

Metodyka i standardy

Digitalizacji dokumentów archiwalnych

Marek Zieliński
Instytut Józefa Piłsudskiego w Ameryce
138 Greenpoint Ave, Brooklyn, NY 11222

Wersja 1.4.5

Spis treści

Standardy.....	2
Struktura Archiwum.....	2
Archiwa.....	2
Archiwum Instytutu Józefa Piłsudskiego w Ameryce.....	3
Zespoły, Jednostki i Dokumenty.....	3
Zdjęcia.....	3
Multimedia, prasa i inne zbiory.....	4
Procedura digitalizacji.....	4
Przygotowanie materiałów.....	4
Skanowanie.....	6
Rozdzielczość.....	6
Duże arkusze (mapy, gazety).....	6
Nazywanie plików.....	6
Dokumenty.....	6
Zdjęcia.....	8
Zapis metadanych.....	9
Dokumenty.....	9
Metadane.....	10
Zobacz też (odnośnik).....	12
Najczęściej używane pola opisów.....	13
Dostosowania specjalne.....	14
Obsługa programu Dspace.....	17
Wpisywanie danych.....	17
Przeglądanie (sprawdzanie).....	21
Edytowanie.....	21

Standardy

Jest wiele standardów i przyjętych praktyk przy digitalizacji archiwów. Nasza metodyka będzie oparta w miarę możliwości o dostępne standardy. W szczególności polecane jest opanowanie dwóch publikacji dotyczących procesu skanowania i opisu dokumentów:

- [*Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images*](#) U.S. National Archives and Records Administration (NARA), June 2004. Jest to opis techniczny procesu skanowania, tworzenia oryginałów cyfrowych, kontroli jakości itp.
- [*Dublin Core Metadata Best Practices*](#), CDP Metadata Working Group, September 2006, który zawiera opis podstawowych pól zestawu Dublin Core, ich znaczenie i zastosowanie.

Do opisu materiałów archiwalnych będziemy używać zestawu Dublin Core, oraz dodatkowych pól które będą potrzebne do organizacji danych. W opisach, jak również w programie do wpisywania danych stosujemy prefix *dc* dla zestawu DublinCore (np. dc.coverage.temporal) oraz prefix *pls* dla metadanych używanych w Instytucie (np. pls.archiveID).

Inne standardy, które mogą być użyte w przyszłości przy wymianie i konwersji danych to [EAD](#) mający zastosowanie w tworzeniu hierarchii opisu archiwalnego, i [TEI](#) mający zastosowanie w oznaczaniu (tagging) tekstu. Dspace posiada również narzędzia do eksportu danych z użyciem [METS](#) / [MODS](#).

Struktura Archiwum

Archiwa

Dla najwyższego elementu w hierarchii (identyfikator archiwum) wskazane jest użycie rejestru zewnętrznego, w celu uniknięcia duplikatów. Takim rejestrem może być rejestr MAB (w chwili obecnej nieistniejący) lub rejestr NDAP. Naczelną Dyрекcją Archiwów Państwowych (NDAP) zajmuje się między innymi katalogowaniem zasobów archiwalnych w instytucjach poza granicami kraju i stworzyła de-facto częściowy rejestr archiwów. Inne archiwa które posiadają własne katalogi, mogą używać innych oznaczeń. Całkowita rezygnacja z oznaczenia poziomu archiwum nie jest polecana.

Rejestr archiwów polonijnych NDAP (częściowy)

- 701 Instytut Józefa Piłsudskiego w Ameryce
- 702 Polski Instytut Naukowy w Nowym Jorku.
- 704 Stowarzyszenie Weteranów Armii Polskiej w Ameryce
- 708 Centralne Archiwum Polonii w Orchard Lake
- 709 Instytut Józefa Piłsudskiego w Londynie

W dalszej części będziemy używać powyższych identyfikatorów aby stworzyć unikalne nazwy plików przy skanowaniu dokumentów.

Biblioteka Kongresu Stanów Zjednoczonych posiada również [rejestr](#) archiwów, bibliotek itp. Jako alternatywę, można używać takiego kodu. Przykładowe kody są podane poniżej.

US-nynyjpi Instytut Józefa Piłsudskiego w Ameryce
US-nnpia Polski Instytut Naukowy w Nowym Jorku.

Kody takie przyznają organizacje narodowe, a koordynuje je [ISO/ISIL](#). Są to globalnie unikalne kody organizacji typu GLAM. Jeśli adna instytucja nie posiada takiego kodu, może o niego wystąpić do odpowiedniej geograficznie organizacji.

Niezależnie od wybranego schematu, krytycznym jest używanie jednolitego schematu dla wszystkich zasobów archiwalnych danego archiwum.

Archiwum Instytutu Józefa Piłsudskiego w Ameryce

Archiwum zawiera różnego rodzaju materiały, gromadzone w różnych okresach. Część kolekcji była zbierana w Polsce, część otrzymana w USA z różnych źródeł. Kolekcje mogły być zbierane przez jedną osobę lub składane z wielu źródeł. Od kilku lat archiwiści z Naczelnej Dyrekcji Archiwów Państwowych porządkują i katalogują kolekcje, stosując w miarę możliwości jednolite zasady i terminologię. Większość skatalogowanych już dokumentów da się opisać według reguł wymienionych poniżej.

Zespoły, Jednostki i Dokumenty

Zebrane materiały dokumentowe i inne podzielone są na [Zespoły Archiwalne](#). Zespół odpowiada w przybliżeniu pojęciu [Fonds](#) w archiwistyce USA. Zespół jest dalej podzielony na Jednostki Archiwalne, które mogą odpowiadać teczkom, jakkolwiek jednostka może zawierać większą ilość (pod)teczek. Wewnątrz jednostki strony dokumentów (nie dokumenty!) są paginowane, tj. nadawane im są kolejne numery stron, ciągle w danej jednostce. Jeśli jednostka podzielona jest na (pod)teczki, numeracja stron jest kontynuowana od pod-teczki do pod-teczki. [Spis zespołów i jednostek](#) w archiwum Instytutu jest dostępny na stronie Internetowej Instytutu. Uwaga: definicja „jednostki archiwalnej” jest dość szeroka, i nie odpowiada pojęciu „item” w archiwistyce anglosaskiej. W archiwach Instytutu jednostki typowo odpowiadają teczkom („files”), grupom teczek, pojedynczym księgom itp. W przypadku zdjęć, jednostka archiwalna przypisana została albo pojedynczemu zdjęciu, albo ich grupie.

Pojęcie **dokumentu** („item”) używane przy opisywaniu zbiorów zakłada typowo większą granularność od „jednostki”. W tezcze może być wiele dokumentów, które opisujemy osobno, np. pojedynczy list, list z komentarzami, list i odpowiedź, odezwa, wycinek prasowy. W niektórych przypadkach można stosować zgrupowanie związanych ze sobą materiałów.

Zdjęcia

Fotografie zgrupowane są w zespół fotografii (o numerze 151). Dodatkowo, fotografie mogą się znajdować w poszczególnych zespołach archiwalnych. Fotografie nie zostały oryginalnie podzielone na grupy, a numery *jednostek archiwalnych* przypisywane były albo poszczególnym fotografiom albo ich zestawom.

W celu uporządkowania zasobu fotografii a także ułatwienia przenoszenia i dodawania zdjęć, ustalony zostaje dodatkowy poziom hierarchii, Seria (Series).

- **Seria (series)** to zgrupowanie na najwyższym poziomie poniżej zespołu. Serie obejmują

powiązane tematycznie zestawy zdjęć. W chwili obecnej w archiwum Instytutu Piłsudskiego w Ameryce wyodrębnionych jest 14 serii:

- 1 Archiwum osobowe
- 2 Józef Piłsudski
- 3 Legiony, inne formacje wojskowe, wydarzenia do 1918 roku
- 4 II Rzeczpospolita
- 5 II wojna światowa
- 6 Powojenne, demobilizacja
- 7 Sztuka
- 8 „Solidarność” i opozycja antykomunistyczna
- 9 Miasta, architektura
- 10 Polonia
- 11 Albumy
- 12 Archiwum Wschodnie
- 13 Michał Potulicki "Germany on War and Peace, 1741-1944, a pictorial history".
- 14 Instytut Józefa Piłsudskiego w Ameryce

Takie zgrupowanie powoduje, że w przypadku zdjęć hierarchia nazewnictwa plików jest nieco inne – patrz poniżej.

Multimedia, prasa i inne zbiory.

Kolekcja audiowizualna jest podobnie jak fotograficzna podzielona na serie. Prasa jest rozdzielona na część opracowaną w katalogu biblioteki, i część archiwalną. Temat ten opisany jest w osobnym dokumencie.

Procedura digitalizacji

Przygotowanie materiałów

- Archiwa w danym **zespole** (kolekcji, fonds) podzielone są na jednostki archiwalne. Są to typowo teczki, aleteczka może być jeszcze podzielona na foldery, zależnie od tego, skąd pochodzi. Pracujemy nad jedną jednostką archiwalną albo teczką na raz. Jest ważne, aby zarezerwować sobie dosyć czasu, żeby zeskanować całą jednostkę na jedno posiedzenie. Może to trwać jeden lub kilka dni, ale powinno się dążyć do skończenia jednostki lub teczki przed rozpoczęciem następnej.
- Należy unikać skanowania pojedynczych dokumentów ze środka teczki.

1) Wybór zespołów.

Wybór zespołów lub ich części dokonywany jest przez personel Instytutu (Dyrektor lub dyżurna archiwistka/archiwista), zgodnie z programem digitalizacji. Należy zachować środki ostrożności przy przenoszeniu kruchych dokumentów. W przypadkach starych i delikatnych dokumentów zaleca się pracować w rękawiczkach.

2) Paginacja

Strony powinny być numerowane w sposób ciągły wewnątrz jednostki. Należy upewnić się, że stosuje się właściwą paginację, szczególnie jeśli dokumenty były opisywane i paginowane kilkakrotnie, i jest kilka systemów numeracji. W razie wątpliwości skonsultować z personelem Instytutu. Jeśli paginacji brakuje, należy się skonsultować z personelem Instytutu i uzupełnić paginację przed rozpoczęciem skanowania.

- a) **Dokumenty:** Strony dokumentów są numerowane w sposób ciągły wewnątrz jednostki. Należy zwrócić uwagę na to, że jednostka może się rozciągać na kilka fizycznych teczek / folderów.
- b) **Zdjęcia:** Zdjęcia początkowo były numerowane w sposób ciągły w ramach całego zespołu, przy czym numer był nadawany albo pojedynczemu zdjęciu albo ich grupie. Po wprowadzeniu serii, każda seria może mieć osobną numerację. W chwili obecnej nie ma planów zmieniania numeracji istniejących zasobów.

3) **Dokumenty i strony**

Dokumenty mogą mieć więcej niż jedną stronę. Skanujemy każdą stronę, a opisujemy cały dokument i zaznaczamy jego strony przy opisie. Do opisu dołączamy skany wszystkich stron. Opisujemy zawsze dokumenty o ciągłej numeracji, nie wybierając poszczególnych stron ze środka grupy. Jeśli istnieje potrzeba odniesienia się do powiązanego dokumentu który znajduje się w innym miejscu (w innej grupie stron, w innej jednostce itp.), używamy pól odniesienia („Zobacz też”) - patrz poniżej.

4) **Arkusze pracy**

Przed rozpoczęciem skanowania zapisujemy informację w arkuszu danych (patrz wzór poniżej). Głównym celem arkusza jest zanotowanie postępu prac, a także ułatwienie powrotu do pracy nad jednostką / teczką po przerwie. W arkuszu zaznaczamy strony zeskanowane i strony opisane, z notatką (inicjałami) operatora. Arkusz należy powielić, jeśli brakuje miejsca. W arkuszu należy też wpisywać notatki i uwagi (patrz następny punkt). Arkusz pracy przechowujemy w jednostce lub teście, której on dotyczy.

5) **Tabela postępów pracy**

Do notowania postępów pracy stosujemy arkusz rozliczeniowy w Google Drive, do którego dostęp uzyskują pracujący. W arkuszu „Digitalizacja” notujemy postęp prac, która jednostka / folder jest „użyty”, i które strony zostały już opracowane, aby ułatwić innym wybranie następnej sekcji. Dane te są ważne, gdyż wiele osób może pracować nad danym zespołem, i w ten sposób łatwiej uniknąć duplikacji czy opuszczenia fragmentów kolekcji.

6) **Ubytki i problemy ze skanowaniem**

Jeśli brakuje strony, jest podwójna paginacja, itp. należy ten fakt zanotować na arkuszu i poinformować o tym personel Instytutu. Jeśli nie uda się zeskanować jakiejś strony (problemy z rozmiarem, itp), należy to też zanotować i powiadomić personel Instytutu. Należy zawsze dążyć do uzyskania kompletnego zestawu stron.

7) **Duplikaty**

Zdarza się, że pojawiają się duplikaty dokumentów. Skanujemy i opisujemy dokument tylko raz (wybierając w miarę możliwości kopię lepszej jakości). Wyjątkiem od tej reguły jest przypadek, w którym kopie różnią się od siebie, np. jedna z nich ma ręczne adnotacje itp. - w takim przypadku należy skanować obie wersje, a w opisie zastosować odpowiednią adnotację.

Skanowanie

Rozdzielczość

- Skanujemy w formacie **TIFF, z rozdzielczością 600 dpi**, w kolorze.
- Zdjęcia mogą wymagać większej rozdzielczości, jeśli są w małym formacie (np. odbitki stykowe, itp). Rozdzielczość 600 dpi stosuje się do zdjęcia w formatach 6 x 9 cali lub większym. Jeśli negatyw lub odbitka ma rozmiar mniejszy, stosujemy proporcjonalnie większą rozdzielczość.
- Pracujemy z **oryginałem cyfrowym** (oryginał skanu w formacie tiff) i z **kopią cyfrową** (typowo zmniejszona kopia w formacie jpg). Oryginały są przeznaczone do przechowania archiwalnego. Kopie cyfrowe są używane jako kopie robocze do prezentacji na stronie internetowej, przy opisywaniu itp.

Duże arkusze (mapy, gazety)

Jeśli arkusz nie mieści się na płycie skanera, skanujemy go w częściach. Najczęściej pojawia się to przy skanowaniu map, dużych płacht gazet itp. Przy skanowaniu części zawsze stosujemy nakładkę tak, żeby fragment na brzegu pojawiał się w kolejnych skanach. Jest to szczególnie ważne w przypadku map, które będą potem sklejjane w całość przy pomocy programu, który rozpoznaje nakładające się fragmenty

Kolejne skany otrzymują kolejny numer podstrony, zaczynając od 001 – patrz strony i podstrony poniżej.

Nazywanie plików

Dokumenty

Do nazywania plików używamy hierarchii odpowiadającej w przybliżeniu hierarchii zasobów Instytutu. Ta hierarchia używana jest w większości zespołów dokumentów, z wyjątkami opisanymi poniżej.

Archiwum

Zespół (kolekcja, fonds)

Jednostka (teczka)

Strona (kolejna strona dokumentu w jednostce)

Podstrona (tylko jeśli potrzebna)

Zastosowany system nazywania plików daje nam dwie użyteczne cechy. Pliki są unikalne w ramach Instytutu, a dzięki użyciu uniwersalnego identyfikatora, również w ramach większej grupy polskich archiwów. Nazwa pliku pozwala na zidentyfikowanie oryginalnej strony, której jest obrazem. Numeryczna hierarchia ułatwia sortowanie i kontrolę jakości pracy, a także odnajdywanie zagubionych plików i poprawianie błędów w trakcie pracy.

1. Nazwa pliku jest numeryczna i składa się z segmentów odpowiadających numeracji poszczególnych elementów np. 701-008-012-034-002.tif , 701-151-10022-003-001.tif. Numery te zapisywane są też w odpowiednich polach bazy danych dla dokumentu w DSpace.
2. Dla większości danych segmenty są 3-cyfrowe. Spodziewamy się, że ilość cyfr dla archiwum (701 w klasyfikacji NDAP) i zespołów (posiadamy ok 160 zespołów) jest wystarczająca.

Sposób katalogowania zdjęć przez pracowników NDAP spowodował, że liczba **jednostek** w zespole 151 przekracza 20,000, tak więc w tym zespole jednostka jest liczbą pięciocyfrową. Podobnie jest w przypadku jednostek, w których liczba **stron** przekracza 999. Należy określić lub oszacować liczbę stron i w razie potrzeby użyć liczby 4 cyfrowej lub większej. Jeśli w trakcie pracy okaże się, że wymagane jest zwiększenie liczby cyfr, można użyć narzędzia, np. *Bulk Rename Utility* do hurtowej zmiany liczby znaków, tak, aby wszystkie pliki w danej jednostce miały tę samą liczbę znaków w oznaczeniu strony.

Strony i podstrony

Podstrony dodajemy typowo w dwóch przypadkach:

- Jeśli strona nie daje się zeskanować w całości (ze względu na rozmiar, stan, uszkodzenia itp)
- Jeśli pojawia się strona, która nie otrzymała numeru podczas paginacji. Może to być osobna strona, odwrotna strona kartki, itp. Jeśli braki w paginacji są duże, należy całą jednostką spaginować od nowa; małe defekty można poprawić używając podstron.

Uwaga: Jeśli używamy podstron, **wszystkie** podstrony danej strony powinny mieć własny numer podstrony, a więc strona **034-001, 034-002, 034-003** (a nie strona 034, 034-001 !).

Okładki

- Okładka foldera z nalepką opisującą pozycję katalogu (zespół, jednostkę) otrzymuje numer strony 000.
- Jeśli jest więcej okładek (np. oryginalne foldery z notatkami, czego dotyczą), otrzymują one numery kolejnych podstron, 000-001, 000-002, ... itp. W opisie lub uwagach podajemy, co zawierała dana okładka.

Opis i przykłady numeracji plików:

Archiwum:	Zespół (Patrz spis zespołów)	Jednostka	Strona	Podstrona (opcjonalnie)
Dokumenty				
701	008	012	034	002
Zawsze 701 dla Instytutu Józefa Piłsudskiego w Ameryce.	008 to Powstania Śląskie	012 to Milicja Górnośląska w Szczakowej...	Paginacja jest już zrobiona. Skonsultować personel Instytutu, jeśli jest podwójna paginacja lub jej brakuje. Użyć strony 000 do skanu nalepki teczki. Jeśli jest kilka teczek w zespole, skany otrzymują numery 000-001, 000-002 itp. Jeśli liczba stron przekracza 999, należy jednolicie stosować kod czterocyfrowy (0001,	W razie potrzeby, jeśli stworzenie obrazu strony wymaga więcej niż jednego skanu. Podstrony używa się także do wykonywania korekt, jeśli zmiana paginacji całej jednostki jest niepraktyczna.

Archiwum:	Zespół (Patrz spis zespołów)	Jednostka	Strona	Podstrona (opcjonalnie)
			0002 itp)	
Nazwy pól w programie DSpace				
pls:archiveID	pls:fondsID	pls:folderID	pls.page.first pls.page.last (liczby bez wiodących zer!)	(podstrony są tylko w pliku, nie używamy podstron w opisie)

Zdjęcia

Do nazywania plików używamy hierarchii odpowiadającej hierarchii zasobów zdjęć, z wyodrębnioną grupą (subfonds).

Archiwum

Zespół (kolekcja, fonds)

Seria (series)

Jednostka (jedno zdjęcie, jeden album lub mała grupa podobnych zdjęć)

Strona (Zdjęcie w jednostce, jeśli wyróżnione)

Podstrona (tylko jeśli potrzebna, np. jeśli odwrotna strona zdjęcia nie ma numeru, a wymaga zeskanowania)

Archiwum:	Zespół (Patrz spis zespołów)	Seria	Jednostka (item)	Strona	Podstrona (opcjonalnie)
Zdjęcia					
701	151	002	0003	001	
Zawsze 701	151 to Zespół (Obecnie 151 to jedyny w Instytucie zespół fotografii)	002 to identyfikator serii „Józef Piłsudski”	Liczba zdjęć (jednostek) w grupie określa liczbę znaków	Jeśli jednostką jest album, można numerować poszczególne zdjęcia albo całe strony, jeśli układ graficzny stanowi całość.	Być może nie jest potrzebna
Nazwy pól w programie DSpace					
pls:archiveID	pls:fondsID	pls:seriesID	pls:itemID	pls.page.first pls.page.last (liczby bez wiodących zer!)	(podstrony są tylko w pliku, nie używamy podstron w opisie)

Przebieg skanowania

1) Skan → oryginał cyfrowy

Układamy arkusz na szkle, odsuwając go od brzegów stołu skanera, aby skanować cały dokument (włącznie z brzegami). Wybieramy „Prescan” i ustawiamy ramkę tak, aby widoczne były brzegi dokumentu. Ważne jest, aby ułożyć dokument równoległe do brzegów skanera. Każda dodatkowa obróbka, np. prostowanie, wprowadza dodatkowe błędy do oryginału cyfrowego.

- a) Wybieramy miejsce do zapisania gotowego pliku. Pliki przechowujemy na dysku sieciowym (np. Kufer), w katalogu oryginałów, w podkatalogach oznaczonych tak jak numeracja Zespołu i Jednostki.
- b) Wpisujemy nazwę (prefiks) wg reguł powyżej (przykładowo 701-008-014-). Ostatnie 3 cyfry – numer strony – program dodaje sam, posuwając licznik o jeden przy kolejnym skanie. Należy upewnić się, że numer ten odpowiada paginacji skanowanej strony. W miarę potrzeb można ręcznie zmienić nazwę pliku na dysku. Przykłady:

{701-008-012-} {034} strona 34 bez podstron.
prefiks licznik

{701-008-012-035-} {002} strona 35 z podstroną 002
prefiks licznik

2) Obróbka

Dokonyjemy obróbki w miarę potrzeby. Typowe przeróbki to

- a) **Obracanie:** Dokumenty mogą wymagać obrócenia o 90 stopni, jeśli tylko tak mieszczą się w skanerze. Należy użyć programu **Obróbka Skanów** do obracania dokumentów (patrz niżej).
- b) Prostowanie, przycinanie, poprawa jasności i wysycenia. Używamy programu GIMP do obróbki oryginału, tylko jeśli zachodzi taka potrzeba.

3) Tworzenie kopii cyfrowej

Tworzymy podstawowa kopie cyfrową do pracy w formacie JPG, z jakością 90 – 95% i oryginalna rozdzielczością zredukowana o 50%, zachowując oryginalne dane EXIF i IPTC. Korzystamy z programu **Obróbka Skanów** do generowania kopii cyfrowych (patrz niżej). Przy robieniu kopii cyfrowych jednocześnie przygotowujemy kopie do wyświetlania na stronie www.

4) Kontrola

Używając klawisza „[Sprawdź Katalogi](#)” programu **Obróbka Skanów** należy skontrolować pracę.

Zapis metadanych

Podstawowe metadane należy zapisywać w trakcie skanowania (patrz notatki w arkuszu digitalizacji). Po skończeniu skanowania serii dokumentów – typowo jednostki można przystąpić do zapisu metadanych. Można do tego używać innego komputera. Zapisujemy dane dotyczące całych **dokumentów, zdjęć** (lub innych rodzajów obiektów), nie poszczególnych jego stron.

Dokumenty

Dokument jest typowo zgrupowaniem stron w układzie odpowiadającym intencji twórcy tego dokumentu. Przykładowo może to być:

- Pojedynczy list
- Wielostronicowy memoriał
- Rękopis książki
- List i jego odpis (kopia) z poprawkami
- Dokument z załącznikami
- List przewodni i załączone do niego dokumenty
- Dokument i jego kolejne wersje
- Dokument i dodatkowe notatki dodane później, dotyczące jego treści.
- Itp.

Czasami w archiwum znajdują dokonane przez autora wybory lub kolekcje dokumentów. W takim przypadku możemy potraktować jako dokument całą kolekcję lub też wyodrębnić jej poszczególne składniki, w zależności od charakteru tego wyboru. W razie wątpliwości należy skonsultować z dyżurnym archiwistą.

Do zapisu danych używamy programu **DSpace**.

Metadane

Poniżej opisane są niektóre z danych (metadanych), które należy albo można przypisać dokumentowi. Pochodzą one z dwóch przestrzeni nazw (po angielsku *namespaces*): Dublin Core (**dc**) i własnego zestawu Instytutu (**pls**).

Archiwum (pls.archiveID)

Zawsze 701 dla Instytutu Piłsudskiego – jest to numer nadany przez Naczelną Dyрекcję Archiwów Państwowych.

Zespół (pls.fondsID)

Numer zespołu według katalogu Instytutu (można sprawdzić na [stronie](#) Internetowej). Występuje też na naklejce pudełka i teczki. Zawsze trzy-cyfrowy (wypełnić zerami w razie potrzeby)

Seria (pls.seriesID)

Seria tematyczna używanej dla fotografii. Patrz spis serii

Jednostka (pls.folderID)

Dokumenty: Jednostka jest typowo folderem, ale może mieć pod-foldery. Paginacja jest zawsze ciągła i unikalna (tj. nie ma dwóch stron o tym samym numerze) w zakresie jednostki.

Obiekt (pls.itemID)

Zdjęcia: Jednostka jest typowo pojedynczym zdjęciem, albumem albo może też być grupą zdjęć. Jeśli jest grupa, to kolejne numery nadawane są poszczególnym zdjęciom, odpowiadając paginacji.

Tytuł (dc.title)

- Zapis tytułu jest obowiązkowy.

- Tytuł to nazwa dokumentu. Pierwszy wybór to nazwa dokumentu przypisana przez jego twórcę, wydawcę. Jeśli brakuje, użyj tytułu nadrzędnego, lub znaleziony w innych źródłach. Jeśli tworzysz, **tytuł powinien opisywać zawartość np. „List Jana Kowalskiego do Urzędu Rady Ministrów”** a nie „List”. Używamy dużej litery tylko na początku tytułu (i w słowach własnych wewnątrz). Można zapisać też alternatywny tytuł, lub tytuł w innym języku.

Tytuł alternatywny (dc.title.alternative)

- Alternatywny tytuł. Jeśli oryginalny tytuł jest w języku innym niż polski, alternatywny może być tłumaczeniem.

Data dokumentu (dc.dateOriginal)

- Opcjonalnie
- Data oryginalnego stworzenia dokumentu. Zakres dat można podać niżej w polu Daty

Autor (dc.contributor.author)

- Opcjonalnie
- Osoba lub jednostka (instytucja) odpowiedzialna za stworzenie dokumentu

Daty (dc.coverage.temporal)

- Opcjonalnie
- Wpisujemy daty występujące w dokumencie i daty, których dokument dotyczy.
- Nie powtarzamy daty dokumentu.
- **Format yyyy-mm-dd (1928-12-27). Można wpisać tylko rok (1928) albo rok i miesiąc (1928-12).** Uwaga: nie wpisywać nic innego poza pojedynczą datą. Zakres data notujemy wpisując dwie daty: początkową i końcową

Miejsca (dc.coverage.spatial)

- Wpisujemy miejsca wymienione w dokumencie lub których dokument dotyczy.

Osoby (pls.person)

- Opcjonalnie
- Wpisujemy osoby wymienione w dokumencie.
- Nie powtarzamy autora.

Język (dc.language.iso)

- Opcjonalnie
- Język dokumentu (użyć oficjalnych skrótów)

Pierwsza strona (pls.page.first)

- Pierwsza strona dokumentu według przyjętej paginacji. Jest to pole cyfrowe, i może zawierać tylko liczby. Tylko pojedyncza liczba, bez wiodących zer. Jeśli wymagany jest komentarz, wpisujemy w pole „Opis”.

Ostatnia strona (pls.page.last)

- Ostatnia strona dokumentu według przyjętej paginacji. Jest to pole cyfrowe, i może zawierać tylko liczby. Tylko pojedyncza liczba, bez wiodących zer. Jeśli wymagany jest komentarz,

wpisujemy w pole „Opis”.

Typ (dc.type)

- To jest ważna kategoryzacja, która pozwoli z czasem na grupowanie dokumentów w różne kolekcje. Wymaga listy (spisu) typów. Typy mogą być różne dla różnych kolekcji (np. List, Nekrolog, Rozprawa, Obwieszczenie, Artykuł, Rozkaz, Spis, itp).

Słowa Kluczowe (dc.subject)

- Wpisać słowa kluczowe dotyczące danego dokumentu. Nie powtarzać nazwisk, dat i miejsc uprzednio wpisanych w inne pola.

Streszczenie (dc.description.abstract)

- Wpisać krótkie streszczenie pozycji, o czym jest, opis zdjęcia itp.

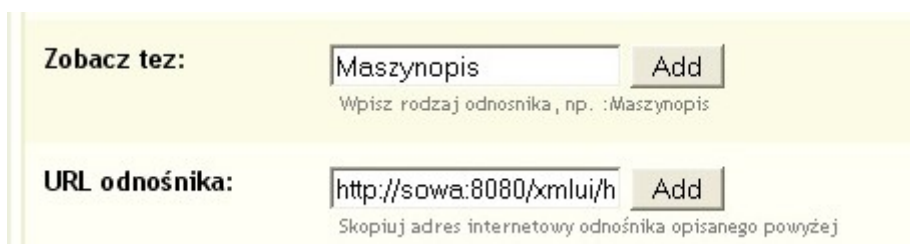
Opis (dc.description)

- Pozycję Opis można użyć, jeśli informacja nie pasuje do Streszczenia. Również dodatkowe uwagi można tu umieścić.

Zobacz też (odnośnik)

W celu umożliwienia wprowadzania linków do wewnętrznych lub zewnętrznych źródeł, wprowadzone zostały dwa pola **Zobacz też (dc.relation)** oraz **URL Odnośnika (dc.relation.uri)**. Pola te używamy do wskazania powiązanych materiałów albo w kolekcji Instytutu albo dostępnych w Internecie.

Przykładem może być odnośnik do tego samego dokumentu, przepisanego na maszynie (może być z komentarzem), który znajduje się w innej teczce (jednostce). Inny przykładem może być link do zewnętrznej strony (np. artykułu w Wikipedii) która rozwija lub opisuje temat opisany w dokumencie.



The image shows a screenshot of a web form with two sections. The first section is labeled 'Zobacz też:' and contains a text input field with the value 'Maszynopis' and an 'Add' button. Below the input field is the text 'Wpisz rodzaj odnośnika, np. :Maszynopis'. The second section is labeled 'URL odnośnika:' and contains a text input field with the value 'http://sowa:8080/xmlui/h' and an 'Add' button. Below this input field is the text 'Skopiuj adres internetowy odnośnika opisanego powyżej'.

Ilustracja: Pola „Zobacz też” i „URL odnośnika” w DSpace

Zobacz też (dc.relation)

- W pole to wpisujemy **opis linku** powiązania, np. „Dokument przepisany z komentarzem”, „Maszynopis”, „Artykuł w Wikipedii”, „Artykuł dotyczący kampanii” itp.

URL Odnośnika (dc.relation.uri)

- W pole to wpisujemy **hiperłącze** (hiperlink) prowadzące do danego źródła. Hiperłącze może mieć postać
 - wewnętrzną, prowadzącą do dokumentu w DSpace, np. <http://sowa:8080/xmlui/handle/701.180/3511>
 - zewnętrzną, prowadzącą do dokumentu w Internecie, np.

http://pl.wikipedia.org/wiki/J%C3%B3zef_Pi%C5%82sudski

- w każdym wypadku pole to powinno zawierać kompletne hiperłącze, prowadzące po kliknięciu do powiązanego materiału.

Wielokrotne odnośniki

Ze względu na ograniczenia DSpace, wpisujemy osobno pola **opisu linku** i **hiperłącza**, chociaż istnieje między nimi zależność jeden do jednego. Jeśli chcemy wpisać więcej niż jeden odnośnik, co jest dopuszczalne, mamy dwie opcje. Uwaga: Należy używać tylko opcji 1 lub 2, nie należy ich mieszać.

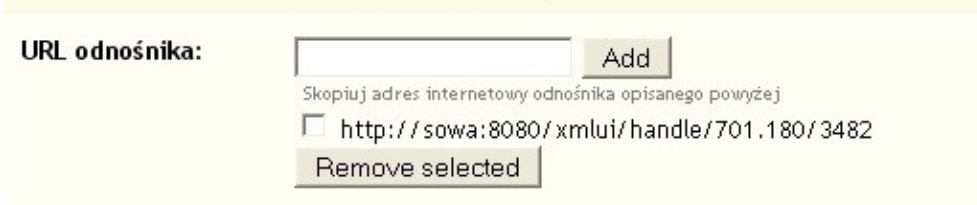
1. Używamy jednego

opisu linku i wielu **hiperłącza**; w tym wypadku wszystkie linki będą miały ten sam opis.



2. Używamy nowego

opisu linku do każdego **hiperłącza**. W tym wypadku, musimy podać dokładnie tyle opisów, ile dodaliśmy hiperłącza, i w tej samej kolejności.



Ilustracja 1: Pola 'Zobacz też' i 'Url odnośnika' po dodaniu pierwszej pary opisu i hiperłącza

Najczęściej używane pola opisów

Nazwa	Pole danych	Przykład	Opis
Archiwum	pls.archiveID	701	Archiwum, zawsze 701 dla Instytutu Piłsudskiego
Zespół	pls.fondsID	003	Zespół archiwalny w katalogu Instytutu Piłsudskiego, zawsze 3 cyfry (wypełniamy zerami na początku, jeśli potrzeba)
Jednostka	pls.folderID	012	Numer jednostki w zespole, zawsze 3 cyfry (wypełniamy zerami na początku, jeśli potrzeba)
Tytuł	dc.title		Wpisz główny tytuł pozycji. Albo oryginalny tytuł, albo, jeśli brak, stworzony przez opisującego
Tytuł alternatywny	dc.title.alternative		Jeśli pozycja ma tytuł alternatywny, wpisz go tu.
Data dokumentu	dc.dateOriginal	1913-12-24	Data powstania oryginalnego dokumentu
Autor	dc.contributor.author		Autor(rzy) tej pozycji
Instytucja - autor	dc.contributor		Autor jeśli jest to instytucja
Pierwsza strona	pls.page.first	23	Numer pierwszej strony spaginowanego dokumentu. Wpisujemy tylko jedną liczbę, bez znaków przestankowych i komentarzy. Nie wpisujemy podstron
Ostatnia strona	pls.page.last	25	Numer ostatniej strony spaginowanego dokumentu. Wpisujemy tylko jedną liczbę, bez znaków przestankowych i komentarzy. Nie

			wpisujemy podstron. Jeśli dokument ma jedną stronę (nawet, jeśli ma kilka podstron), wpisujemy tę samą liczbę w oba pola (pierwsza i ostatnia strona)
Język	dc.language.iso	pl	Wybierz główny język pozycji. Jeśli nie ma języka w spisie, zgłoś administratorowi i wybierz "inny". Jeśli jednostka nie ma specyficznego języka (np. zdjęcie), wybierz "N/A".
Miejsca	dc.coverage.spatial		Miejsca dotyczące lub wymienione w tej pozycji
Daty	dc.coverage.temporal	1810-07-30 1810-07 1810	Daty (zakres dat) dotyczące tej pozycji. Format yyyy-mm-dd np. 1810-07-30. Tylko jedna data w polu (<u>nie zakres dat</u>) i tylko w formacie podanym powyżej.
Osoby	pls.person		Osoby wymienione w dokumencie.
Typ	dc.type		Wybierz typ zawartości tej pozycji. Aby wybrać więcej niż jedną wartość z listy, trzymaj naciśnięty klawisz "Ctrl" lub "Shift".
Słowa kluczowe	dc.subject		Wpisz słowa lub frazy kluczowe.
Streszczenie	dc.description.abstract		Wpisz streszczenie pozycji
Opis	dc.description		Wpisz opis i inne uwagi
Opis linku	dc.relation		Opis linku powiązania do innego dokumentu, wskazanego przez hiperłącza
Hiperłącze	dc.relation.uri		Hiperłącze do materiału do którego nawiązuje opis linku (dc.relation)
Wydawca	dc.publisher		Enter the name of the publisher of the previously issued instance of this item.
Identyfikatory	dc.identifier		If the item has any identification numbers or codes associated with
Przypis	dc.identifier.citation		Enter the standard citation for the previously issued instance of this item.
Data wydania	dc.date.issued		Please give the date of previous publication or public distribution
Nr Serii / Jednostki	dc.relation.ispartofseries		Enter the series and number assigned to this item by your community.

Dostosowania specjalne

Często okazuje się, że kolekcje mają specjalne potrzeby dotyczące zapisu danych. W miarę potrzeb będziemy dodawać dane i modyfikować sposób zapisu, aby najlepiej opisać dany rodzaj dokumentów.

Zespół 008: Powstania Śląskie

Nazwiska wpisujemy uwzględniając numer strony.

- W pod-pole "Last name" wpisujemy: **Numer strony**, przecinek, Nazwisko. Opcjonalnie: Numer strony, przecinek, Nazwisko, przecinek, stopień wojskowy. Po przecinku dajemy spację.
- W pod-pole „First name” wpisujemy imię lub imiona, rozdzielone spacjami.

Przykłady:

17, Kowalski, Pułkownik	Artur
17, Raczewski, Starszy Sierżant	Jan
17, [Komornicki]	
17, Gajda; [Gajda]	Franciszek Józef
itp.	

Program zamieni kolejność, dając imię najpierw, więc w spisie na stronie będzie to widać tak

17, Artur Kowalski, Pułkownik
17, Jan Raczewski, Starszy Sierżant
17, Komornicki
17, Franciszek Józef Gajda; [Gajda]

W przykładzie powyższym, wszystkie nazwiska znajdują się w oryginalnym dokumencie na jednej stronie (stronie 17).

Zespół 151: Kolekcja Fotografii

W przypadku kolekcji fotografii, *jednostką* nie jest folder, tylko pojedyncza fotografia ewentualnie ich grupa. Zdjęcia zostały tak wcześniej skatalogowane.

Aby ułatwić grupowanie zdjęć, wprowadzamy dodatkowe pole (*seria*) – fotografia będzie należeć do jednej z tych serii.

Skanowanie stron fotografii

Jeśli rewers fotografii zawiera istotną informację, ręczny zapis, etc., należy zeskanować również drugą stronę. Jeśli jest tam tylko opis zdjęcia, należy ten tekst wpisać w pole „Opis”

Opisywanie fotografii

W zasadzie opisujemy każde zdjęcie, traktując je jako osobny dokument. Jednakże można opisać całą grupę (w tej samej *jednostce*), jeśli jest to zdjęcie tej samej grupy osób, i różni się tylko szczegółami ujęcia. Jeśli pojawiają się inne osoby, lub szczegóły opisu, wpisujemy je osobno.

Przykład:

Opisujemy zdjęcia o numerach:

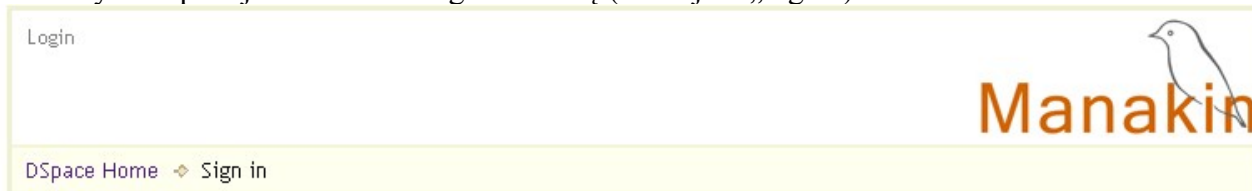
701.151.003.10022.001, 701.151..003.10022.002, 701.151.003.10022.003 i traktujemy je jako osobne dokumenty. W pole „*Jednostka*” wpisujemy za każdym razem 10022, a w oba pola: *Pierwsza strona* i *Ostatnia strona* wpisujemy 1 dla pierwszego zdjęcia, 2 dla drugiego i 3 dla trzeciego:

<i>Jednostka</i>	<i>Pierwsza strona</i>	<i>Ostatnia strona</i>
10022	1	1
10022	2	2
10022	3	3

Obsługa programu Dspace

Wpisywanie danych

Pierwszym etapem jest zawsze zalogowanie się (kliknij na „login”)



DSpace Home ♦ Sign in

Wybieramy „Communities and Collections” z menu, i klikamy na zespół, który będziemy opracowywać (Jeśli go brakuje, prosimy administratora o dodanie Zespołu).



Browse

- All of DSpace
 - [Communities & Collections](#)
 - [By Issue Date](#)
 - [Authors](#)
 - [Titles](#)
 - [Subjects](#)

Aby dodać nową pozycję, wybieramy *Submit a new item to this collection*

Jeśli pracowaliśmy już nad jakąś pozycją i nie dokończyliśmy pracy, możemy wybrać *Submissions*, co pokaże nam niedokończone pozycje, nad którymi pracowaliśmy.

Arkusze wprowadzania danych zaczyna się od pytań:



Item submission

Initial Questions → Describe → Describe → Upload → Review → License → Complete

Initial Questions

Multiple titles:
The item has more than one title, e.g. a translated title

Published:
The item has been published or publicly distributed before

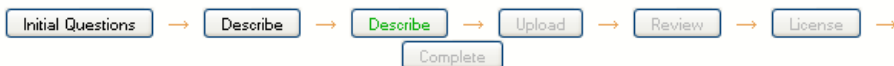
Save & Exit Next >

Będziemy przechodzić kolejno przez wszystkie etapy. Zakreślamy pierwszy kwadracik, jeśli dokument ma alternatywne tytuły, drugi jeśli potrzeba będzie wpisać dane publikacji. Uwaga: jeśli zmienimy zdanie, można wrócić do tych pytań w każdym etapie wpisywania, poprzez wybranie etapu (w ramkach w górnej części ekranu). Po naciśnięciu klawisza *Next* > Pojawia się arkusz danych:

Obowiązkowe dane to **Archiwum, Zespół i Jednostka, Tytuł, Pierwsza strona i Ostatnia strona**. Pola Archiwum i Jednostka powinny być już wypełnione. Wpisujemy numer pierwszej i ostatniej strony dokumentu (lub tę sama stronę w oba pola, jeśli strona jest tylko jedna). Autor może być osoba, instytucją lub można użyć obu pól. Naciskamy klawisz *Add* aby dodać więcej niż jedną pozycję danego rodzaju. Po zakończeniu naciskamy *Next* > aby przejść do następnej strony. W każdej chwili możemy też nacisnąć „Save & Exit” i wrócić później do pracy w tym samym miejscu.

Kolejna strona daje nam możliwość dodania danych:

Item submission



Describe Item

Daty:
Year Month Day
Daty (zakres dat) dotyczące tej pozycji

Miejsca:
Miejsca dotyczące lub wymienione w tej pozycji

Osoby:
Last name, e.g. Smith First name(s) + "Jr", e.g. Donald Jr
Osoby wymienione w pozycji

Typ:
Animation
Artykuł
Książka
Rozdział
Zestaw danych
Dokument
Wybierz typ zawartości tej pozycji. Aby wybrać więcej niż jedną wartość z listy, trzymaj naciśnięty klawisz "Ctrl" lub "Shift".

Słowa kluczowe:
Wpisz słowa lub frazy kluczowe.

Streszczenie:
Wpisz streszczenie pozycji

Opis:
Wpisz opis i inne uwagi

Identyfikatory:
If the item has any identification numbers or codes associated with it, please enter the types and the actual numbers or codes below.

Nr Serii / Jednostki:
Series Name Report or paper No.
Enter the series and number assigned to this item by your community.

< Previous

Save & Exit

Next >

Na tej stronie możemy wpisać dodatkowe osoby, miejsca i daty, typ dokumentu (można wybrać więcej niż jeden typ), dodać słowa kluczowe, streszczenie i opis dokumentu, oraz identyfikatory, jeśli takie istnieją. Nie powtarzamy danych, które już wpisaliśmy (jako autora, daty dokumentu itp.).

Następna strona pozwala nam podłączyć pliki. Wybieramy kolejne pliki (jpg, 300dpi) dokumentu.

Upload File(s)

File:

Please enter the full path of the file on your computer corresponding to your item. If you click "Browse...", a new window will allow you to select the file from your computer.

File Description:

Optionally, provide a brief description of the file, for example "Main article", or "Experiment data readings".

I naciskamy „Upload file & add another”. Powtarzamy, aż wszystkie strony zostaną dołączone.

W następnym etapie możemy sprawdzić, czy wszystko wygląda w porządku, i zrobić poprawki, jeśli trzeba:

Review Submission

Initial Questions

Multiple titles: Yes

Published: Yes

Describe Item

Archiwum: 701

Zespół: 23

A w ostatnim przypisujemy obiektowi standardową licencją Instytutu Piłsudskiego.

Distribution License

There is one last step: In order for DSpace to reproduce, translate and distribute your submission worldwide, you must agree to the following terms.

Grant the standard distribution license by selecting 'I Grant the License'; and then click 'Complete Submission'.

NOTE: PLACE YOUR OWN LICENSE HERE This sample license is provided for informational purposes only.

NON-EXCLUSIVE DISTRIBUTION LICENSE

By signing and submitting this license, you (the author(s) or copyright owner) grants to DSpace University (DSU) the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the

Przeglądanie (sprawdzanie)

Większość kolekcji skonfigurowana jest tak, aby opracowane dokumenty (submissions) przechodziły przez proces przeglądania (*review*), zwykle przez innych członków zespołu.

Proces przeglądania składa się z 3 etapów (kroków) i nie wszystkie kroki muszą być uaktywnione w każdej kolekcji. Osoby przypisane do kolejnych kroków przeglądu otrzymują e maile o wprowadzeniu dokumentu, i mogą przejść dokument do przeglądu. Etapy przeglądu to:

1. Zaakceptuj / odrzuć.
2. Zaakceptuj / odrzuć / edytuj.
3. Edytuj metadane.

Każdy z etapów może mieć przypisanych tych samych lub innych edytorów. Po zalogowaniu edytor może obejrzeć gotowe do przeglądu wpisy i wybrać te, które chce skontrolować.

Edytowanie

Jeśli zakończyliśmy wpisywanie danych i nacisnęliśmy klawisz „Complete submission”, dokument został przekazany w następny etap procesu (workflow). Edytor w następnym etapie może wrócić do opisu w trybie edycji lub zwrócić dokument oryginalnemu autorów do poprawek.

Dokument zaaprobowany do kolekcji w Dspace może być również modyfikowany. W tej postaci pola opisane są ich oryginalnymi nazwami Dublin Core.



„Metodyka i standardy Digitalizacji dokumentów archiwalnych” by [Marek Zielinski](#) is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).