



Warszawa, dn. 8 grudnia 2025 r.

PROTOKÓŁ ZE WSTĘPNYCH KONSULTACJI RYNKOWYCH

prowadzonych na podstawie art. 84 Ustawy Pzp

Zamawiający w dniu 16 października 2025 r. ogłosił na stronie internetowej **informację o zamiarze przeprowadzenia wstępnych konsultacji rynkowych w celu przygotowania postępowania o udzielenie zamówienia publicznego na system składający się z repozytorium oraz portalu on-line dla zasobów historii mówionej.**

Do udziału we wstępnych konsultacjach rynkowych zgłosił się podmiot: **Mobile MS Sp. z o.o., ul. Chochola 23, 91-230 Łódź** (dalej: Wykonawca). Wstępne konsultacje rynkowe odbyły się w dniu **20 listopada 2025 r.** o godz. 14:00 w formie spotkania on-line na platformie **MS Teams**.

W spotkaniu uczestniczyli:

1) ze strony Zamawiającego:

- Piotr Bastian
- Grzegorz Deniszczuk
- Justyna Podlasińska
- Karolina Tabak

2) ze strony Partnera Projektu – Centrum Archiwistyki Społecznej

- Ewa Majdecka
- Wojciech Budzisz

3) Wykonawcy:

- Łukasz Ściebior
- Jakub Jakubiak

Zamawiający na wstępie wyjaśnił, że wstępne konsultacje rynkowe służą przygotowaniu postępowania, którego przedmiotem będzie **dostarczenie systemu składającego się z repozytorium oraz portalu on-line dla zasobów historii mówionej**

Zamawiający podczas konsultacji chciał uzyskać informacje nt.:

- 1) Możliwości agregacji i integracji danych z bardzo różnych instytucji oraz techniczne sposoby łączenia, przesyłania i standaryzowania materiałów historii mówionej (w tym zgodność z Europeana, Kroniką itd.).
- 2) Narzędzi wspierające opracowanie materiałów – szczególnie automatyzacja transkrypcji, generowanie wstępnych metadanych, segmentacja treści i inne funkcje AI ułatwiające żmudną pracę instytucji kultury.
- 3) Rozwiązań ułatwiających użytkownikom dostęp i eksplorację treści – m.in. prezentacja na mapie, podział wywiadów na krótkie segmenty, różne widoki narracyjne, mechanizmy wyszukiwania i przeglądania dużych zbiorów.
- 4) Doświadczenia i kompetencje wykonawcy w zakresie dużych projektów historii mówionej, ich wizualizacji, projektowania interfejsów oraz pracy z dużymi zbiorami danych i AI – aby ocenić, na ile proponowane podejścia odpowiadają potrzebom planowanej platformy.

Dotyczy projektu pn. „Narratio. Cyfrowe repozytorium historii mówionej” realizowanego z dofinansowaniem z Funduszu Europejskiego na Rozwój Cyfrowy 2021-2027 w ramach działania 2.3 „Cyfrowa dostępność i ponowne wykorzystanie informacji”

Zamawiający podejmuje środki mające na celu zapobieżenie zakłócenia konkurencji z uwagi na udział Wykonawcy w konsultacjach polegające na zamieszczeniu protokołów ze wstępnych konsultacji rynkowych na stronie internetowej Zamawiającego.

Szczegółowy zakres konsultacji zawiera Załącznik nr 1 do Protokołu.

Na tym wstępne konsultacje rynkowe zostały zakończone.

Protokół opracowany przez:

- Karolina Tabak – Dział Projektów Cyfrowych

ZAŁĄCZNIK NR 1 DO PROTOKOŁU

INFORMACJE UZYSKANE PRZEZ ZAMAWIAJĄCEGO PODCZAS WSTĘPNYCH KONSULTACJI RYNKOWYCH W CELU PRZYGOTOWANIA POSTĘPOWANIA O UDZIELENIE ZAMÓWIENIA PUBLICZNEGO NA SYSTEM SKŁADAJĄCY SIĘ Z REPOZYTORIUM ORAZ PORTALU ON-LINE DLA ZASOBÓW HISTORII MÓWIONEJ

W ramach konsultacji omówiono zagadnienia dotyczące projektowania, tworzenia i utrzymania systemów do gromadzenia, opracowania i udostępniania zasobów historii mówionej. Wymiana informacji miała charakter ogólny i dotyczyła standardowych praktyk rynkowych, możliwych wariantów architektury, integracji, procesu opracowania, bezpieczeństwa oraz podejścia do projektowania doświadczeń użytkownika.

1. Architektura infrastrukturalna i przechowywanie danych

Omówiono założenia infrastruktury Zamawiającego, zakładające środowisko hybrydowe z silnym naciskiem na **on-premise**, w tym m.in. klastry wysokiej dostępności, macierze dyskowe oraz „chłodną archiwizację” (np. taśmy LTO). Zwrócono uwagę na potrzebę rozdzielania warstw systemu (repozytorium „master”/pliki wzorcowe vs. warstwa udostępnieniowa/online) oraz na wyzwania skalowania ruchu dla treści audio-wideo.

2. Skalowalność, CDN i testy wydajnościowe

W rozmowie wskazano, że typowy ruch może być umiarkowany, natomiast istotnym ryzykiem są **piki popularności** (np. w dniu uruchomienia lub kampanii komunikacyjnych). Omówiono podejście polegające na wykorzystaniu usług CDN/WAF (np. Cloudflare) oraz alternatywne warianty techniczne. Wykonawca rekomendował podejście oparte o **scenariusze i testy wydajnościowe** (obciążeniowe), a decyzję o uruchomieniu dodatkowych usług (np. CDN do streamingu) oprzeć o wyniki testów i założone parametry ruchu.

3. AI i wyszukiwanie semantyczne

Omówiono kierunek wykorzystania AI przede wszystkim do **inteligentnego wyszukiwania** (kontekstowego), a nie budowy chatbota. Wskazano możliwość realizacji semantycznego wyszukiwania w oparciu o:

- przetwarzanie wsadowe (np. nocne) „cięższych” operacji (analiza, embeddingi),
- przechowywanie reprezentacji wektorowych (np. w bazie wektorowej / rozszerzeniach typu pgvector),
- odpytywanie bazy wektorowej w czasie rzeczywistym w celu ograniczenia obciążenia infrastruktury.

Wykonawca potwierdził praktyczne doświadczenia z podobnymi mechanizmami (wektoryzacja słowników/hasłowania i dobór sugestii na podstawie podobieństwa).

4. Transkrypcja, modele mowy i podejście hybrydowe

Wskazano potrzebę automatyzacji transkrypcji z uwzględnieniem specyfiki materiałów historii mówionej (różna jakość nagrań, dykcja, szumy). Zarekomendowano przewidzenie etapu **pilotażu/testów porównawczych** (2–3 popularne rozwiązania) na realnych próbkach danych Zamawiającego. Omówiono możliwość zastosowania podejścia **hybrydowego**:

- przetwarzanie przez API dostawców komercyjnych w sytuacjach, gdy jest to dopuszczalne i opłacalne,
- przetwarzanie na infrastrukturze Zamawiającego (self-host) dla materiałów wrażliwych/prawnie ograniczonych, z możliwością automatycznego wyboru ścieżki (workflow) zależnie od oznaczeń/praw do materiału.

5. Integracje i standardy wymiany danych (w tym OSA, Europeana, Kronika)

Rozmawiano o integracji zasobów z systemem **OSA** oraz o potrzebie standaryzacji wymiany danych. Jako sensowny kierunek wskazano oparcie integracji o protokół **OAI-PMH** (zarówno dla eksportu, jak i importu), co może ułatwić:

- integrację z Europeana (w uzgodnionym zakresie – w szczególności metadane opisowe, z linkowaniem do źródła),
- integrację z innymi agregatorami i systemami dziedzinowymi,
- ograniczenie integracji „szytych na miarę” dla każdego systemu osobno.

Wskazano również wątek potencjalnej integracji z rozwiązaniami bibliotecznymi (np. dLibra) – do weryfikacji technicznej po stronie interesariuszy.

6. Dwukierunkowa synchronizacja: logika, walidacja, aktualizacje

Omówiono założenie **dwukierunkowej synchronizacji** między repozytorium/agregatorem a OSA. W rozmowie wskazano, że:

- inicjowanie synchronizacji może następować poprzez **akcję użytkownika** w OSA (zgłaszanie zestawów/obiektów do synchronizacji),
- walidacja może obejmować automatyczne sprawdzenia techniczne (formaty, kompletność, zgodność pól), natomiast weryfikacja merytoryczna nie musi być elementem obowiązkowym,
- należy przewidzieć sposób obsługi aktualizacji po synchronizacji (spójność danych, oznaczanie nieaktualnych pochodnych – np. transkrypcji – w przypadku podmiany pliku źródłowego).

7. Pliki wzorcowe, bezpieczeństwo i mechanizmy typu WORM/WARM

Dyskutowano kwestię przechowywania **dużych plików wzorcowych** oraz ryzyka ich przypadkowej lub złośliwej podmiany. Wskazano na zasadność wdrożenia mechanizmów klasy **WORM/WARM** (write once / ograniczona edycja), np. z:

- „oknem czasowym” pozwalającym na korekty (np. 1–2 dni),
- późniejszym wymogiem akceptacji administratora dla zmian, co zmniejsza ryzyko operacyjne i wzmacnia kontrolę nad zasobem archiwalnym.

8. Segmentacja, prezentacja treści i funkcje portalu on-line

W części dotyczącej udostępniania i UX wskazano przykładowe kierunki rozwiązań:

- **ścieżki tematyczne/narracyjne** budowane na bazie obiektów, osób, miejsc, wątków,
- **kolekcje** (kuratorские oraz potencjalnie użytkowników – z moderacją przed publikacją),
- widoki oparte o byty (osoby, miejsca) i ich powiązania z treściami,
- prezentacja geograficzna (mapy, filtry geoprzestrzenne), szczególnie przy materiałach, gdzie występuje wiele nazw geograficznych,
- wyszukiwanie w metadanych i w treści (w tym w transkrypcjach), jako rozszerzenie klasycznych filtrów katalogowych,
- uwzględnienie obiektów towarzyszących (fotografie, dokumenty) jako kontekstu do nagrań.

9. Projektowanie UX/UI i model współpracy – iteracyjność i badania użytkowników

Wykonawca rekomendował, aby przy projektach tej skali i nowatorstwa w Polsce:

- zapewnić **badania z użytkownikami** co najmniej na etapie prototypowania/makiet oraz kolejną iterację po wdrożeniu lub po zasileniu treścią,
- unikać „sztywnych”, kaskadowych zapisów ograniczających możliwość iteracji,
- przewidzieć elastyczność na korekty wynikające z badań (w tym budżet i tryb wprowadzania zmian),
- rozważyć współpracę z podmiotem specjalizującym się w UX (jako uzupełnienie kompetencji wykonawcy software’owego),
- oprzeć testy na zdefiniowanych personach i zróżnicowanych grupach docelowych.

10. Uwagi wdrożeniowe i organizacyjne

Zwrócono uwagę, że decyzje dotyczące doboru narzędzi (AI, CDN, modele transkrypcji) są silnie zależne od szybko zmieniającego się rynku, dlatego racjonalne jest:

- opisywanie wymagań w sposób funkcjonalny (co system ma osiągnąć),
- przewidywanie etapów pilotażowych/porównawczych,
- projektowanie architektury umożliwiającej wymianę komponentów bez przebudowy całości systemu.