Załącznik nr 1 do rozeznania

**Opis Przedmiotu zamówienia**

**Nazwa przedmiotu zamówienia**

„Tak rodziła się niepodległość - życie codzienne Bydgoszczy w cyfrowej odsłonie. Etap 2”, w ramach programu Kultura cyfrowa 2025 - digitalizacja czasopisma „Ostdeutsche Presse” ze zbiorów Wojewódzkiej i Miejskiej Biblioteki Publicznej im. dr. W. Bełzy w Bydgoszczy.

**Przedmiot zamówienia**

1. Przedmiot zamówienia składa się z dwóch zadań:

1. Zadanie 1 - usługa digitalizacji 1 887 publikacji ciągłych - czasopisma „Ostdeutsche Presse” z lat 1915-1921 o łącznej ilości skanów - 11 976,
2. Zadanie 2 - usługa rozpoznania tekstu w 5 225 publikacjach czasopisma „Ostdeutsche Presse” z lat 1890-1900 oraz 1915-1921 o łącznej ilości stron - 42 081.

Usługi polegają na:

* 1. Wykonaniu na sprzęcie Wykonawcy 11 976 plików skanów publikacji w formacie TIFF w siedzibie Wykonawcy. Wymagania jakościowe dotyczące plików TIFF zostały zawarte w Załączniku nr 1.
  2. Konwersji wykonanych plików TIFF do plików prezentacyjnych w formacie JPEG. Wymagania jakościowe dotyczące sposobu konwersji do formatu prezentacyjnego oraz wykonania plików JPEG zostały zawarte w Załączniku nr 2.
  3. Konwersji wykonanych oraz przekazanych Wykonawcy plików TIFF do plików prezentacyjnych w formacie PDF. Wymagania jakościowe dotyczące sposobu konwersji do formatu prezentacyjnego oraz wykonania plików PDF zostały zawarte w Załączniku nr 2.
  4. Rozpoznaniu tekstu na zdigitalizowanych stronach kolejnych publikacji drukowanego czcionką gotycką (szwabacha) w języku niemieckim uwzględniającego graficzno-logiczny układ tekstu, osadzeniu go w plikach prezentacyjnych PDF oraz zapisaniu pod postacią tekstowych plików XML zgodnie z wymogami standardu ALTO v. 4.0 lub nowszej. Wymagania jakościowe dotyczące sposobu wykonania rozpoznania tekstu opisano w Załączniku nr 3.
  5. Wykonaniu zestawów metadanych deskryptywnych opisujących każdą zdigitalizowaną publikację oraz odpowiadających im kompletów metadanych strukturalnych, administracyjno-technicznych i konserwatorskich opisujących pliki każdej strony każdej zdigitalizowanej publikacji. Wykonane metadane zostaną osadzone w plikach prezentacyjnych PDF, w plikach macierzystych TIFF oraz w plikach JPEG, jako metadane wewnętrzne, jak i zapisane w postaci tekstowych plików xml zgodnych z wymogami standardów METS, MODS, MARCxml, PREMIS, PAIS, MIX. Wymóg dokumentowania za pomocą metadanych zewnętrznych dotyczy również pliku miniatury i testowego pliku skanu. Wymagania jakościowe dotyczące sposobu wykonania metadanych zawarte są w Załączniku nr 4.
  6. Opracowaniu i wykonaniu zgodnie z wymogami języka XML plików wielojęzycznych metadanych deskryptywnych, umożliwiających w pełni zautomatyzowane publikowanie plików prezentacyjnych PDF wraz z tymi metadanymi w systemie biblioteki cyfrowej dLibra w wersji 6.
  7. Opracowaniu i wykonaniu – w oparciu o pliki skanów i powierzone pliki eksportu z katalogu OPAC – wielojęzycznych metadanych deskryptywnych w formacie MARC-21 i użycia ich podczas wykonywania plików metadanych zewnętrznych METS.
  8. Wykonaniu pliku zestawienia zbiorczego w formacie arkusza kalkulacyjnego. Sposób wykonania pliku zestawienia zbiorczego opisany jest w Załączniku nr 5.
  9. Wykonany przedmiot zamówienia zapisany zostanie przez Wykonawcę na nośniku udostępnionym przez Zamawiającego.
  10. Dodatkowo przedmiot zamówienia przeniesiony zostanie do repozytorium cyfrowego Biblioteki Narodowej w celu bezpiecznej archiwizacji publikacji. Przeniesienia dokona Zamawiający. Przedmiot zamówienia musi spełniać **Warunki przekazywania obiektów cyfrowych do BN** podane w niniejszym linku: <https://www.bn.org.pl/uslugi/centrum-kompetencji/warunki-przekazywania-obiektow-cyfrowych-do-bn>
  11. Odpowiedzialność oraz wszelkie koszty związane z dostarczeniem zbiorów oraz nośników cyfrowych do siedziby Wykonawcy w celu wykonania usługi digitalizacji oraz zwrot zbiorów i nośników cyfrowych do siedziby Zamawiającego po wykonaniu usługi ponosi Wykonawca
  12. Wykaz obiektów bibliotecznych przeznaczonych do zdigitalizowania zawiera Załącznik nr 6.
  13. Sposób zapisu wykonanych plików oraz pożądana struktura katalogów zawarte są w załączniku nr 7.

2. Wymagania organizacyjne:

1. Wykonanie skanów czasopisma w formacie TIFF musi się odbywać w siedzibie Wykonawcy i za pomocą sprzętu Wykonawcy.
2. Wydanie i zwrot zbiorów przeznaczonych do digitalizacji potwierdzone zostanie protokołem zdawczo-odbiorczym.
3. Przekazanie zbiorów do digitalizacji może nastąpić po podpisaniu umowy. Termin wydania przedmiotu zamówienia zostanie ustalony z Zamawiającym.
4. W czasie trwania usługi pełna odpowiedzialność (w tym finansowa) za przekazane zbiory spoczywa na Wykonawcy, przejmuje on odpowiedzialność za zbiory dziedzictwa kulturowego z chwilą podpisania protokołu odbioru od Zamawiającego na czas transportu oraz wykonania usługi. Wykonawca zobowiązany jest do zwrotu przekazanych zbiorów w stanie niepogorszonym.
5. Zamawiający udostępni nośniki cyfrowe z publikacjami zawierającymi pliki macierzyste TIFF czasopisma „Ostdeutsche Presse” z lat 1890-1900, na których następnie Wykonawca umieści przetworzone pliki cyfrowe. Odpowiedzialność oraz wszelkie koszty związane z dostarczeniem nośników cyfrowych do siedziby Wykonawcy w celu wykonania usługi digitalizacji oraz ich zwrotu po wykonaniu usługi do Zamawiającego ponosi Wykonawca.
6. Wykonawca przekaże Zamawiającemu:

* co najmniej 20% (tj. min. 2 395) gotowych plików macierzystych TIFF,
* co najmniej dwa kompletne obiekty - numery czasopisma (zawierające pliki TIFF, JPEG, PDF, ALTO, METS, plik testowy oraz miniaturę)

w celu weryfikacji **do 01.08.2025 r.** Dopuszcza się udostępnienie plików na serwerze Wykonawcy do pobrania lub dostarczenie ich na nośniku cyfrowym.

1. Wykonawca w celu zapoznania się z charakterem i zakresem zadania – tj. dokładnym stanem technicznym i ilością - może przeprowadzić wizję lokalną przedmiotu objętego zamówieniem. Dokonanie wizji możliwe będzie po uprzednim zgłoszeniu do osoby ze strony Zamawiającego upoważnionej do kontaktowania się z wykonawcami w kwestiach merytorycznych. Przeprowadzenie wizji jest możliwe - od poniedziałku do piątku w godz. 8:00 - 15:00.
2. Wykonawca winien najpóźniej w dniu podpisania umowy z Zamawiającym przedstawić ważną, aktualną polisę ubezpieczeniową. Opłacona polisa, a w przypadku jej braku inny dokumentu potwierdzający, że wykonawca jest ubezpieczony od odpowiedzialności cywilnej w zakresie prowadzonej działalności związanej z przedmiotem zamówienia. Suma gwarancyjna min. 100.000,00 na jedno i wszystkie zdarzenia ubezpieczeniowe.
3. Wykonawca może powierzyć wykonanie części niniejszego zamówienia podwykonawcom. W takim przypadku zobowiązany jest do wykazania w formularzu ofertowym części zamówienia, której wykonanie zamierza powierzyć podwykonawcom. Powierzenie wykonania części zamówienia podwykonawcom nie zwalnia wykonawcy z odpowiedzialności za należyte wykonanie tego zamówienia.

3. Wymagania dot. gwarancji:

Na przedmiot zamówienia Wykonawca udzieli min. 24 miesięcznej gwarancji. Bieg terminu gwarancji i rękojmi rozpoczyna się z dniem podpisania protokołu ostatecznego odbioru przedmiotu umowy lub od daty potwierdzenia usunięcia wad w razie ich stwierdzenia przy odbiorze końcowym przedmiotu umowy.

**4. Termin realizacji umowy:**

Wymagany termin wykonania przedmiotu zamówienia: **do dnia 23.09.2025 r.**

**Załączniki**

Załącznik nr 1 - Wymagania jakościowe dotyczące plików TIFF   
Załącznik nr 2 - Wymagania jakościowe dotyczące plików JPEG, PDF oraz pliku miniatury graficznej  
Załącznik nr 3 - Wymagania dotyczące wykonania rozpoznania tekstu oraz plików ALTO

Załącznik nr 4 - Metadane dotyczące wykonanych obiektów cyfrowych   
Załącznik nr 5 - Plik zestawienia zbiorczego wykonanych plików cyfrowych

Załącznik nr 10 - Wykaz obiektów bibliotecznych przeznaczonych do digitalizacji oraz stan zachowania   
Załącznik nr 11 - Nazewnictwo folderów dyskowych i plików cyfrowych

Załącznik nr 1

**Załącznik nr 1 - Wymagania jakościowe dotyczące plików TIFF**

1. Format plików macierzystych : TIFF, jednostronicowy.
2. Głębia barwy każdego pliku skanu : RGB 24 bitowy.
3. Optyczna rozdzielczość skanowania: 400 ppi. Przez powierzchnię skanowania rozumie się całkowitą powierzchnię, jaką zajmuje zeskanowana strona. Wykonawca, jeżeli uzna to za stosowne, przykładowo w celu osiągnięcia wymaganego poziomu rozpoznania OCR może dla wybranych stron wykonać skanowanie z rozdzielczościami wyższymi niż określone w niniejszym punkcie, a które to należy rozumieć, jako rozdzielczości minimalne. Dopuszczalnymi wartościami rozdzielczości optycznej skanowania są wyłącznie : 400, 450 i 600 ppi.
4. Kompresja pliku macierzystego **:** dopuszczalne jest zastosowanie kompresji bezstratnej LZW.
5. Wszystkie strony wchodzące w numerację, nawet puste, podobnie jak okładki oraz strony nienumerowane, muszą być zeskanowane.
6. Wymiary zeskanowanej strony w pliku TIFF (szerokość, wysokość) muszą być takie same jak wymiary strony obiektu papierowego. Przeskalowania wymiarów plików TIFF   
   są niedopuszczalne.
7. Współczynnik Gamma **:** wartość zalecana 2,10. Wykonawca, mając na uwadze, jakość wykonywanego skanu, może dokonywać zmian zgodnie z posiadaną wiedzą,   
   w przedziale wartości 1.60÷2.30.
8. Plik skanu powinien wiernie zachowywać kolorystykę oryginału papierowego (zapis skanów bez rozdzielania profilu wyświetlania na pole predefiniowane i treść skanu).
9. Plik skanu powinien być wykonany z jednolitym i równomiernym oświetleniem całej powierzchni skanowanej strony.
10. Jakiekolwiek inne operacje na plikach macierzystych są zabronione. Dotyczy to m.in. operacji zmiany rozdzielczości (downsampling/upsampling), redukcji głębi kolorów, zmiany jasności, manipulacji poziomem kontrastu.
11. Zawartość metadanych EXIF generowana podczas skanowania powinna pozostać   
    w plikach TIFF bez zmian. Ponadto, w każdym pliku TIFF należy zapisać wewnętrzne metadane uniwersalne, zgodnie z wymogami określonymi w Załączniku nr 8.
12. Pod skanowane obiekty biblioteczne (bez względu na rodzaj skanowanego obiektu) oraz pod każdą ze skanowanych stron należy podkładać beżową podkładkę tak,   
    by obszar marginesów w plikach skanów nie zawierał czarnych fragmentów stołu skanera.
13. Dla każdego woluminu obraz skanu każdej strony zawierającej przynajmniej jedno słowo - należy poziomować względem tekstu zawartego na tej stronie. Dla stron zawierających tekst w pionie oraz w poziomie, poziomowanie prowadzić należy względem tego kierunku,   
    w którym na stronie umieszczono większą część tekstu. Maksymalne dopuszczalne odchylenie wierszy tekstu od poziomu nie może przekroczyć 1,5°.
14. Marginesy wokół zeskanowanej strony :
    * co najwyżej 2.0 mm wokół obrzeży nie graniczących ze stroną sąsiadującą obiektu papierowego,
    * co najwyżej 2.0 mm wzdłuż obrzeża graniczącego ze stroną sąsiadującą obiektu papierowego.
    * Powyższy wymóg nie dotyczy skanów rozkładanych kart lub wklejonych   
      do obiektów map, które powinny być kadrowane do swojego własnego wymiaru**.**
15. Zawartością pliku skanu jest cała powierzchnia zeskanowanej strony oraz powierzchnia zajęta przez marginesy o rozmiarach określonych powyżej.
16. W przypadku obiektów graficznych (zdjęć, ilustracji, drzeworytów, map, itp.) zaprezentowanych w całości na dwóch sąsiadujących stronach, takie dwie strony należy zeskanować do pojedynczego pliku skanu.
17. Jeżeli treścią zeskanowanej strony jest w całości tekst lub tabela wartości odwrócona   
    o 90°, plik skanu należy zapisać na nośniku z rotacją o 90° w taki sposób,   
    aby przeglądarka wyświetlała wiersze tekstu poziomo i nie była wymagana – celem naturalnego zapoznawania się z treścią strony – operacja obrotu strony.
18. Jeżeli zagięcie strony powoduje w jakimkolwiek stopniu przesłanianie jej treści, wymagane jest niwelowanie tych zagięć przed wykonaniem skanowania. W przypadku stron, dla których operacja ta może spowodować trwałe ich uszkodzenie, Wykonawca - po potwierdzeniu tego faktu przez Zamawiającego - może odstąpić od jej realizacji.
19. Pliki macierzyste każdego zdigitalizowanego obiektu należy zapisać w oddzielnym folderze dyskowym. Sposób nazewnictwa folderów oraz plików TIFF opisano   
    w Załączniku nr 7.
20. Zachowanie oryginalnego układu (porządku) stron jest obligatoryjne i realizowane przez stosowne nazywanie kolejnych plików skanów.
21. Kopie plików TIFF muszą być przechowywane przez Wykonawcę do końca okresu gwarancyjnego i na wniosek Zamawiającego udostępniane mu w celu skopiowania   
    na nośniki będące własnością Zamawiającego.
22. Dla każdej publikacji zwartej należy wykonać plik skanu testowego, który zawierać będzie stronę tytułową pierwszej skanowanej publikacji (numeru czasopisma) oraz dołączony wzorzec kolorów (np. Adobe RGB 1998). Wzorzec kolorów należy umieścić na skanowanej stronie. Programy udostępniające składowe koloru wskazanego kursorem myszki, powinny w obszarze bieli i czerni w użytym wzorcu wskazywać wartości zawarte w przedziałach : dla bieli R 250÷255, G 250÷255, B 250÷255, a dla czerni R 0÷5, G 0÷5, B 0÷5. Publikacje, dla których plik skanu testowego nie potwierdzi spełnienia niniejszego wymogu, nie będą podlegać odbiorowi przez Zamawiającego. Sposób nazewnictwa i lokalizację zapisu plikówskanu testowego opisany jest w Załączniku nr 7.

Załącznik nr 2

**Załącznik nr 2 - Wymagania jakościowe dotyczące plików JPEG, PDF oraz pliku miniatury graficznej**

**Pliki w formacie prezentacyjnym JPEG:**

1. Rozmiar poszczególnych stron publikacji cyfrowej zapisanych w formacie JPEG (szerokość   
   i wysokość) musi być identyczny jak rozmiar plików TIFF, z których je wykonano.
2. Głębia barwy poszczególnych stron publikacji cyfrowej zapisanych w formacie JPEG musi być taka sama jak głębia barw plików TIFF, z których strony je wykonano.
3. Rozdzielczość poszczególnych stron publikacji cyfrowej zapisanych w formacie JPEG musi być taka sama jak rozdzielczość plików TIFF, z których je wykonano.
4. Pliki w formacie JPEG nie mogą zawierać żadnych dodatkowych treści (reklam, znaków wodnych identyfikujących Wykonawcę).
5. Konwersję plików TIFF do formatu JPG należy wykonać z kompresją 80%.
6. Pliki w formacie JPEG każdego zdigitalizowanego obiektu należy zapisać w oddzielnym folderze dyskowym. Sposób nazewnictwa folderów oraz plików JPEG opisano w Załączniku nr 7.
7. We wszystkich plikach JPEG należy zintegrować uniwersalne metadane wewnętrzne identyfikujące i opisujące publikację. Sposób wykonania - zapisywanych w plikach JPEG - metadanych opisany jest w Załączniku nr 8.

**Pliki w formacie prezentacyjnym PDF:**

1. Sposób zapisu plików pojedynczej publikacji : plik wielostronicowy zawierający wszystkie strony zdigitalizowanego obiektu.
2. Każda strona pliku PDF wykonana będzie w oparciu o segmentację zawartości   
   na warstwy wizualne : treści i tła oraz ukrytą warstwę rozpoznanego tekstu.
3. Rozmiar poszczególnych stron publikacji zapisanych w formacie PDF (szerokość   
   i wysokość) musi być identyczny jak rozmiar plików TIFF, z których je wykonano.
4. Głębia barwy poszczególnych stron publikacji zapisanych w formacie PDF musi być taka sama jak głębia barw plików TIFF, z których strony te wykonano**.**
5. Za należycie wykonaną stronę w formacie PDF a zawierającą wyłącznie tekst, uważa się stronę   
   o precyzji segmentacji przynajmniej 95% (spośród każdych 100 liter lub liczebników, co najwyżej 5 liter/liczebników znajduje się warstwie tła).
6. Publikacje w formacie PDF nie mogą zawierać żadnych dodatkowych treści (reklam, znaków wodnych identyfikujących Wykonawcę czy też ukrytych hiperłącz).
7. Publikacje w formacie PDF w podstawowej przeglądarce producenta powinny wyświetlać się   
   z powiększeniem „Dopasuj do okienka”.
8. Pojedyncza publikacja w formacie PDF zapisana będzie w oddzielnym folderze dyskowym. Sposób nazewnictwa folderów, pliku każdej publikacji opisany jest w Załączniku nr 7.
9. We wszystkich publikacjach PDF należy zintegrować uniwersalne metadane identyfikujące   
   i opisujące publikację. Sposób wykonania zapisywanych w publikacjach metadanych opisany jest w Załączniku nr 8.
10. Kopie plików PDF muszą być przechowywane przez Wykonawcę do końca okresu gwarancyjnego.

**Miniatura graficzna – wymogi techniczne wykonania.**

1. Dla każdego obiektu cyfrowego należy wykonać plik miniatury graficznej w formacie plików PNG,
2. Dłuższy bok miniatury będzie posiadać 300 pikseli,
3. Treścią miniatury będzie strona tytułowa obiektu, gdy obiekt nie posiada strony tytułowej, treścią miniatury będzie okładka lub pierwsza strona zawierająca tekst.
4. Pliki miniatur każdego zdigitalizowanego obiektu należy zapisać w oddzielnym folderze dyskowym. Sposób nazewnictwa folderów oraz plików PNG opisano w Załączniku nr 7.

Załącznik nr 3

**Załącznik nr 3 - Wymagania dotyczące wykonania rozpoznania tekstu OCR oraz plików ALTO**

1. Każda strona, każdego obiektu cyfrowego, zawierająca co najmniej jedno słowo drukowane, poddana będzie rozpoznaniu OCR,
2. Wymóg wykonania rozpoznania OCR nie dotyczy stron zawierających wyłącznie : grafiki, fotografie, pismo odręczne lub zapis nutowy.
3. Wymagana skuteczność prawidłowego rozpoznania OCR słów na pojedynczej stronie: przynajmniej 80%,
4. Niższy próg trafności rozpoznania OCR dopuszczony jest jedynie dla stron zawierających wzory   
   i wyrażenia matematyczne, fizyczne, chemiczne i/lub odwrócone o 90° względem poziomu zestawienia tabelaryczne,
5. Rozpoznanie OCR należy prowadzić wyłącznie z uwzględnieniem układu graficzno-logicznego strony (pełne treści kolejnych artykułów wydrukowanych w kolejnych obszarach/kolumnach strony) i w taki też sposób rozpoznany tekst musi być zapisany w plikach prezentacyjnych   
   w formacie PDF,
6. Rozpoznany tekst pojedynczej strony należy zachować jako prosty plik tekstowy zawierający wyłącznie rozpoznane na stronie słowa. W takich plikach tekstowych należy stosować kodowanie liternictwa UTF-8. Pliki tekstowe zapisane będą w folderze „TEXT”.
7. Ponadto, rozpoznany tekst strony należy zapisać jako tekstowy plik XML odpowiadający wymogom standardu ALTO w wersji 4.0 lub nowszej. Plik zapisany będzie w folderze „ALTO”,
8. Sposób nazewnictwa pliku dla każdej strony publikacji opisany jest w Załączniku nr 7.

**Wykonanie pliku ALTO dla rozpoznanego tekstu pojedynczej strony.**

Każdy wykonany plik ALTO musi zawierać :

* + - 1. Współrzędne kolejnych słów wyrażonych w jednostce inch/1200 lub w pikselach,
      2. Określenie z nazwy pliku obrazu, którego dotyczy rozpoznany tekst (element fileName),
      3. Określenie identyfikatora pliku obrazu, a nadanego w pliku METS (element fileIdentifier),

1. Określenie nadanego w pliku METS identyfikatora obiektu cyfrowego, którego dotyczą : w/w plik obrazu i plik ALTO (element documentIdentifier),
2. Określenie użytego oprogramowania (elementy contentGeneration oraz postOperation),
3. Wyznaczenie wartości elementów TopMargin, LeftMargin, RightMargin, BottomMargin,
4. Przypisanie – w oparciu o metadane deskryptywne - atrybutowi LANG w kontenerach „ComposedBlock”, „TextBlock” lub ew. „TextLine” kodu języka, którego dotyczy rozpoznany tekst.
5. Podział każdego słowa na sylaby należy zapisywać w atrybucie HYP, któremu jako wartość, należy przypisać znak myślnika, a nie numer kodu znaku podziału, co było znamiennym dla starszych wersji schematu ALTO,
6. Dla elementów string, których zawartością jest początkowa lub końcowa część słowa podzielonego na sylaby, należy przypisać odpowiednie wartości do atrybutów SUBS\_CONTENT oraz SUBS\_TYPE.

Załącznik nr 4

**Załącznik nr 8 - Metadane dotyczące wykonanych obiektów cyfrowych**

Dla każdego wykonanego obiektu cyfrowego należy opracować metadane deskryptywne (opisowe)   
na podstawie metadanych, które Zamawiający przekaże Wykonawcy w postaci tekstowego pliku metadanych zapisanych zgodnie ze standardem MARC21.

Dla każdej zdigitalizowanej strony, każdego wykonanego obiektu należy opracować metadane administracyjno-techniczne i konserwatorskie.

Na podstawie opracowanych metadanych, dla każdego obiektu cyfrowego należy wykonać pod postacią tekstowego pliku xml zestaw metadanych zewnętrznych odpowiadających standardom METS, MODS, MARCXML, MIX, PREMIS, ALTO. Ponadto, dla każdego obiektu cyfrowego wykonane będą pliki deskryptorów definiujące typy obiektów. Wymaganym jest, by pliki te wykonać zgodnie   
z wymogami standardu PAIS.

Plik metadanych METS musi zawierać lokalne odniesienia do plików binarnych zawartości cyfrowych (TIFF oraz JPEG) i rozpoznanego tekstu (ALTO) ze ścieżkami typu URN definiowanymi relatywnie względem położenia pliku METS.

Wszystkie używane w plikach metadanych METS przestrzenie nazw powinny być zaimportowane   
w nagłówku pliku METS oraz powinny posiadać dostępne w Internecie pliki schematów pozwalających na ich walidację.

W pliku METS musi wystąpić sekcja nagłówka informacyjnego <mets:metsHdr>.

W pliku METS musi wystąpić sekcja <mets:structMap … TYPE="physical"> definiująca strukturę fizyczną obiektu.

W pliku METS obiektu, dla którego wykonano rozpoznanie OCR, musi wystąpić sekcja <mets:structMap … TYPE="logical"> definiująca strukturę logiczną obiektu.

Dla każdego pliku TIFF, JPEG oraz pliku PDF wyznaczona będzie suma kontrolna MD5   
i wraz z jego rozmiarem oraz datą utworzenia umieszczone zostaną w pliku METS, w sekcji <mets:fileSec>.

W sekcji pliku METS <mets:amdSec> należy umieścić metadane techniczne zgodnie   
ze specyfikacją standardu MIX (jako sekcje <mets:techMD>) oraz metadane dotyczące proweniencji, zgodnie ze specyfikacją standardu PREMIS w sekcji <mets:digiprovMD>,

Zapis metadanych technicznych należy zoptymalizować umieszczając wartości wspólne dla wszystkich stron obiektu w jednej sekcji, a w sekcjach kolejnych (dla każdej strony) wartości nie powtarzające się. Oczekiwanymi są wartości złożonych metadanych: BasicDigitalObjectInformation, BasicImageInformation, ImageCaptureMetadata oraz ImageAssessmentMetadata.

Plik metadanych METS zawierać musi metadane opisujące status praw autorskich np. jako sekcja <mets:rightsMD> lub jako sekcja zgodna ze specyfikacją standardu PREMIS. Należy przyjąć,   
że wszystkie wykonane obiekty cyfrowe należą do domeny publicznej.

Metadane deskryptywne odnoszące się do obiektu źródłowego będą zawarte w sekcjach <mets:dmdSec> zgodnie z wymogami standardów MODS oraz MARCXML.

W pliku METS, w sekcji metadanych odpowiadających standardowi MODS oczekuje się wartości złożonych metadanych: titleInfo, name, typeOfResource, genre, originInfo, language, physicalDescription, note, subject, relatedItem, location oraz accessCondition.

W pliku METS, w sekcji metadanych odpowiadających standardowi MARCXML oczekuje się wszystkich wartości metadanych, które zostały przekazane Wykonawcy w postaci tekstowego pliku metadanych zapisanego zgodnie ze standardem MARC21.

W aplikacji pozwalającej na edycję plików xml, plik metadanych METS powinien wyświetlać się podobnie jak zaprezentowany poniżej fragment przykładowego pliku METS (sekcja metadanych deskryptywnych).

[<mets:dmdSec STATUS="**current**" ID="**DM\_MODS**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

[<mets:mdWrap LABEL="**Rekord metadanych deskryptywnych**" MDTYPE="**MODS**" MIMETYPE="**text/xml**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

[<mets:xmlData><mods:mods version="**3.5**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

[<mods:titleInfo displayLabel="**Tytuł publikacji**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:title>Głos Krajny, R: 1932, Nr 18.</mods:title>

<mods:subTitle>orędownik powiatowy.</mods:subTitle>

</mods:titleInfo>

[<mods:name type="**personal**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:namePart type="**family**">Malicki</mods:namePart>

<mods:namePart type="**given**">Wacław</mods:namePart>

<mods:namePart type="**date**">1886-1939</mods:namePart>

<mods:displayForm>Wacław Malicki (1886-1939). Wyd.</mods:displayForm>

[<mods:role>](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:roleTerm xml:lang="**pl**" type="**code**">Wyd.</mods:roleTerm>

<mods:roleTerm type="**text**" authority="**marcrelator**">Publisher</mods:roleTerm>

<mods:roleTerm type="**code**" authority="**marcrelator**">pbl</mods:roleTerm>

</mods:role>

</mods:name>

<mods:typeOfResource>text</mods:typeOfResource>

<mods:genre authority="**marcgt**">periodical</mods:genre>

<mods:genre xml:lang="**pl**" displayLabel="**Typ źródła**" authority="**local**">czasopismo</mods:genre>

[<mods:originInfo eventType="**publication**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

[<mods:place supplied="**yes**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:placeTerm type="**text**" authority="**marccountry**">Poland</mods:placeTerm>

<mods:placeTerm type="**code**" authority="**marccountry**">pl</mods:placeTerm>

</mods:place>

[<mods:place>](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:placeTerm xml:lang="**pl**" type="**text**">Wyrzysk</mods:placeTerm>

</mods:place>

[<mods:place>](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:placeTerm xml:lang="**pl**" type="**text**"

valueURI="**http://id.loc.gov/authorities/names/nr96019510.html**"

authorityURI="**http://lccn.loc.gov/nr96019510**">Nakło</mods:placeTerm>

</mods:place>

<mods:publisher>Wyd. Wacław Malicki.</mods:publisher>

<mods:dateOther encoding="**marc**">1932</mods:dateOther>

<mods:issuance>serial</mods:issuance>

</mods:originInfo>

[<mods:language>](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:languageTerm type="**text**" authority="**iso639-2b**"

valueURI="**http://id.loc.gov/vocabulary/iso639-2/pol**"

authorityURI="**http://id.loc.gov/vocabulary/iso639-2**">Polish</mods:languageTerm>

<mods:languageTerm type="**code**" authority="**rfc3066**">pl</mods:languageTerm>

</mods:language>

[<mods:physicalDescription displayLabel="**Opis fizyczny**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:form authority="**marcform**">print</mods:form>

<mods:internetMediaType>image/vnd.djvu</mods:internetMediaType>

<mods:internetMediaType>image/jpeg</mods:internetMediaType>

<mods:note type="**physical details**">31.9 x 47.8 cm.</mods:note>

<mods:extent unit="**pages**">4</mods:extent>

<mods:digitalOrigin>digitized other analog</mods:digitalOrigin>

</mods:physicalDescription>

<mods:note xml:lang="en" displayLabel="**Additional data**">Numbering: R.12 (1931) - R.20 (1939).</mods:note>

<mods:note xml:lang="**en**" displayLabel="**Additional data**">Frequency: Semiweekly</mods:note>

<mods:note xml:lang="**pl**" displayLabel="**Adnotacje**">Numeracja: R. 12 (1931) - R. 20 (1939).</mods:note>

<mods:note xml:lang="**pl**" displayLabel="**Adnotacje**">Częstotliwość: Dwa razy w tygodniu</mods:note>

[<mods:relatedItem><mods:titleInfo>](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:title>Stanowi kontynuację: Orędownik Powiatu Wyrzyskiego, (1929-1931)</mods:title>

</mods:titleInfo></mods:relatedItem>

[<mods:relatedItem displayLabel="**Sygnatura publikacji**" type="**original**"><mods:titleInfo>](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:title>R 829.1932 (2°)</mods:title>

</mods:titleInfo></mods:relatedItem>

[<mods:location>](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:url>[http://opac.ksiaznica.torun.pl/cgi-bin/wspd\_cgi.sh/wo\_ropis.p?IDRek=1282504289101</mods:url](http://opac.ksiaznica.torun.pl/cgi-bin/wspd_cgi.sh/wo_ropis.p?IDRek=1282504289101%3c/mods:url)>

</mods:location>

[<mods:location xml:lang="**en**" displayLabel="**Location of Original**">](file:///J:\WiMBP\_Kompilacja%2001\METS\R%20829.1932\Glos_Krajny_Oredownik_Powiatowy\1932\Nr018\Obiekt-WiMBP-018522.xml)

<mods:physicalLocation valueURI="**http://id.loc.gov/authorities/names/n96095765.html**" authorityURI= "**http://lccn.loc.gov/n96095765**">Wojewódzka i Miejska Biblioteka Publiczna (Bydgoszcz, Poland)

</mods:physicalLocation></mods:location>

<mods:accessCondition xml:lang="**en**">public domain</mods:accessCondition>

<mods:accessCondition xml:lang="**pl**">domena publiczna</mods:accessCondition>

<mods:accessCondition xml:lang="**pl**">Bez ograniczeń</mods:accessCondition>

</mods:mods>

</mets:xmlData>

</mets:mdWrap>

</mets:dmdSec>

We wszystkich plikach skanów oraz plikach w formacie PDF i JPEG należy zapisać metadane wewnętrzne. Dozwolonymi są wyłącznie atrybuty zgodne z wymogami standardów EXIF oraz XMP.  
  
Oczekiwane zestawy metadanych wewnętrznych w plikach TIFF oraz PDF przedstawiono poniżej   
w postaci przykładowych okienek aplikacji udostępniających zapoznawanie się metadanymi plików różnych typów.

Załącznik nr 5

**Załącznik nr 5 - Plik zestawienia zbiorczego wykonanych plików cyfrowych**

1. Plik zestawienia zbiorczego zawiera ilościowo-jakościowe zestawienie wszystkich wykonanych obiektów cyfrowych. Wykonany powinien być jako zestawienie tabelaryczne i zapisany   
   w formacie arkusza Excela lub Open Office’a**.** W zestawieniu zbiorczym winnyznaleźć się opisy wszystkich zeskanowanych i zdigitalizowanych obiektów bibliotecznych**.**
2. W arkuszu, dla każdego wykonanego obiektu cyfrowego, należy umieścić następujące informacje:

* Nadany obiektowi numer,
* Typy (rodzaje wykonanych plików cyfrowych),
* Nazwa pliku przedstawiającego pierwszą, ostatnią i tytułową stronę obiektu,
* Ilość zdigitalizowanych stron, plików skanów, skanów testowych, łącznej ilości plików TIFF, JPEG oraz plików PDF. Dla powyższych plików - pod zestawieniem tabelarycznym - należy podać ich zsumowane wartości łączne,
* Uśredniona wartość szerokości i wysokości stron obiektu (wklejone do obiektu mapy, wykroje posiadające znacząco inny rozmiar niż pozostałe strony, nie są brane pod uwagę podczas wyznaczania wartości średniej),
* Pełny tytuł, sygnatura, data wydania, miejsce wydania, wydawca, opis i język obiektu,
* Numer obiektu w obrębie rocznika oraz częstotliwość wydawania,
* Trzy wartości znamienne metadanym MODS – typ zasobu, Wydanie/Emisja, Rodzaj dokumentu.

1. Sposób nazewnictwa Pliku zestawienia zbiorczego opisany jest w Załączniku nr 7.
2. Przykładowy wygląd fragmentu arkusza zawierającego zestawienie zbiorcze, przedstawiono na kolejnej stronie.

Załącznik nr 6

**Załącznik nr 6 - Wykaz obiektów bibliotecznych przeznaczonych do digitalizacji oraz stan zachowania (dotyczy zadania nr 1)**

Zbiór przeznaczony do digitalizacji liczy 1 887 numerów czasopisma „Ostdeutsche Presse” z lat 1915-1921 o łącznej ilości 11 976 stron. Zasób składa się z 19 woluminów czasopism oprawionych   
w kwartały lub półrocza. Wymiary woluminów wraz z oprawami: szerokość: 350-380 mm, wysokość: 480-510 mm, głębokość: 40-80 mm. W skład wydawnictwa wchodzą też dodatki do numerów czasopism o wymiarach A4 i mniejszych.

Stan zachowania przeważającej części zbioru jest zadowalający. Zdarzają się jednak zagięcia, naderwania papieru bądź oderwane od całości luźne karty. Skanowanie wymaga pieczołowitości, dbałości i staranności, aby nie dopuścić do powstania uszkodzeń, mogących pojawić się z powodu braku ostrożności przy przewracaniu kart czy podczas rozkładania oprawnych woluminów.

W związku z głębokim wszyciem marginesu wewnętrznego, w celu osiągnięcia zadowalającej jakości skanów oraz dokładniejszego rozpoznania tekstu, sznury poszczególnych woluminów zbioru czasopism zostały rozcięte i poluzowane. Wymagana jest szczególna dokładność tak, aby nie dopuścić   
do wypadnięcia kart.

Wykaz woluminów:

|  |  |  |
| --- | --- | --- |
| Lp. | Rocznik | Kwartał |
| 1 | 1915 | 1 |
| 2 | 1915 | 2 |
| 3 | 1915 | 3 |
| 4 | 1915 | 4 |
| 5 | 1916 | 1-2 |
| 6 | 1916 | 3-4 |
| 7 | 1917 | 1-2 |
| 8 | 1917 | 3-4 |
| 9 | 1918 | 1 |
| 10 | 1918 | 2 |
| 11 | 1918 | 3-4 |
| 12 | 1919 | 1-2 |
| 13 | 1919 | 3 |
| 14 | 1919 | 4 |
| 15 | 1920 | 1 |
| 16 | 1920 | 2 |
| 17 | 1920 | 3 |
| 18 | 1920 | 4 |
| 19 | 1921 | 1 |

Załącznik nr 7

**Załącznik nr 7 - Nazewnictwo folderów dyskowych i plików cyfrowych**

Każdy typ wykonanych plików cyfrowych należy zapisać w folderach dyskowych o ściśle określonej strukturze:

1. Pliki TIFF zapisane będą w folderach zgodnie z formułą:

\Repozytorium cyfrowe\Rocznik\Numer\TIFF\

np. Repozytorium cyfrowe\1901\Nr001\TIFF\

1. Pliki JPEG zapisane będą w folderach zgodnie z formułą:

\Repozytorium cyfrowe\Rocznik\Numer\JPEG\

1. Pliki PDF zapisane będą w folderach zgodnie z formułą:

\Repozytorium cyfrowe\Rocznik\Numer\PDF\

1. Pliki ALTO zapisane będą w folderach zgodnie z formułą:

\Repozytorium cyfrowe\Rocznik\Numer\ALTO\

1. Pliki METS zapisane będą w folderach zgodnie z formułą:

\Repozytorium cyfrowe\Rocznik\Numer

1. Pliki skanów testowych zapisane będą zgodnie z formułą:

\Repozytorium cyfrowe\Rocznik\Numer\Tester\

1. Arkusz zestawienia zbiorczego zapisany będzie zgodnie z formułą:

\Repozytorium cyfrowe\

Dysk twardy z przekazywanymi plikami cyfrowymi może zawierać tylko i wyłącznie foldery określone wg powyższych formuł. Wymóg ten nie dotyczy folderów i/lub plików automatycznie utworzonych przez system operacyjny lub zapisanych fabrycznie na zewnętrznym dysku twardym przez jego producenta.

Wykonanym plikom cyfrowym należy nadać nazwy zgodnie z poniższymi warunkami :

1. Nazwa pliku składa się z części podstawowej wyrażonej za pomocą liczebnika złożonego   
   z siedmiu cyfr, wypełnionego od lewej strony zerami oraz – w przypadku plików jednostronicowych – liczebnika wskazującego kolejność obrazu tego pliku w dokumencie cyfrowym. Drugi liczebnik - współtworzący nazwę pliku - wyrażony jest za pomocą trzech cyfr   
   z wypełnieniem zerami od lewej strony. Należy poprzedzić go znakiem myślnika,
2. Liczebnik podstawowej części nazwy pliku, w zasobie archiwum cyfrowego Zamawiającego jest unikalnym numerem zwiększanym o jeden z każdym dodanym plikiem do archiwum cyfrowego,
3. W przypadku plików wielostronicowych, drugi liczebnik w nazwie pliku nie wystąpi,   
   a część podstawowa występować będzie pod nazwą „Publikacja” (.PDF) wraz z dodaniem liczebnika odpowiadającego numerowi obiektu.
4. Każdy plik miniatury graficznej obiektu posiadać będzie nazwę „Miniaturka.png”. Należy zapisać go w folderze poszczególnego numeru czasopisma.
5. Każdy plik zewnętrznych metadanych METS będzie posiadać nazwę opartą o formułę „Obiekt-WiMBP-NrObiektu.xml”, gdzie człon „NrObiektu” oznacza numer wykonanego obiektu cyfrowego wyrażonego za pomocą liczebnika złożonego z sześciu cyfr, wypełnionego od lewej strony zerami,
6. Każdy plik skanu testowego zawierającego próbnik kolorów posiadać będzie nazwę „Plik\_testowy\_Rocznik\_Oznaczenie woluminu.tif”, np. „Plik\_testowy\_1901\_01. Należy zapisać go w folderze „Tester”,
7. Rozszerzenia nazw plików wyrażone będą wyłącznie małymi literami alfabetu, np. .tif, .jpeg, .pdf, .png, .txt, .xml,
8. Inne rozszerzenia dla niektórych typów plików, np. .tiff, .jpeg nie będą użyte,
9. Nazwa (wartość) jaka nadana będzie pierwszemu zdigitalizowanemu plikowi oraz pierwszemu wykonanemu plikowi METS przekazane zostanie Zamawiającemu po rozpoczęciu realizacji przedmiotu zamówienia
10. Przykłady nazw plików odpowiadających powyższym wymogom Zamawiającego :

0005125-001.tif

0743237-001.jpeg

0743238-002.jpeg

0743239-003.xml

Obiekt-WiMBP-072405.xml

Publikacja072405.pdf

1. Arkusz zestawienia zbiorczego należy nazwać „Zestawienie zbiorcze.xlsx”. Jeżeli użyto starszej wersji MS Office lub pakietu Open Office, powyższe rozszerzenie nazwy arkusza należy odpowiednio zmienić.