated the TAPE project, relating to audio-visual archives in Europe, financed by the Culture 2000 program. In previous years, much of the Committee's work had been taken up with SEPIA (Safeguarding European Photographic Image for Access), concerning photographic collections. This project had allowed for the recognition of many "niche" photographic collections, stored in institutions of cultural memory. The task of the SEPIA program was to complete activities related to the conservation and digitization of collections of historical photographs, as well as defining the role of new technologies for the management of these collections. It is possible to find more information about this program on the site <http://www.knaw.nl/ecpa/sepia>.

Within the framework of SEPIA, the Working Groups for Model Description, drew up a model for the description of a SEPIADES photograph. The group was made up of 5 partners (The Municipal Museum in Stockholm, The Norwegian Offices for Archives, Libraries and Museums, the National Library of Spain, The Finnish Museum of Photography, and The European Committee on Preservation and Access). SEPIADES is a multi-level model similar to IAAD(G), (ISAD, 1999), which derives from the collection, through groups to the single unit. They decided on such a resolution, assuming that in the case of photography, a separate description of every entity is almost impossible, and not always necessary (De Lusenet, 2005).

Apart from 21 basic elements, SEPIADES contains over 400 suggested data elements for describing photographs, with references to literary sources online and elsewhere. In the final report, the Working Group working on the SEPIADES model, exchanging the fundamental properties of the model, recommended the use of the Dublin Core standard. This standard was perceived as being particularly useful for the exchange of metadata amongst various institutions. The SEPIADES model is based on the format XML, in which all records are written, which facilitates browsing in any editor, and the export of records to Dublin Core XML files. The model uses the protocol OAI-PMH, in accordance with the Open Access initiative which facilitates access to collections held in other institutions. It provides also a tool for full-text search.

The main task of the SEPIADES model, similarly to other standards for cataloguing photographs, is to arrive at a professional, complete description of the photograph. A good description of a photograph enables its search potential, thus making it visible to scientists, librarians, archivists, as well as casual users. The description contains not only information about the author and the contents of the photograph, but also information about the photographed object, about the mode of acquiring the photograph, its physical properties, its physical characteristics and condition, ownership rights, such as limited access, and so on. Only when provided with a full description, to include additional information, can the photograph become an object that can be searched for and interpreted.

The quality of the description increases if certain norms are observed. In practice various other standards are used for cataloguing photographs, such as ISAD (G), ISBD, AACR (Anglo-American, 2006), and sometimes principles are drawn up for a given purpose and not based on any standard. In a number of European countries solutions which aim at the synchronization of cataloguing photographs have been developed. In Norway, in museums and cultural-historical institutions, a system was implemented by the name of “Feltkatalogen”. Dutch institutions apply the FOTIOS system, developed by the Dutch Photographic Association, whereas Swedish institutions of memory use the “Dataelementkatalogen” system.

In Poland the richest photograph collection is owned by the Archive of Mechanical Documentation in Warsaw, which from February 2008 has been renamed as the National Digital Archive (NAC). This institution is carrying out the ZoSIA project (www.nac.gov.pl), (integrated archival information system) as based on open standards – ISAD(G), EAD (Encoded, b.d.). In the NAC a Polish version of the International Standard of the Notation of Information Archival Stores – EAD_PL. NAC has placed approximately 80 thousand photographs online. These are principally photographs from the period of the Second Polish Republic and photographs of the Polish Section of Radio Free Europe. Data given in the description is: call number of the photograph (reflecting the name of the archival team), the place and description of the event commemorated in a photo, date when taken, surnames of visible persons in the photo (in the case of group photographs – the surnames of the foregrounded persons) and the surname of the photographer, place names technical aspects concerning the original (negative, positive, sizes, colour etc.). A broadening of the description of the photograph is planned, pertaining to additional data concerning persons featured (profession, function, position, academic title, military rank, pseudonym etc.), keywords, status of the image (original, reproduction),

information about the copyright, remarks (e.g. concerning the technical condition of the original) (Pašek, 2006).

Library circles principally use the MARC standard, whereas archives use the ISAD standard, which is applied for the description of archival collections (fitted to the file containing records). For the description of archival photographs, the EAD standard is used, supported by the Library of the Congress in the USA. There is also another alternative for the standardized description of the photograph, which isn't associated with a need for the implementation of some system or package. This solution arose with the development of digital photography at the beginning of the 21st century, coupled with mass use by news agencies. The already existing standards were at odds with the need to rapidly describe photographs. MARC was too complicated, whereas ISAD necessitated the indicating of a given collection to which the photograph belonged.

From 2001 Adobe Acrobat and Adobe Photoshop (also and in other graphics programs) have offered the possibility of saving metadata directly in files in the XMP standard. Description of the contents of the file in accordance with XMP may be encoded in the PDF file, as well as in graphic formats such as JPEG, GIF or TIFF. The description can be made in any standard, as well as in Dublin Core or MARC. However this kind of solution is not appropriate neither for libraries nor archives, where linking the description directly with the graphic file allows the user to obtain information about the contents of the file. Limits on the description only to the "internal" description following the loss of such a file means that the information as to what was lost is lost also.

The establishment of a unified standard for the description of photographs for all institutions of memory is not a simple matter and may not be an appropriate solution. "The Team for Digitization", established in 2006 by the Minister for Culture and National Heritage initially looked to a "branch" solution: MARC for libraries and EAD (encrypted ISAD) for archives, and Dublin Core as an acceptable standard for organising information pertaining to the compiling of collections created as a result of the digitization process.

Making recourse to the same norm does not mean that all descriptions created by various institutions will be the same, because the same norm will always be subject to interpretation. A key issue is the interpretation of these rules. Attitudes to cataloguing photographs depends on the role that an institution in possession of collections actually fills. It is also dependent on the demands of users, or more pointedly those whom these said institutions perceive their recipients to be.

Photographs are documents which are particularly difficult to describe on account of the problem with obtaining contextual information. In many cases an expert must be able assess the technical aspects of the photograph, whereas the cataloguer must also possess a scholarly knowledge needed for a correct interpretation. All photographs and iconographic materials in general must be analyzed on an individual basis and then catalogued and described in a way that provides recipients with all possible retrievable information. The description of the same object can be different depending on what aspect of the photograph the cataloguer chose to focus on. Problems arise (irrespective of the accepted standard, whether it will be Core Dublin or EAD), when it comes to an interpretation of the basic attributes such as “date” and “photographer”.

SEPIA

Safeguarding European Photographic Images for Access

Date?

- Date visual image: 1658
- Date physical image: date of exposure: 1980
- Date physical image: publishing date: 1981

Creator?

- Vermeer, J., person, painter
- Jones, M., person, photographer
- Smith, R., person, scan operator



About SEPIADES--> A long and winding road...

Fig. 1. Variants of description of photograph of the painting “The Milkmaid”: J. Vermeer (Klijn, 2003)

Problems with the interpretation of fields was well illustrated by Edwin Klijn (ECPA) at a seminar dedicated to the SEPIA project (Klijn, 2003). The object of description is a photograph of the famous painting by Johannes Vermeer “The Milkmaid” (fig. 1).

The questions which arise here are: What is the date of the recreated object: 1658 – what year was the photographic image made, 1980 – developed (copies made), 1981 – when was the photograph published? Who is the creator of the described object: the painter of the image, the photographer who took the photograph, or perhaps the operator of the scanner who digitalized the photograph of the object?

OF TECHNOLOGY AND CATALOGUING THEM FOR THE DIGITAL LIBRARY OF WARSAW UNIVERSITY OF TECHNOLOGY

Similar problems to those indicated in the earlier part of this article are to be found also in the Digital Library of the Warsaw University of Technology (DL WUT), operating within the dLibra system (created and developed by the Poznan Supercomputing and Networking Center).

The Main Library of the Warsaw University of Technology (WUT) decided to make its photographic collections available to a wider group of users. And so, in 2008 it began digitizing its photographic collections. The collection is composed of, among others, photographic materials, including 40 thousand positives (from the end the 19th the century to 1976) – mainly black and white, as well as 18 thousand diapositives from the end of the 1950s to 1976 (many of which are in colour). No inventories or catalogues remain extant,- In the case of diapositives, card indices have survived – alphabetical and subject.

The digitization of photographic collections in the Main Library of the WUT has two purposes: the archival protection of photographic collections and the safe availability of documents. It was decided that photographic collections located in the Main Library, hitherto little used, would be presented to potential users, and that those photographs with adjudged legal status, would be published in the DL WUT (<http://bcpw.bg.pw.edu.pl>). The remaining photographs would be scanned and made available locally, in the workroom of the National Library Heritage Collection The selection of parameters for scanning (resolution, file format) depends on the allocation of the digital document.

Photographs for the archival store are digitalized with resolution 600 dpi in the uncompressed TIFF file format. Photographs to be made available in the digital library are being converted to the JPEG format for resolutions 300 dpi, which gives files a much lower size but with a sufficient quality for the image. Photographs for the archive are not subject to any doctoring, however the photographs presented in the digital library are sometimes retouched slightly when it comes to sharpening, or evening out tones and levels. Scanning of photographs was performed in greyscale with 8 bit/px – 256 levels of grey.

For photographs which were rather blurry, in a number of cases we applied a sharpening mask, not-exceeding 100% (a ray of sharpening not exceeding 1.3 pixels). In several cases photographs were brightened. Miniatures are added to the photograph descriptions, constituting approximately 10% of the size of the picture. On photographs a watermark is added – an imprint basically – in the right bottom, with contents “© Copyright year, the DL WUT”, small enough to not spoil the integrity of the photograph itself. Such measures will not fully prevent these photographs from being put

to further unauthorized use, but may they discourage their misuse on the Internet.

Collection “Iconography”, presenting photograph, graphics, postcards, maps from the collection of the Main Library of the WUT, was established in July 2008. The first photographic documents were photographs which were museum exhibit items presented at the exhibition “Not only a Technical Book”, organised by both the Main Library and the Museum of the WUT in 2008. Attempts at describing the objects with metadata, due to an absence of norms and standards, were generally intuitive and based on experience taken from the cataloguing of other types of documents.

All exhibited items constituting the property of the Museum of the WUT are catalogued in the traditional way. Describing such a document (a photograph of the museum object) in the digital library _was_ simpler than describing a haphazard photograph. If the object was to be exhibited in other exhibitions, it was given a name which was used and confirmed in many primary catalogues. Among the objects were medals, written documents such as student record books, congratulatory addresses, or payroll lists of the Polytechnic Institute. One photograph for example is that of student uniforms between the years 1896-1915. Exhibit items, where the original was a print, were at first catalogued in the NUKAT catalogue in the MARC 21 format. For creating the bibliographical description, guidelines typical for different types of documents were used, such as for prints and electronic documents (Lenartowicz, ed., 2003; Sanetra, 2003). Presented below are two examples of catalogued museum exhibits in the DL WUT:

Example 1

TITLE – Indeks studenta Politechniki Warszawskiej Jana Strzeszewskiego przyjętego na Wydział Architektury [The student record book of Jan Strzeszewski, accepted into the Architecture Department of Warsaw University of Technology].

CREATOR – Politechnika Warszawska. Wydział Architektury [Warsaw University of Technology. Architecture Department].

SUBJECT – eksponat muzealny; historia Politechniki Warszawskiej; indeksy studentów; Politechnika Warszawska [museum exhibit; history of the Warsaw University of Technology; student record books; the Warsaw University of Technology.

DESCRIPTION – origin of title: Title non-authentic as well as information: it includes 2 pages of index.

PUBLISHER – Warsaw: the Warsaw University of Technology.

DATE – 1917.

RELATION – Object from the Museum of WUT (information of subcollection, related to exhibits from the exhibition).

SOURCE – http://gate.bg.pw.edu.pl/F?func=direct&l_base=wtu01&doc_number=000165237 (link to original source, from which the described document derives in the digital library). In other words to the bibliographic description in the local catalogue ALEPH BGPW; the description is a faithful copy of records from the NUKAT catalogue; in this field information pertaining the holding place of the original is located: The original is located in the Warsaw University of Technology Museum.

RIGHTS – <http://bcpw.bg.pw.edu.pl/dlibra/docmetadata?id=574&from=&dirids=1> (fields contain information about legal ownership of original. Sometimes aside from the Main Library of the Warsaw University of Technology legal title may also be claimed by a descendant or institution. Legal title to the source of the object is owned by the Museum of WUT).

Example 2

TITLE – Fotografia daru dla cara Mikołaja II za wyrażenie zgody na budowę Instytutu Politechnicznego [Photograph of a gift to Tsar Nicholas II for having granted permission for the building of the Warsaw University of Technology].

SUBJECT – eksponat muzealny; historia Politechniki Warszawskiej; Politechnika Warszawska; historia szkolnictwa polskiego [museum exhibit; the history of the Warsaw University of Technology; the Warsaw University of Technology; the history of Polish schooling].

DATE – 19th century.

DESCRIPTION – Includes information about the contents of the photograph as well as historical data: Photograph of cake stand made in silver by the artistic workshop “Brothers Łopieńscy”. The photograph dates from the 19th century, a gift from Ms Józefina Łopieńska (1996) to the Museum of WUT.

TYPE – photograph

SOURCE – In the event that an object does not have as yet a description in either the local catalogue or the NUKAT catalogue, information is entered on the object’s place of holding: Original located in the Museum of WUT RELATION – Objects from the Museum of WUT.

RIGHTS – the Museum of WUT

<http://bcpw.bg.pw.edu.pl/dlibra/docmetadata?id=562&from=&dirids=1>.

Each museum exhibit item received subject keywords (museum exhibit item, history of the Warsaw University of Technology), objects supplemented with words best describing the object (e.g.: history of Polish schooling, student record books). The majority of exhibition objects do not have descriptions in the NUKAT catalogue, which are to be created following the publication

of instructions for MARC 21 for photographic documents catalogued by Krystyna Sanetra. They will then be added to the Central Catalogue.

Currently, much of the collection is made up of: the digitalized and catalogued photographs of Warsaw artist-photographer Henryk Poddębski, the outstanding documentary photographer from the first half of the 20th century, and are included in the series Polish Photography Before 1939. Other works making up the collection are: the photographs of Czesław Olszewski, the Warsaw photographer of modernist architecture; as well as documentary photographer of the post-war generation such as Edward Falkowski, Leonarda Jarzembski or and Leonard Sempoliński. The Society for the Care of Monuments of the Past brought together documentary photographers, who participated in inventory projects aimed at immortalising monuments and other objects important from the perspective of the city's history. The main theme was the city and its architecture, with particular reference to places and buildings which were subject to transformation, decay, or reconstruction. An awareness of irreversible change and the importance of recording and describing phenomena constituted the essence of these documentary photographs. Post-war photographers linked to Warsaw were guided by the need to capture the radical transformations taking place within the city. These photographs, which recorded the images of objects and places undergoing change or about to disappear, gained special significance in the context of the preservation of both memory and the historical identity. When presenting them to the public, one should place great emphasis on the collection and record all pertinent details.

When working on Poddębski's photographs, it was necessary to overcome problems typical for documents of this type, which crop up not only in digital libraries, but also in traditional libraries. In spite of the fact that Polish norms have yet to be published, a great deal of effort is being put into the description process, as based on the guidelines issued by both the National Library of Poland and those provided by the NUKAT Centre (Pietrzak, comp., 2008; Sanetra, n.d.).

Problems with the formal cataloguing of iconographic documents begin with their definition. According to BN guidelines, an iconographic document is a "two-dimensional, image, motionless, not intended for projection by any portable technique". This definition was based on the French norms and principles as published by The Library of Congress (FD from 44-077, 1977; Betz, 1982). In this light of this fact, it is possible to doubt the contents of the new collection of the DL WUT, since the digital version of an iconographic document is an image intended for projection. Moreover, amongst the many types of iconographic documents, we may list: graphics, sketches, paintings, photographs (positives and negatives), graphic matrices, posters, invitations, postcards, and medals.

From among the iconographic documents, numbering 18,000 items, we may look in vain for diapositives.

Given our experience with creating iconographic collections for the DL WUT, it seems that an iconographic document should be defined as a graphic document, whose contents are presented in the form of the image. The present definition divides a iconographic document between projection and non-projection. Such a division seems sees the digitalization of library collections as artificial. An elaborating of the definition has arisen from the MARC21 format, which points to different ways of treating projection documents. However, it doesn't changing the fact that a iconographic document once scanned, becomes an "image destined for projection", but by way of its digitization has not ceased to be an iconographic document.

Bundling iconographic document with electronic documents is the result of some misunderstanding. On account of its idiosyncrasies, an iconographic document i.e. its presentation in the form of an image, irrespective of the form it takes, remains an iconographic document.

The descriptive and subject cataloguing of such a document is very difficult, and is preceded by the laborious accumulation of essential information. A photograph is generally without text. The librarian must process visual information in order to generate bibliographical description. Data gleaned from the document, constitutes the basis for creating the record through the cataloguing and supplementing of information from outside the basic source.

DUBLIN CORE AND MARC

Descriptions of the photograph in the DL WUT are made in the Dublin Core format, and applicable to the description of electronic documents. The descriptions intended for the Central NUKAT Catalogue are created in the MARC21 format.

Dublin Core, the format most often used for the description of metadata used in digital libraries, consists of 15 elements called attributes. A number of them are universally known, since they appear in other formats of bibliographical description, e.g. Author, Title, Subject Description, Publisher. Part of the attributes contains technical data, such as the size of the file, and the identifier or the format. However, the remaining elements are intended for information typical for a document presented in digital libraries, e.g. copyright. Below are a detailed set of fields of the format Dublin Core (Dublin Core, 2009) and description in the DL WUT.

TITLE – includes the name of source, which is formally known.

CREATOR – Unit for creating contents of source. It includes proper and corporative names. In DL WUT the keyword is accepted in the form established in the file of sample entries of the catalogue NUKAT. In this field we giving the name of the creator of the photograph.

SUBJECT – generally includes keywords, subject headings (acceptable are also symbols of the Universal Decimal Classification describing contents of the source).

DESCRIPTION – This element can contain a free text determining the content of the document, such as the abstract, the table of contents, information about illustrations and the edition. In the case of the DL WUT in this field you will generally find historical data, dates for the construction of the object, and information about the creators of the object – in other words, the architects and others.

PUBLISHER – the entity responsible for making the source accessible.

CONTRIBUTOR – entity responsible for co-creating/contributing to the source. Similarly as in the attribute Author (Creator) entries are accepted in the form established in the file of the source sample subject words of the NUKAT catalogue. As a value of this attribute, one enters the names of the photographic studios, printers, the authors of the accompanying documents, editors, creators of the collection etc.

DATE – Date of publishing the source. This element of the format contains the date of the making of the original (in this case of the photograph). It is most often the date from a negative. The date of the copy which we have in collection is given in the field Description.

TYPE – defining the category of source content

FORMAT – contains the name of the format, in which the digital form of the source was written (in the case of dLibry the program alone recognizes the format and fills in the field automatically).

IDENTIFIER – Identifier of the digital store. Most often contains the URL address. In case of dLibry digital libraries it is a unique URL address for every document.

SOURCE – sending to the source from which the digital figure of the document comes from. In the case of the DL WUT this is sending of the source to bibliographical description in the local ALEPH catalogue, which is a faithful copy of the NUKAT record.

LANGUAGE – Language of contents of the source. In principle in DL WUT, this field is not applied because of the fact that on our photographs there is no text.

RELATION – sending to sources which remain in relation with that described in the DL WUT. This element contains the name of the series or sub-collection, where we place our photographs. E.g. Polish photograph before 1939.

COVERAGE – range of the content of source.

RIGHTS – information about legal ownership transfer of source. This attribute is very important in the case of digital libraries, because it determines the legal status of the document. In the case of the DL WUT, it happens that apart from the Main Library or the Museum of the Warsaw Technical University, descendants or heirs may also retain legal title.

The most difficult element of the bibliographical description of an iconographic document is the title zone. Data for this field should appear in the form which appears in the document. It can be a text placed by the author on the obverse, on the reverse, on the original cover, or on the wrapping of the document. A text written by the author is acceptable, from the negative. If we do not have such data, we look for inscriptions made by subsequent owners. They often contain the descriptions of persons and objects or places described in the photograph. We also look for titles in accompanying documents or other copies, if any do exist. Later we look to specialist literature in catalogues, bibliography, biographical dictionaries, encyclopedia, or the Internet.

In the case of H. Poddębski's photograph collection, the title is described as based on the image recorded in the photograph. If the need arises, every title is supplemented with a geographical name. One does not add any specification of the type of document, nor does one apply square brackets. The situation is different in the NUKAT catalogue. In the case of Henryk Poddębski's photographs where the text is missing, all information included in the fields of the bibliographical description is included in square brackets and supplemented by remarks in fields 5 XX. In the digital library the attribute Source refers the user to the description in the MARC21 format in the NUKAT catalogue. For this reason we can depart from widely accepted principles, wherein we assume that the description of the document in the digital library should be user-friendly. The substantial amount of square brackets does not support this notion as they are incomprehensible to the majority of users. In a number of descriptions in the local catalogue (in the future the NUKAT catalogue) in the title zone the four brackets appear, such as:

245 \a[The Cathedral of the Birth of Mary, The Blessed Virgin in Sandomierz] \h [Photographic Document] : \b [interios] / \c [Henryk Poddębski].

http://gate.bg.pw.edu.pl/F?func=direct&l_base=wtu01&doc_number=000169977.

The average user of digital libraries is generally not interested in those regulations for cataloguing which have been adopted by libraries. The description most often assesses the usefulness of information included in the bibliographical record. As a result, a greater emphasis is placed on the reliability of the information (identification of all elements of the description, such as: author, title, year of issue, details about the object) found in the bibliographical description

as opposed to actual formal requirements, such as the numbers and kinds of brackets in individual fields.

SUBJECT CATALOGUING OF ICONOGRAPHIC DOCUMENTS

The next group of issues involve the material cataloguing of a photograph. It would be highly unusual for a user to look for a photograph according to the title as formulated by the cataloguer. Therefore, the attribute Subject and Keyword as in the case of a iconographic document is the most important element of the description. Since the photographic document does not have a search text, we are looking to the DL WUT to endow this attribute with as many phrases as possible. In the case of books, every word from the text can enter the index. However, the iconographic document requires another form of mapping its contents. Subject cataloguing in the Main Library of the WUT is made up of: classification according to the UDC, keywords and the call number of the document, indicating its place on the shelf. In a digital library, for determining the subject of a document, keywords are exclusively used. Entries are taken from the keyword authority file created in DL WUT, but they also allow for a certain amount of flexibility.

In the DL WUT an authority file of keywords is being built in accordance with the methodology for the building of thesauri; but only the method for forming lexical units is used. Keywords have the form of nouns and are most often the terms used in specialist literature. If it is possible, a natural train of words is kept: on top spot the noun is in the nominative case, followed by adjectives, nouns, participles, numerals or prepositional phrases determining the noun. There are exceptions to this rule, where the said train of words is not provided. Some keywords are not included in the file of sample entries. And so in in this way they do not deviate from the rest in terms of its linguistic form. However these are always the words which best describe the contents of the photograph.

In the DL WUT for the descriptions of photographic documents a dozen or so keywords are applied. A different approach is taken with the local Main Library catalogue, which includes copies of records from the NUKAT catalogue, where every keyword entered into the model file has its equivalent in the form of a UDC symbol. Such a connection necessitates certain limitations which are missing in the Digital Library, since documents are only catalogued thematically with the help of keywords. No restrictions are applied in quantitative terms. It is recognized that the more keywords assigned to the document, then the easier it is to identify it. Moreover, the dLibra software has a very helpful tool which offers the possibility of creating a dictionary of synonyms.

In the case of the DL WUT, this mechanism completes the records for so-called rejected forms of terms, which in traditional catalogues fulfill the role of so-called reference marks “see”. In model files created in the MARC21 format, these are variants of keywords which are to be found in the fields 4XX. Popularly, they are determined as tropes, e.g.:

MAIN HEADING: Politechnika Warszawska [Warsaw University of Technology].

DICTIONARY OF SYNONYMS: PW [WUT], the Warsaw University of Technology, Państwowa Wyższa Szkoła Techniczna [National College of Technology], Warszawski Instytut Politechniczny im. Mikołaja II [The Nicholas II Warsaw Polytechnical Institute].

PHOTOGRAPHIC DESCRIPTIONS IN EUROPEANA

When discussing issues associated with standards for photographic descriptions for digital libraries, we must mention the principles applied in Europeana, a service which ensures access to multilingual collections of European digital libraries, spearheaded by the European Commission within the framework of the program “i2010: Digital Libraries” and overseen in the EDLnet.

The Europeana prototype has been operating since December 2008. The combination of descriptive metadata, traditionally treated differently in libraries, museums and archives, proved an issue which the creators of Europeana had to resolve. It was decided to treat descriptive metadata as search metadata. The format used was Dublin Core with certain modifications. New elements were added, such as: the User tag (public tags created by registered users), Unstored (a type of basket for useful information which does not fit into other fields), Object (internal use only) and labels of the IsShownBy relation (URI link to the digital object) and IsShownAt (contextual presentation of the object, in the parent digital library).

Also, a specification was drawn up to copy information from bibliographical descriptions of digitized objects to metadata typical for Europeana (Specification determined in the document for the Metadata Elements for the European Prototype) (*Specification...*, 2008). Proposed is the keeping of all XML attributes, which included the initial description without changes. The use of the xml:lang attribute is recommended wherever possible: determining language, and in which metadata they are saved. However the element dc:language determines the language in which an object (document) is written. Every description must contain 5 compulsory fields: the date of the authentic document, link to the digital object, name of the institution storing the original, the title and the object type (text, image, motion picture, sound). Giving the language is recommended, and

also in the case of picture documents if they contain some text (placard, poster).

In the case of a photograph bereft of all notes, stamps, etc., this attribute is ignored. The fields which the user can search for are: title, author, date, subject (keywords). One should remember that Europeana does not store digital objects, but shows them in a brief general description. In turn, it provides a link to the digital object, or the miniature of the object and refers to the description in the parent library of the object.

An exemplary photograph description (Tram in the Warsaw ghetto) in Europeana contains 7 attributes (fig. 2). The number of keywords is striking, albeit this is not the norm for objects in Europeana.

In the description of this photograph used by La Réunion des musées nationaux (Rmn) a Polish reader may be surprised by the lack of data in the data field (in Europeana) and the generalized date – the 20th century – in the original description (fig. 3). The empty field “Date” in Europeana results perhaps from the conditions of the conversion, permitting only the determined format of the input. It is worth recalling that in the document Specification for the Metadata Elements for the Europeana Prototype, the contents of the field “Date” were given as the date of the creation of the original object (analogue or digital born), rather than the date of the digitization of the object.

Original description (source in the parent entity), aside from data, like the call number, includes also information about legal title: the keywords are: antisémitisme, étoile de David, ghetto, scène de rue, Seconde Guerre mondiale, tramway, transport en commun.



Tramway dans le Ghetto de Varsovie

Twórca: Heydecker Joe ,

Opis: photographie (technique),

Institution: Culture.fr/collections

Język: fr

Źródło: Allemagne, Berlin, BPK

Prawa: Diffusion uniquement sur le territoire français. (C) BPK, Berlin, Dist RMN - Joe Heydecker

Mniej

Temat: TRAMWAY DANS LE GHETTO DE VARSOVIE PHOTOGRAPHIES ALLEMAGNE BERLIN BPK CHEMIN DE FER 20EME SIECLE XXEME SIECLE XXI SIECLE DAVID ETOILE ETOILE JAUNE DEUXIEME GUERRE MONDIALE 2EME GUERRE MONDIALE GUERRE 1939 1945 COURANT DE PENSEE PHOTO VIE INTELLECTUELLE ET SCIENTIFIQUE THEMATIQUE TRANSPORT FERROVIAIRE PHOTOGRAPHIE TECHNIQUE GUERRE EPOQUE CONTEMPORAINE ETOILE DE DAVID SIECLE TRANSPORT EN COMMUN SYMBOLE RELIGIEUX TRANSPORT FUNICULAIRE GENRE ICONOGRAPHIQUE DESCRIPTION ICONOGRAPHIQUE TRANSPORT PAR CABLE TRAMWAY SCENE DESCRIPTION ICONOGRAPHIQUE EVENEMENT POLITIQUE RELIGION HEYDECKER JOE POLITIQUE 20E SIECLE THEORIE COURANT DE PENSEE TECHNIQUE TECHNIQUE EVENEMENT ANTISEMITISME SCENE DE RUE PERIODE CONTEMPORAINE SECONDE GUERRE MONDIALE TRAIN GHETTO EVENEMENT MILITAIRE RACISME

Date: 0

Fig. 2 Description of photograph: “Tramway dans le Ghetto de Varsovie” in Europeana

<http://www.europeana.eu/portal/full-doc.html?query=varsovie&start=42&startPage=37&uri=>

<http://www.europeana.eu/resolve/record/03903/29837FD87BE8FE9BFFB98301EC7B70385F0E34DB&view=table&pageId=bd#>

Cote cliché :	07-523493
N° d'inventaire :	Fonds : Photographies
Titre :	Tramway dans le Ghetto de Varsovie
Auteur :	Heydecker Joe
Droits d'auteur :	(C) Bildarchiv Preußischer Kulturbesitz
Crédit photographique :	Diffusion uniquement sur le territoire français. (C) BPK, Berlin, Dist RMN
/ Joe Heydecker	
Période :	20e siècle
Technique/Matière :	photographie (technique)
Localisation :	Allemagne, Berlin, BPK
Mots-clés :	antisémitisme . étoile de David . ghetto . scène de rue . Seconde Guerre mondiale . tramway . transport en commun

Fig. 3. Description of photograph “Tramway dans le Ghetto de Varsovie” in La Réunion des musées nationaux (Rmn)

<http://www.photo.rmn.fr/cf/htm/CPicZ.aspx?E=2C6NU0CW9CGO>

Amongst the numerous keywords placed in field “Subject” and describing “Tramway dans le Ghetto de Varsovie” there is an absence of denotations, inseparably linked with the historical period and the very event commemorated in the photo. There is an absence of phrases such as: fascism, Naziism, German occupation of Poland, persecution of Jews. This example shows that in spite of numerous keywords, the description does not always deliver complete information about the object to the average user, who has found his way to the digital library by way of the open Internet.

The problem with integrating the format of the description (at least in part by encompassing compulsory fields) is typical not only for photographs, but for all digital objects made available

through Europeana. The goal of Europeana is to facilitate access to the multilingual and multicultural heritage of Europe and to preserve its cultural legacy for future generations. In order to achieve this aim a unified standard is essential for all cultural institutions – libraries, archives, museums. Every cultural institution with digital collections may become a Europeana partner provided that the technical requirements are put in place. Polish libraries, archives and museums have such a possibility. Here we return to an issue stated at the beginning of this article: the need for all institutions of memory to arrive at a unified standard for the description of photographs.

NEW CHALLENGES FOR LIBRARIANS

Regulations for cataloguing will not alleviate the array of problems that libraries face when formally and materially cataloguing iconographic documents, particular given that these same regulations devote little time to difficult and ambiguous cases which outnumber the difficulties met when cataloguing books. A lack of data may always lead to the visual misinterpretation of an iconographic document. It is necessary to accept that the cataloguer will require the help of a subject specialist, such as an art historian. The correct identification of an object in a photograph is sometimes very difficult and time-consuming. It often happens that collecting data for a bibliographic description can take up to three weeks. The library cataloguer or the editor of the digital library must be ready to accept help from external sources, such as network communities. Such solutions are offered by the Web 2.0, which have made an immense contribution in recent years.

There is no doubt that supplying a bibliographical description with a substantial amount of keywords is needed. The question is whether controlled vocabulary is needed, when it is a lexis written in a dictionary of synonyms. Perhaps a solution would be a dictionary based on enquiries entered by users in digital library search engines. For example, incorrect forms of enquiries would find their way to a synonym dictionary (such an idea appeared at one time on the Polish Library 2.0 Forum)

Taking into consideration the specificity of digital libraries which the user usually finds via a search engine such as Google, one should provide complete information about the digital object, and in the case of a photograph, where a distinct feature is the absence of text, there must be very specific information about the context of the image. This is related to the necessity of breaking from traditional and dogmatic modes of describing documents.

In digital libraries, the source document, the electronic document, as well as the object are described in one record. This is particularly visible in the case of a photograph, where it is important to provide as much information about the photograph as possible. Such a solution is often defined as a feature typical for libraries based on dLibra software. However this solution is not associated with the software, but with the specificity of digital libraries, which largely rely on scanned documents, as opposed to digital born documents. For the user they are simply copies of authentic documents.

In digital times with universal access to information via the Internet, the needs of the user must be the starting point for digital libraries. These needs are often elusive in the absence of either research or tools which would allow us to determine such needs. Libraries, as well as other institutions of memory, have to contend with the unification of standards followed by their adaptation to the newer demands associated with the ever-wider access to digital collections.

In looking to link the collections of cultural institutions and creating a global library, it will be necessary to unify both standards and the aggregation of data. This goal will not be achieved, however, by making recourse to once-determined norms and provisions. Standards are needed for the cohesion and accuracy of a catalogue. But the fundamental argument for applying a standard should be the possibility of searching scattered databases and enjoying universal access to data held in other institutions. In order to develop better solutions, joint actions on the part of institutions of memory are essential. An example may be an initiative taken by Dutch libraries, archives and museums, which in January 2009 entered into an agreement with associations governing copyright pertaining to the digitization and granting of access to collections of national heritage (FOBID, 2009).

Standards need to be constantly developed and adapted to the needs of users, even though they may not be able to formulate these said needs. Librarians, or the creators of digital collections, must accept that the methods for classifying and cataloguing documents, developed over the centuries, must evolve. A new challenge will be posed by the Semantic Web, using the analysis of user actions, considering classification and search semantic associations. Solutions of this type are already being applied, such as in the Digital Library of Gdańsk Technical University (the future Pomeranian Digital Library), where JeromeDL software is using the advanced services of the semantic network. The technology of semantic knowledge, raising the value and functionality of the digital library, will be one of the possible responses to the needs of users and librarians, who must be open to new technologies and ideas.

BIBLIOGRAPHY

Anglo-American Cataloguing Rules (2006). [online]; AACR [access: 20.05.2009]. Available in World Wide Web: <<http://www.aacr2.org>>.

Betz, Elizabeth W., compiled by (1982). *Graphic materials: rules for describing. Original items and historical collections* [online]; Library of Congress [access: 20.05.2009]. Available in World Wide Web: <<http://www.loc.gov/rr/print/gm/GraMatWP8.pdf>>.

Chojnacki, Piotr; Morawski, Radosław (2008). Digitalizacja i opracowanie fotografii z zasobu IPN w systemie ZEUS [Digitizing and cataloguing photographs from the stores of IPN in the Zeus system]. In: *Konferencja naukowa „Fotografia w nowoczesnym archiwum”* [Academic Conference: “The photograph in the modern archive]. Warszawa: Wydaw. Instytutu Pamięci Narodowej, pp. 39-47.

De Lusenet, Yola (2005). Przechowywane fotografie, Doświadczenia programu SEPIA w latach 1999-2003 [Preserved photographs. Experiences of the SEPIA program in the years 1999-2003.]. In: *ARCHEION, TCVIII*, pp. 109-120.

Dublin Core [online] (1999). Dublin Core Metadata Initiative [access: 20.05.2009]. Available in World Wide Web: <<http://dublincore.org/documents/1999/07/02/dces/>>.

Encoded Archival Description [online]; [access: 20.05.2009]. Available in World Wide Web: <<http://www.loc.gov/ead/ead.html>>.

AFN (1997). FD Z 44-077 Septembre 1997 Documentation – Catalogue de l’image fixe – Redaction de la description bibliographique. Paris: Association Francaise de Normalisation.

FOBID Netherland Library Forum [online]; FOBID [access: 20.05.2009]. Available in World Wide Web: <<http://sitegenerator.bibliotheek.nl/fobid/overig29/overig50.asp>>.

Forum Biblioteka 2.0 [Library 2.0 Forum] (b.d.). *Tezaurus dla bibliotek cyfrowych; Hasła formalne, przedmiotowe, rzeczowe* [Thesaurus for digital libraries: formal and subject headings] [online]; Biblioteka 2.0 [access: 20.05.2009]. Available in World Wide Web: <<http://forum.biblioteka20.pl/>>.

Górny, Mirosław; Skutecki, Jakub (2005). Problemy udostępniania zbiorów ikonograficznych w bibliotekach cyfrowych. [Issues with opening access to photographic collections in digital libraries] In: *e-publication of the Institute IniB UJ* [online], no 1. Institute of Information and Book Studies, Jagiellonian University [access: 20.05.2009]. Available in World Wide Web: <http://www.inib.uj.edu.pl/wyd_iinb/s3_z1/gorny.pdf>.

ISAD(G) (1999). *Międzynarodowy Standard Opisu Archiwalnego* [The International Standard for Archival Description]

General Part. Version 2. [online]; Archiwum Główne Akt Dawnych w Warszawie [access: 20.05.2009]. Accessible in World Wide Web: <<http://www.agad.archiwa.gov.pl/isad/ISADv2PL.pdf>>.

Klijn, Edwin; Sesink, Laurents (2003). SEPIA working group on descriptive models and tools. In: *An International SEPIA Conference at the Finnish Museum of Photography* [online]; European Commission on Preservation and Access [access: 20.05.2009]. Available in World Wide Web: <<http://www.knaw.nl/ecpa/sepia/conferencePapers/Klijn.pdf>>.

Klijn, Edwin; De Lusenet, Yola. (2005). SEPIADES. Katalogowanie zbiorów fotograficznych [SEPIADES. Cataloguing photographic collections]. *ARCHEION, TCVIII*, pp. 121-136.

Communique (2005). *Communique of the Committee to the European Parliament, Council, the European Economic-Social Committee and the Committee of the i2010 Regions: COM Digital Libraries*. 465 final version [online]; Brussels: cCommission of the European Union, 30 September, 2005. European Commission [access: 29.01.2009]. Available in World Wide Web: <http://ec.europa.eu/information_society/activities/digital_libraries/doc/communication/pl_comm_digital_libraries.pdf>.

Kucharska, Jolanta B.; Miller, Maria; Wornbard, Małgorzata (2009). Różne aspekty opisu dokumentów ikonograficznych w Bibliotece Cyfrowej Politechniki Warszawskiej na przykładzie kolekcji fotografii Henryka Poddębskiego [Various aspects of the description of iconographic documents in the Digital Library of the Warsaw University of Technology University on the example of the collection of photographs of Henryk Poddębski]. In: *Konferencja naukowa „Polskie biblioteki cyfrowe 2008”* [Scientific Conference “Polish digital libraries 2008”], Poznań: Wydaw. Polskiej Akademii Nauk, pp. 53-61.

Lenartowicz, Maria, red. (2005). *Format MARC21 rekordu bibliograficznego dla książki* [online]. Wersja ze zmianami redakcyjnymi Marii Lenartowicz. Narodowy Centralny Katalog Biblioteczny [MARC21 format of the bibliographical record for the book [online]. Version with the editorial changes of Maria Lenartowicz. National Central Library Catalogue]; Centrum NUKAT [access: 20.05.2009]. Available in World Wide Web: <http://www.nukat.edu.pl/nukat/book/aneks_bib-070202b.pdf>.

MARC Standards (2008). [online]; Library of Congress [access: 20.05.2009]. Available in World Wide Web: <<http://www.loc.gov/marc/>>.

Metadane (b.d.). Wikipedia [online]; [access: 20.05.2009]. Accessible in World Wide Web: <<http://pl.wikipedia.org/wiki/Metadane>>.

Nahotko, Marek (2004). *Metadane* [Metadata]. Kraków: Wydaw. Uniwersytetu Jagiellońskiego.

Narodowe Archiwum Cyfrowe (2008). [online]; [access: 20.05.2009]. Available in World Wide Web: <<http://www.nac.gov.pl/cms/main.php?did=2>>.

Pątek, Krzysztof (2006). *Doświadczenia Archiwum Dokumentacji Mechanicznej w dziedzinie digitalizacji zasobu materiałów archiwalnych* [Experiences of the archive for mechanical documentation in the digitization of the stores of archive materials]. Papers given by archivists [online]; Directorship of National Archives [access: 20.05.2009].

Available in World Wide Web: <http://www.archiwa.gov.pl/images/stories/file/pdf/krzysztof_patek.pdf>.

Pietrzak, Agata, oprac. (2008). *Format MARC 21 rekordu bibliograficznego: dokumenty ikonograficzne: wersja 2008* [Format Marc 21 of the bibliographical record: iconographic documents: version 2008]. Warszawa: Wydaw. Biblioteki Narodowej.

Płoszajski, Grzegorz, red. (2009). *Standardy w procesie digitalizacji obiektów dziedzictwa kulturowego* [online]; [Standards in the digitization process of objects of cultural heritage]. Warszawa: Biblioteka Główna Politechniki Warszawskiej [access: 20.05.2009]. Available in World Wide Web: <<http://bcthe.warsaw.edu.pl/dlibra/docmetadata?id=1262>>.

PN (2009). PN-N-01152-4 April 2009 Opis bibliograficzny Cz. 4: Dokumenty ikonograficzne [Bibliographic Description. Part 4. Iconographic documents].

Sanetra, Krystyna (2003). *Format MARC21 rekordu bibliograficznego dla dokumentu elektronicznego* [Format MARC21 of a bibliographical record for electronic documents]. Warszawa: Wydaw. Stowarzyszenia Bibliotekarzy Polskich.

Sanetra, Krystyna (b.d.). *Format Marc21 rekordu bibliograficznego dla dokumentu ikonograficznego* [Format Marc 21 of bibliographical record: iconographic documents: version 2008] [private document].

Schmidt, Kazimierz (2008). *Standardy tak, ale jakie i czy za wszelką cenę?* [Yes to standards... but at any price?]: Academic conference „Fotografia w nowoczesnym archiwum” [Photography in the modern archive] Warszawa: Wydaw. Instytutu Pamięci Narodowej, pp. 31-36.

SEPIA Safeguarding European Photographic Images for Access [online]; European Commission on Preservation and Access [access: 20.05.2009]. Accessible in World Wide Web: <<http://www.knaw.nl/ecpa/sepia/>>.

Specification for the Metadata Elements for the Europeana Prototype V 2.0, 28/08/2008 (2008) (online). Europeana [access: 20.05.2009]. Available in World Wide Web: <http://dev.europeana.eu/public_documents/Specification_for_metadata_documents_in_the_Europeana_prototype.pdf>.

Ślaska, Katarzyna (2009). Europeana i polityczne zamierzenia Biblioteki Narodowej w zakresie digitalizacji [Europeana and the political aims of the National Library of Poland in field of digitization]. In: *XIII Konferencja z cyklu Digitalizacja „Od digitalizacji zaawansowanej do dojrzałej”* [XIIIth Conference from the digitizing cycle ‘From advanced to mature digitization’] Warszawa: Wydaw. Centrum Promocji Informatyki, pp. 85-108.

Wróblewska, Magdalena (2008). Obrazy miejsca – dawne fotografie Warszawy [Pictures of place – of photographs of Warsaw]. In: Warsaw University. Department Pamphlet, March, pp. 28-30.

