

Załącznik 1.4.

Wymagania jakościowe dotyczące wykonania przedmiotu zamówienia.

Pliki w formacie prezentacyjnym PDF.

- Sposób zapisu plików pojedynczej publikacji w formacie PDF: plik wielostronicowy zawierający wszystkie strony zdigitalizowanego obiektu.
- Rozmiar poszczególnych stron publikacji zapisanych w formacie PDF (szerokość i wysokość) musi być identyczny jak rozmiar plików TIFF, z których je wykonano.
- Głębia barwy poszczególnych stron publikacji zapisanych w formacie PDF musi być taka sama jak głębia barw plików TIFF, z których strony te wykonano.
- Publikacje w formacie PDF nie mogą zawierać żadnych dodatkowych treści (reklam, znaków wodnych identyfikujących Wykonawcę czy też ukrytych hiperłączy).
- Publikacje w formacie PDF w podstawowej przeglądarce producenta powinny wyświetlać się z powiększeniem „Dopasuj do okienka”.
- Pojedyncza publikacja w formacie PDF zapisana będzie w oddzielnym folderze dyskowym. Sposób nazewnictwa folderów, pliku każdej publikacji opisany jest w załączniku 1.8.
- W przypadku pozycji współprawnych (tzw. „klocki”), każde dzieło samoistne wydawniczo tworzy pojedynczą publikację w formacie PDF. W przypadku obiektów, których konstrukcja rodzi wątpliwości, co do sposobu utworzenia PDF, kwestię tę jednorazowo należy uzgodnić z zamawiającym.
- Publikacje, które w treści mają wyodrębnione części i/lub rozdziały należy wyposażyć w odpowiednie zakładki i strukturę spisu treści. Sposób wykonania panelu ze spisem treści opisany jest w załączniku 1.6.
- Każda strona każdej publikacji poddana będzie rozpoznaniu OCR. Wyjątkiem są okładki lub strony publikacji zawierające tekst z użyciem wyłącznie czcionek ozdobnych, strony nie zawierające ani jednego znaku, strony zawierające wyłącznie zdjęcia, ilustracje, wykroje lub mapy i co najwyżej numer strony oraz strony zapisane pismem odręcznym.
- OCR wykonany będzie w technologii uwzględniającej separację obszarów obrazu (tekst, elementy graficzne) oraz wysoki stopień kompresji z zastosowaniem odpowiedniego stopnia kompresji dla każdego z obszarów pliku wejściowego oraz rozpoznaniem tekstu (OCR) dla tekstowych obszarów obrazu. Wykonawca przed realizacją przedstawi próbne pliki PDF wykonane dla wskazanej grupy materiałów celem akceptacji przez Zamawiającego.
- Wymagana skuteczność rozpoznania OCR: przynajmniej 84% przy gorszej jakości graficznej tekstu wydrukowanego czcionką łacińską, przynajmniej 90% przy jej dobrej jakości graficznej tekstu wydrukowanego czcionką łacińską oraz przynajmniej 80% w przypadku tekstu drukowanego czcionką gotycką.
- O klasyfikacji stron oryginału do kategorii „gorsza jakość” lub „dobra jakość” przesądza Zamawiający, uzależniając ją od następujących parametrów:
 - na stronie występują widoczne wypłowienia (zwłaszcza obszarów z tekstem),
 - strona posiada pofałdowania obniżające liniowość oraz równoległość kolejnych wierszy tekstu, a w których to wierszach rozpoznanie OCR jest niższe,
 - strona nosi ślady jej mechanicznych zgięć, a w których to miejscach rozpoznanie OCR jest znacznie niższe,
 - na stronie występują drobne braki, rozdarcia lub sklejenia ze stroną sąsiadującą,



- widoczne są braki fragmentów pojedynczych liter w tekście (niska jakość czcionki drukarskiej) - zarówno zeskanowanej strony, jak i w oryginale papierowym.

Ocena jakości stron oryginału obiektów leży po stronie Zamawiającego.

- Niższy próg trafności rozpoznania OCR dopuszczony jest jedynie dla stron zawierających wzory i wyrażenia matematyczne, fizyczne, chemiczne i/lub odwrócone o 90° zestawienia tabelaryczne oraz strony zawierające tekst zarówno czcionkami alfabetu łacińskiego jak i czcionkami cyrylicznymi, arabskimi, hebrajskimi i / lub greckimi.
- Rozpoznanie OCR należy prowadzić wyłącznie z uwzględnieniem układu graficznego kolejnych stron i w taki też sposób rozpoznany tekst należy zintegrować z publikacjami PDF (zapisać w poszczególnych plikach). Rozpoznawanie tekstu bez uwzględnienia szpał, kolumn lub innych wydzielonych na stronach regionów jest niedopuszczalne.
- We wszystkich publikacjach PDF należy zintegrować uniwersalne metadane identyfikujące i opisujące publikację. Sposób wykonania zapisywanych w publikacjach metadanych opisany jest w załączniku **1.5**.
- Pierwszą stroną każdej publikacji w formacie PDF będzie metryczka obiektu. Sposób wykonania metryczki opisany jest w załączniku **1.6**.
- Kopie plików PDF muszą być przechowywane przez Wykonawcę do końca okresu gwarancyjnego i na wniosek Zamawiającego udostępniane mu w celu skopiowania na nośniki będące własnością Zamawiającego.