

# **POLITECHNIKA BIAŁOSTOCKA**

## **WYDZIAŁ INFORMATYKI**

**KATEDRA: Mediów Cyfrowych i Grafiki Komputerowej**

## **PRACA DYPLOMOWA INŻYNIERSKA**

**TEMAT: Opracowanie portalu internetowego do oceny jakości dźwięku**

**WYKONAWCA: Marcin Milewski**

**PODPIS: .....**

**PROMOTOR: dr hab. inż. Sławomir Zieliński**

**PODPIS: .....**

**BIAŁYSTOK 2017**

### Karta dyplomowa

<b>POLITECHNIKA BIAŁOSTOCKA</b>  Wydział Informatyki  Katedra Mediów Cyfrowych i Grafiki Komputerowej	Studia stacjonarne <hr style="border: 0; border-top: 1px dotted black;"/>  I stopnia inżynierskie <hr style="border: 0; border-top: 1px dotted black;"/>	Nr albumu studenta 93582
		Rok akademicki 2016/2017
		Kierunek studiów informatyka  Specjalność -
<b>Marcin Milewski</b>		
<b>Temat pracy dyplomowej:</b>  <div style="text-align: center; font-weight: bold;"> Opracowanie portalu internetowego do oceny jakości dźwięku </div>		
<b>Zakres pracy :</b>  <ol style="list-style-type: none"> <li>1. Opracowanie przeglądu literatury dotyczącego narzędzi do oceny jakości dźwięku i obrazu</li> <li>2. Wybór metody oceny jakości dźwięku</li> <li>3. Opracowanie portalu internetowego</li> <li>4. Wykorzystanie opracowanego portalu do testów oceniających jakość dźwięku</li> <li>5. Udokumentowanie uzyskanych wyników</li> </ol>		
<b>Słowa kluczowe:</b> opracowanie portalu internetowego		
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> dr hab. inż. Sławomir Zieliński    <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">podpis promotora</p> </div> <div style="width: 45%;"> dr hab. Zenon Sosnowski    <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">podpis kierownika katedry</p> </div> </div>		
<div style="display: flex; justify-content: space-between;"> <div style="width: 30%;">   31.01.2016  <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">Data wydania tematu pracy dyplomowej - podpis promotora</p> </div> <div style="width: 30%;"> 28.02.2017  <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">Regulaminowy termin złożenia pracy dyplomowej</p> </div> <div style="width: 30%;"> <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">Data złożenia pracy dyplomowej - potwierdzenie dziekanatu</p> </div> </div>		
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">Ocena promotora</p> </div> <div style="width: 45%;"> <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">Podpis promotora</p> </div> </div>		
<div style="display: flex; justify-content: space-between;"> <div style="width: 30%;"> <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">Imię i nazwisko, stopień/ tytuł recenzenta</p> </div> <div style="width: 30%;"> <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">Ocena recenzenta</p> </div> <div style="width: 30%;"> <hr style="border: 0; border-top: 1px dotted black;"/> <p style="text-align: center; font-size: small;">Podpis recenzenta</p> </div> </div>		

**Thesis topic:** Development of an Internet portal for audio quality assessment.

## SUMMARY

The aim of this study was the development of an Internet portal for audio quality assessment. The web-based application enables users to create auditory tests based on the MUSHRA and ABX methods. Application allows to perform remote auditory tests, export obtained results to CSV or JSON files and to perform on demand some basic statistics checks.

The purpose and scope of this study are presented in the first chapter. The second chapter contains a review of multimedia quality assessment methodologies and their classification. The third and fourth chapters are concerned with the multimedia objective and subjective quality assessment methods respectively, in particular those used for audio and image evaluation.

In the fifth chapter, the existing Internet-based tools for audio and image quality assessment are reviewed. Moreover, this chapter contains considerations and rationale for the selection of the implemented test method, with the concluding outcome presented at the end of the chapter.

In the sixth chapter, the process of the development of the web-based application is described. The functional and extra-functional requirements, the use case diagram, the system architecture, the authorization and authentication details and the data structures are presented. The technical specifications for the audio playback and the implemented user interfaces for the MUSHRA and ABX Test methods are shown.

The seventh chapter contains the experimental details of the remotely performed auditory test based on the MUSHRA method, using the developed Internet portal. In this chapter the test material, the conditions and the instructions for the participants are presented.

In the last chapter the results of the performed test are discussed. The procedure and criteria of data screening are described. As a final outcome of the study, a comparison of the results between the remotely and laboratory performed auditory test is shown.

Marcin Milewski  
imię i nazwisko studenta  
93582  
nr albumu  
Informatyka, studia stacjonarne  
kierunek i forma studiów  
dr hab. inż. Sławomir Zieliński  
promotor pracy dyplomowej

Białystok, dnia 23.02.2017r.

#### OŚWIADCZENIE

Przedkładając w roku akademickim 2016/2017 Promotorowi Panu Sławomirowi Zielińskiemu pracę dyplomową pt.: „Opracowanie portalu internetowego do oceny jakości dźwięku”, dalej zwaną pracą dyplomową,

**oświadczam, że:**

- 1) praca dyplomowa stanowi wynik samodzielnej pracy twórczej,
- 2) wykorzystując w pracy dyplomowej materiały źródłowe, w tym w szczególności: monografie, artykuły naukowe, zestawienia zawierające wyniki badań (opublikowane, jak i nieopublikowane), materiały ze stron internetowych, w przypisach wskazywałem/am ich autora, tytuł, miejsce i rok publikacji oraz stronę, z której pochodzą powoływane fragmenty, ponadto w pracy dyplomowej zamieściłem/am bibliografię,
- 3) praca dyplomowa nie zawiera żadnych danych, informacji i materiałów, których publikacja nie jest prawnie dozwolona,
- 4) praca dyplomowa dotychczas nie stanowiła podstawy nadania tytułu zawodowego, stopnia naukowego, tytułu naukowego oraz uzyskania innych kwalifikacji,
- 5) treść pracy dyplomowej przekazanej do dziekanatu Wydziału Informatyki jest jednakowa w wersji drukowanej oraz w formie elektronicznej,
- 6) jestem świadomy/a, że naruszenie praw autorskich podlega odpowiedzialności na podstawie przepisów ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t. j.: Dz.U. z 2006 r. Nr 90, poz. 631 z późn. zm.), jednocześnie na podstawie przepisów ustawy z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym (t. j.: Dz. U. z 2012 r. poz. 572 z późn. zm.) stanowi przesłankę wszczęcia postępowania dyscyplinarnego oraz stwierdzenia nieważności postępowania w sprawie nadania tytułu zawodowego,
- 7) udzielam Politechnice Białostockiej nieodpłatnie licencji na korzystanie z pracy dyplomowej w celu realizacji przeprowadzenia procedury antyplagiatowej przyjętej w Uczelni oraz na przekazanie pracy do Ogólnopolskiego Repozytorium Prac Dyplomowych, jak również udostępnianie i przechowywanie jej w Bibliotece Politechniki Białostockiej przez okres 50 lat od obrony pracy dyplomowej.

  
czytelny podpis studenta

Na podstawie art. 24 ust. 1 ustawy z dnia 29 sierpnia 1997 roku o ochronie danych osobowych (j.t. z 2014 r., poz. 1182 z późn. zm.) informuję, że administratorem danych jest Politechnika Białostocka, ul. Wiejska 45A, 15-351 Białystok. Dane będą przetwarzane w celach realizacji procedury antyplagiatowej przyjętej w Politechnice Białostockiej i nie będą udostępniane odbiorcom danych w rozumieniu art. 7 pkt 6 ustawy o ochronie danych osobowych. Osobie, której dane dotyczą, przysługuje prawo dostępu do treści swoich danych oraz ich poprawiania. Podanie danych jest obowiązkowe (art. 167b ustawy z dnia 27 lipca 2005 roku Prawo o szkolnictwie wyższym j.t. Dz.U. z 2012 r., poz. 572 z późn. zm.).

## SPIS TREŚCI

<b>Wykaz skrótów.....</b>	<b>7</b>
<b>1. Wstęp.....</b>	<b>8</b>
<b>2. Ocena jakości treści multimedialnych.....</b>	<b>10</b>
<b>3. Obiektywne metody oceny jakości treści multimedialnych .....</b>	<b>15</b>
3.1. Podstawowe bloki obliczeniowe metryk opartych na ludzkim systemie postrzegania .....	16
3.1.1. Dekompozycja sygnału .....	16
3.1.2. Detekcja cech wspólnych i artefaktów .....	17
3.1.3. Wykrycie poziomu zauważalnej zmiany w postrzeganiu (JND) .....	17
3.1.4. Modelowanie uwagi wizualnej (VA) .....	18
3.2. Obiektywne metody oceny jakości obrazu i wideo .....	18
3.2.1. ITU-R BT.1683 .....	18
3.2.2. ITU-R BT.1907 .....	21
3.3. Obiektywne metody oceny jakości dźwięku .....	22
3.3.1. ITU-R BS.1387 (PEAQ) .....	22
3.3.2. ITU-T P.862 i ITU-T P.862.2 .....	23
3.3.3. ITU-T P.863 (POLQA) .....	23
<b>4. Subiektywne metody oceny jakości treści multimedialnych .....</b>	<b>25</b>
4.1. Ocena jakości obrazu .....	25
4.2. Ocena jakości dźwięku .....	28
4.2.1. ITU-R BS.1116 .....	28
4.2.2. ITU-R BS.1534 (MUSHRA) .....	29
4.2.3. Test ABX .....	31
<b>5. Internetowe narzędzia do oceny jakości dźwięku i obrazu oraz wybór implementowanej metody testowej.....</b>	<b>32</b>
5.1. Wykaz istniejących narzędzi .....	32
5.2. Wybór implementowanej metody testowej .....	34
<b>6. Opracowanie portalu internetowego do oceny jakości dźwięku.....</b>	<b>36</b>
6.1. Wymagania funkcjonalne i нефункционалне (pozaфункционалне) .....	36
6.2. Diagram przypadków użycia .....	38
6.3. Architektura aplikacji .....	41
6.4. Wykorzystane platformy programistyczne i narzędzia pracy .....	43

6.4.1. Platformy programistyczne .....	43
6.4.2. Narzędzia .....	44
6.5. Proces implementacji.....	44
6.5.1. Struktura bazy danych i opis struktur danych.....	44
6.5.2. Proces rejestracji użytkownika, uwierzytelniania i autoryzacji .....	49
6.5.3. Kreator testów odsłuchowych.....	50
6.5.4. Odtwarzacz testów odsłuchowych .....	51
6.5.5. Menadżer testów .....	56
6.6. Testy poprawności implementacji .....	56
<b>7. Wykorzystanie opracowanego portalu do testów oceniających jakość dźwięku ....</b>	<b>58</b>
7.1. Użyte próbki dźwiękowe .....	58
7.2 Cechy testu i instrukcja użytkownika .....	59
<b>8. Udokumentowanie uzyskanych wyników .....</b>	<b>60</b>
<b>9. Podsumowanie .....</b>	<b>69</b>
<b>Literatura .....</b>	<b>71</b>
<b>Spis rysunków .....</b>	<b>77</b>
<b>Spis tabel.....</b>	<b>78</b>
<b>Spis kodów źródłowych .....</b>	<b>79</b>
<b>Wykaz załączników.....</b>	<b>80</b>
<b>Załącznik 1 Instrukcja użytkownika opracowanego portalu internetowego .....</b>	<b>Z1.1</b>
<b>Załącznik 2 Instrukcja instalacji opracowanego portalu internetowego .....</b>	<b>Z2.1</b>

## WYKAZ SKRÓTÓW

SNR	– (z ang. <i>Signal-to-Noise Ratio</i> ) stosunek sygnału do szumu
PSNR	– (z ang. <i>Peak Signal-to-Noise Ratio</i> ) szczytowy stosunek sygnału do szumu
THD	– (z ang. <i>Total Harmonic Distortion</i> ) współczynnik zawartości harmonicznych
ITU	– (z ang. <i>International Telecommunication Union</i> ) Międzynarodowy Związek Telekomunikacyjny
HVS	– (z ang. <i>Human Visual System</i> ) ludzki układ wzrokowy
HAS	– (z ang. <i>Human Auditory System</i> ) ludzki układ słuchowy
MSE	– (z ang. <i>Mean Squared Error</i> ) błąd średniokwadratowy
MAE	– (z ang. <i>Mean Absolute Error</i> ) średni błąd bezwzględny
PDM	– (z ang. <i>Perception-Driven Metrics</i> ) metryki oparte o modele psychofizyczne ludzkiego systemu postrzegania
SD	– (z ang. <i>Signal-Driven</i> ) metryki oparte o aspekty sygnału z uwzględnieniem ludzkiego systemu postrzegania
VA	– (z ang. <i>Visual Attention</i> ) uwaga wizualna
JND	– (z ang. <i>Just-Noticeable Difference</i> ) poziom zauważalnej zmiany w postrzeganiu
FR	– (z ang. <i>Full Reference</i> ) pełen wzorzec odniesienia
NR	– (z ang. <i>No Reference</i> ) brak wzorca odniesienia
RR	– (z ang. <i>Reduced Reference</i> ) zredukowany wzorzec odniesienia
CSF	– (z ang. <i>Contrast Sensitivity Function</i> ) funkcja czułości kontrastu
VQM	– (z ang. <i>Video Quality Metric</i> ) całkowita miara zniekształcenia wideo
HDTV	– (z ang. <i>High Definition Television</i> ) telewizja wysokiej rozdzielczości
FFT	– (z ang. <i>Fast Fourier Transform</i> ) szybka Transformacja Fouriera
MOV	– (z ang. <i>Model Output Variable</i> ) wyjściowa zmienna modelu
IIR	– (z ang. <i>Infinite Impulse Response</i> ) nieskończona odpowiedź impulsowa
NB	– (z ang. <i>Narrowband</i> ) sygnał wąskopasmowy
SWB	– (z ang. <i>Super-wideband</i> ) sygnał super-szerokopasmowy
MOS	– (z ang. <i>Mean Opinion Score</i> ) średnia wartość ocen
IQR	– (z ang. <i>Interquartile range</i> ) rozstęp ćwiartkowy

## 1. Wstęp

Jakość dźwięku jest jednym z najważniejszych czynników uwzględnianych przy tworzeniu dźwiękowych lub multimedialnych urządzeń i systemów nadawczych, czego przykładami są między innymi internetowe serwisy muzyczne, radio cyfrowe (Digital Audio Broadcasting, Digital Radio Mondiale), telefony komórkowe, a także algorytmy kompresji dźwięku [10].

Do oceny jakości dźwięku stosowane są metody obiektywne, jak i subiektywne – opierające się na indywidualnym wrażeniu słuchacza. Klasyczne metody obiektywne, takie jak pomiar stosunku sygnału do szumu (SNR) lub współczynnika zawartości harmonicznego (THD), nie są zbyt wiarygodne do oceny postrzeganego przez słuchacza dźwięku. Problem niskiej wiarygodności staje się jeszcze bardziej widoczny, gdy metody te stosowane są do oceny nowoczesnych kodeków, które są zarówno nieliniowe i niestacjonarne [10]. Standardowe metody oceny obiektywnej, zalecane przez Międzynarodowy Związek Telekomunikacyjny (ITU), opisane są między innymi w rekomendacjach ITU-R BS.1387 (do oceny jakości fonii) [25] i ITU-T P.863 (do oceny jakości mowy) [32].

Metody subiektywne niwelują wadę metod obiektywnych, jaką jest niezdolność do stwierdzenia, co faktycznie odczuwa końcowy słuchacz [22]. Przykładami subiektywnych metod ustandaryzowanych przez ITU są: ITU-R BS.1116 – zalecana do oceny systemów, w których zaburzenia jakości dźwięku są bardzo małe [28], a także ITU-R BS.1534-3 (MUSHRA) – zalecana do oceny jakości dźwięków, pomiędzy którymi istnieją średniej wielkości zaburzenia [26]. Do oceny, czy pomiędzy dwoma dźwiękami zachodzą różnice, wykorzystany może być test ABX [53].

### Cel i zakres pracy

Celem niniejszej pracy jest opracowanie portalu internetowego umożliwiającego zdalne przeprowadzenie subiektywnych testów odsłuchowych. Portal ten ma umożliwić użytkownikowi tworzenie odsłuchowych testów subiektywnych, z wykorzystaniem własnoręcznie spreparowanych próbek dźwiękowych i w oparciu o wybrane metody. Portal ma umożliwić również wzięcie udziału w testach innych użytkowników,



sprawdzenie podstawowych statystyk dotyczących stworzonych testów i wyeksportowanie uzyskanych rezultatów do pliku.

Zakres pracy obejmował:

1. opracowanie przeglądu literatury dotyczącego narzędzi do oceny jakości dźwięku i obrazu,
2. wybór metody oceny jakości dźwięku,
3. opracowanie portalu internetowego,
4. wykorzystanie opracowanego portalu do testów oceniających jakość dźwięku,
5. udokumentowanie uzyskanych wyników.

## 2. Ocena jakości treści multimedialnych

Wraz z ewolucją systemów multimedialnych, wymagania użytkowników nie ograniczają się tylko do samych dostępności usług. Użytkownicy co raz bardziej oczekują, aby dostarczana usługa spełniała ich wymagania związane z jakością treści. Systemy audiowizualne stają się także co raz bardziej złożone i pozwalają na nowe możliwości prezentowania treści, czego przykładem może być rzeczywistość rozszerzona, a także środowiska immersyjne [50].

Ocena jakości jest jednym z ważnych problemów w dziedzinie przetwarzania multimedii. Na jakość treści multimedialnych ma wpływ wiele czynników, takich jak: pozyskiwanie samej treści, przetwarzanie, kompresja, transmisja, interfejs końcowy, dekodowanie itp. Stopień zniekształcenia jakości treści multimedialnych postrzegany przez użytkownika zależy również od takich czynników jak np. indywidualne zainteresowania, oczekiwania co do jakości i własnego doświadczenia, typ interfejsu końcowego oraz jego właściwości. Należy zauważyć, że w przypadku treści audiowizualnej, jakość obrazu może wpłynąć na postrzeganą ocenę jakości dźwięku i vice-versa [12, 46].

Metodologie oceny jakości treści multimedialnych możemy podzielić na subiektywne – opierające się na indywidualnych wrażeniach użytkownika końcowego i obiektywne – opierające się np. na modelach przybliżających ludzkie postrzeganie. Jako że człowiek jest ostatecznym odbiorcą i interpretatorem otrzymywanych treści, ocena jakości metodą subiektywną jest najbardziej dokładna. Ocena taka służy też jako punkt odniesienia dla obiektywnych metod oceny jakości. Przeprowadzanie eksperymentów badania jakości metodami subiektywnymi wymaga od grupy ludzi słuchania lub oglądania materiału testowego, a następnie poddania go ocenie. Pomimo dokładnych wyników, wymagają one znacznego wysiłku i czasu, a także kosztu finansowego [12].

Ocena jakości metodami obiektywnymi w porównaniu z metodami subiektywnymi, nie wymaga tak znacznych wysiłków, ponieważ eliminuje konieczność zaangażowania człowieka jako osoby oceniającej. Metody te są lepsze pod względem powtarzalności, nie pozwalają jednak stwierdzić, co faktycznie odczuwa końcowy słuchacz. Ocena subiektywna jest wciąż wymagana do budowy i nauki algorytmów oceny metod obiektywnych, jest także często stosowana do weryfikacji ich dokładności [12, 22].

Zestawienie wybranych subiektywnych i obiektywnych metod oceny jakości treści multimedialnych, standaryzowanych przez ITU, a także obliczane przy ich użyciu metryki, przedstawiono kolejno w Tabeli 2.1 i Tabeli 2.2 [81].

**Tabela 2.1.** Zestawienie wybranych, standaryzowanych przez ITU metod subiektywnej oceny treści multimedialnych [81]

<b>Zastosowanie</b>	<b>Standard</b>	<b>Rekomendowani uczestnicy</b>	<b>Metryka**</b>
Mowa	ITU-T P.800 ITU-T P.830	Niespecjaliści	MOS DMOS
Fonia (ang. <i>Audio</i> )	ITU-R BS.1116	Specjaliści	Wartości różnic między ocenianymi próbkami
	ITU-R BS.1534 (MUSHRA)	Specjaliści	Wartości średnie każdej z ocenianych próbek
Obraz	ITU-R BT.500	Specjaliści i niespecjaliści	Różnica ocen punktów pomiędzy warunkami wzorcowymi a warunkami testowymi
Multimedia*	ITU-T P.910 ITU-R BT.1788	Niespecjaliści Niespecjaliści	DV Wartości średnie ocen

\* Multimedia – definiowane jako połączenie różnych form mediów, takich jak fonia, wideo, tekst, obraz, faks i telefonia [33].

\*\* Metryka – system powiązanych ze sobą miar, ułatwiający ujęcie ilościowe jakiejś cechy [65].

**Tabela 2.2.** Zestawienie wybranych, standaryzowanych przez ITU metod obiektywnej oceny treści multimedialnych [81]

Zastosowanie	Standard	Metryka
Mowa	ITU-T P.862 (PESQ)	MOS-LQO
	ITU-T P.862.2 (Szerokopasmowa)	MOS-LQO*
	ITU-T P.863 (POLQA)	MOS-LQOn MOS-LQQ <sub>sw</sub>
Fonia	ITU-R BS.1387 (PEAQ)	ODG
Obraz	ITU-R BT.1683 ITU-R BT.1907	VQM MOS Q
Multimedia	ITU-T J.247	OMOS DMOS <sub>p</sub> VQM

\* Inna niż w standardzie ITU-T P.862.

Klasyfikację metod obiektywnych przeprowadzić można ze względu na sposób obliczania metryki jakości. Metryki te sklasyfikować można następująco [12]:

1. tradycyjne – oparte na samej dokładności sygnału,
2. oparte w różnym stopniu na ludzkim systemie postrzegania.

Do grupy metryk tradycyjnych zalicza się metryki takie jak błąd średniokwadratowy (MSE), średni błąd bezwzględny (MAE), stosunek sygnału do szumu (SNR), czy szczytowy stosunek sygnału do szumu (PSNR). Metryki te są szeroko stosowane do oceny jakości lub zniekształceń sygnału wizyjnego. Są one jednak krytykowane za słaby związek z faktycznie postrzeganą jakością obrazu przez użytkownika. Wynika to z następujących powodów [12, 74]:

1. nie każda zmiana lub zniekształcenie sygnału jest zauważalne przez użytkownika przy odbiorze,
2. nie każdy obszar sygnału otrzymuje od użytkownika taki sam poziom uwagi,

3. postrzegane zmiany w sygnale nie są wprost proporcjonalne do faktycznego poziomu zmian lub zniekształceń.

Metryki bazujące w różnym stopniu na ludzkim systemie postrzegania podzielić można na [12]:

1. oparte o modele psychofizyczne ludzkiego systemu postrzegania – PDM
2. oparte o aspekty sygnału z uwzględnieniem ludzkiego systemu postrzegania – SDM.

Metryki PDM budowane są na podstawie właściwości psychofizycznych i wiedzy psychologicznej dotyczących ludzkiego układu wzrokowego (HVS) lub też ludzkiego układu słuchowego (HAS). Przykładem funkcji wykorzystywanej w metrykach oceny jakości obrazu, modelującej ludzką czułość wzroku zależną od częstotliwości sygnału wizualnego jest funkcja czułości kontrastu (CSF). Model HVS implementować może również postrzeganie kolorów, lokalny kontrast, czasoprzestrzenny CSF, zjawisko maskowania wizualnego itd. W przypadku oceny jakości mowy lub dźwięku, w modelu HAS implementuje się m.in. różne właściwości ucha zewnętrznego i środkowego, a także zjawisko maskowania psychoakustycznego [12, 77].

Ocena w metrykach SDM dokonywana jest w związku z ekstrakcją i analizą określonych cech samego sygnału, bądź artefaktów w nim zawartych. Cechami tymi mogą być cechy statystyczne, podobieństwo strukturalne, zniekształcenia jasności lub barw, czy zniekształcenia wprowadzane na różnych etapach przetwarzania, transmisji, bądź kompresji np. szumy blokowe. Metryki te uwzględniają również ludzki system postrzegania. Wykonywane jest to poprzez analizę zawartości i zniekształceń sygnału, a nie oddolne modelowanie systemu postrzegania, tak jak w metrykach PDM. Przez to, że nie wymagają one modelowania HAS i HVS, są mniej skomplikowane i wymagają mniejszej mocy obliczeniowej. Metryki te, w ostatnich latach stały się coraz bardziej popularne [12, 77].

Metody obiektywne podzielić też można ze względu na dostępność próbki wzorcowej (referencyjnej) tj. treści względem której dokonywana jest ocena. Wyróżnia się metody, które do obliczenia metryki zakładają, że [12, 45, 67]:

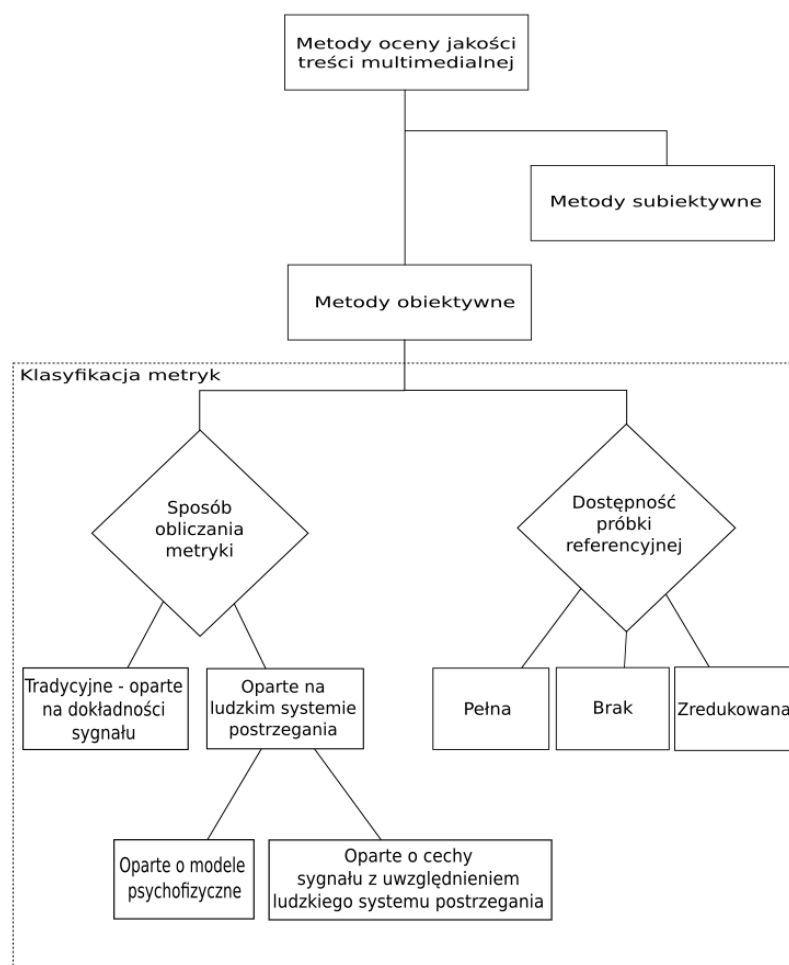
1. Próbka referencyjna jest w pełni dostępna (FR) – do przeprowadzenia oceny, wymagane jest, aby cała oryginalna treść multimedialna była dostępna. Oryginalna treść jest używana jako odniesienie podczas oceny.
2. Brak jest próbki referencyjnej (NR) – metody nie wymagają próbki referencyjnej do przeprowadzenia oceny.

3. Próbką referencyjną jest dostępna tylko w części (RR) – wymagana jest tylko część oryginalnej treści lub informacje mogące częściowo ją opisać.

Większość metryk opartych w różnym stopniu na ludzkim systemie postrzegania to metryki FR – wymagające kompletnego sygnału wzorcowego. Ogólnie mówiąc, metryki FR są bardziej dokładne w mierzeniu jakości w porównaniu do metryk RR i NR, ponieważ do przeprowadzenia oceny posiadają więcej informacji [12].

Każda z klas znajduje zastosowania w innych obszarach. Metryki FR są najbardziej odpowiednie do pomiarów jakości treści multimedialnych, które nie są przeprowadzane w czasie rzeczywistym. Przykładami takich pomiarów mogą być testy laboratoryjne, czy pomiary mające na celu ulepszanie kodeków, gdzie warunki pomiarów mogą być łatwo kontrolowane, a szczegółowa i precyzyjna analiza liczy się bardziej niż natychmiastowy rezultat. Metryki NR i RR są bardziej odpowiednie do monitorowania usług, gdzie pomiary w czasie rzeczywistym i zgłaszanie alarmów są kluczowe. Metryki RR jednakże, wciąż wymagają dostępu do próbki referencyjnej w jakimś stopniu [77].

Klasyfikację metod oceny jakości treści multimedialnych przedstawiono na Rys. 2.1.



**Rys. 2.1.** Klasyfikacja metod oceny jakości treści multimedialnych [12]

### 3. Obiektywne metody oceny jakości treści multimedialnych

Celem badań nad obiektywnymi metodami oceny jakości jest dostarczenie takich metryk jakości, na podstawie których automatycznie przewidzieć można postrzeganą przez użytkownika jakość treści multimedialnych [51,74].

Obiektywne metryki jakości mają znaczenie m.in. w następujących zastosowaniach [65, 74]:

1. W systemie kontroli jakości – do monitorowania, a następnie automatycznego dostosowywania jakości. Przykładem może być system pozyskujący obraz lub wideo, który wykorzystuje metrykę jakości do automatycznego dostosowania swoich parametrów tak, aby jakość pozyskiwanej treści była jak najlepsza.
2. Do oceny algorytmów i systemów przetwarzania sygnałów. Metryka jakości może posłużyć do wyboru najkorzystniejszego do określonego zadania, jednego z wielu możliwych systemów bądź algorytmów przetwarzania sygnałów.

Tradycyjne metryki jakości, bazujące na dokładności samego sygnału takie jak MSE, SNR i PSNR to metryki FR [67]. MSE i PSNR są podstawowymi miarami mierzącymi różnicę pomiędzy dwoma sygnałami. Wynik może być rozumiany jako stopień podobieństwa lub jako moc sygnału błędu pomiędzy dwoma sygnałami [45].

MSE zdefiniowany jest poniższym wzorem [45]:

$$MSE(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 \quad (1)$$

gdzie:

$x, y$  – dwa skończone i dyskretne sygnały, każdy o równej długości  $N$ ,

$N$  – długość sygnału,

$x_i, y_i$  – wartość  $i$ -tej próbki, kolejno sygnału  $x$  i  $y$ .

Miarę PSNR można uzyskać z wartości MSE następująco [45]:

$$PSNR(x, y) = 10 \log_{10} \frac{L^2}{MSE(x, y)}, \quad (2)$$

gdzie:

$L$  – maksymalna możliwa wartość sygnału np. jeżeli rozdzielczość bitowa sygnału wynosi 8 bitów to  $L = 2^8 - 1 = 255$

### 3.1. Podstawowe bloki obliczeniowe metryk opartych na ludzkim systemie postrzegania

Podstawowymi blokami obliczeniowymi metryk opartych na ludzkim systemie postrzegania są bloki odpowiadające za [12, 49]:

1. dekompozycję sygnału (np. dekompozycja obrazów w inną przestrzeń barw),
2. detekcję cech wspólnych (np. kontrast, ruch) i artefaktów (np. szumów blokowych, rozmycia),
3. wykrycie poziomu zauważalnej zmiany w postrzeganiu. (JND – ang. *Just-noticeable distortion*). Poziom zauważalnej zmiany w postrzeganiu jest to poziom maksymalnej zmiany lub zniekształcenia jakie można wprowadzić do treści multimedialnej (np. obrazu lub dźwięku), której nie zauważy większość obserwatorów.
4. modelowanie uwagi wizualnej (VA – ang. *Visual Attention*).

Większość metryk PDM używa dekompozycji sygnału poprzez ekstrakcję jego cech. Blok detekcji cech wspólnych i artefaktów jest kluczowy dla wielu metryk SDM. Odbierany przez użytkownika efekt powszechnych artefaktów daleko wykracza poza możliwość ich interpretacji z wykorzystaniem tradycyjnych metryk, takich jak MSE i PSNR. Bloki JND i VA mogą być używane niezależnie lub razem z poprzednimi blokami. Ich celem jest ocena stopnia postrzeganych różnic [12, 49].

#### 3.1.1. Dekompozycja sygnału

W przypadku obrazu i wideo, dekompozycja sygnału oznacza dekompozycje treści wizualnej na wiele różnych przestrzeni i kanałów np. czasu, częstotliwości, barw. Dekompozycja taka jest konieczna do dalszej oceny i wynika z wielu cech HVS



m.in. oddzielny sposób przetwarzania treści czarno-białej i kolorowej, czy różne ścieżki przetwarzania treści wizualnej (zależnie od typu ruchu). Treść wizualna jest też inaczej przetwarzana przez HVS dla różnych częstotliwości sygnału i orientacji obrazu. Podział na oddzielne kanały służy więc łatwiejszej symulacji odpowiedzi HVS. Dekompozycja sygnałów dźwiękowych bazuje na takich właściwościach układu słuchowego, jak postrzeganie głośności czy częstotliwości. Sygnały dźwiękowe dzielone są m.in. na wiele różnych pasm służących łatwiejszemu modelowaniu poszczególnych cech HAS, takich jak różne właściwości ucha zewnętrznego i środkowego, zjawisk jednoczesnego i wstecznego maskowania, czy post-maskowania itp. [12].

### **3.1.2. Detekcja cech wspólnych i artefaktów**

Proces detekcji cech sygnału i artefaktów jest powszechnie stosowany do oceny jakości treści wizualnych. Duże znaczenie dla odbieranej treści wizualnej mają cechy kontrastowe obrazu, takie jak jasność, kolor, orientacja, tekstura, ruch itp. HVS jest bardziej czuły na kontrast treści, niż na bezwzględną amplitudę sygnału. W przypadku sygnałów dźwiękowych wiele cech i miar wydobywanych jest na podstawie parametrów takich jak głośność, modulacja amplitudowa, maskowanie czy adaptacja. Artefakty, czyli pewne zniekształcenia strukturalne, powszechnie pojawiają się przy różnych metodach kompresji, a także w procesie transmisji. W rezultacie mogą powodować irytujący efekt u obserwatorów. Powszechne artefakty powstające przy kodowaniu sygnału, takie jak szumy blokowe, rozmycia, uszkodzenia krawędzi, czy artefakty pojawiające się przy ostrych przejściach w sygnale (ang. *Ringing artifacts*) ignorowane są przez metryki tradycyjne od strony efektów postrzegania u odbiorcy [10, 12].

### **3.1.3. Wykrycie poziomu zauważalnej zmiany w postrzeganiu (JND)**

Nie każda zmiana w sygnale jest zauważalna przy odbiorze. JND to poziom maksymalnej zmiany lub zniekształcenia, jakie można wprowadzić do treści multimedialnej (np. obrazu lub dźwięku), której nie zauważy większość obserwatorów. Wykrycie takiego poziomu jest potrzebne, aby stwierdzić które różnice mogą być zignorowane w procesie oceny, a które nie mogą. W przypadku obrazów i wideo do wykrycia JND wykorzystuje się dyskretną transformację kosinusową. Dla sygnałów dźwiękowych istnieje kilka typów JND, np. amplitudowy i częstotliwościowy. Wartość amplitudowego JND dla zwykłego słuchacza to około 1dB. Zdolność HAS do wykrywania zmian amplitudy i częstotliwości jest zależna od częstości i ilości tych zmian [12].

### **3.1.4. Modelowanie uwagi wizualnej (VA)**

Modelowanie uwagi wizualnej odnosi się do wybiórczego postrzegania treści wizualnej przez człowieka. Nie każdy obszar sygnału otrzymuje od użytkownika taki sam poziom uwagi. Jest to związane z tym że ludzki system postrzegania wybiera do szczegółowej analizy tylko części sygnału. Modelowanie VA w ogólności jest łatwiejsze dla wideo, niż obrazu. Obserwator mający wystarczająco dużo czasu na obserwację statycznego obrazu, ostatecznie zauważy wiele początkowo niewidocznych zniekształceń. W przypadku sygnału wideo, każda klatka wyświetlana jest obserwatorowi przez bardzo krótki czas, a dodatkowo ruch kamery lub obiektu może skupiać uwagę obserwatora. W porównaniu z sygnałami wizualnymi, istnieje znacznie mniejsza ilość badań nad modelowaniem uwagi słuchowej. Jednym z zaproponowanych modeli uwagi słuchowej jest model składający się z sieci neuronowej oscylatorów [12].

## **3.2. Obiektywne metody oceny jakości obrazu i wideo**

Standardowe metody do obiektywnej oceny jakości wideo (sekwencji obrazów), jak również statycznego obrazu, wymienione są w następujących rekomendacjach ITU [29, 30]:

1. ITU-R BT.1683
2. ITU-R BT.1907

### **3.2.1. ITU-R BT.1683**

Rekomendacja ITU-R BT.1683 [29] standaryzuje metody do szacowania odbieranej jakości sygnału wideo w systemie transmisji jednokierunkowej i ma zastosowanie do sygnałów w paśmie podstawowym (sygnałów transmitowanych bez modulacji, tj. bez żadnego przesunięcia w dziedzinie częstotliwości [47]) dla klas wideo TV0-TV3 i klasy multimedialnej MM4 zdefiniowanych w [34].

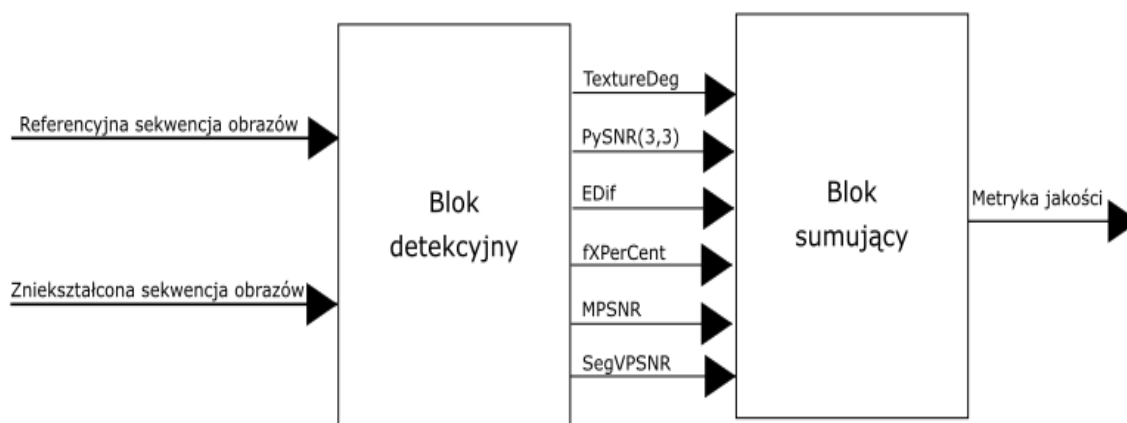
Rekomendacja ITU-R BT.1683 definiuje cztery modele do szacowania jakości treści wizualnej, mające zastosowanie m.in. do [29]:

- oceny jakości kodeków i przeprowadzania testów akceptacyjnych,
- monitorowania jakości treści (potencjalnie w czasie rzeczywistym),
- pomiarów jakości systemów przechowywania lub transmisji, wykorzystujących techniki kompresji i dekompresji treści.

Definiowane modele w rekomendacji ITU-R BT.1683 to [29]:

1. BTFR
2. Edge PSNR (EPSNR)
3. CPqD-IES
4. NTIA General Model

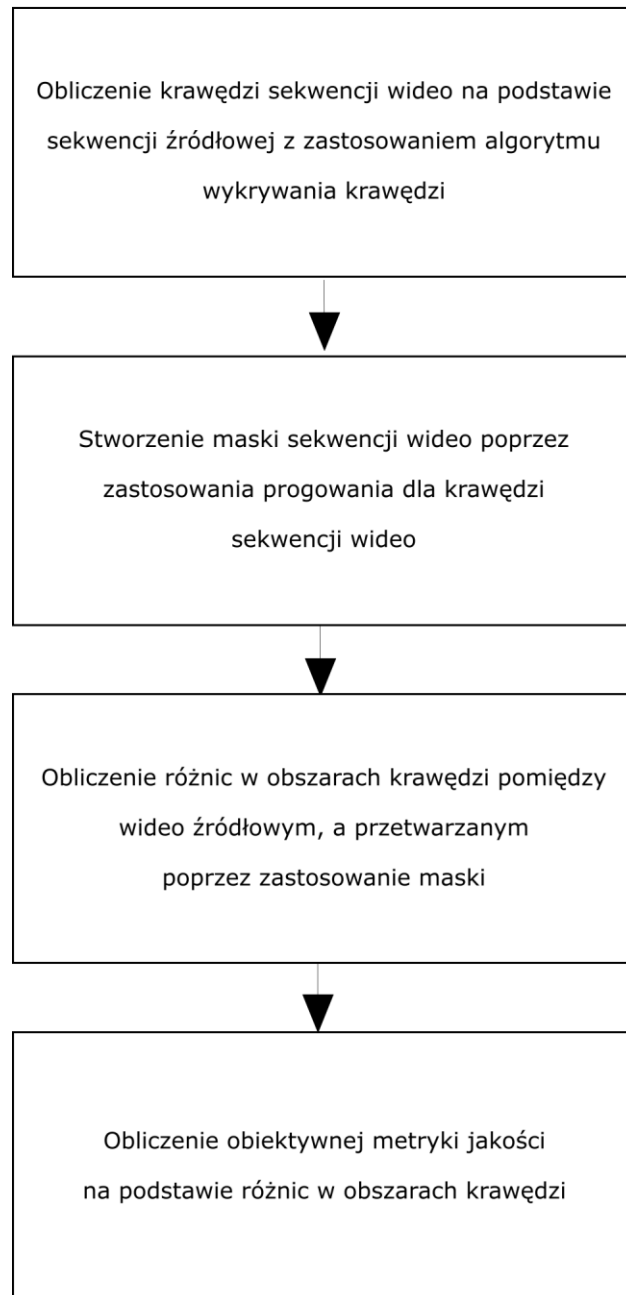
BTFR zapewnia automatyczne narzędzie do oceny jakości wideo, bazując swą ocenę na modelu HVS. Do przeprowadzenia oceny, narzędzie to wymaga pełnej dostępności próbki referencyjnej (FR). Algorytm BTFR składa się z bloku detekcyjnego i bloku sumującego. W bloku detekcyjnym, na podstawie referencyjnej (niezniekształconej) i zniekształconej sekwencji obrazów, obliczane są znaczące przy postrzeganiu przez HVS parametry. Parametry te, są następnie sumowane z odpowiednią dla siebie wagą w bloku sumującym, dając szacunkową miarę jakości na wyjściu. Ogólny schemat działania algorytmu przedstawiony jest na Rys. 3.1 [29].



**Rys. 3.1.** Schemat blokowy algorytmu BTFR [29]

Model EPSNR (Rys. 3.2) również wymaga pełnej dostępności próbki referencyjnej. Model wykorzystuje zaobserwowane zjawisko dużej czułości HVS na zniekształcenia wokół krawędzi. Zaobserwowano również, że algorytmy kompresujące często produkują artefakty przy krawędziach. Model ten jest łatwy w implementacji, wystarczająco szybki do monitorowania w czasie rzeczywistym i sprawdza się przy wielu różnych zaburzeniach obrazu [29].

**Diagram blokowy EPSNR**



**Rys. 3.2.** Schemat blokowy algorytmu EPSNR [29]

CPqD-IES to algorytm oceny obrazu bazujący na segmentacji. Segmentacja polega na dzieleniu naturalnych scen w obrazie na oddzielne powierzchnie, zbiory krawędzi i regiony teksturowa. Następnie zbiór parametrów obiektywnych przypisywany jest każdemu z tych kontekstów i dokonuje się obliczenia relacji pomiędzy parametrami obiektywnymi, a wynikami testów subiektywnych, zastosowanych do zbioru naturalnych scen przetworzonych przez kodek MPEG-2. Relacja pomiędzy każdym parametrem obiektywnym, a jego subiektywnym poziomem zniekształcenia, przybliżana jest z zastosowaniem funkcji logistycznych (ang. *Logistic function*) dając w efekcie

estymowany poziom zniekształcenia dla danego parametru. Ostateczny rezultat miary jakości dokonywany jest poprzez kombinację liniową estymowanych poziomów zniekształcenia, gdzie waga danego poziomu zależy od jego istotności statystycznej [29].

NTIA General Model zaprojektowany został, aby dostarczyć metrykę jakości ogólnego przeznaczenia dla różnych systemów wideo obejmujących szeroki zakres jakości i przepływności sygnałów. Procesy które wykonuje algorytm to kolejno: próbkowanie strumienia obrazów referencyjnego i przetwarzanego, kalibracje obu strumieni, ekstrakcje cech mających znaczenie przy postrzeganiu, obliczenie parametrów jakościowych i obliczenie miary VQM [29].

### 3.2.2. ITU-R BT.1907

Rekomendacja ITU-R BT.1907 [30] zawiera obiektywną metodę FR oceny jakości wideo do zastosowania przy sygnałach HDTV. Metoda ta oparta jest o ludzki system postrzegania. Wykorzystując modele psychowizualne i poznawcze, ma na celu estymację oceny subiektywnej. Wyjściowa metryka obliczana jest w następujący sposób [30]:

1. sekwencje wideo są wstępnie przetwarzane - dla każdej klatki usuwany jest szum poprzez wykorzystanie filtru; każda klatka jest też poddawana podpróbkowaniu,
2. dokonywane jest dopasowanie okien czasowych pomiędzy sekwencją referencyjną i przetwarzaną (zniekształconą),
3. dokonywane jest dopasowanie okien przestrzennych (ang. *Spatial Frame*) pomiędzy sekwencją referencyjną i przetwarzaną,
4. obliczane są cechy lokalnej jakości przestrzennej tj. podobieństwo lokalne i lokalna różnica, bazująca na ludzkim postrzeganiu,
5. dokonywana jest analiza rozkładu statystycznego lokalnych podobieństw i różnic,
6. wykonywany jest pomiar globalnego współczynnika zniekształceń przestrzennych na podstawie artefaktów blokowych,
7. wykonywany jest pomiar globalnego współczynnika zniekształceń czasowych na podstawie oceny globalnej, lokalnej intensywności ruchu i czasu wyświetlania klatki,
8. metryka jakości obliczana jest na podstawie nieliniowej funkcji agregującej powyższe parametry,
9. w celu uniknięcia błędnej estymacji jakości w przypadku względnie dużych różnic przestrzennych, powyższe etapy wykonywane są dla trzech różnych

orientacji poziomych i pionowych sekwencji wideo. Wyjściową, estymowaną miarą jakości jest najwyższy wynik wśród różnych pozycji przestrzennych.

### 3.3. Obiektywne metody oceny jakości dźwięku

Standardowe metody obiektywne i narzędzia do oceny jakości dźwięku opisują poniższe rekomendacje ITU [25, 31, 32, 35]:

1. ITU-R BS.1387 (PEAQ) – do oceny jakości fonii
2. ITU-T P.862 i ITU-T P.862.2 (PESQ) – do oceny jakości mowy
3. ITU-T P.863 (POLQA) – do ocena jakości mowy

#### 3.3.1. ITU-R BS.1387 (PEAQ)

PEAQ jest algorytmem modelującym psychoakustyczne właściwości słuchu HAS. W algorytmie zaimplementowano dwa modele słuchowe [25]:

1. bazujący na FFT,
2. bazujący na zestawie filtrów.

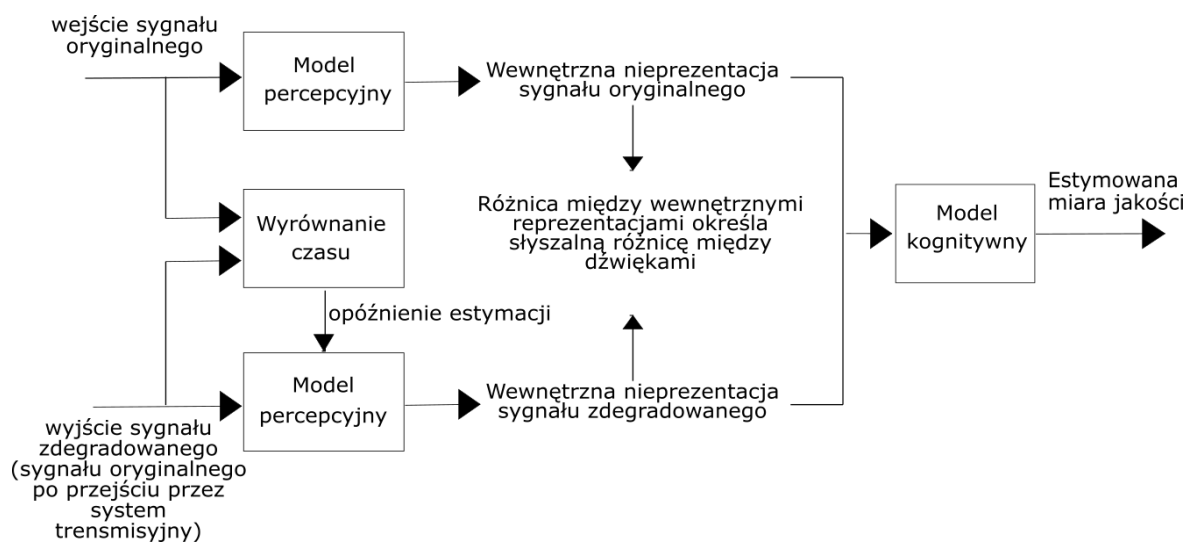
Dla modelu bazującego na FFT, sygnały (referencyjny i zniekształcony) są najpierw transformowane w dziedzinę częstotliwości, następnie, widmo amplitudowe używane jest do dalszego przetwarzania w blokach modelujących cechy ucha zewnętrznego i środkowego. W dalszym etapie składowe częstotliwościowe są też grupowane w częstotliwościowe pasma krytyczne postrzegane przez HAS. Grupy te są następnie przetwarzane przez blok modelujący cechy ucha wewnętrznego w celu uzyskania wzorców wysokości dźwięku (ang. *Pitch Patterns*). Wzorce te poddawane są następnie blokom modelującym maskowanie jednoczesne i maskowanie przednie [10, 12, 25].

W przypadku modelu bazującego na zestawie filtrów, sygnał dekomponowany jest na różne pasma częstotliwościowe przez 40 filtrów (częstotliwości środkowe w tych pasmach mają zakres od 50Hz do 18kHz). Następne etapy przetwarzania przeprowadzane są jak modelu bazującym na FFT. W modelu tym dodatkowo zaimplementowano zjawisko maskowania wstecznego [10, 12, 25].

Rekomendacja ITU-R BS.1387 definiuje dwie wersje algorytmu tj. podstawową i zaawansowaną. Wersja podstawowa do obliczenia estymowanej miary jakości wykorzystuje 11 parametrów (MOV), obliczanych w modelu FFT; model bazujący na zestawie filtrów nie jest w niej wykorzystywany. Wersja zaawansowana wykorzystuje parametry MOV z obu modeli [25].

### 3.3.2. ITU-T P.862 i ITU-T P.862.2

Rekomendacja ITU-T P.862 definiuje obiektywną metodę przewidywania subiektywnej jakości dla sygnału mowy w wąskopasmowych (do 3.1 kHz) zestawach telefonicznych i wąskopasmowych kodekach mowy. Definiowana metoda znana jest pod skróconą nazwą PESQ. PESQ porównuje sygnał referencyjny  $x(t)$  ze zniekształconym sygnałem  $y(t)$ , który jest rezultatem przejścia sygnału  $x(t)$  przez system transmisyjny. Rezultatem jest estymowana, subiektywna miara jakości, którą uzyskaliby sygnał zniekształcony w testach subiektywnych. Do porównania obu sygnałów wykorzystywany jest model percepcyjny i kognitywny [31]. Schemat blokowy metody PESQ przedstawiony jest na Rys. 3.3.



**Rys. 3.3.** Schemat blokowy metody PESQ [31]

Rekomendacja ITU-T P.862.2 definiuje rozszerzenie metody PESQ do zastosowania dla sygnałów o zakresie częstotliwości 50-7000Hz (szerokopasmowych). Rozszerzenie zmienia filtr wejściowy stosowany dla obu sygnałów w wąskopasmowej metodzie PESQ na filtr IIR [35]

### 3.3.3. ITU-T P.863 (POLQA)

Rekomendacja ITU-T P.863 opisuje obiektywną metodą do przewidywania jakości mowy dla sygnałów wąskopasmowych NB (300-3400Hz) i super-szerokopasmowych SWB (50 - 14000Hz). Rekomendacja wspiera dwa tryby operacyjne: jeden dla sygnałów NB, drugi dla SWB. Definiowany w rekomendacji algorytm to algorytm FR. Algorytm porównuje sygnał referencyjny  $x(t)$  ze zniekształconym sygnałem  $y(t)$ , który jest

rezultatem przejścia sygnału  $x(t)$  przez system transmisyjny. Rezultatem jest estymowana, subiektywna miara jakości, którą uzyskałby sygnał zniekształcony w testach subiektywnych. Pierwszym krokiem w algorytmie POLQA jest podział obu sygnałów na bardzo małe przedziały czasowe – ramki. Następnie obliczane jest opóźnienie każdej ramki referencyjnej względem zniekształconej. Na podstawie wykrytych opóźnień estymowana jest częstotliwość próbkowania obu sygnałów względem siebie. W przypadku znaczącej różnicy częstotliwości próbkowania, sygnał o większej częstotliwości próbkowania poddany jest decymacji, a opóźnienia obliczone zostają ponownie. Następnie sygnały zostają porównane przy użyciu modelu percepcyjnego. Wartości opisujące różnicę pomiędzy oboma sygnałami są przekazywane do komponentu modelującego procesy kognitywne. Poddawane są tam dalszemu przetwarzaniu. Wskaźniki obliczane w komponentcie kognitywnym są następujące [32]:

1. wskaźnik odpowiedzi częstotliwościowej (FREQ),
2. wskaźnik szumu (NOISE),
3. wskaźnik pogłosu pomieszczenia (REVERB),
4. trzy wskaźniki opisujące różnicę pomiędzy sygnałami w przestrzeni tworzonej przez czas, głośność i wysokość dźwięku.

Otrzymane wskaźniki są na końcu łączone, dając wynikową miarę jakości [32].



## 4. Subiektywne metody oceny jakości treści multimedialnych

Wzorcem odniesienia miary jakości treści multimedialnej są eksperymenty subiektywne, metody te są więc najbardziej wiarygodne. Przeprowadzanie eksperymentów badania jakości metodami subiektywnymi wymaga od grupy ludzi (zazwyczaj 15-30 osób) słuchania lub oglądania materiału testowego, a następnie poddania go ocenie. Jedną z powszechnych miar subiektywnie postrzeganej jakości jest MOS – używana szczególnie w przypadku oceny jakości treści wizualnych oraz oceny jakości mowy. Etykiety jakości wykorzystywane w ocenach typu MOS przedstawiono w Tabeli 4.1. [7, 12, 77].

Jako że każda osoba ma inne zainteresowania i oczekiwania wobec treści i jej jakości, różnice, w jej subiektywnych ocenach nie mogą być całkowicie wyeliminowane. Dla eksperymentów subiektywnych czynniki te minimalizowane są poprzez dużą próbę słuchaczy, precyzyjne instrukcje, wstępny trening i przeprowadzanie testów w środowisku kontrolowanych. Jednakże, pamiętać należy, że uzyskana miara jakości jest bardziej miarą opisywaną przez rozkład statystyczny, niż przez dokładną jej wartość [77].

**Tabela 4.1.** Skale ocen jakości i zniekształceń [7]

Ocena	Jakość	Zniekształcenie
5	Doskonała	Niezauważalne
4	Dobra	Zauważalne
3	Dostateczna	Nieco irytujące
2	Słaba	Irytujące
1	Zła	Bardzo irytujące

### 4.1. Ocena jakości obrazu

Standaryzowane przez ITU metody subiektywne i warunki przeprowadzania eksperymentów oceny jakości obrazu opisuje rekomendacja ITU-R BT.500 [27].

Rekomendacja ITU-R BT.500 opisuje następujące metody [27]:

1. DSIS (ang. *Double-stimulus Impairment Scale*) – dwa bodźce, wykorzystywana skala zniekształcenia,
2. DSCQS (ang. *Double-stimulus Continuous Quality-scale*) – dwa bodźce, wykorzystywana ciągła skala oceny,
3. metody SS (ang. *Single-stimulus*) – jeden bodziec,
4. metody porównywania pojedynczo prezentowanych bodźców (ang. *Stimulus-comparision methods*)
5. SSCQE (ang. *Single Stimulus Continuous Quality Evaluation*) – ciągła ocena na podstawie jednego bodźca
6. SCSCE (ang. *Simultaneous Double Stimulus for Continuous Evaluation*) – ciągła ocena na podstawie dwóch bodźców jednocześnie.

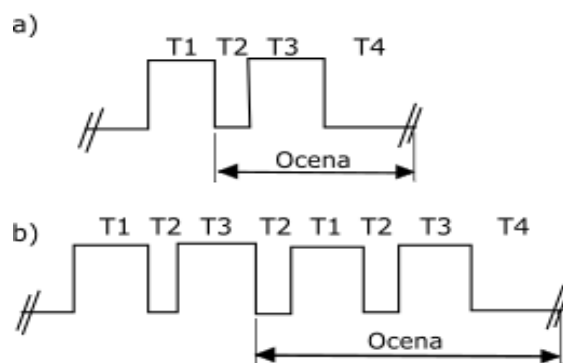
Zalecane użycie metod do odpowiedniego problemu przedstawia Tabela 4.2. [27].

Rekomendacja zawiera też wytyczne dotyczące środowiska laboratoryjnego, m.in. dotyczące jasności, rozdzielczości i rozmiarów wyświetlacza, czy oświetlenia w pomieszczeniu. Wytyczne te mają na celu modelowanie uśrednionego pomieszczenia konsumenta i warunków domowych. Rekomendowani uczestnicy mogą być ekspertami (tzn. osobami mającymi wiedzę o powszechnych artefaktach obrazu) lub nie [27].

Przykład prezentacji materiału dla uczestnika testu w metodzie DSIS został przedstawiony na Rys. 4.4. Rekomendacja opisuje dwa warianty prezentacji. W wariancie I (Rys. 4.1. a) obraz referencyjny i obraz testowy prezentowane są tylko raz, zaś w wariancie II (Rys. 4.1. b) obraz referencyjny i obraz testowy prezentowane są dwa razy.

**Tabela 4.2.** Zalecane użycie metod dla danego problemu [27]

<b>Problem</b>	<b>Metoda</b>
Pomiar jakości systemów względem próbki wzorcowej	DSCQS
Pomiar odporności systemów na zniekształcenia	DSIS
Szacowanie jakości systemów (brak próbki referencyjnej)	Skalowanie przy użyciu skali ciągłej (ang. <i>Ratio-scaling</i> ), skalowanie przy użyciu klas (kategorii) jakości (ang. <i>Categorical scaling</i> ) (metody w badaniu)
Porównanie jakości zamiennych systemów	Porównanie bezpośrednie, Skalowanie przy użyciu skali ciągłej, skalowanie przy użyciu klas (kategorii) jakości (metody w badaniu)
Identyfikacja czynników mających wpływ na postrzeganie systemów do różnicowania i miar ich percepcyjnego wpływu	W badaniu
Określenie punktu, w którym zniekształcenia zaczynają być dostrzegalne	Metoda wymuszonego wyboru, metoda dostosowania (metody w badaniu)
Określenie czy systemy są różnie postrzegane	Metoda wymuszonego wyboru (w badaniu)
Pomiar jakości kodowania obrazu stereoskopowego	DSCQS
Pomiar dokładności (ang. <i>fidelity</i> ) między dwoma sekwencjami wideo	SDSCE
Porównanie odporności narzędzi na błędy	SDSCE



Fazy prezentacji:

T1= 10s Obraz referencyjny

T2= 3s Obraz o barwie średnio-szarej (200mV)

T3= 10s Próbką testowa

T4= 5-11s Obraz o barwie średnio-szarej

**Rys. 4.1.** Fazy prezentacji materiału testowego w metodzie DSIS:

a) wariant I, b) wariant II [27]

## 4.2. Ocena jakości dźwięku

Standardowe metody subiektywne i warunki przeprowadzania eksperymentów oceny jakości dźwięku opisane są w następujących rekomendacjach [24]:

1. ITU-R BS.1116,
2. ITU-R BS.1534 (MUSHRA).

Do oceny, czy zachodzą różnicę pomiędzy dwoma dźwiękami, wykorzystany może być też test ABX [53].

### 4.2.1. ITU-R BS.1116

Rekomendacja ITU-R BS.1116 definiuje metodę do oceny jakości systemów fonii, w których zaburzenia jakości dźwięku są tak małe, że niemożliwa byłaby ich detekcja bez rygorystycznej kontroli warunków eksperymentu i właściwej analizy statystycznej. Zalecanymi uczestnikami testów są eksperci (specjaliści) tzn. osoby mające wiedzę i umiejętności pozwalające na rozpoznawanie małych zniekształceń między dźwiękami. Rekomendacja zawiera też wytyczne dotyczące m.in. pokoju odsłuchowego, panelu testowego i wymagań elektroakustycznych [28].

Metoda testowa jest określana jako podwójnie „ślepa”<sup>1</sup>, trójbodźcowa, z ukrytą próbką referencyjną. W preferowanej i najbardziej czulej wersji metody, jeden bodziec (próbka) jest w danej chwili wybierany spośród trzech bodźców ("A", "B", "C"). Jawna próbka referencyjna dostępna jest zawsze jako bodziec "A". Próbka referencyjna i próbki podlegające ocenie są jednocześnie dostępne i przypisywane losowo podczas każdej próby jako "B" lub "C". Zadaniem uczestnika testu jest ocena zniekształceń próbki "B" względem "A", a także ocena zniekształceń próbki "C" względem "A", zgodnie z 5 stopniową skalą zniekształceń podaną wcześniej w Tabeli 4.1.

Ostateczny rezultat obliczany jest w następujący sposób [28]:

$$Z_i = \frac{(x_i - x_{si})}{s_{si}} \cdot s_s + x_s, \quad (3)$$

gdzie:

$Z_i$  – znormalizowany rezultat,

$x_i$  – wynik oceny  $i$ -tego uczestnika,

$x_{si}$  – średni wynik oceny  $i$ -tego uczestnika w sesji  $s$ ,

$x_s$  – średni wynik oceny wszystkich uczestników w sesji  $s$ ,

$s_s$  – odchylenie standardowe oceny dla wszystkich uczestników w sesji  $s$ ,

$s_{si}$  – odchylenie standardowe oceny  $i$ -tego uczestnika w sesji  $s$ .

#### 4.2.2. ITU-R BS.1534 (MUSHRA)

Rekomendacja ITU-R BS.1534 definiuje metodę do oceny jakości systemów fonii, w których zaburzenia pomiędzy dźwiękami można określić jako zaburzenia średniej wielkości. Metoda ta została z powodzeniem przetestowana i ukazano, że jest odpowiednia do oceny systemów pośredniej jakości dźwięku, dając dokładne i wiarygodne rezultaty. Rekomendacja zawiera też wytyczne dotyczące m.in. pokoju odsłuchowego, panelu testowego i wymagań elektroakustycznych [26].

Metoda testowa, zwana MUSHRA, określana jest jako podwójnie „ślepa”, wielobodźcowa, z ukrytą próbką referencyjną i ukrytymi, specjalnie zniekształconymi próbkami - kotwicami (ang. *Anchor*), mającymi za zadanie zwiększenie stabilności oceny.

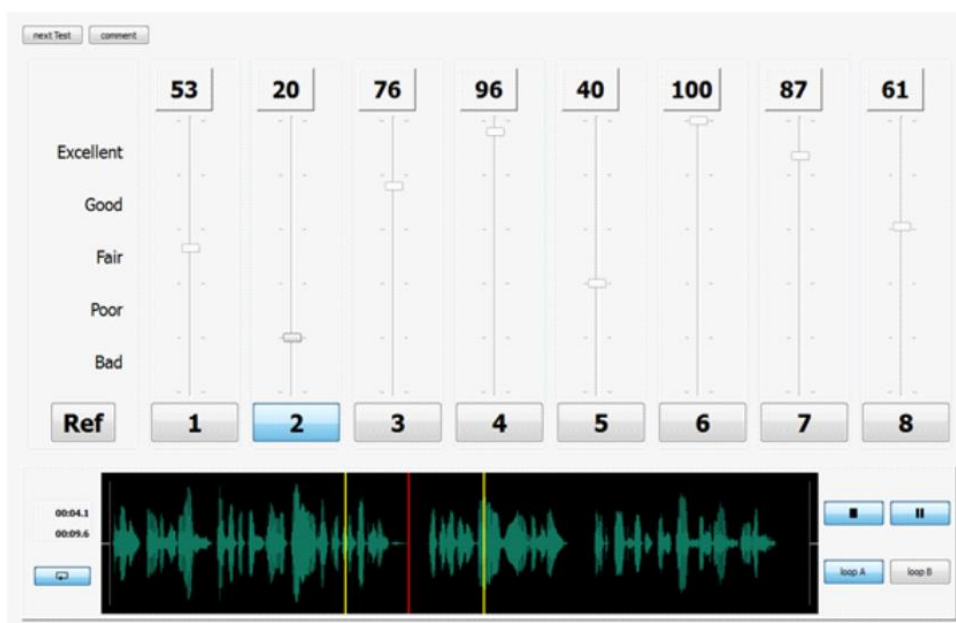
<sup>1</sup> W metodzie podwójnie „ślepej” (ang. *Double-blind*) w danym momencie czasu zarówno uczestnik testu i osoba nadzorująca test nie wie jaka próbka jest odtwarzana, co wynika z losowego przyporządkowania próbek.

Jawna próbka referencyjna jest zawsze dostępna. Stosowana skala oceny to skala 100 stopniowa, ciągła, z określonymi przedziałami. Podczas testu uczestnik może dowolnie przełączać się pomiędzy wszystkimi bodźcami tj. próbkami testowymi (ang. *Systems Under Test*), a jawną referencją lub między jedną, a drugą próbką testową. Zaleca się, aby łączna ilość bodźców (sygnałów) nie przekraczała 12 (np. 9 próbek testowych, 1 ukrytej referencji, 1 kotwicy niskoczęstotliwościowej, 1 kotwicy średniczęstotliwościowej). Zalecana długość trwania każdej próbki to około 10 sekund [26].

Rekomendowane jest użycie dwóch kotwic. Kotwica pierwsza (niskoczęstotliwościowa) i kotwica druga (średniczęstotliwościowa) to próbka referencyjna po przejściu przez filtr dolnoprzepustowy o częstotliwości odcięcia  $f_c$ , kolejno 3.5 kHz i 7 kHz. Charakterystyka filtru o  $f_c = 3.5$  kHz powinna być następująca [26]:

- maksymalne zafalowanie (ang. *Ripple*):  $\pm 0.1$  dB,
- minimalne tłumienie przy częstotliwości 4 kHz = 25 dB,
- minimalne tłumienie przy częstotliwości 4.5 kHz = 50dB.

Podobnie jak w przypadku metody opisanej w rekomendacji ITU-R BS.1116, wciąż zalecanymi słuchaczami powinni być eksperci [26]. W przypadku interfejsu słuchacza zaleca się, aby uczestnik nie mógł ocenić próbki innej, niż aktualnie odsłuchiwana, ze względu na możliwość przypadkowego pomylenia się i dokonania oceny niewłaściwej próbki. Przykładowy interfejs użytkownika przedstawiony jest na Rys. 4.2.



**Rys. 4.2.** Przykładowy interfejs użytkownika w metodzie MUSHRA [26]

Rezultaty przeprowadzonych testów powinny zawierać [26]:

- opis materiału testowego,
- ilość uczestników (testujących),
- graficzną prezentację wyników tj. wykresy pudełkowe pokazujące IQR median próbek testowych i średnie próbek testowych w 95% przedziale ufności.

#### 4.2.3. Test ABX

Test ABX jest to podwójnie „ślepy” test wymuszonego wyboru. Procedura testowa przeprowadzana jest w następującej kolejności [53]:

1. słuchaczowi prezentowany jest sygnał "A",
2. słuchaczowi prezentowany jest sygnał "B",
3. prezentowany jest dźwięk "X", losowo wybierany jako "A" lub "B",
4. ostatecznie słuchacz ma za zadanie wskazać, czy dźwięk "X" brzmi bardziej jak "A" lub "B".

Do odrzucenia hipotezy  $H_0$  (o tym, że między dwoma dźwiękami nie zachodzi statystycznie znacząca różnica) użyć można testu chi-kwadrat Pearson'a lub testu dwumianowego. W przypadku testu chi-kwadrat Pearson'a, wartość statystyki  $\chi^2$  obliczana jest w następujący sposób [11]:

$$\chi^2 = \frac{4 \cdot (x - n/2)^2}{n}, \quad (4)$$

gdzie:

- $x$  – liczba poprawnych wskazań,  
 $n$  – liczba prób.

## 5. INTERNETOWE NARZĘDZIA DO OCENY JAKOŚCI DŹWIĘKU I OBRAZU ORAZ WYBÓR IMPLEMENTOWANEJ METODY TESTOWEJ

W ostatnich latach ocena jakości treści multimedialnych z zastosowaniem subiektywnych testów przeprowadzanych przez Internet zyskała na popularności. Przeprowadzane eksperymenty subiektywnej oceny jakości z wykorzystaniem Internetu prowadziły do porównywalnych rezultatów, jak w eksperymentach laboratoryjnych. Przy eksperymentach internetowych, popularność zdobywa również wykorzystanie platform delegujących przeprowadzenie takich testów bardzo szerokiej, zwykle nieznanej grupie ludzi (ang. *Crowdsourcing*). Przykładami takich platform są *Amazon Mechanical Turk* i *Microworkers*. Rozwiązanie takie pozwala na uzyskanie dużej próby wyników przy względnie małym koszcie finansowym [23, 64].

W rozdziale tym przedstawiony zostanie wykaz istniejących narzędzi do internetowej oceny jakości dźwięku i obrazu. Analiza funkcjonalności i zastosowanych rozwiązań w istniejących narzędziach, przyczynić się może do dokładniejszego sformułowania wymagań funkcjonalnych i pozafunkcjonalnych portalu, którego opracowanie przedstawiono w rozdziale 6. Dokonany zostanie również wybór implementowanej w portalu metody do oceny jakości dźwięku.

### 5.1. Wykaz istniejących narzędzi

Jednym z istniejących narzędzi umożliwiających zdalne przeprowadzenie subiektywnych eksperymentów oceny jakości dźwięku jest platforma programistyczna BeagleJS. Platforma ta umożliwia stworzenie i przeprowadzanie testów odsłuchowych w oparciu o metody MUSHRA i test ABX. Platforma ta nie posiada kreatora testów i nie umożliwia integracji z bazą danych [40, 41].

Kolejnym narzędziem umożliwiającym tworzenie i przeprowadzanie testów z wykorzystaniem metod MUSHRA, ITU-R BS.1116, a także metody porównań parami, jak ABX, jest platforma programistyczna webMUSHRA. Narzędzie to umożliwia dodawanie prób treningowych, czy też automatyczną generację kotwic. Konfiguracja wykonywana jest w oparciu o pliki w formacie YAML. Do przeprowadzania testów



wykorzystywany jest implementowany przez przeglądarki interfejs Web Audio API. Zaimplementowany w narzędziu interfejs MUSHRA przedstawiony został na Rys. 5.1. Narzędzie obecnie nie jest w pełni dostępne publicznie [5, 73].

Zaawansowaną platformą umożliwiającą przeprowadzanie eksperymentów zdalnych z zastosowaniem takich metod jak MUSHRA, ABX Test, AB Test jest WAET (ang. *Web Audio Evaluation Toolkit*). W platformie tej możliwe jest także zastosowanie wielu różnych skal ocen, m.in. 5 stopniową skalę oceny ITU-R, ACR (ang. *Absolute Category Rating*), DCR (ang. *Degradation Category Rating*), CCR (ang. *Comparision Category Rating*). Platforma posiada graficzny kreator testów oraz pliki konfiguracyjne. Rezultaty testów przechowywane są w plikach XML [37].

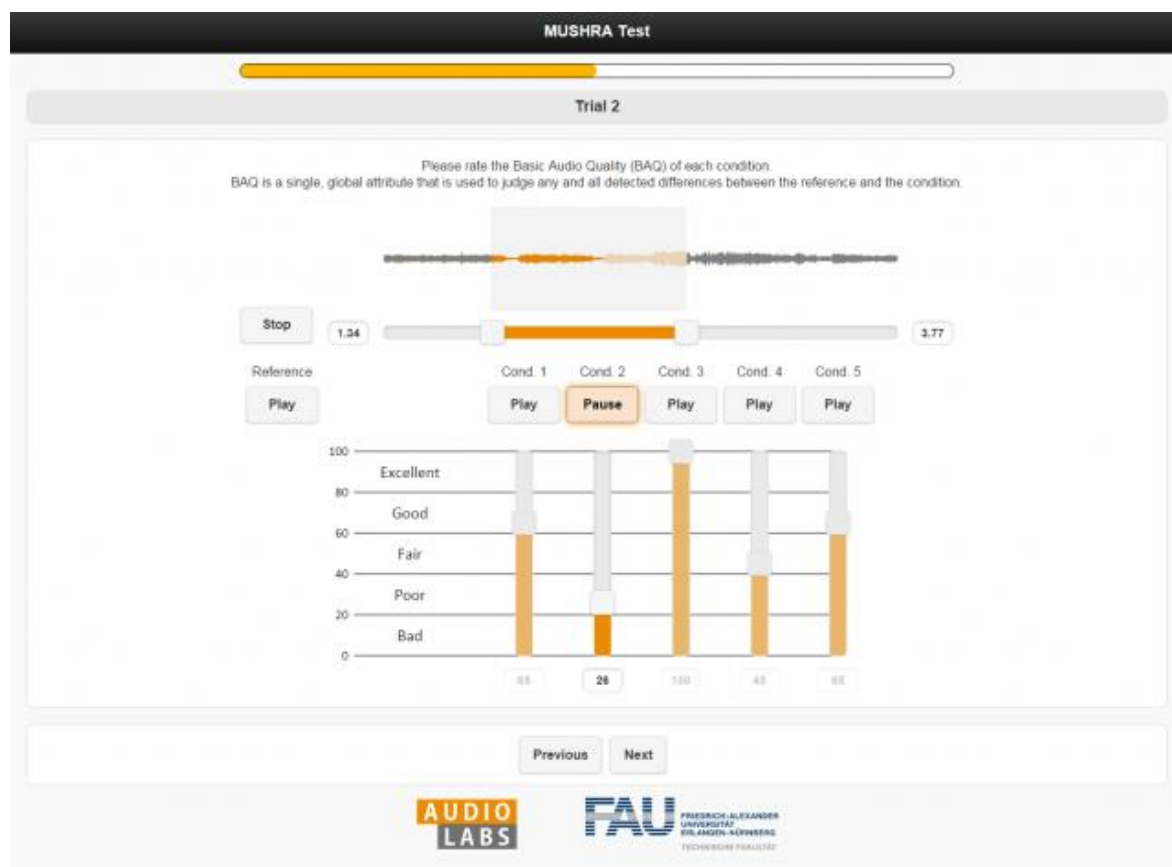
CrowdMOS to zestaw skryptów i narzędzi pozwalających na ocenę jakości obrazów i dźwięków z wykorzystaniem platformy Amazon Mechanical Turk. CrowdMOS umożliwia zastosowanie metod oceny jakości, takich jak ACR, DCR i MUSHRA. Do odsiewu niewiarygodnych uczestników (np. oceniających treść losowo – w celu szybkiego otrzymania zapłaty za pracę) stosowany jest 95% przedział ufności dla średniej ocen [23, 39].

Quadrant of Euphoria jest internetowym portalem "crowdsourcingowym" służącym do oceny jakości obrazu, wideo i dźwięku. Jedyną możliwą do zastosowania metodą oceny jakości jest binarna metoda porównania parami. Do odrzucania rezultatów niewiarygodnych użytkowników stosowana jest relacja przechodnia wyrażana jako TSR (ang. *Transitivity Satisfaction Rate*). W przypadku gdy uczestnik preferuje bodziec testowy A od bodźca B i bodziec B od bodźca C, powinien też preferować bodziec testowy A od bodźca C. Jeżeli warunek ten nie zachodzi dla założonej ilości ocen użytkownika, uzyskane przez niego rezultaty są odrzucane [13, 23].

QualityCrowd2 to platforma programistyczna umożliwiająca tworzenie testów subiektywnych do oceny jakości obrazu, wideo i dźwięku. Platforma oferuje własny język skryptowy do tworzenia testów. Możliwymi do zastosowania metodami oceny jakości są: metoda ACR oraz metody, podczas których prezentowany jest jeden lub dwa bodźce w danej chwili. Do zastosowania są też różne skale tj. skala ciągła i dyskretna [23, 61].

WESP+ jest platformą programistyczną, która umożliwia tworzenie testów do oceny jakości obrazu, wideo, dźwięku, a także oceny efektów sensorycznych (ang. *sensory effects*). Możliwymi do zastosowania metodami testowymi są: ACR, ACR-HR, DSCQE i metoda prezentowania dwóch bodźców jednocześnie (do oceny efektów sensorycznych) [23, 59, 43].

Wszystkie wyżej wymienione narzędzia, oprócz portalu Quadrant of Euphoria, są platformami programistycznymi (ang. *Framework*) [37] i nie stanowią gotowego portalu internetowego. Platformy te wymagają od potencjalnego twórcy testu, własnoręcznego skonfigurowania i uruchomienia systemu. Portal Quadrant of Euphoria umożliwia przeprowadzenie testów z wykorzystaniem tylko jednej metody tj. binarnej metody porównywania parami [13] i aktualnie dostępny jest jedynie w wersji demonstracyjnej [1].



Rys. 5.1. Interfejs metody MUSHRA w platformie webMUSHRA [5]

## 5.2. Wybór implementowanej metody testowej

Ze względu na internetowy charakter portalu, a przez to niekontrolowane warunki przeprowadzania eksperymentów oceny subiektywnej, m.in. niejednolity i zazwyczaj niewysokiej jakości sprzęt dźwiękowy, możliwość uczestnictwa osób, które nie są ekspertami, rozpatrywana metoda nie powinna dotyczyć oceny jakości bardzo małych zaburzeń dźwięku. Metoda, służąca do oceny bardzo małych zaburzeń dźwięku to metoda opisana w standardzie ITU-R BS.1116 [28]. W związku z powyższym argumentem, implementację tej metody w portalu odrzucono. Zgodnie z rekomendacjami ITU, metodą

do oceny pośrednich i dużych zaburzeń dźwięku jest MUSHRA, opisana w standardzie ITU-R BS.1534 [26]. Jest to obecnie powszechnie wykorzystywana metoda do oceny jakości dźwięku. Metoda ta została wykorzystana m.in. do wyboru najlepszego algorytmu do radia cyfrowego DAB [48], czy też oceny jakości popularnych kodeków dźwięku, zastosowanych zarówno do kodowania dźwięku przy niskiej przepływności sygnału (od 16 kbit/s do 64 kbit/s) [17], jak i wysokiej (od 64 kbit/s do 1500 kbit/s), przy wielu kanałach dźwięku [16]. Metoda MUSHRA została też użyta do oceny nowego kodeku dźwięku przestrzennego MPEG-H [38]. Z racji powszechności użycia metody MUSHRA, istnieje duże prawdopodobieństwo, że potencjalny użytkownik portalu chciałby ją wykorzystać m.in. ze względu na możliwość łatwego porównania wyników. Jako metodę implementowaną w portalu wybrano zatem metodę MUSHRA.

Jako drugą metodę do zaimplementowania w opracowanym w niniejszej pracy portalu wybrano metodę ABX. Test ABX jest prostą metodą oceny jakości wykorzystywaną przez wielu twórców algorytmów przetwarzania dźwięku [79] i społeczności entuzjastów dźwięku np. do porównania jakości urządzeń odtwarzania dźwięku [78]. Jest to jedna z historycznie pierwszych metod wykorzystywanych do oceny jakości dźwięku [15]. Test ABX pozwolić może na obalenie hipotezy  $H_0$ , że pomiędzy dwoma dźwiękami nie zachodzą statystycznie znaczące różnice [11].

## **6. OPRACOWANIE PORTALU INTERNETOWEGO DO OCENY JAKOŚCI DŹWIĘKU**

### **6.1. Wymagania funkcjonalne i нефunkcjonalne (pozafunkcjonalne)**

Wymagania funkcjonalne dotyczące portalu internetowego do oceny jakości dźwięku są następujące:

1. Możliwość założenia konta użytkownika i uwierzytelniania w systemie.
2. Możliwość tworzenia grup testowych dla użytkowników uwierzytelnionych. W skład grupy testowej wchodzić może wiele testów odsłuchowych, wykorzystujących metody MUSHRA lub ABX.
3. W przypadku pojedynczego testu wykorzystującego metodę MUSHRA, możliwość użycia od 3 do 12 próbek dźwiękowych.
4. Możliwość określenia ilości powtórzeń dla każdego testu, przy jego tworzeniu.
5. Możliwość wzięcia udziału w testach, zarówno dla użytkowników uwierzytelnionych i niewierzytelnionych (anonimowych).
6. Możliwość podglądu podstawowych statystyk dotyczących własnych testów (ilość rezultatów, średnie ocen wraz z 95% przedziałami ufności).
7. Możliwość eksportu uzyskanych rezultatów własnych testów do pliku w formacie CSV i JSON.
8. Możliwość oznaczenia własnego testu jako ukończony lub usunięcia go.
9. Autoryzowany dostęp do zasobów (np. dostęp do rezultatów testu tylko dla jego twórcy).

Wymagania нефunkcjonalne zdefiniowane zostały następująco:

1. Nie przechowywanie haseł do kont użytkowników w postaci tekstu jawnego w bazie danych.
2. Nie wprowadzanie żadnej utraty jakości próbek dźwiękowych, zarówno podczas wysyłania ich przez użytkownika w celu zapisu do bazy danych, podczas samego zapisu do bazy danych, jak i podczas odtwarzania ich przez uczestnika testu.
3. Ograniczenie maksymalnego rozmiaru przesyłanych próbek dźwiękowych.

4. Architektura aplikacji pozwalająca na łatwą skalowalność portalu w przypadku zwiększenia ilości uczestników w systemie, a w związku z tym zwiększenia ilości operacji zapisu, odczytu i transmisji. Architektura nie powinna też utrudniać dodawania nowych funkcjonalności do portalu.
5. Wykorzystanie implementowanego przez przeglądarki interfejsu programistycznego do odtwarzania dźwięku – Web Audio API [73]. Definiuje to dozwolone formaty dla próbek dźwiękowych (Tabela 6.1) i specyfikacje techniczną portalu dotyczącą odtwarzanego dźwięku (Tabela 6.2).

**Tabela 6.1.** Wspierane formaty dla popularnych przeglądarek internetowych i ich wersji wydania [52]

<b>Format pliku dźwiękowego</b>	<b>Chrome</b>	<b>Firefox</b>	<b>Opera</b>	<b>Safari</b>
WAV (PCM)	Tak	Tak (3.5)	Tak (10.50)	Tak
MP3	Tak	Tak*	Tak	Tak (3.1)
Vorbis w Ogg	Tak	Tak (3.5)	Tak (10.50)	Nie
Opus w Ogg	Tak (27.0)	Tak (15.0)	Brak danych	Brak danych
FLAC	Nie	Tak	Brak danych	Brak danych

\* Ze względu na prawa patentowe, kodek MP3 nie jest wbudowany w przeglądarkę Firefox. Wsparcie polega na możliwości wykorzystania tego kodeka, jeśli jest zainstalowany w systemie użytkownika [52].

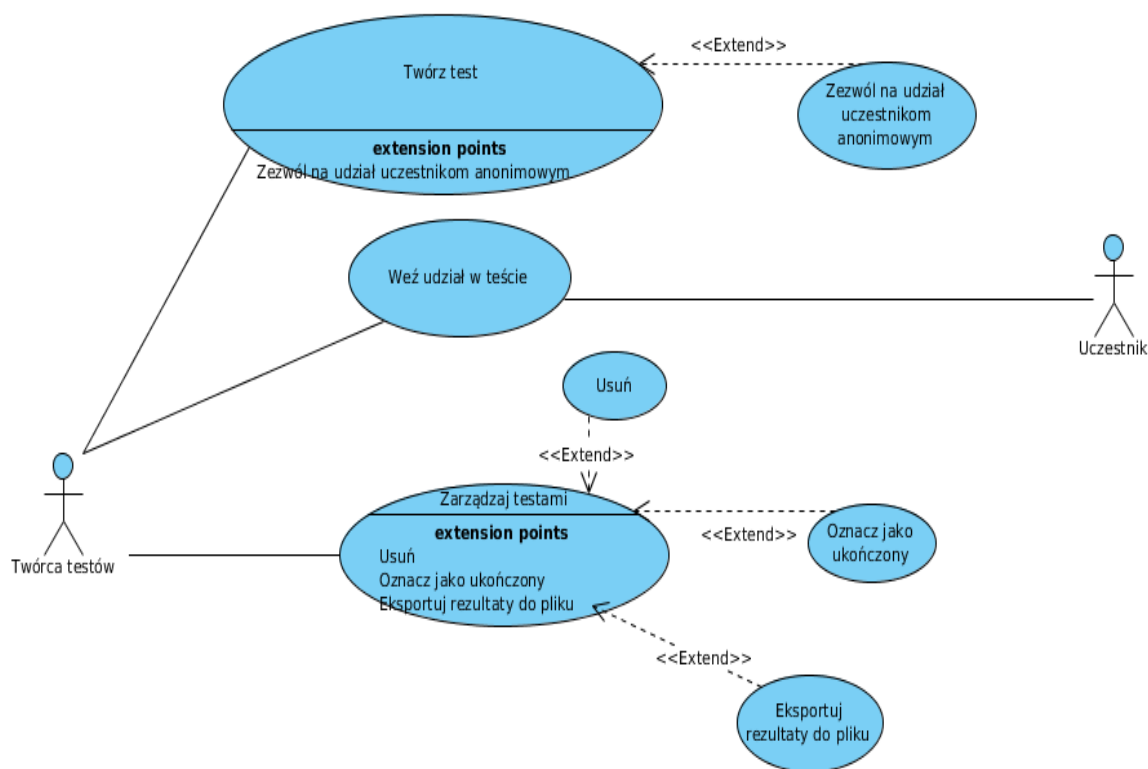
**Tabela 6.2.** Parametry odtwarzania dźwięku w opracowanym portalu [40]

<b>Parametr</b>	<b>Wartość</b>
Ilość kanałów	Do 32 (lub więcej*)
Częstotliwość próbkowania liniowego sygnału dźwiękowego PCM	Zakres od 8000 Hz do 96000 Hz (lub więcej*)
Rozdzielczość bitowa dla sygnału dźwiękowego PCM	Do 32 bitów

\* Specyfikacja określa tylko dolny przedział (wymagane minimum) dla implementowanego przez przeglądarki internetowe interfejsu Web Audio API [40].

## 6.2. Diagram przypadków użycia

Realizowane przez system funkcję i ich interakcje z aktorami opisuje diagram przypadków użycia, przedstawiony na Rys. 6.1.



**Rys. 6.1.** Diagram UML przypadków użycia [praca własna]

W systemie wyróżniono dwóch aktorów:

1. Twórca testów – jest to zarejestrowany użytkownik (posiadający konto użytkownika w portalu), który stworzył test odsłuchowy.
2. Uczestnik – jest to zarejestrowany lub niezarejestrowany użytkownik, który może wziąć udział w teście odsłuchowym. W przypadku uczestnika niezarejestrowanego, brać on może udział tylko w testach w których twórca zezwolił na udział użytkownikom anonimowym.

Dla każdego przypadku użycia, sporządzono opis zawierający informacje na temat uczestniczących aktorów, podstawowego i alternatywnego ciągu zdarzeń, zależności czasowych i wartości uzyskiwanych przez aktorów po zakończeniu przypadku użycia.

Przypadek użycia: **Twórz test:**

1. uczestniczący aktorzy: Twórca testów (Twórca),
2. podstawowy ciąg zdarzeń:
  - 2.1. System wyświetla regulamin dla Twórcy i wstępny formularz w którym podać należy m.in. cel badań i szacunkowy czas trwania testów. W formularzu jest też możliwość zezwolenia na udział w testach uczestnikom anonimowym (niezalogowanym).
  - 2.2. Po uprzedniej akceptacji regulaminu i wypełnieniu formularza wstępnego, system wyświetla formularz pozwalający na stworzenie grupy testowej.
  - 2.3. Twórca wypełnia i zatwierdza formularz.
3. alternatywny ciąg zdarzeń:
  - 3.1. W przypadku niespełnienia określonych wymogów walidacyjnych formularza, Twórca powiadamiany jest o tym stosownym komunikatem.
4. zależności czasowe:
  - 4.1. częstotliwość wykonania: nieokreślona.
  - 4.2. przewidywane spiętrzenia: brak.
  - 4.3. typowy czas realizacji: 10 minut.
  - 4.4. maksymalny czas informacji: nieokreślony.
5. wartości uzyskiwane przez aktorów po zakończeniu przypadku użycia:
  - 5.1. Adres URL pozwalający na wzięcie udziału w danym zestawie testowym. Adres ten, Twórca może przesłać potencjalnym uczestnikom.
  - 5.2. Wpis grupy testowej dla Twórcy w menadżerze zarządzania testami.

Przypadek użycia: **Weź udział w teście:**

1. uczestniczący aktorzy: Uczestnik, Twórca,
2. podstawowy ciąg zdarzeń:
  - 2.1. Uczestnikowi, po kliknięciu w odnośnik prowadzący do grupy testów (uprzednio otrzymany od Twórcy, bądź wyszukany i wybrany z listy testów), prezentowany jest regulamin i krótki formularz mający na celu charakteryzację danego uczestnika pod względem m.in. płci i wieku.
  - 2.2. Po akceptacji regulaminu i zatwierdzeniu formularza, automatycznie następuje pobieranie próbek dźwiękowych na urządzenie Uczestnika wymaganych do przeprowadzenia pierwszego testu z grupy.

- 2.3. Po pobraniu próbek, system wyświetla interfejs stosowny do danej metody testowej i Uczestnik wykonuje test odsłuchowy zgodnie z zamieszczoną instrukcją. Jeżeli dana grupa testowa, zawiera więcej niż jeden test, pobierane są w tym momencie próbki dźwiękowe do przeprowadzenia następnego testu.
- 2.4. W przypadku ukończenia danej próby (iteracji), Uczestnik zatwierdza rezultaty. Jeżeli test wymaga więcej niż jedną iterację, następuje losowanie nowej kolejności próbek dźwiękowych i wykonywana jest kolejna iteracja.
- 2.5. Po zakończeniu wszystkich iteracji danego testu, rezultaty są zapisywane do bazy danych. Jeżeli grupa testowa zawiera kolejny test, następuje powrót do kroku 2.3.
- 2.6. Po zakończeniu wszystkich testów z danej grupy testowej, Uczestnikowi wyświetlana jest informacja z podziękowaniem za udział i prośbą o wypełnienie formularza, w którym istnieje możliwość zgłoszenia uwag, propozycji i błędów. W formularzu zawarte są także pytania dotyczące prawidłowego działania interfejsu.
3. alternatywny ciąg zdarzeń:
  - 3.1. Uczestnik rezygnuje w trakcie wykonywania testu. W przypadku rezygnacji przed zapisem rezultatów do bazy danych (co odbywa się po wykonaniu ostatniej iteracji danego testu) uzyskane rezultaty są tracone.
4. zależności czasowe:
  - 4.1. częstotliwość wykonania: nieokreślona,
  - 4.2. przewidywane spiętrzenia: nieokreślono,
  - 4.3. typowy czas realizacji: 15 minut,
  - 4.4. maksymalny czas realizacji: nieokreślony.
5. wartości uzyskiwane przez aktorów po zakończeniu przypadku użycia:
  - 5.1. Wpis uzyskanych rezultatów dla wykonanej grupy testowej Twórcy.



Przypadek użycia: **Zarządzaj testami:**

1. uczestniczący aktorzy: Twórca testów (Twórca),
2. podstawowy ciąg zdarzeń:
  - 2.1. Po wybraniu zakładki menu Menadżer Testów, prezentowana jest Twórcy lista należących do niego grup testowych.
  - 2.2. Po wybraniu danej grupy testowej, system wyświetla szczegółowe informacje na jej temat, w tym m.in. wszystkie testy, nazwy próbek dźwiękowych, ich rezultaty w postaci podstawowych statystyk. Twórca ma też możliwość wyeksportowania rezultatów danego testu do pliku i zapisania go na swoim urządzeniu. Twórca może oznaczyć daną grupę testową jako ukończoną (np. gdy zebrał wystarczającą liczbę wyników). Ma też możliwość usunięcia danej grupy testowej z bazy danych.
3. alternatywny ciąg zdarzeń: brak,
4. zależności czasowe:
  - 4.1. częstotliwość wykonania: nieokreślono,
  - 4.2. przewidywane spiętrzenia: po wykonaniu grupy testowej przez uczestnika,
  - 4.3. typowy czas realizacji: nieokreślono,
  - 4.4. maksymalny czas realizacji: nieokreślono.
5. wartości uzyskiwane przez aktorów po zakończeniu przypadku użycia:
  - 5.1. W przypadku eksportu rezultatów do pliku – plik z rezultatami na urządzeniu fizycznym,
  - 5.2. W przypadku usunięcia grupy testowej – usunięcie grupy testowej z bazy danych systemu,
  - 5.3. W przypadku zakończenia grupy testowej – zablokowanie możliwości uczestnictwa w danej grupie testowej.

### 6.3. Architektura aplikacji

W celu łatwiejszego skalowania aplikacji w przyszłości i dodawania nowych funkcjonalności wybrano architekturę bazującą na mikroservisach – małych, autonomicznych usługach (serwisach) pracujących razem [54]. Aplikację wykonano również w oparciu o wzorzec architekuralny REST (ang. *Representational State Transfer*) [19]. Komunikacja pomiędzy serwisami odbywa się z wykorzystaniem protokołu HTTP (ang. *Hypertext Transfer Protocol*) [18]. Formatem wymiany danych jest JSON

(ang. *JavaScript Object Notation*) [9]. Na zaimplementowaną architekturę aplikacji składają się następujące mikroserwisy [54]:

1. serwis konfiguracyjny,
2. serwis katalogujący,
3. serwis bramy,
4. serwis API (ang. *Application Programming Interface*),
5. serwis UI (ang. *User Interface*),
6. serwis uwierzytelniający,
7. serwis administratorski, służący do zarządzania i monitorowania działania pozostałych serwisów.

Zadaniem serwisu konfiguracyjnego jest dostarczanie konfiguracji startowej pozostałym serwisom. Dzięki takiemu podejściu, dane konfiguracyjne np. dane niezbędne do połączenia do bazy danych, czy nazwa i port dla danego serwisu scentralizowane są w jednym miejscu i mogą być łatwo zmieniane [54]. Pozostałe serwisy muszą znać adres serwisu konfiguracyjnego, aby możliwe było zapytanie o dane konfiguracyjne podczas ich uruchamiania.

Serwis katalogujący dostarcza dynamiczną usługę katalogowania mikroservisów. Każdy serwis podczas swojego uruchomienia rejestruje się w serwisie katalogującym, podając dane potrzebne do wyszukania serwisu, takie jak nazwę hosta i port. Rejestrujący się serwis podaje też nazwę własną, która zestawiona zostaje z danymi potrzebnymi do wyszukania go w sieci. Dzięki takiemu rozwiązaniu mikroserwisy nie muszą znać dokładnych adresów pozostałych, aby się z nimi porozumieć. Wystarczy, że znają tylko nazwę własną serwisu i mają dostęp do serwisu katalogującego [54, 56].

Serwis bramy pełni funkcję odwrotnego proxy dla zapytań pochodzących z serwisu UI tzn. na podstawie adresu otrzymanego zapytania, serwis bramy decyduje gdzie zapytanie przekazane ma być dalej [3]. Dzięki temu serwis UI nie kontaktuje się bezpośrednio z pozostałymi serwisami, a z jednym, co ułatwia m.in. jego implementację. Serwis bramy kontaktuje się z resztą serwisów w imieniu klienta.

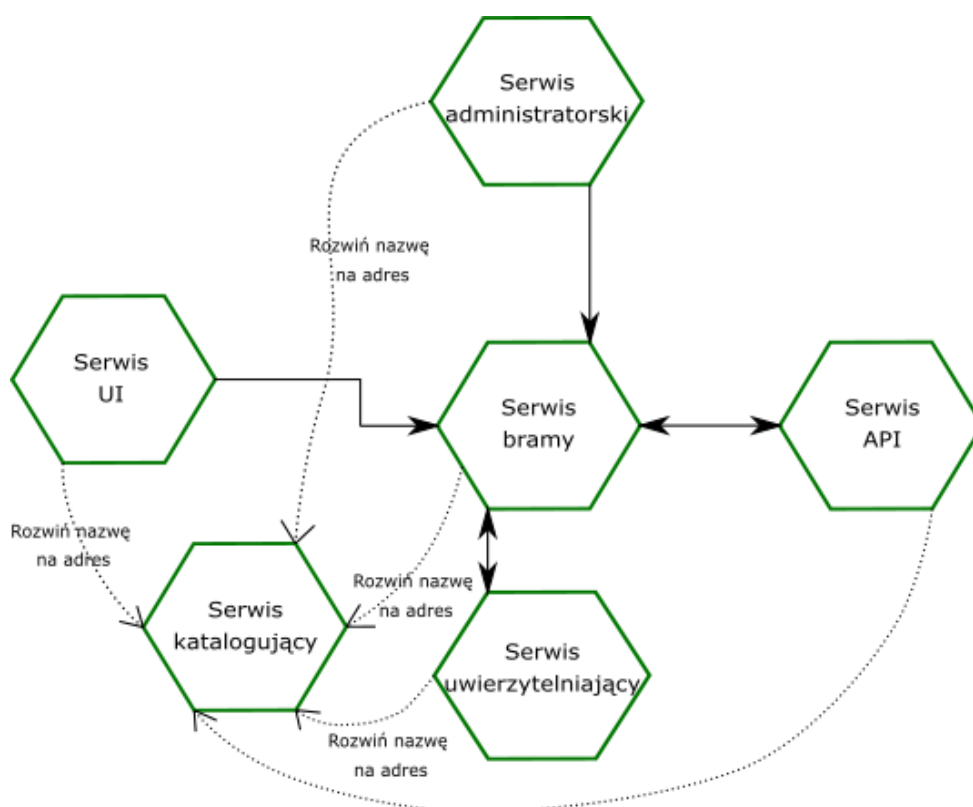
W serwisie API zaimplementowano warstwę dostępu do danych i logikę serwerową aplikacji. Serwis API odpowiada też za autoryzację dostępu do danych. Serwis ten wystawia publiczny interfejs programistyczny aplikacji, który umożliwia dostęp do zasobów i operacje na nich.

W serwisie UI zaimplementowano interfejs użytkownika aplikacji. Serwis UI korzysta z interfejsu programistycznego aplikacji serwisu API i serwisu uwierzytelniającego.

Serwis uwierzytelniający odpowiada za uwierzytelnienie użytkownika na podstawie przesłanego zapytania z loginem i hasłem.

Serwis administratorski służy do zarządzania i monitorowania działania pozostałych serwisów.

Architekturę aplikacji przedstawiono na Rys. 6.2. [44].



**Rys. 6.2.** Architektura aplikacji, opracowano na podstawie [44]

## 6.4. Wykorzystane platformy programistyczne i narzędzia pracy

### 6.4.1. Platformy programistyczne

Do implementacji aplikacji wykorzystano następujące platformy programistyczne:

1. Spring Boot (wersja 1.4.1) – do budowy części infrastrukturalnej i warstwy dostępu do danych (ang. *Backend* [69]). Na platformę zdecydowano się ze względu na to, że umożliwia łatwą, przeprowadzaną w sposób deklaracyjny

konfigurację, wstrzykiwanie zależności oraz posiada wiele modułów służących zabezpieczeniu aplikacji, a także umożliwieniu pracy w środowisku rozproszonym [57].

2. Angular 2 (wersja 2.4.1) – do budowy warstwy prezentacyjnej portalu (ang. *Frontend* [69]) opartej o implementowane przez przeglądarki procesory HTML i CSS, a także język JavaScript. Platformę wybrano, ze względu na możliwość łatwiejszej pracy z typowanym językiem TypeScript, tworzenie uniwersalnych komponentów i wykorzystanie modelu Single Page Application [21].

Jako system zarządzania relacyjną bazą danych wykorzystano PostgreSQL w wersji 9.5 [71].

W celu uniknięcia początkowej, żmudnej konfiguracji mającej na celu integrację między platformą Spring Boot i Angular 2, jako szkielet dla aplikacji wykorzystano Spring Boot Angular 2 Starter [44] .

#### **6.4.2. Narzędzia**

Jako systemu kontroli wersji użyto programu Git w wersji 2.7.4. Git jest rozproszonym systemem kontroli wersji, czyli systemem rejestracji zmian w plikach, pozwalającym m.in. na późniejsze przywoływanie do dowolnej z nich [66].

Do tworzenia i modyfikowania kodu źródłowego wykorzystano zintegrowane środowisko programistyczne IntelliJ IDEA w wersji 16.3 [36].

Do zarządzania bibliotekami i budowy modułów użyto programu Apache Maven w wersji 3.3.9 [4] i Webpack w wersji 2.2.4 [75].

### **6.5. Proces implementacji**

Pełny kod źródłowy opracowanego portalu zamieszczony jest na płycie CD dołączonej do pracy. Zarówno instrukcja użytkownika, jak i instrukcja instalacji są dołączone do pracy, kolejno jako załącznik 1 i załącznik 2.

#### **6.5.1. Struktura bazy danych i opis struktur danych**

Wszystkie tabele bazy danych posiadają pole *id* będące kluczem podstawowym każdej z nich. Wszystkie tabele główne tzn. takie, które nie są tabelami łączącymi posiadają pole *uuid* (obecnie niewykorzystywane). Pole to ma być globalnym identyfikatorem danego rekordu, w przypadku posiadania więcej niż jednej bazy danych.

Większość tabel, których zawartość ma być nadzorowana, przypisywana danemu użytkownikowi lub może być przez użytkownika modyfikowana posiada następujące pola:

1. *created\_by* – nazwa użytkownika tworzącego rekord (zarządzane automatycznie),
2. *created\_date* – data utworzenia rekordu (zarządzane automatycznie),
3. *updated\_by* – nazwa użytkownika modyfikującego rekord (zarządzane automatycznie),
4. *updated\_date* – data modyfikacji (zarządzane automatycznie),
5. *version* – pole wersji, wykorzystywane przy operacjach transakcyjnych w trybie optymistycznego blokowania [60] (zarządzane automatycznie).

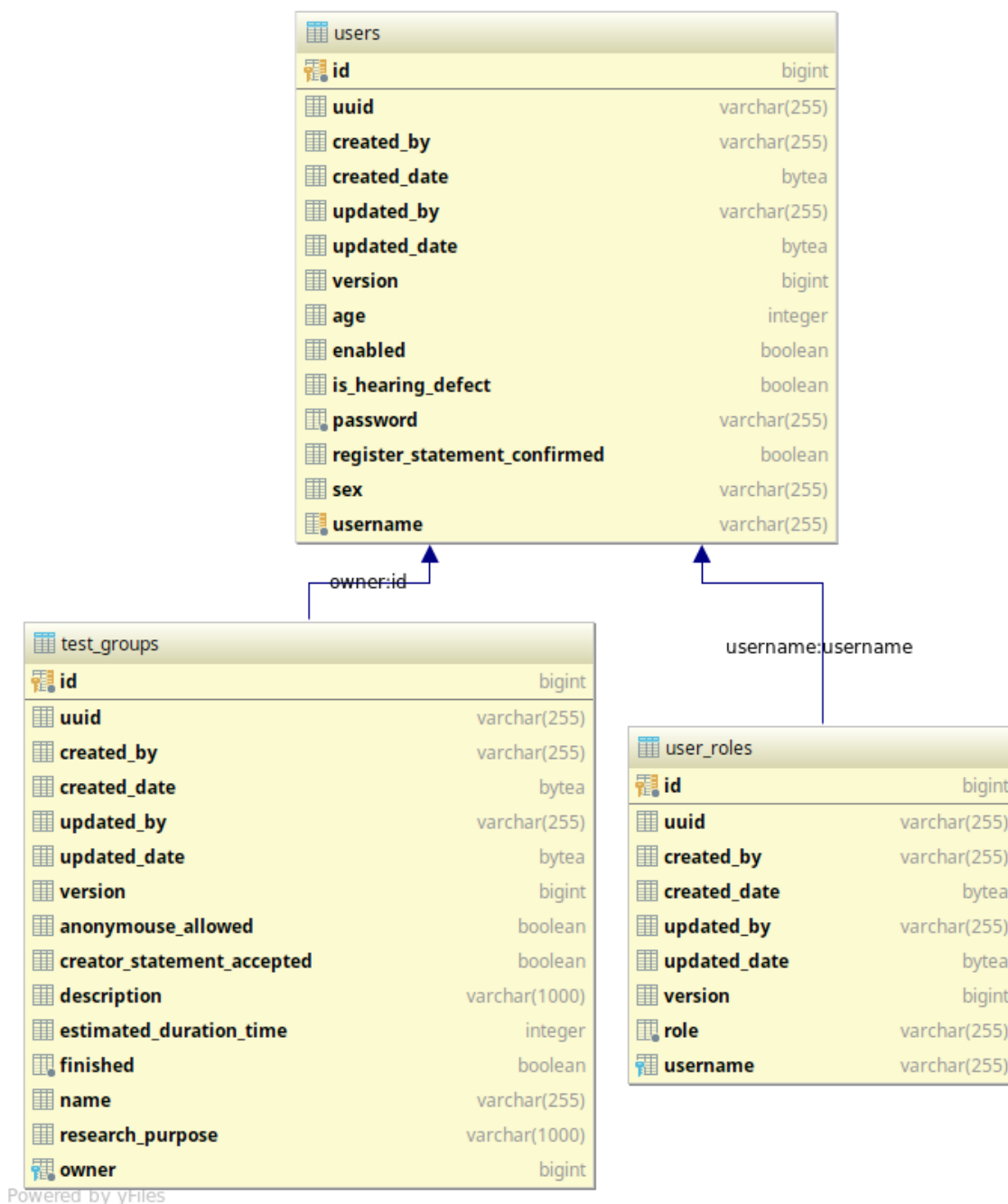
Tabela użytkowników (*users*) zawiera dane zarejestrowanych użytkowników, takie jak login i hasło. Tabela zawiera też pola takie jak wiek i płeć (Pola te nie są wymagane przy rejestracji, aczkolwiek istnieje możliwość podania ich w formularzu rejestracyjnym. Dane te są wymagane przed rozpoczęciem każdego testu). Pole dotyczące zaakceptowania regulaminu przy rejestracji, dotyczącego danych osobowym, pełni dodatkową rolę walidacyjną po stronie serwerowej. Tabela połączona jest relacją jeden do wielu z tabelą ról (*roles*) i z tabelą grup testów (*test\_groups*). W tabeli ról przechowywane są role użytkowników w serwisie (użytkownik – rola domyślna lub administrator).

W tabeli grup testów (*test\_groups*) przechowywane są grupy testowe. Pojedyncza grupa testowa składa się z pól takich jak nazwa grupy, szacunkowy czas trwania (w minutach), opis grupy i cel badań. Przechowane są tam też informacje czy testy z danej grupy mogą być wykonywane przez użytkowników anonimowych (niezalogowanych) i czy twórca testów zaakceptował regulamin twórcy (pole walidacyjne po stronie serwera). Grupa testowa połączona jest relacją jeden do wielu z tabelą testów (*audio\_tests*). Pola tabeli *audio\_tests* opisują pojedynczy test odsłuchowy, składający się z nazwy, opisu (instrukcji), kolejności w grupie i ilości iteracji. Z tabelą *audio\_tests* połączone są w relacji jeden do jednego tabele *abx\_tests* i *mushra\_tests*, definiujące typ danego testu; w relacji jeden do wielu z tabelami grupy rezultatów (*group\_results*) i próbek dźwiękowych (*test\_samples*). Każda próbka dźwiękowa składa się z nazwy pliku, klucza (nazwy bodźca) i zawartości pliku (poprzez relację jeden do jednego z tabelą *resources*).

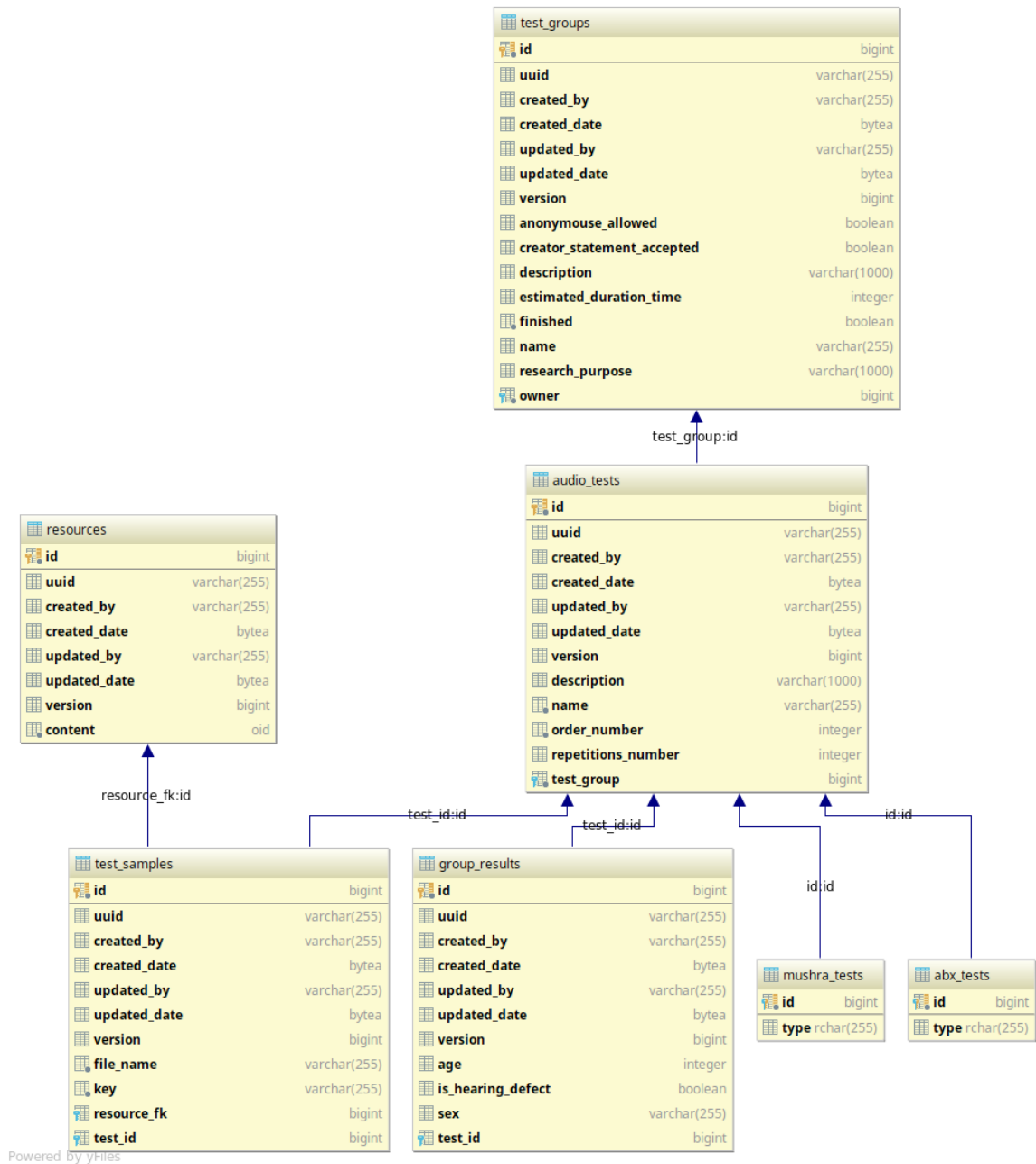
Tabela *group\_results* poprzez relację jeden do wielu z tabelą *test\_results* zawiera grupy rezultatów. Grupa rezultatów to rezultaty wszystkich iteracji (powtórzeń) danego wykonania testu. W skład grupy rezultatów wchodzi cechy charakteryzujące wykonawcę testu takie jak wiek, płeć, a także pojedyncze rezultaty. Każdy pojedynczy rezultat składa

się z numeru iteracji i odpowiednio: dla testu MUSHRA – z grupy wyników dla każdej próbki (nazwa bodźca i ocena), dla testu ABX – z odpowiedzi wykonującego i prawidłowej odpowiedzi.

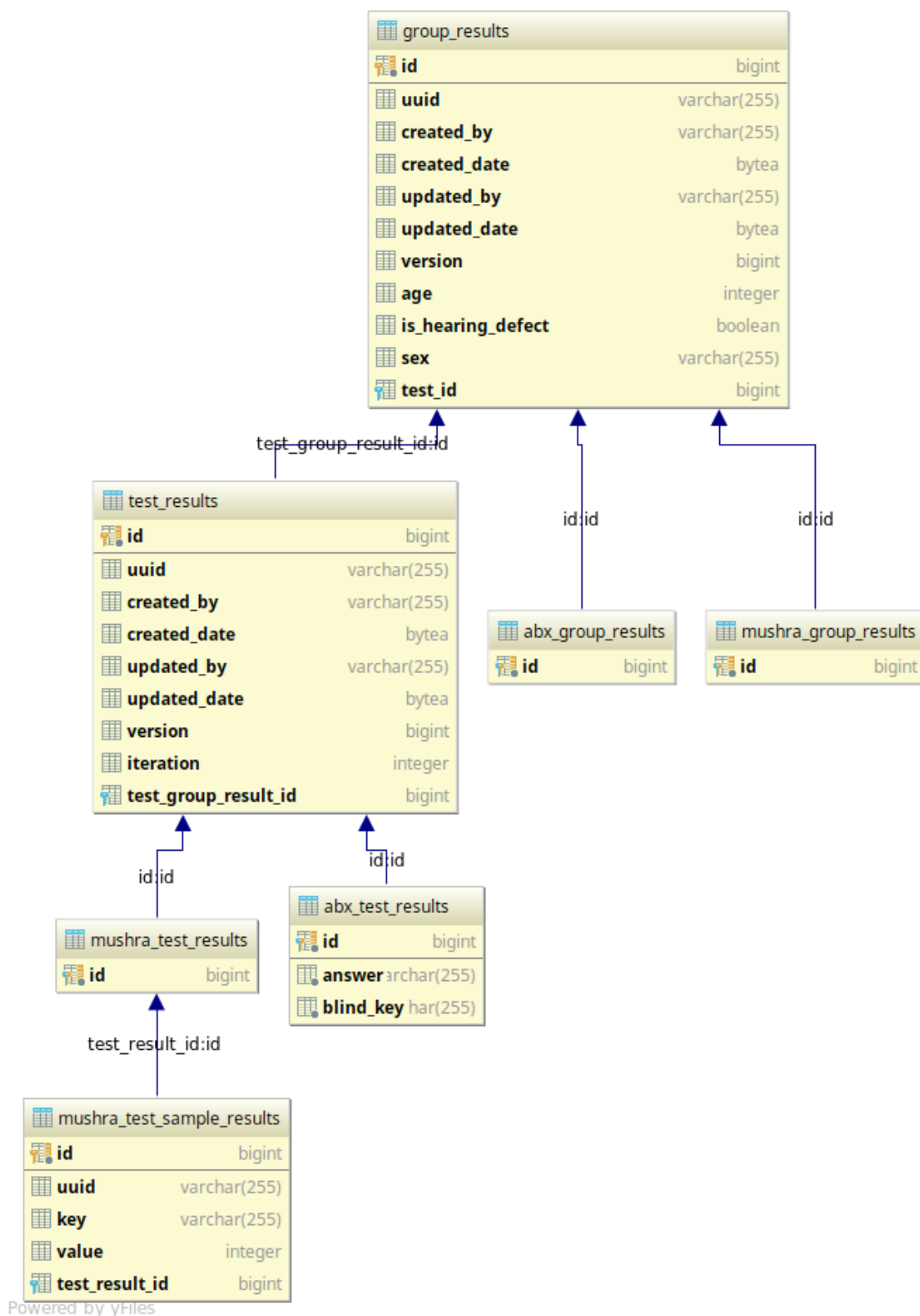
Diagramy ER (ang. *Entity-Relationship Diagram* [14]) tabel bazy danych przedstawione są na Rys. 6.3 (diagram tabeli użytkowników), Rys. 6.4 (diagram tabeli grup testów), Rys. 6.5 (diagram tabeli grup rezultatów).



Rys. 6.3. Diagram ER tabeli użytkowników [praca własna]



Rys. 6.4. Diagram ER tabeli grup testów [praca własna]



Rys. 6.5. Diagram ER tabeli grup rezultatów [praca własna]



### 6.5.2. Proces rejestracji użytkownika, uwierzytelniania i autoryzacji

Rejestracja nowego użytkownika w serwisie polega na przesłaniu zapytania HTTP metodą POST (jedna z metod protokołu HTTP [18]), zawierającego dane wymagane do późniejszego jego uwierzytelnienia (login i hasło), na adres URL `/users` serwisu API. Po poprawnej walidacji danych, tworzone jest konto nowego użytkownika. Hasło użytkownika przy zapisie do bazy danych zamieniane jest algorytmem bcrypt [58] na skrót (ang. *Hash*) z dołączonym ciągiem zaburzającym (ang. *Salt*). W Kodzie źródłowym 6.1 przedstawiono procedurę tworzenia i zapisu do bazy danych konta użytkownika.

Proces uwierzytelniania wykorzystuję standard OAuth 2.0 [55] i token JWT (ang. *JSON Web Token*) [6]. Użytkownik chcący uwierzytelnić się w serwisie wysyła login i hasło do serwisu uwierzytelniania. W przypadku poprawnych danych uwierzytelniających, serwis zwraca użytkownikowi jednorazowe hasło – token JWT. Serwis uwierzytelniający podpisuje token algorytmem HMAC SHA256 [42] z wykorzystaniem własnego klucza prywatnego RSA [62] i samej treści tokena. Serwis API, znając klucz publiczny RSA serwisu uwierzytelniającego, weryfikuje poprawność podpisu. Rozwiązanie takie uniemożliwia podrobienie tokena (w praktyce należałoby zdobyć klucz prywatny serwisu uwierzytelniającego). Dowolna modyfikacja jego treści (np. dodanie nowej roli) spowoduje niepoprawną weryfikację podpisu. Token zapisywany jest w pamięci przeglądarki użytkownika i wysyłany z każdym zapytaniem o określony zasób do serwisu API. Weryfikowana jest też data ważności tokena. Następnie, serwis API udostępnia (przesyła) dany zasób lub wykonuje zadaną operację, bądź nie.

Autoryzacja do zasobów bazuje na uwierzytelnieniu użytkownika (użytkownik uwierzytelniony – zalogowany, ma możliwość tworzenia testów, niezalogowany takiej możliwości nie ma). Dodatkowo autoryzacja do zasobu który jest przypisany tylko do danego użytkownika, jak np. stworzona grupa testowa i prawo do jej usunięcia, wykonywana jest poprzez sprawdzenie, czy przesyłana wraz z zapytaniem nazwa użytkownika (zawarta w tokenie JWT) odpowiada wartości pola `created_by` obiektu.

## Kod źródłowy 6.1. Procedura tworzenia konta użytkownika w serwisie API

[praca własna]

---

```
/* Po odbiorze zapytania HTTP z metodą POST (adres /users) i poprawnej
walidacji danych, tworzone jest konto nowego użytkownika, po czym następuje
jego zapis do bazy danych. Przed zapisem do bazy danych hasło użytkownika
zamieniane jest na skrót. Po poprawnym zapisie do bazy danych, metoda zwraca
odpowiedź HTTP ze statusem 201 (CREATED) i zawartością stworzonego
obiektu konta. */
@RequestMapping(method = RequestMethod.POST, path = "/users")
@ResponseBody
public ResponseEntity<UserEntity> create(@RequestBody UserDto userDto) {
    UserEntity user = userService.save(userDto);
    return ResponseEntity
        .created(URI.create("/users/" + user.getId()))
        .body(user);
}
/* Metoda warstwy serwisowej, tworząca konto użytkownika i wykonująca zapis
do bazy danych. */
@Override
@Transactional
public UserEntity save(UserDto userDto) {
    UserEntity user = new UserEntity(userDto.getUsername(),
        passwordEncoder.encode(userDto.getPassword()),
        Sets.newHashSet(RoleType.ROLE_USER),
        userDto.getRegisterStatementConfirmed(),
        new PersonalData(userDto.getSex(), userDto.getAge()));
    return userRepository.save(user);
}
/* Obiekt odpowiadający wymaganej zawartości przesyłanych danych
z formularza przez użytkownika. Walidacja poprawności danych zadana jest
w sposób deklaratywny. Pola nazwy użytkownika i hasła nie mogą być puste
i muszą mieć odpowiednio co najmniej 4 i 6 znaków. */
@Value
public class UserDto {
    @NotNull @Size(min = 4) String username;
    @NotNull @Size(min = 6) String password;
    ...
}
```

---

### 6.5.3. Kreator testów odsłuchowych

Kreator testów odsłuchowych pozwala na stworzenie grupy testowej składającej się z wielu testów. Użytkownik, po akceptacji regulaminu twórcy, ma do wyboru dwie metody testowe tj. MUSHRA i Test ABX, dla każdego z tworzonych testów. W przypadku metody

MUSHRA może wybrać do 12 próbek dźwiękowych. Użytkownik ma możliwość wybrania próbek dźwiękowych (plików) ze swojego urządzenia fizycznego.

Aby nie przysyłać jednego formularza o bardzo dużej wielkości (razem ze wszystkimi próbkami dźwiękowymi), próbki dźwiękowe wysyłane są do serwisu API, bezpośrednio po ich każdorazowym wybraniu przez użytkownika. Po poprawnym zapisaniu ich w pamięci trwałej na serwerze, użytkownik otrzymuje unikalny identyfikator próbki (*uuid*). Po poprawnym wypełnieniu formularza (walidacji po stronie klienta), obok danych niezbędnych do stworzenia grupy testowej, takich jak nazwa grupy, nazwy testów, instrukcje itd., przesyłane są identyfikatory próbek. Po poprawnej walidacji otrzymanego zapytania (walidacji po stronie serwera), pliki, razem z grupą testową zapisywane są w bazie danych. Wielkość pojedynczej próbki ograniczona jest do 100 MB.

#### **6.5.4. Odtwarzacz testów odsłuchowych**

Do zaimplementowania interfejsu metod testowych wykorzystano komponenty HTML z bibliotek Angular Material [20], jQuery UI [70] i Bootstrap [8]. Wykorzystano także komponent Audio HTML5 i Web Audio API [73].

Interfejs stosowny do danej metody testowej wyświetlany jest po pobraniu próbek dźwiękowych dla danego testu i po ustaleniu losowo, kolejności przypisania próbek do widocznych dla uczestnika przycisków. W przypadku kiedy grupa testowa posiada więcej niż jeden test, próbki dla testu kolejnego pobierane są podczas wykonywania testu aktualnego. Rozwiązanie takie zmniejsza czas oczekiwania na możliwość wykonania kolejnego testu z grupy. Każda pobrana próbka zapisywana jest w pamięci przeglądarki jako obiekt typu BLOB (ang. *Binary Large Object*) z właściwym sobie typem (np. wav).

Wspólna część interfejsu (taka sama dla każdej metody testowej) to część odtwarzacza plików dźwiękowych. Interfejs ten pozwala na zmianę stanu głośności, zatrzymanie bądź wznowienie odtwarzania, ograniczenie przedziału odtwarzania z lewej i prawej strony, zatrzymanie lub wznowienie odgrywania próbki dźwiękowej, włączenie lub wyłączenie trybu synchronicznego odtwarzania (przejście pomiędzy dźwiękami odbywa się z zachowaniem aktualnej pozycji w czasie odtwarzanej próbki; nie powoduje odtwarzania nowej próbki od początku przedziału) i włączenia lub wyłączenia zapętlania odtwarzania.

Dla każdej pobranej próbki dźwiękowej, tworzony jest osobny obiekt posiadający klucz (nazwę bodźca), źródło pliku dźwiękowego (adres obiektu BLOB), element Audio HTML5, i lokalny wzmacniacz głośności. Definicja takiego obiektu przedstawiona

jest w Kodzie źródłowym 6.2. Dodatkowo każdy taki obiekt jest duplikowany (posiada wszystkie dane takie same, zmieniany jest mu tylko klucz np. z "A" na "A\_DUP").

**Kod źródłowy 6.2.** Definicja obiektu AudioObject [praca własna]

---

```
export interface AudioObject {  
  key: string,  
  audioSource: MediaElementAudioSourceNode  
  audioElement: HTMLAudioElement  
  gainNode: GainNode  
}
```

---

W przypadku wciśnięcia przycisku o danej nazwie (np. "A") wyszukiwany jest obiekt próbki z kluczem o wartości "A". Istnieją trzy główne przypadki stanu dla odtwarzacza:

1. żadna próbka dźwiękowa nie była wcześniej odtwarzana,
2. próbka dźwiękowa jest odtwarzana, a wciśnięty bodziec ma spowodować odtworzenie innej próbki i wyłączenie aktualnie odtwarzanej (przejście),
3. odtwarzana próbka kończy się i następuje ponowne jej odtworzenie od początku przedziału.

W pierwszym przypadku, wyszukana próbka dźwiękowa odtwarzana jest wraz z liniowym zwiększaniem poziomu głośności w czasie. Wzmocnienie zmienia się w przeciągu czasu (w przybliżeniu)  $currentTime + 60ms$  (wartość parametru *fadeIn*) od momentu wybrania bodźca, od  $-\infty$  dB do 0 dB.

W drugim przypadku pomiędzy dwoma dźwiękami następuje przenikanie. W tym samym momencie, wzmocnienie dźwięku aktualnie odtwarzanego, liniowo maleje od 0 dB do  $-\infty$  dB, w czasie 60ms (wartość parametru *fadeOut*), zaś wzmocnienie dźwięku wybranego, liniowo wzrasta od  $-\infty$  dB do 0 dB. Sytuacja przejścia przedstawiona jest na Rys. 6.6. Kod źródłowy 6.3 przedstawia proces przełączania próbek, podczas gdy włączone jest odtwarzanie synchroniczne.

**Kod źródłowy 6.3.** Proces przełączania próbek z zachowaniem czasu aktualnie  
odgrywanej próbki [praca własna]

---

```
...
// aktualny czas kontekstu odtwarzania
let ctxCurrentTime = this.audioContext.currentTime;
//zachowanie czasu aktualnie odgrywanej próbki
nextAudioObject.audioElement.currentTime =
currentAudioObject.audioElement.currentTime;
// ustawienie liniowej zmiany wzmocnienia dźwięku próbki następnej,
// od poziomu minimalnego do maksymalnego, w czasie fadeInTime
nextAudioObject.gainNode.gain.
  linearRampToValueAtTime(1.0,ctxCurrentTime + this.params.fadeInTime);
// ustawienie liniowej zmiany wzmocnienia dźwięku próbki aktualnie
// odgrywanej,
//od poziomu maksymalnego do minimalnego, w czasie fadeOutTime
currentAudioObject.gainNode.gain.
  linearRampToValueAtTime(0.0,ctxCurrentTime + this.params.fadeOutTime);
// włączenie odtwarzania próbki następnej
nextAudioObject.audioElement.play();
// wstrzymanie odtwarzania próbki aktualnej (z opóźnieniem)
this.doDelayedPause(currentAudioObject.audioElement,
this.params.fadeOutTime + 100);
...
```

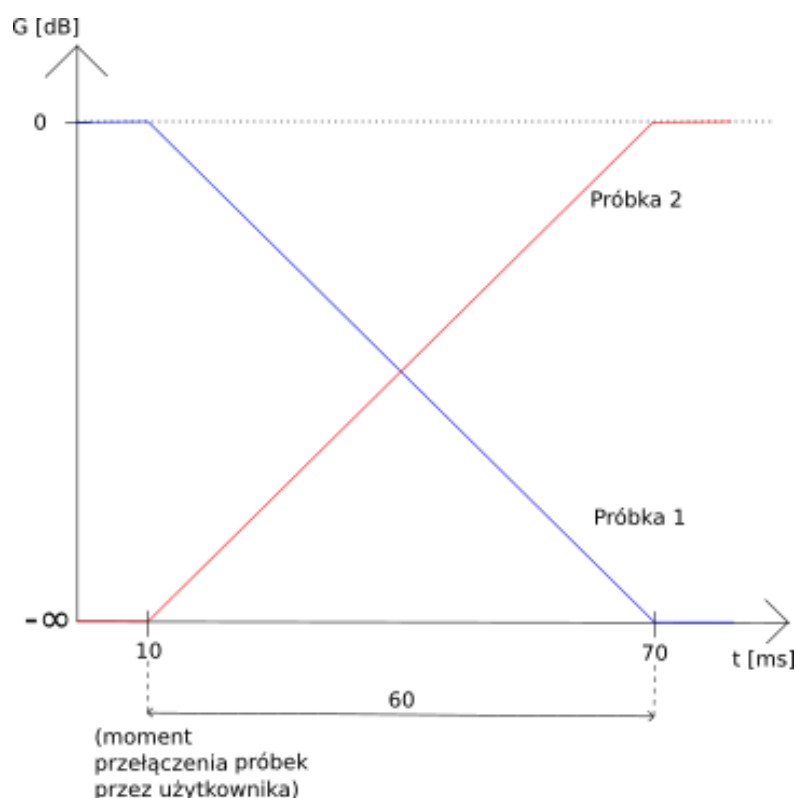
---

Do obsługi przypadku trzeciego wykorzystywana jest funkcja wykonująca się w czasie odtwarzania dowolnej próbki co około 10ms (wartość parametru *intervalTime*). Funkcja ta próbkuje aktualny czas odtwarzanego dźwięku i uaktualnia pasek postępu w interfejsie użytkownika, jak i czas odtwarzania. W przypadku, gdy czas odtwarzanego dźwięku jest większy niż wartość definiowana przez prawy koniec przedziału, pomniejszona o wartość parametrów *fadeOut* i *intervalTime*, następuję (w przypadku włączonego zapętlania):

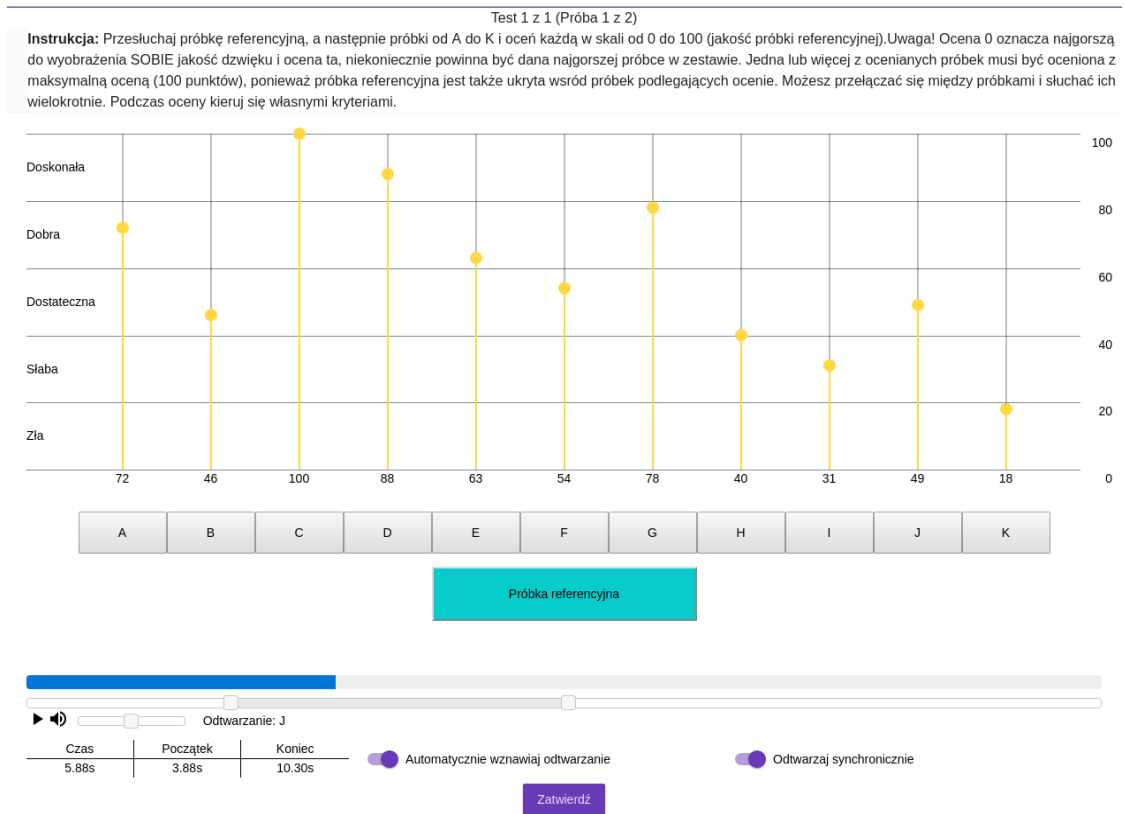
1. wzmocnienie aktualnie odtwarzanej próbki liniowo maleje od 0 dB do  $-\infty$  dB w czasie *fadeOut*,
2. czas odtwarzania duplikatu próbki ustawiany jest na początek przedziału odtwarzania, zaś jego wzmocnienie liniowo wzrasta od  $-\infty$  dB do 0 dB w czasie *fadeIn* (dzieje się to w tym samym czasie co krok pierwszy).

Pomimo, że przenikanie dźwięków przy przełączaniu nie jest rekomendowane w najnowszej edycji standardu MUSHRA, zdecydowano się na nie z powodu niejednostajnego efektu chwilowego włączenia i wyłączenia dźwięku, który utrudniał

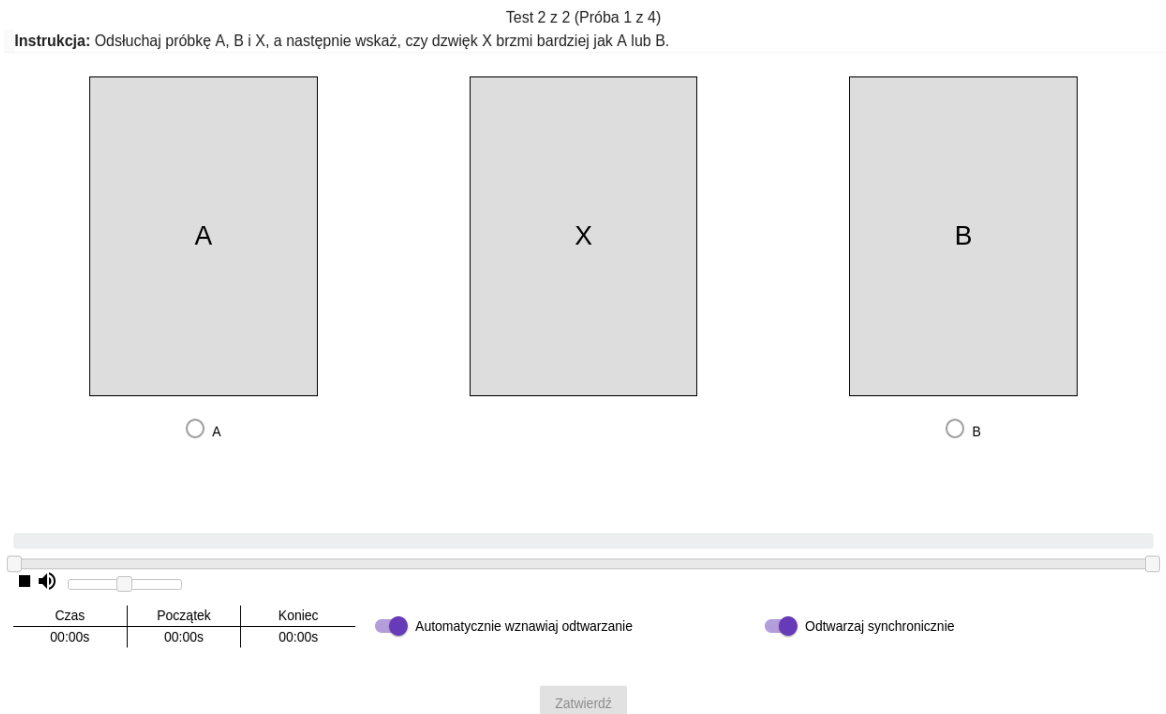
przeprowadzanie testów (co szerzej opisano w rozdziale 6.6). W standardzie ITU-R BS.1534-3 (MUSHRA), przy przełączaniu dźwięków, rekomendowane jest stopniowe wyciszanie (w czasie 5 ms) i stopniowe zwiększanie poziomu głośności (w czasie 5 ms) z wykorzystaniem funkcji kosinusoidy "podniesionej" (ang. *Raised-cosine*) [26]. Wartości parametrów *fadeIn* i *fadeOut* dobrane zostały eksperymentalnie, tak aby wyeliminować wyżej wymieniony efekt. Zaimplementowany interfejs dla metody MUSHRA, przedstawiono na Rys. 6.7, zaś dla testu ABX na Rys. 6.8.



**Rys. 6.6.** Wykres przedstawiający zmianę wzmocnienia dźwięków podczas synchronicznego przełączania próbek [praca własna]



**Rys. 6.7.** Zaimplementowany interfejs metody MUSHRA [praca własna]



**Rys. 6.8.** Zaimplementowany interfejs testu ABX [praca własna]

### 6.5.5. Menadżer testów

Menadżer testów pozwala użytkownikowi na sprawdzenie podstawowych statystyk dotyczących stworzonego przez siebie testu, eksport rezultatów do pliku CSV i JSON, a także usunięcie grup testowych lub oznaczenie danej grupy jako ukończona.

Do obliczania statystyk posłużono się funkcjami agregującymi bazy danych i bibliotekami mctad.js [68] (po stronie użytkownika, do obliczania przedziałów ufności dla wartości średnich ocen – dla testu MUSHRA i statystyki  $\chi^2$  – dla testu ABX) i Apache Commons Math [2] (po stronie serwera, do obliczania statystycznej istotności otrzymanych rezultatów w rozkładzie dwumianowym dla testu ABX [11]). Sposób w jaki prezentowane są użytkownikowi statystyki dotyczące testu ABX przedstawiono na Rys. 6.9.

Próbki

Rezultaty

Rezultaty niesformatowane

Odpowiedzi: 12

Prawidłowe odpowiedzi: 8

Procent prawidłowych odpowiedzi: 66.67%

Prawdopodobieństwo uzyskania tylu lub więcej prawidłowych odpowiedzi w rozkładzie dwumianowym B(12, 0.5): 0.19385

Skumulowana statystyka Pearsona $\chi^2$	$\chi^2$ dla p = 0.05	$\chi^2$ dla p = 0.01
1.333	3.841	6.635

Próbka	Odpowiedzi	Nieprawidłowe odpowiedzi	Prawidłowe odpowiedzi	Procent prawidłowych odpowiedzi
A	8	2	6	75.0%
B	4	2	2	50.0%

**Rys. 6.9.** Prezentacja statystyk testu ABX w opracowanym portalu [praca własna]

## 6.6. Testy poprawności implementacji

Do przetestowania poprawności zapisu plików dźwiękowych i właściwego dopasowania do nich numerów identyfikacyjnych, wykorzystano testy jednostkowe w kodzie źródłowym.

Poprawność wykonywania testów odsłuchowych m.in. wykonanie zadanej przez tworzącego test użytkownika ilości iteracji, losowania nowej kolejności próbek, wysyłania rezultatów, obliczania statystyk, jak i eksportu rezultatów do pliku przeprowadzono następująco:

1. Dla metody MUSHRA stworzono i wykonano kilkakrotnie test składający się z 4 próbek dźwiękowych kolejnych głosek: "A" (próbka referencyjna), "B", "C", "D"



i 4 iteracji. Przy wykonywaniu testu próbka brzmiąca, tak jak głoska A oceniona miała być zawsze na 100, B na 80, C na 60 i D na 40.

2. Dla testu ABX stworzono i wykonano kilkakrotnie test składający z próbki dźwiękowej głoski "A" i "B". Przy wykonywaniu testu, w przypadku gdy bodziec X brzmiał jak głoska "A", zaś bodziec "B" również tak jak głoska "A", przypisywano X jako bodziec B.

Dla obu metod powyższe testy wykonano, sprawdzając poprawność właściwego losowania kolejności, po czym sprawdzono poprawność przypisywania wartości w zakładce statystyk menadżera testów. Sprawdzono również poprawność eksportu rezultatów do pliku, poprzez porównanie ich z wartościami w bazie danych.

Podczas weryfikacji poprawności działania wykonywania testów odsłuchowych, zauważono irytujący efekt chwilowego włączenia i wyłączenia dźwięku przy przełączaniu między próbkami. Efekt ten w dużym stopniu wyeliminowano poprzez zaimplementowanie liniowej zmiany głośności w czasie, zwiększenie interwału potrzebnego do mierzenia czasu, jak i zastosowanie szeregu optymalizacji, m.in. zmniejszenie ilości odwoływań do DOM (obiektowego modelu strony tworzonej przez przeglądarkę) i zamknięcie sekcji kodu odpowiadających za przełączanie i zapętlanie, w bloku pętli *while* przerywanej na koniec sekcji. Dzięki takiemu rozwiązaniu, wykonanie tych sekcji kodu nie jest przeplatane z instrukcjami wykonywanymi interwałowo (potrzebne do odmierzania czasu), co miałoby miejsce z powodu jednowątkowego wykonywania programu JavaScript przez przeglądarki internetowe [72].

## 7. Wykorzystanie opracowanego portalu do testów oceniających jakość dźwięku

Opracowany portal posłużył do przeprowadzenia testu jakości dźwięku z wykorzystaniem metody MUSHRA.

### 7.1. Użyte próbki dźwiękowe

Test stworzono, z wykorzystaniem podzbioru 11 próbek dźwiękowych wykorzystanych wcześniej przez Zielińskiego [82] w laboratoryjnych testach obejmujących ocenę jakości dźwięku z użyciem metody MUSHRA. Wykorzystane próbki zostały przefiltrowane z użyciem filtru Czebyszewa typu I, rzędu 13, z maksymalnym zafalowaniem wynoszącym 0.1 dB. Czas trwania każdej próbki to 20 sekund. Dla każdej próbki, częstotliwość próbkowania wynosiła 48 kHz, a rozdzielczość bitowa 16 bitów. Wykaz użytych próbek dźwiękowych przedstawiono w Tabeli 7.1 [82].

**Tabela 7.1** Wykaz użytych w teście próbek dźwiękowych z testu laboratoryjnego *B*, przeprowadzonego przez Zielińskiego [82]

Numer próbki	Kodowanie próbki w systemie	Częstotliwość odcięcia filtru dolnoprzepustowego [kHz]
1 (jawna i ukryta próbka wzorcowa)	A	20
2	B	10
3	C	8
4	D	7
5	E	6.5
6	F	6
7	G	5.5
8	H	5
9	I	4.5
10	J	4
11	K	3.5

## **7.2 Cechy testu i instrukcja użytkownika**

Dla każdego uczestnika, test składał się z 3 prób. Próba pierwsza traktowana była jako próbna, tzn. rezultaty w niej otrzymane nie były brane pod uwagę przy obliczeniu statystyk. Instrukcja użytkownika była następująca: "Przesłuchaj próbkę referencyjną, a następnie próbki od A do K i oceń każdą w skali od 0 do 100 (jakość próbki referencyjnej). Uwaga! Ocena 0 oznacza najgorszą do wyobrażenia sobie jakość dźwięku i ocena ta, niekoniecznie powinna być dana najgorszej próbce w zestawie. Jedna lub więcej z ocenianych próbek musi być oceniona z maksymalną oceną 100 punktów, ponieważ próbka referencyjna jest także ukryta wśród próbek podlegających ocenie. Możesz przełączać się między próbkami i słuchać ich wielokrotnie. Podczas oceny konsekwentnie kieruj się własnymi kryteriami i zachowaj pełną koncentrację"

Każdy uczestnik testu był też powiadomiony o ilości prób i o tym, że próba pierwsza traktowana jest jako próba treningowa. Informacja ta była widoczna przed rozpoczęciem testu.

## 8. Udokumentowanie uzyskanych wyników

Test został wykonany 46 razy, zaś całkowita ilość osób, które wykonały test to około 43 (co najmniej 3 osoby powtórzyły test, po stwierdzeniu, że nie zastosowały się należycie do instrukcji). Osoby biorące udział w teście opisać można jako osoby nie będące specjalistami, ponieważ nie uczestniczyły wcześniej w żadnym formalnym teście odsłuchowym. Ze względu na to, że test wykonywany był zdalnie, warunki eksperymentu nie były kontrolowane w żaden sposób.

Przykład wyeksportowanych danych dla wykonanego testu do pliku w formacie CSV (po usunięciu danych mogących zostać potraktowane jako dane osobowe), przedstawiono na Rys. 8.1.

Analizę danych przeprowadzono w środowisku *RStudio* [63] z wykorzystaniem języka *R* i m.in. biblioteki *dplyr* [76] (do filtrowania danych).

```
index,iteration,sample,value
0,1,G,30
0,1,A,100
0,1,E,52
0,1,B,100
0,1,K,9
0,1,H,23
0,1,F,35
0,1,I,15
0,1,J,5
0,1,C,53
0,1,D,65
0,2,C,100
0,2,F,65
0,2,J,21
0,2,D,67
0,2,B,70
0,2,H,49
0,2,I,16
0,2,E,46
0,2,K,12
0,2,A,100
0,2,G,41
0,3,B,73
0,3,K,14
0,3,H,33
0,3,J,14
0,3,G,49
0,3,I,11
0,3,C,100
0,3,D,60
0,3,E,41
0,3,A,100
0,3,F,36
```

**Rys. 8.1.** Przykładowe rezultaty w formacie CSV (*index* – numer porządkujący wykonanie testu, *iteration* – numer próby, *sample* – kod próbki, *value* – ocena) [rezultaty własne]

Zgodnie z założeniem, dla każdego uczestnika, uzyskane rezultaty z pierwszej próby zostały odrzucone i nie wykorzystano ich do dalszej analizy. Rezultaty nie zawierające ocen z pierwszej próby zostały dalej poddane odsiewowi (ang. *Screening*), w celu wykluczenia danych pochodzących od osób, które m.in. nie zastosowały się do instrukcji. Odsiew rezultatów został dokonany pod względem trzech następujących kryteriów, bazujących na kryteriach zastosowanych przez Zielińskiego w pracy [82]:

1. Zdolność do wykrycia próbki wzorcowej – średnia ocena dla ukrytej próbki referencyjnej (nr 1) powinna wynosić ponad 97. Uwzględniono tu możliwość popełnienia niezamierzonej pomyłki wynikającej z błędnego przeświadczenia użytkownika o przesunięciu suwaka oceny na najwyższą pozycję (ocena 100), podczas gdy faktycznie znajdował się on nieco niżej.
2. Kryterium wykrycia drugiej pod względem jakości próbki dźwiękowej w zestawie (próbki nr 2) – ocena dla próbki nr 2 w zestawie powinna być mniejsza niż ocena dla próbki wzorcowej dla obu prób branych pod uwagę.
3. Kryterium spójności – średni kwadrat błędu z testu analizy wariancji ANOVA, obliczony dla każdego słuchacza oddzielnie, powinien być mniejszy niż 20 (względem 100 punktowej skali oceny). Czynniki klasyfikującymi w teście ANOVA były próbki o numerach: 2, 3, 4, 7, 11. Dla każdego czynnika klasyfikującego ilość obserwacji wynosiła 2 (wartość ocen dla próbek dźwiękowych z próby drugiej i trzeciej).

Zastosowanie każdego następnego kryterium odbywało się na zbiorze rezultatów odsianych przez kryterium poprzedzające. Po zastosowaniu ostatniego kryterium, wynikowy zbiór danych zawierał 24 wykonania testu z 46 na samym początku (do dalszej analizy wzięto zatem około 52% łącznej ilości wykonań). Tak duży odsetek błędnie wykonanych testów mógł mieć związek z brakiem kontroli nad warunkami eksperymentalnymi, jak i brakiem motywacji ze strony niektórych uczestników do zastosowania się do wszystkich zaleceń instrukcji. Podsumowanie odsiewu rezultatów przedstawiono w Tabeli 8.1. Kod źródłowy 8.1 przedstawia procedurę odsiewu wstępnego zbioru danych dla dwóch początkowych kryteriów.

**Tabela 8.1.** Podsumowanie odsiewu rezultatów w teście internetowym [rezultaty własne]

Łączna ilość wykonań testu	Ilość wykonań niepełniających kryterium nr 1	Ilość wykonań niepełniających kryterium nr 2	Ilość wykonań niepełniających kryterium nr 3	Ilość wykonań testu wzięta do dalszej analizy
46	15 (32.6%)	5 (16.1%)	2 (7,7%)	24 (52.2%)
<b>Pozostała ilość</b>	31	26	24	

**Kod źródłowy 8.1.** Procedura odsiewu rezultatów (dwa początkowe kryteria)

[praca własna]

**#Odrzucenie pierwszej próby**

```
data <- data %>%
  filter (iteration > 1)
```

**#Przefiltrowanie zgrupowanych rezultatów po kodzie próbki wzorcowej i jej średniej ocenie z pozostałych dwóch prób. Średnia ocena ma być #większa niż 97**

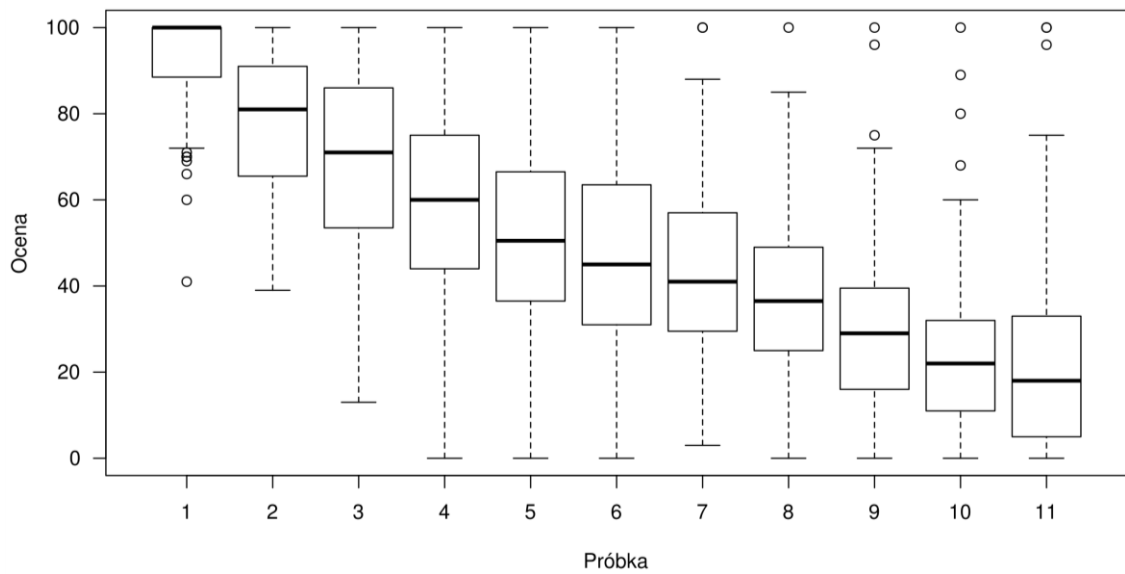
```
stage1 <- data %>%
  group_by(index, sample) %>%
  filter(sample == 'A' & mean(value) > 97)
```

**#Przefiltrowanie pozostałych rezultatów ze względu na wykrycie drugiej pod względem jakości próbki dźwiękowej w każdej próbie danego testu (oprócz pierwszej)**

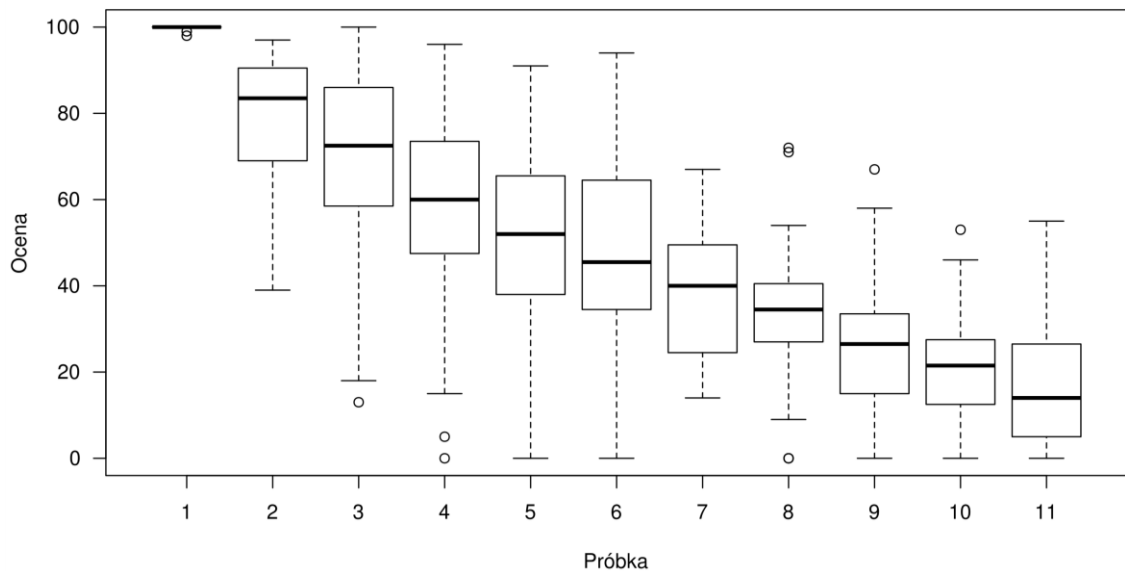
```
stage2 <- data %>%
  filter(index %in% stage1$index) %>%
  group_by(index) %>%
  arrange(index, iteration, sample) %>%
  mutate(disc_passed = discriminate_index(sample, value, iteration)) %>%
  filter(disc_passed == TRUE) %>%
  select(index, iteration, sample, value)
```

Wykresy pudełkowe przedstawiające rozstęp ćwiartkowy (zakres 50% wszystkich obserwacji położonych centralnie w rozkładzie) i mediany ocen dla próbek w zestawie, przed i po procedurze odsiewu przedstawione są na odpowiednio na Rys. 8.2 a i b. Na wykresie pudełkowym prezentującym zbiór danych po procedurze odsiewu (Rys. 8.2 b) można zauważyć, że zmniejszone zostało zróżnicowanie uzyskanych rezultatów względem zbioru danych przed odsiewem, szczególnie dla próbek dźwiękowych o numerach 1, 7 i 8 (Rys. 8.2 a).

a)



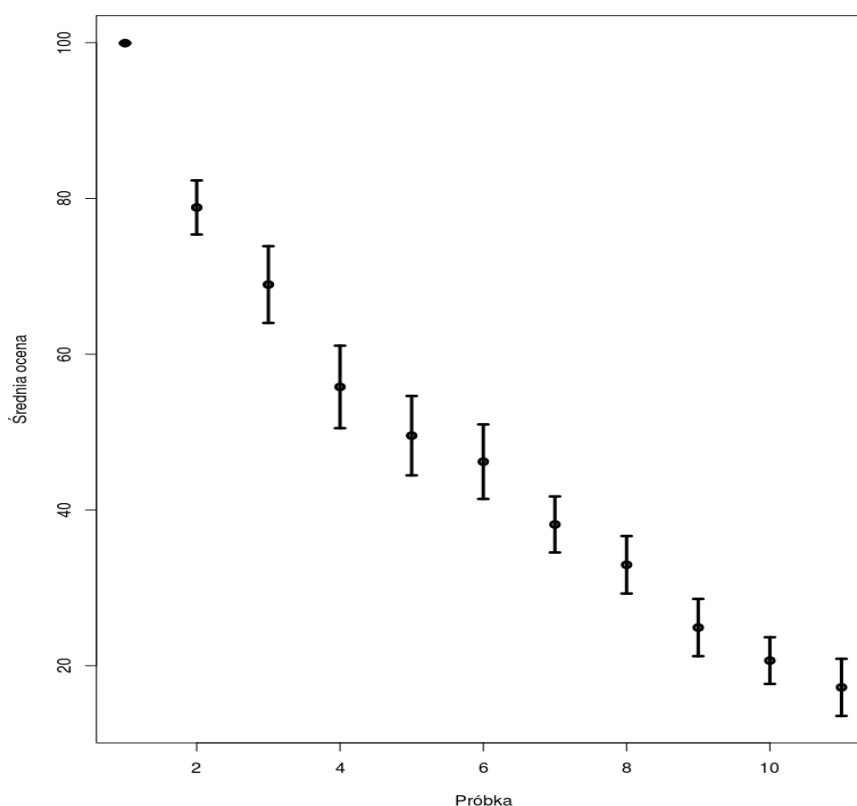
b)



**Rys. 8.2.** Wykresy pudełkowe dla rezultatów przed odsiewem (a) i dla rezultatów po odsiewie danych (b) [praca własna]

Dla wynikowego zbioru danych obliczono następnie średnie wartości ocen dla każdej próbki dźwiękowej w zestawie testowym. Średnie wartości ocen i odpowiadające im 95% przedziały ufności przedstawiono na Rys. 8.3. Na tym wykresie zauważyć można, że wartości średnie ocen maleją wraz ze spadkiem wartości częstotliwości odcięcia filtru dolnoprzepustowego. Wynika stąd, że jakość dźwięku maleje wraz z zawężeniem pasma akustycznego dźwięku (próbka nr 1 odpowiada częstotliwości

odcienia wynoszącej 20 kHz, podczas gdy próbka nr 11 reprezentuje dźwięk o paśmie zawężonym do 3,5 kHz).



**Rys. 8.3.** Średnie wartości ocen i związane z nimi 95 % przedziały ufności dla wynikowego zbioru rezultatów [praca własna]

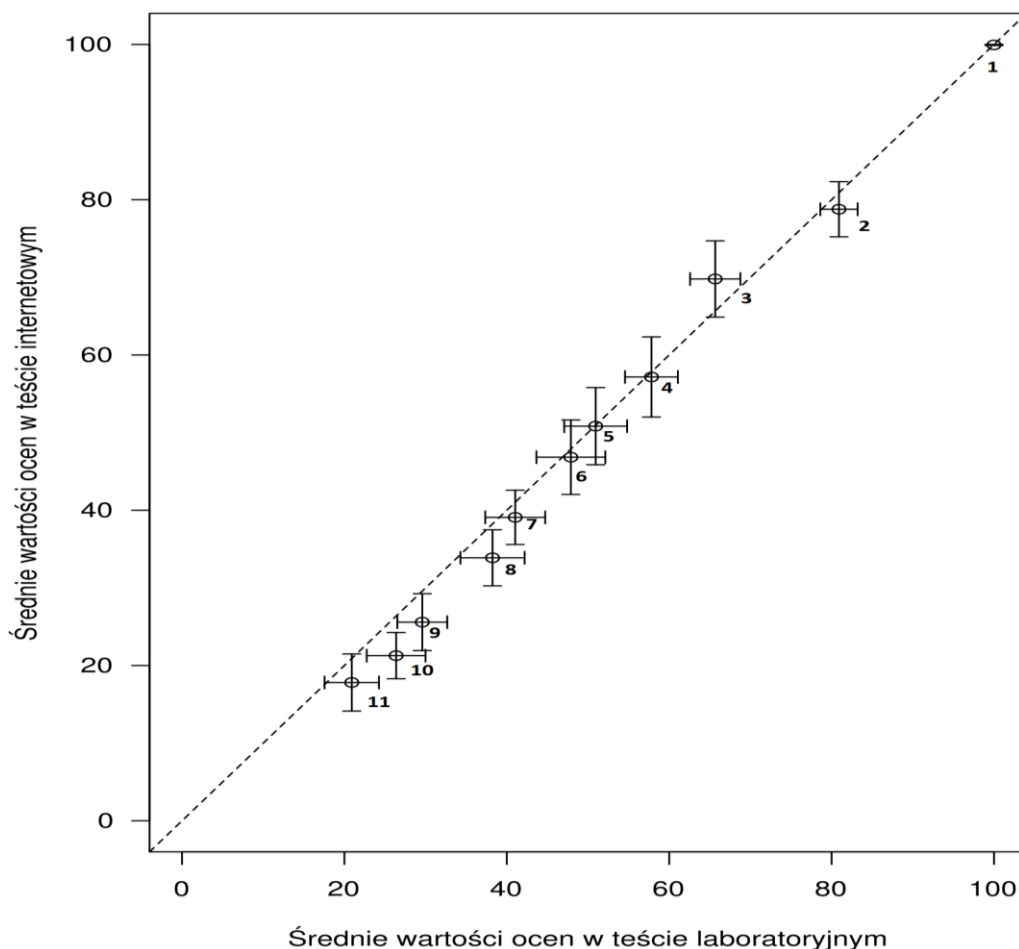
W celu porównania metody internetowej i laboratoryjnej wykorzystano również te same, opisane powyżej, kryteria odsiewu. Kryteria te zastosowano na zbiorze rezultatów uzyskanym w teście laboratoryjnym *B*, przeprowadzonym przez Zielińskiego [82] – wykorzystano tam ten sam zestaw próbek dźwiękowych (Tabela 8.2). Następnie sporządzono wykres punktowy, ilustrujący średnie wartości ocen uzyskanych dla testu laboratoryjnego i testu internetowego, wraz z 95% przedziałami ufności (Rys. 8.4). Wyznaczono również współczynnik korelacji Pearson’a pomiędzy oboma zbiorami wartości średnich (Tabela 8.3).



**Tabela 8.2.** Podsumowanie odsiewu rezultatów testu laboratoryjnego *B*,  
przeprowadzonego przez Zielińskiego [82, kryteria własne]

<b>Łączna ilość wykonań testu (ilość niezależnych słuchaczy)</b>	<b>Ilość wykonań niepełniających kryterium nr 1</b>	<b>Ilość wykonań niepełniających kryterium nr 2</b>	<b>Ilość wykonań niepełniających kryterium nr 3</b>	<b>Ilość wykonań testu wzięta do dalszej analizy</b>
30	2 (6.7%)	0 (0%)	1 (3.5%)	27 (90%)
<b>Pozostała ilość</b>	28	28	27	

Porównując wyniki zaprezentowane w Tabeli 8.1 oraz Tabeli 8.2, zauważyć należy, że procentowo odrzucono dużo więcej rezultatów w teście internetowym, niż w teście laboratoryjnym. W porównaniu do testu internetowego, gdzie do dalszej analizy uwzględniono zaledwie 52% wszystkich wykonań, w teście laboratoryjnym było to aż 90%. Metoda internetowa wprowadziła zatem (procentowo) znacznie więcej rezultatów z błędnie wykonanego testu, niż metoda laboratoryjna. Jednakże, zastosowanie procedury przesiewowej pozwoliło oczyścić oba zbiory danych z rezultatów pochodzących z błędnych wykonań.



**Rys. 8.4.** Wykres punktowy ilustrujący średnie wartości ocen dla testu laboratoryjnego i testu internetowego oraz związane z nimi 95% przedziały ufności [praca własna]

Na wykresie przedstawionym na Rys. 8.4 każdy punkt oznacza średnią wartość oceny dla danej próbki dźwiękowej. Rzutując prostopadłe punkt wykresu i poziome przedziały ufności na oś OX odczytuje się wartość średnią oceny danej próbki (i związany z nią 95% przedział ufności) uzyskanej w teście laboratoryjnym. W przypadku rzutu prostopadłego punktu wykresu i pionowych przedziałów ufności na oś OY odczytuje się wartość średnią oceny próbki (i związany z nią 95% przedział ufności) uzyskanej w teście internetowym. Przerywana linia widoczna na wykresie reprezentuje linię odniesienia  $Y=X$ . Punkt znajdujący się na tej linii oznacza, że wartość średnia oceny danej próbki jest taka sama dla testu laboratoryjnego i testu internetowego. W przypadku kiedy przedziały ufności nakładają się na linię przerywaną, oznacza to, że wartość oczekiwana oceny danej próbki zawiera się (z prawdopodobieństwem 95%) zarówno w rezultatach otrzymanych z testu internetowego i testu laboratoryjnego (definiowanych przez przedziały ufności). Pociąga to za sobą też obserwację, że obie metody dały dla danej próbki statystycznie identyczny rezultat.

Należy zauważyć, że oba zbiory danych są ze sobą w dodatniej korelacji liniowej – wraz ze wzrostem jakości próbki dźwiękowej, średnia wartość oceny rośnie w przypadku obu testów: laboratoryjnego i internetowego. Zauważyć też można, że wszystkie średnie wartości ocen dla próbek dźwiękowych o słabej jakości (próbki o numerach 6 – 11) znajdują się poniżej linii odniesienia  $Y=X$ . Oznacza to, że uzyskane średnie ocen dla tych próbek są większe w teście laboratoryjnym, niż w teście internetowym. Rozkład taki wynikać może z zastosowania w testach laboratoryjnych lepszej jakości sprzętu odtwarzającego dźwięk. Zastosowany sprzęt odtwarzający [82] jest lepszy pod względem jakości odtwarzania, niż sprzęt który posiadała znaczna większość osób do wykonywania testu internetowego. W związku z tym dźwięk słabej jakości na lepszym sprzęcie mógł być oceniony wyżej, niż słabej jakości dźwięk na słabym sprzęcie odtwarzającym.

Należy zauważyć również, że dla próbek o wyższej jakości (oprócz próbki referencyjnej, tj. dla próbek o numerach 5 – 2) przedziały ufności są większe dla testu internetowego, niż laboratoryjnego. Większy rozrzut danych dla próbek wysokiej jakości również wskazywać może na użycie gorszej jakości sprzętu odtwarzającego w teście internetowym, niż w teście laboratoryjnym.

W przypadku większości próbek dźwiękowych przedziały ufności nakładają się na linię odniesienia  $Y=X$ , co oznacza, że uzyskane rezultaty dla testu internetowego są identyczne w granicach błędu statystycznego, w stosunku do rezultatów uzyskanych w teście laboratoryjnym. Potwierdza to względnie wysoką precyzję (powtarzalność) testu internetowego względem testu laboratoryjnego dla danego zestawu próbek.

**Tabela 8.3.** Współczynnik korelacji Pearson’a dla średnich ocen z obu testów oraz odpowiadający mu 95% przedział ufności [praca własna]

Współczynnik korelacji Pearson’a	95% przedział ufności dla współczynnika korelacji
$r(9) = 0.9967, p < 0.001$	0.9869, 0.9992

Obliczony współczynnik korelacji Pearson’a dla średnich wartości ocen z testu internetowego i testu laboratoryjnego, którego wartość jest bliska jedności, również potwierdza wysoką, dodatnią zależność liniową pomiędzy uzyskanymi średnimi wartościami ocen w obu eksperymentach.

Niniejszy wynik, jak również wyniki wcześniej zaprezentowanych analiz pokazują, że internetowa metoda oceny jakości dźwięku może stanowić odpowiednik metody laboratoryjnej.

Rezultaty przeprowadzonego testu i skrypty wykorzystane do analizy statystycznej zamieszczone są na dołączonej do pracy płycie CD.

## 9. Podsumowanie

Celem pracy było opracowanie internetowego portalu do oceny jakości dźwięku. Opracowany portal umożliwia potencjalnym użytkownikom zdalne tworzenie i uczestnictwo w testach odsłuchowych opartych o metody MUSHRA i ABX. W portalu zaimplementowano także następujące funkcjonalności: możliwość sprawdzenia podstawowych statystyk dotyczących stworzonego testu i eksport uzyskanych rezultatów do pliku w formacie CSV i JSON.

Aplikacja została opracowana w architekturze mikroservisów i w oparciu o architekturę REST. Komunikacja między poszczególnymi serwisami odbywa się z wykorzystaniem protokołu HTTP, zaś formatem wymiany danych jest JSON. Podstawowe platformy programistyczne, które zostały użyte do opracowania aplikacji to Spring Boot i Angular 2. Dane są przechowywane w relacyjnej bazie danych PostgreSQL.

Aplikację starano się opracować w taki sposób, aby rozwijanie jej w przyszłości, dodawanie nowych funkcjonalności, a także przystosowanie do zwiększonej ilości użytkowników i operacji transmisji nie stanowiło dużego problemu.

Celem przyszłych prac nad stworzonym portalem może być jego integracja z popularnymi platformami delegującymi zadania (ang. *Crowdsourcing*). Ze względu na duży odsetek osób, które nie oceniły co najmniej jednej próbki na 100 punktów, planowane jest dodanie funkcji opcjonalnego wyświetlania słuchaczowi stosownego komunikatu, przed zatwierdzeniem przez niego dokonanych ocen.

Opracowany portal internetowy rozwiązuje niektóre problemy związane z odsłuchowymi testami subiektywnymi, m.in. konieczność zgromadzenia grupy osób w konkretnym miejscu i czasie. Dodatkowo, poprzez swój sieciowy charakter portal ułatwia dotarcie do większej ilości osób i umożliwia przeprowadzenie testów odsłuchowych w różnych miejscach jednocześnie.

Rezultaty uzyskane w teście internetowym są bardzo zbliżone do rezultatów uzyskanych w teście laboratoryjnym [82]. W teście laboratoryjnym, w porównaniu do testu internetowego, zauważalny jest dużo mniejszy odsetek osób, których rezultaty zostały odrzucone i nie uwzględnione w dalszej analizie. Zastosowanie metody internetowej wprowadza więc znacznie większą ilość błędnie wykonanych testów. Powodem tego mogą

być niekontrolowane warunki eksperymentu, a także brak motywacji ze strony niektórych uczestników do poprawnego zastosowania się do wszystkich zaleceń instrukcji. Jednakże zastosowanie odpowiedniej procedury przesiewowej pozwala skutecznie oczyścić początkowy zbiór danych.

Obliczony współczynnik korelacji Pearson'a jest dodatni i wskazuje na wysoką zależność liniową uzyskanych średnich wartości ocen w obu eksperymentach. Precyzja (powtarzalność) metody internetowej względem metody laboratoryjnej jest wysoka dla badanego zestawu próbek dźwiękowych, co widoczne jest na wykresie punktowym przedstawiającym zestawienie średnich wartości ocen obu testów. Wskazuje to na możliwą przydatność portalu internetowego i testów internetowych służących subiektywnej ocenie jakości dźwięku.

Wybrane wyniki badawcze uzyskane w ramach niniejszej pracy wraz z opisem technicznym opracowanej aplikacji zgłoszone będą jako współautorski artykuł naukowy na konferencję o nazwie „*16th International Conference on Computer Information Systems and Industrial Management Applications*”, która odbędzie się w Białymstoku w czerwcu bieżącego roku [80].

## Literatura

- [1] Academia Sinica. *Quadrant of Euphoria*. Witryna internetowa. <http://mmnet.iis.sinica.edu.tw/proj/qoe>, stan z 17.02.17
- [2] Apache Software Foundation. *Apache Commons Math*. Witryna internetowa. <http://commons.apache.org/proper/commons-math>, stan z 06.02.17
- [3] Apache Software Foundation. *Forward and Reverse Proxies*. Witryna internetowa: <http://httpd.apache.org>, stan z 06.02.2017
- [4] Apache Software Foundation. *Maven*. Witryna internetowa. <https://maven.apache.org>, stan z 06.02.17
- [5] AudioLabs. *webMUSHRA*. Witryna internetowa. <https://www.audiolabs-erlangen.de>, stan z 06.02.2017
- [6] Auth0. *Java JWT Library*. Witryna internetowa. <https://jwt.io>, stan z 06.02.17
- [7] Belias V.: *A study on Quality of Experience (QoE) for multimedia systems*, Bachelor Thesis on Informatics, Alexander TEI of Thessaloniki, Department of Informatics, 2013
- [8] Bootstrap Core Team. *Bootstrap*. Witryna internetowa. <https://v4-alpha.getbootstrap.com>, stan z 06.02.17
- [9] Bray E.T, Google Inc.: *The JavaScript Object Notation (JSON) Data Interchange Format*, RFC 7159, IETF, 2014
- [10] Campbell D., Jones E., Glavin M.: *Audio quality assessment techniques—A review, and recent developments*, Signal Processing 2009, 89, 1489 -1500
- [11] Carlstrom D. *ABX Company*. Witryna internetowa. <http://djcarlst.provide.net/>, stan z 06.02.2017
- [12] Chen Ch. W., Chatzimisios P., Dagiuklas T., Atzori L.: *Multimedia quality of experience (QoE) current status and future requirements*, Wiley, 2016;
- [13] Chen K.-T., Wu C.-C., Chang Y.-C., Lei C.-L.: *A Crowdsourceable QoE Evaluation Framework for Multimedia Content*, Proceedings of the 17th ACM international conference on multimedia 2009, 491 – 500
- [14] Chen P.: *The entity-relationship model—toward a unified view of data*, ACM Transactions on Database Systems (TODS) 1976, Vol. 1, No. 1, 9 – 36
- [15] Clark D.L: *Ten years of a/b/x testing*, AES Convention 91(3167), AES, 1991

- [16] European Broadcasting Union: *EBU evaluations of multichannel audio codecs*, EBU Tech 3324 Technical Document, Geneva 2007
- [17] European Broadcasting Union: *EBU subjective listening tests on low-bitrate audio codecs*, EBU Tech 3296 Technical Document, Geneva 2003
- [18] Fielding R.T., Gettys J., Mogul J.C, Nielsen H.F., Masinter L., Leach P.J., Berners-Lee T.: *Hypertext Transfer Protocol -- HTTP/1.1*, RFC 2616, IETF, 1999
- [19] Fielding R.T.: *Principled Design of the Modern Web Architecture*, *ACM Transactions on Internet Technology* 2002, Vol. 2, No. 2
- [20] Google. *Angular Material*. Witryna internetowa. <https://material.angular.io>, stan z 06.02.17
- [21] Google. *Angular*. Witryna internetowa. <https://angular.io>, stan z 06.02.17
- [22] Hoeg W., Christensen L., Walker R.: *Subjective assessment of audio quality the means and methods within the EBU*, *EBU Technical Review* 1997, 40 – 50;
- [23] Hoßfeld T., Hirth M., Korshunov P., Hanhart P., Gardlo B., Keimel C., Timmerer C.: *Survey of Web-based Crowdsourcing Frameworks for Subjective Quality Assessment*, 2014
- [24] International Telecommunication Union (ITU): *A guide to ITU-R Recommendations for subjective assessment of sound quality*, Recommendation ITU-R BS.1283-1, 2003
- [25] International Telecommunication Union (ITU): *Method for objective measurements of perceived audio quality*, ITU-R Recommendation BS.1387, Geneva 1998
- [26] International Telecommunication Union (ITU): *Method for the subjective assessment of intermediate quality level of audio systems*, ITU-R Recommendation BS.1534-3, Geneva 2015;
- [27] International Telecommunication Union (ITU): *Methodology for the subjective assessment of the quality of television pictures*, Recommendation ITU-R BT.500-13, 2012
- [28] International Telecommunication Union (ITU): *Methods for the subjective assessment of small impairments in audio systems*, ITU-R Recommendation BS.1116-3, Geneva 2015;
- [29] International Telecommunication Union (ITU): *Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference*, Recommendation ITU-R BT.1683, 2004



- [30] International Telecommunication Union (ITU): *Objective perceptual video quality measurement techniques for broadcasting applications using HDTV in the presence of a full reference signal*, Recommendation ITU-R BT.1907, 2012
- [31] International Telecommunication Union (ITU): *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*, Recommendation ITU-T P.862, 2001
- [32] International Telecommunication Union (ITU): *Perceptual objective listening quality assessment*, Recommendation ITU-T P.863, 2014
- [33] International Telecommunication Union (ITU): *Requirements for an objective perceptual multimedia quality model*, ITU-T Recommendation J.148, 2003
- [34] International Telecommunication Union (ITU): *Subjective audiovisual quality assessment methods for multimedia applications*, Recommendation ITU-T P.911, Annex B, 1998.
- [35] International Telecommunication Union (ITU): *Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs*, Recommendation ITU-T P.862.2, 2007
- [36] JetBrains. *IntelliJ IDEA*. Witryna internetowa. <https://www.jetbrains.com>, stan z 06.02.17
- [37] Jillings N., De Man B., Moffat D., Reiss D. J., Stables R.: *Web Audio Evaluation Tool: A framework for subjective assessment of audio*, 2nd Web Audio Conference, 2016
- [38] Jurgen H., Johannes H., Achim K., Plogsties J.: *MPEG-H 3D Audio – New Standard for Coding of Immersive Spatial Audio*
- [39] Keimel C., Habicht J., Horch C., Diepold K.: *QualityCrowd - A Framework for Crowd-based Quality Evaluation*, Picture Coding Symposium 2012, 245 – 248
- [40] Kraft S. *BeagleJS*. Witryna internetowa. <https://github.com/HSU-ANT/beaglejs>, stan z 06.02.2017
- [41] Kraft S., Zölzer U.: *BeagleJS: HTML5 and JavaScript based Framework for the Subjective Evaluation of Audio Quality*, Linux Audio Conference, 2014, Karlsruhe, Niemcy
- [42] Krawczyk, et. al.: *HMAC: Keyed-Hashing for Message Authentication*, IETF, 1997

- [43] Krumpholz M. *Automated Web based Subjective evaluation Platform (WESP+)*. Witryna internetowa. [https://github.com/maxkr/qoe\\_evaluation\\_platform](https://github.com/maxkr/qoe_evaluation_platform), stan z 17.02.17
- [44] Kucharzyk T. *Spring Boot Angular2 Starter*. Witryna internetowa. <https://github.com/shardis/spring-angular2-starter>, stan z 06.02.2017
- [45] Kudělka Jr. M.: *Image Quality Assessment, WDS'12 Proceedings of Contributed Papers Part I* 2012, 94–99
- [46] Lassalle J., Gros L., Coppin G.: *Combination of physiological and subjective measures to assess quality of experience for audiovisual technologies, Third International Workshop on Quality of Multimedia Experience* 2011, 13-18
- [47] Lathi B.P.: *Modern Digital and Analog Communication Systems*, Holt, Rinehart and Winston, 1983
- [48] Lee S., Lee Y-T., Seo J., Baek M-S., Lim C-H., Park H.: *An Audio Quality Evaluation of Commercial Digital Radio Systems*, IEEE Transactions on Broadcasting 2011, Vol. 57(3), 629 – 635
- [49] Lin. W, Kuo J.C.C.: *Perceptual visual quality metrics: A survey*, Journal of Visual Communication and Image Representation 2011 2011, Vol.22, No. 4, 297 –312
- [50] Martini M. G., Chen Ch. W., Chen Z., Dagiuklas T., Sun L., Zhu X.: *Guest Editorial QoE-Aware Wireless Multimedia Systems, IEEE Journal On Selected Areas In Communications* 2012, vol. 30, no. 7, 1153 – 1156;
- [51] Mohammadi P., Ebrahimi-Moghadam A., Shirani S.: *Subjective and Objective Quality Assessment of Image: A Survey*, 2014
- [52] Mozilla Developer Network. *Media formats supported by the HTML audio and video elements*. Witryna internetowa. [https://developer.mozilla.org/en-US/docs/Web/HTML/Supported\\_media\\_formats](https://developer.mozilla.org/en-US/docs/Web/HTML/Supported_media_formats), stan z 17.02.17
- [53] Munson W. A., Gardner B. M.: *Standardizing Auditory Tests, Acoustical Society of America* 1950, no. 22, 675
- [54] Newman S.: *Building Microservices. Designing Fine-Grained Systems*, O'Reilly Media, 2015
- [55] OAuth 2.0 Community. *OAuth 2.0*. Witryna internetowa. <https://oauth.net/2>, stan z 06.02.17
- [56] Pivotal Software. *Microservice Registration and Discovery with Spring Cloud and Netflix's Eureka*. Witryna internetowa:

- <https://spring.io/blog/2015/01/20/microservice-registration-and-discovery-with-spring-cloud-and-netflix-s-eureka> , stan z 06.02.2017
- [57] Pivotal Software. *Spring Boot*. Witryna internetowa. <https://projects.spring.io>, stan z 06.02.17
- [58] Provos N., Mazieres D., Sutton J.T.: *A Future-Adaptable Password Scheme*, *Proceedings of 1999 USENIX Annual Technical Conference 1999*, 81-92
- [59] Rainer B., Walzl M., Timmerer C.: *A web based subjective evaluation platform*, Fifth International Workshop on Quality of Multimedia Experience 2013, 24 – 25
- [60] RedHat. *Pessimistic and optimistic locking*. Witryna internetowa. <https://docs.jboss.org>, stan z 06.02.17
- [61] Ribeiro F., Florêncio2 D., Zhang C., Seltzer M.: *CROWDMOS: An approach for crowdsourcing mean opinion score studies*, IEEE International Conference on Acoustics, Speech and Signal Processing 2011, 2416 – 2419
- [62] Rivest R., Shamir A., Adleman L.: *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*, Communications of the ACM 1978, 21(2), 120 – 126
- [63] RStudio. *RStudio*. Witryna internetowa. <https://www.rstudio.com>, stan z 25.02.2017
- [64] Schoeffler M., Stöter F.-B., Edler B., Herre J.: *Towards the Next Generation of Web-based Experiments: A Case Study Assessing Basic Audio Quality Following the ITU-R Recommendation BS.1534 (MUSHRA)*, 1st Web Audio Conference 2015
- [65] Serral-Gracià R., Cerqueira E., Curado M., Yannuzzi M., Monteiro E., Masip-Bruin X.: *An Overview of Quality of Experience Measurement Challenges for Video Applications in IP Networks*, 2012
- [66] Software Freedom Conservancy. *Git*. Witryna internetowa. <https://git-scm.com/doc>, stan z 06.02.17
- [67] Sonchhatra P., Kumar S.: *No Reference Image Quality Assessment: Feature Extraction Approach*, *International Journal of Science and Research* 2013, Vol.5, 241-245
- [68] Thaise E. *mctad.js*. Witryna internetowa. <http://erictheise.com/mctad.js>, stan z 06.02.17
- [69] Thapliyal V.: *Difference Between Frontend and Backend MVC*, Witryna internetowa. <http://joomlatuts.net>, stan z 06.02.17
- [70] The jQuery Foundation. *jQuery UI*. Witryna internetowa. <https://jqueryui.com>, stan z 06.02.17

- [71] The PostgreSQL Global Development Group. *PostgreSQL*. Witryna internetowa. <https://www.postgresql.org>, stan z 06.02.17
- [72] W3C. *JavaScript Best Practices*. Witryna internetowa. <https://www.w3.org/wiki>, stan z 06.02.17
- [73] W3C. *Web Audio API*. Witryna internetowa. <https://www.w3.org/TR/webaudio>, stan z 06.02.2017
- [74] Wang Z., Bovik A. C., Lu L.: *Why Is Image Quality Assessment So Difficult?*, *IEEE International Conference on Acoustics, Speech, and Signal Processing* 2002, Vol.4, 3313-3316
- [75] Webpack. *Webpack Module Bundler*. Witryna internetowa. <https://webpack.github.io>, stan z 06.02.17
- [76] Wickham H., et. al. *Package 'dplyr'*. Witryna internetowa. <https://cran.r-project.org/web/packages/dplyr/dplyr.pdf>, stan z 25.02.2017
- [77] Winkler S., Mohandas P.: *The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics*, *IEEE Transactions on Broadcasting* 2008, Vol. 54, No. 3, 660 – 668
- [78] XYZPawel. Blogspot. *Testy ABX*. Witryna internetowa. <http://xyzpawel.blogspot.com/2009/07/testy-abx.html>, stan z 17.02.17
- [79] You Y.: *Audio Coding. Theory and Applications*, Springer Science & Business Media, 2010
- [80] Zieliński S., Milewski M.: *Audio quality assessment over the Internet – A case study*, 16th International Conference on Computer Information Systems and Industrial Management Applications, Białystok, June 16-18, 2017 – artykuł w przygotowaniu.
- [81] Zieliński S., Rumsey F., Bech S.: *Towards Unification of Methods for Speech, Audio, Picture and Multimedia Quality Assessment*, 138th AES Convention proceedings, AES, 2015, Paper number: 9308
- [82] Zieliński S.: *Is a Multi-Slider Interface Layout Responsible for a Stimulus Spacing Bias in the MUSHRA Test?*, *The Journal of Institute of Fundamental Technological of Polish Academy of Sciences, Archives of Acoustics* 2015, Vol. 40, No. 4, 585 – 594

## Spis rysunków

<b>Rys. 2.1.</b> Klasyfikacja metod oceny jakości treści multimedialnych [12] .....	14
<b>Rys. 3.1.</b> Schemat blokowy algorytmu BTFR [29] .....	19
<b>Rys. 3.2.</b> Schemat blokowy algorytmu EPSNR [29].....	20
<b>Rys. 3.3.</b> Schemat blokowy metody PESQ [31] .....	23
<b>Rys. 4.1.</b> Fazy prezentacji materiału testowego w metodzie DSIS: a) wariant I, b) wariant II [27] .....	28
<b>Rys. 4.2.</b> Przykładowy interfejs użytkownika w metodzie MUSHRA [26] .....	30
<b>Rys. 5.1.</b> Interfejs metody MUSHRA w platformie webMUSHRA [5].....	34
<b>Rys. 6.1.</b> Diagram UML przypadków użycia [praca własna] .....	38
<b>Rys. 6.2.</b> Architektura aplikacji, opracowano na podstawie [44] .....	43
<b>Rys. 6.3.</b> Diagram ER tabeli użytkowników [praca własna] .....	46
<b>Rys. 6.4.</b> Diagram ER tabeli grup testów [praca własna] .....	47
<b>Rys. 6.5.</b> Diagram ER tabeli grup rezultatów [praca własna] .....	48
<b>Rys. 6.6.</b> Wykres przedstawiający zmianę wzmocnienia dźwięków podczas synchronicznego przełączania próbek [praca własna].....	54
<b>Rys. 6.7.</b> Zaimplementowany interfejs metody MUSHRA [praca własna] .....	55
<b>Rys. 6.8.</b> Zaimplementowany interfejs testu ABX [praca własna] .....	55
<b>Rys. 6.9.</b> Prezentacja statystyk testu ABX w opracowanym portalu [praca własna] .....	56
<b>Rys. 8.1.</b> Przykładowe rezultaty w formacie CSV ( <i>index</i> – numer porządkujący wykonanie testu, <i>iteration</i> – numer próby, <i>sample</i> – kod próbki, <i>value</i> – ocena) [rezultaty własne] .....	60
<b>Rys. 8.2.</b> Wykresy pudełkowe dla rezultatów przed odsiewem (a) i dla rezultatów po odsiewie danych (b) [praca własna].....	63
<b>Rys. 8.3.</b> Średnie wartości ocen i związane z nimi 95 % przedziały ufności dla wynikowego zbioru rezultatów [praca własna] .....	64
<b>Rys. 8.4.</b> Wykres punktowy ilustrujący średnie wartości ocen dla testu laboratoryjnego i testu internetowego oraz związane z nimi 95% przedziały ufności [praca własna] .....	66

## Spis tabel

<b>Tabela 2.1.</b> Zestawienie wybranych, standaryzowanych przez ITU metod subiektywnej oceny treści multimedialnych [81].....	11
<b>Tabela 2.2.</b> Zestawienie wybranych, standaryzowanych przez ITU metod obiektywnej oceny treści multimedialnych [81].....	12
<b>Tabela 4.1.</b> Skale ocen jakości i zniekształceń [7].....	25
<b>Tabela 4.2.</b> Zalecane użycie metod dla danego problemu [27] .....	27
<b>Tabela 6.1.</b> Wspierane formaty dla popularnych przeglądarek internetowych i ich wersji wydania [52] .....	37
<b>Tabela 6.2.</b> Parametry odtwarzania dźwięku w opracowanym portalu [40].....	37
<b>Tabela 7.1</b> Wykaz użytych w teście próbek dźwiękowych z testu laboratoryjnego <i>B</i> , przeprowadzonego przez Zielińskiego [82].....	58
<b>Tabela 8.1.</b> Podsumowanie odsiewu rezultatów w teście internetowym [rezultaty własne] .....	62
<b>Tabela 8.2.</b> Podsumowanie odsiewu rezultatów testu laboratoryjnego <i>B</i> , przeprowadzonego przez Zielińskiego [82, kryteria własne] .....	65
<b>Tabela 8.3.</b> Współczynnik korelacji Pearson'a dla średnich ocen z obu testów oraz odpowiadający mu 95% przedział ufności [praca własna] .....	67

## Spis kodów źródłowych

<b>Kod źródłowy 6.1.</b> Procedura tworzenia konta użytkownika w serwisie API [praca własna] .....	50
<b>Kod źródłowy 6.2.</b> Definicja obiektu AudioObject [praca własna] .....	52
<b>Kod źródłowy 6.3.</b> Proces przełączania próbek z zachowaniem czasu aktualnie odgrywanej próbki [praca własna] .....	53
<b>Kod źródłowy 8.1.</b> Procedura odsiewu rezultatów (dwa początkowe kryteria) [praca własna] .....	62

## Wykaz załączników

1. Załącznik 1. Instrukcja użytkownika portalu.
2. Załącznik 2. Instrukcja instalacji portalu.
3. Płyta CD zawierająca dwa foldery:
  1. Skompresowany folder o nazwie *kod-zrodlowy.zip*, zawierający pełny kod źródłowy aplikacji wraz ze skryptem pozwalającym na uruchomienie portalu (*start-script.py*).
  2. Folder o nazwie *test*, zawierający rezultaty testu z usuniętymi danymi osobowymi (*dataB.csv*), który przeprowadzony został z wykorzystaniem opracowanego portalu i sieci Internet. Folder ten zawiera też skrypty, które posłużyły do analizy rezultatów (*analysis\_internet.R*, *analysis\_lab.R*, *analysis.R*).



**Załącznik 1**  
**Instrukcja użytkownika opracowanego portalu**  
**internetowego**

Interfejs użytkownika domyślnie dostępny jest pod adresem *http://localhost:8080*. Po wpisaniu tego adresu wyświetli się strona domowa portalu (Rys. 1).

The screenshot shows the home page of the 'Internetowy portal do oceny jakości dźwięku'. The header is purple with a menu icon, the portal name, and buttons for 'Zarejestruj się' and 'Zaloguj'. The main content area is white and contains several paragraphs of text explaining the portal's purpose and usage. It mentions that users can create groups of tests and check statistics. There are links to 'Oświadczenie Uczestnika' and 'Oświadczenie Twórcy'. A section titled 'Zaimplementowane metody testów odsłuchowych:' lists 'ABX Test' and 'Bazująca na specyfikacji ITU-R BS.1534-3 (MUSHRA)'. At the bottom, there is a link to a 'formularza'.

System powstał w wyniku realizacji pracy dyplomowej na Wydziale Informatyki Politechniki Białostockiej

Witaj na portalu Internetowym pozwalającym na tworzenie i uczestnictwo w testach odsłuchowych, pozwalających na ocenę jakości dźwięku

Jako anonimowy użytkownik, możesz wziąć udział tylko w testach w których twórca pozwala na udział użytkownikom anonimowym.

Po utworzeniu konta użytkownika możesz tworzyć własne grupy testowe i sprawdzać statystyki swoich testów, możesz brać też udział we wszystkich testach odsłuchowych.

Aby wziąć udział w dowolnym teście odsłuchowym, jako słuchacz musisz zaakceptować Oświadczenie Uczestnika, a następnie stosować się do instrukcji zamieszczanej przez twórcę do każdego testu.

Jako twórca przed stworzeniem grupy testowej musisz zaakceptować Oświadczenie Twórcy.

[Kliknij aby zobaczyć cały tekst obu oświadczeń.](#)

Rozwijane menu serwisu znajduje się w lewym górnym rogu strony i widoczne będzie po kliknięciu w następującą ikonę: .

Po zalogowaniu się w portalu pojawią się tam też dwie nowe zakładki prowadzące do kreatora testów i menadżera testów.

Jeśli masz jakieś uwagi, propozycje lub zauważyłeś błąd, proszę o wypełnienie krótkiego [formularza](#).

Zaimplementowane metody testów odsłuchowych:

- ABX Test
- Bazująca na specyfikacji ITU-R BS.1534-3 (MUSHRA)

**Rys. 1.** Strona domowa portalu

Użytkownik chcąc mieć możliwość stworzenia testu odsłuchowego musi utworzyć konto użytkownika. W tym celu należy przejść do zakładki *Zarejestruj się* (Rys. 2).

The screenshot shows the 'Zarejestruj się' (Register) page. It has a purple header with the portal name and buttons for 'Zarejestruj się' and 'Zaloguj'. The main content area is white and contains a registration form. The form has fields for 'Nazwa użytkownika \*', 'Hasło \*', 'Powtórz hasło \*', and 'Email \*'. Below these fields is a section titled 'Dodatkowe dane wymagane przed przystąpieniem do każdego testu' which contains fields for 'Płeć \*' and 'Wiek (w latach) \*'. At the bottom of the form is a checkbox labeled 'Wyrażam zgodę na przetwarzanie i składowanie powyższych danych osobowych w ramach portalu.' and a 'Zarejestruj' button.

Nazwa użytkownika \*

Hasło \*

Powtórz hasło \*

Email \*

Dodatkowe dane wymagane przed przystąpieniem do każdego testu

Płeć \*

Wiek (w latach) \*

☐ Wyrażam zgodę na przetwarzanie i składowanie powyższych danych osobowych w ramach portalu.

Zarejestruj

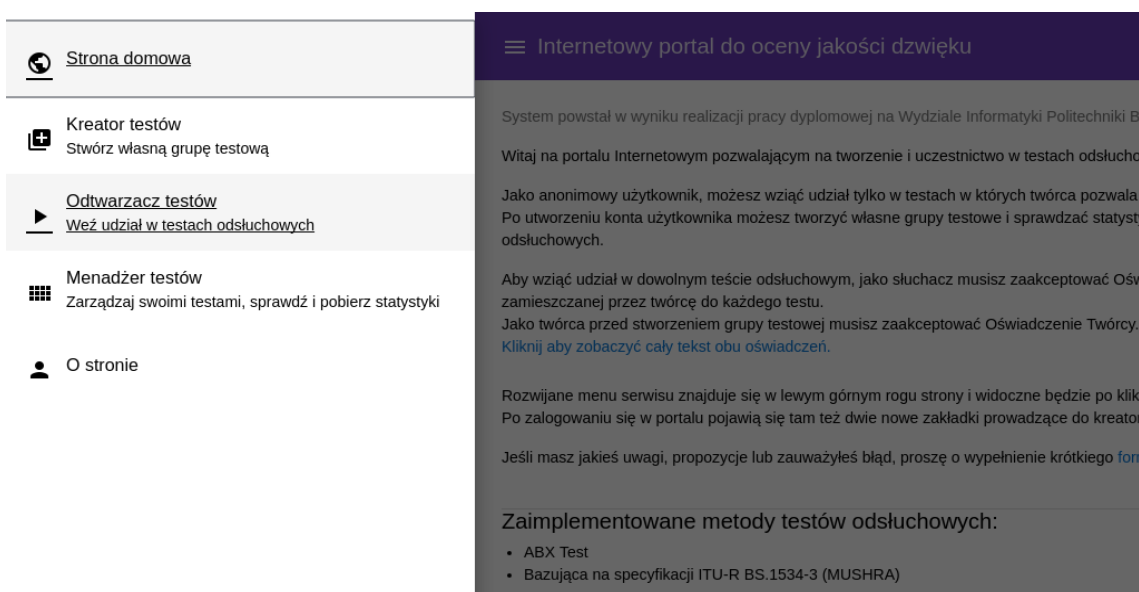
**Rys. 2.** Strona umożliwiająca rejestrację.

Strona umożliwiająca zalogowanie się w serwisie przedstawiona jest na Rys. 3

The screenshot shows the login interface of the 'Internetowy portal do oceny jakości dźwięku'. At the top, there is a purple header bar with a hamburger menu icon on the left and two buttons, 'Zarejestruj się' and 'Zaloguj', on the right. Below the header, the text 'Zaloguj się' is displayed. There are two input fields: 'Login \*' and 'Hasło \*'. At the bottom, there is a grey button labeled 'Zaloguj'.

**Rys. 3.** Strona umożliwiająca zalogowanie się w serwisie

Po pomyślnym utworzeniu konta użytkownika i uwierzytelnieniu, użytkownik może stworzyć test. W tym celu należy otworzyć rozwijane menu serwisu, które znajduje się w lewym górnym rogu strony (Rys. 4) i przejść do zakładki *Kreator testów*. Po otworzeniu tej zakładki, aby móc utworzyć test, użytkownik musi zaakceptować regulamin dla twórcy testu i scharakteryzować test (m.in. podać cel badań) (Rys. 5).



**Rys. 4.** Menu serwisu

Internetowy portal do oceny jakości dźwięku

Użytkownik: marcin Wyloguj

### Kreator testów odsłuchowych

#### Oświadczenie Twórcy Testów

1. Twórca testów oświadcza, że próbki dźwiękowe nie posiadają treści powszechnie uważanych za wulgarne i niezgodne z prawem.
2. Twórca testów oświadcza, że zaspokaja wszystkie licencje, prawa, zgody zamieszczaną treścią zarówno w chwili zamieszczenia, jak i podczas użytkowania portalu. Oświadcza też że nie łamie stosownych praw przechowując utwory w bazie danych portalu.
3. Twórca testów oświadcza, że ponosi wyłączną odpowiedzialność za własne treści oraz skutki zamieszczenia lub opublikowania.
4. Twórca testów oświadcza, że dane osobowe przetwarza będzie wyłącznie w celach przeprowadzanego eksperymentu i nie będzie ich upubliczniał wraz z wynikami.
5. Twórca testów oświadcza, że udział w eksperymencie jest dobrowolny i nieodpłatny, a uczestnik może w każdej chwili zrezygnować bez podania przyczyny. Uczestnik nie ponosi żadnych konsekwencji z tego powodu.
6. Twórca testów oświadcza, że jest to eksperyment poświęcony badaniu jakości dźwięku w zakresie zdefiniowanego przez Twórcę celu badań.

☐ Przeczytałem i w pełni akceptuję Oświadczenie Twórcy Testów

#### Charakterystyka grupy testowej

Cel badań

Szacunkowy czas trwania (w minutach)

☐ Zezwól na uczestnictwo w testach anonimowym użytkownikom

Dalej

**Rys. 5.** Początkowy widok na stronie Kreatora testów

Po wypełnieniu formularza i kliknięciu *Dalej*, użytkownik może stworzyć grupę testową. W skład grupy dodawać może wiele testów. Każdy test może być stworzony w oparciu o metodę MUSHRA lub ABX. Przykład wypełnionego formularza dodawania grupy testowej przedstawiony jest na Rys.6.

Internetowy portal do oceny jakości dźwięku

Użytkownik: marcinWyloguj

### Kreator testów odsłuchowych

Nazwa grupy testowej  
Przykładowa nazwa

Opis grupy testowej  
Przykładowy opis

#### Testy

##### Test 1

Wybierz metodę testową MUSHRA

Ilość próbek 3

Nazwa testu  
Test pierwszy

Instrukcja dla użytkownika  
Przykładowa instrukcja

Ilość powtórzeń  
2

A (Próbka referencyjna)

REFERENCE.WAV

B

B10000.WAV

C

B3500.WAV

Dodaj kolejny test

Utwórz

Rys. 6. Formularz dodawania grupy testowej

Po pomyślnym dodaniu grupy testowej, użytkownik dostaje informację zwrotną o tym, a także adres dostępu do niej pozwalający na uczestnictwo w teście innym osobom (Rys. 7).

Internetowy portal do oceny jakości dźwięku

Użytkownik: marcinWyloguj

Pomyślnie dodano test.

Adres dostępu do tej grupy testowej pozwalający na uczestnictwo w teście innym osobom

<http://207.154.197.64:8080/test-player/player/4453> Kopiuuj

[Przejdź do panelu zarządzania](#)

[Kliknij by uczestniczyć](#)

Rys. 7. Informacja zwrotna po pomyślnym dodaniu grupy testowej.

Uczestnikowi, po wpisaniu adresu testu do wykonania, przedstawiony zostanie wstępny formularz, w którym akceptuje on m.in. regulamin dla słuchacza (Rys. 8).

Internetowy portal do oceny jakości dźwięku

Użytkownik: marcin Wyloguj

Zostałeś zaproszony do uczestnictwa w grupie testów odsłuchowych o nazwie: Przykładowa nazwa

Opis grupy: Przykładowy opis

Twórca: marcin

Cel badań: Testy serwisu

Szacunkowy czas trwania (w minutach): 15

Identyfikator: 4453

### Oświadczenie Słuchacza

- Słuchacz wyraża zgodę na uczestnictwo w eksperymencie poświęconym badaniu jakości dźwięku.
- Słuchacz zapoznał się z charakterystyką testów, akceptuje szacunkowy czas trwania testów i nie ma zastrzeżeń do celu badań.
- Słuchacz akceptuje fakt że po naciśnięciu na przycisk dowolnej próbki dźwiękowej rozpocznie się odtwarzanie dźwięku.
- Słuchacz zapoznał się i akceptuje to że przed rozpoczęciem eksperymentu, a dalej każdego z testów powinien dostosować poziom odtwarzanego dźwięku poprzez maksymalne ściszenie urządzenia fizycznego. W chwili rozpoczęcia tj. po naciśnięciu na przycisk dowolnej próbki dźwiękowej, powinien dostosować poziom odtwarzanego dźwięku poprzez stopniowe zwiększanie głośności urządzenia do komfortowego dla siebie poziomu.
- Słuchacz akceptuje fakt że nie każda próbka musi posiadać zbliżony poziom głośności.
- Słuchacz nie może rościć praw z tytułu uszkodzenia słuchu ani do twórców portalu, twórców testów, ani twórców wykorzystywanych bibliotek.
- Słuchacz przyjmuje do wiadomości i akceptuje, że podczas korzystania z usług portalu, pomimo oświadczenia twórcy i nadzoru administracji może być narażony na treści obraźliwe, nieprzyzwoite lub które użytkownik uzna za niedopuszczalne. Słuchacz zrzeka się z tego tytułu praw do stosownych roszczeń wobec twórców portalu i twórcy testów. Słuchacz powinien zgłosić wszelkie niedopuszczalne treści administracji portalu.
- Słuchacz akceptuje to, że jego rezultaty wraz z danymi osobowymi tj. płeć, wiek, odpowiedź na pytanie osobiste - czy posiada Pan/Pani wadę/ubytek słuchu jak i login uczestnika (jeśli test nie jest przeprowadzany anonimowo) będą wysłane i przechowywane w bazie danych. Ze strony portalu rezultaty wraz z danymi osobowymi będą przetwarzane tylko i wyłącznie na potrzeby wyznaczenia odpowiednich statystyk, przesyłane będą dla twórcy danego zestawu testowego (zarówno w formie przetworzonej jak i w formie nieprzetworzonej, w różnym formacie plikowym). Dalsze przetwarzanie rezultatów wraz z danymi osobowymi reguluje Oświadczenie Twórcy.
- Słuchacz akceptuje to, że przed rozpoczęciem wykonywania testów z danej grupy testowej, z bazy danych serwisu pobrane będą na jego urządzenie pliki dźwiękowe niezbędne do przeprowadzenia danego testu. Dodatkowo w celu optymalizacji, podczas wykonywania testu pobierane będą niezbędne pliki dźwiękowe do przeprowadzenia testu następnego.
- Udział w eksperymencie jest dobrowolny i nieodpłatny, a słuchacz może w każdej chwili zrezygnować bez podania przyczyny. Uczestnik nie ponosi żadnych konsekwencji z tego powodu.

☐ Przeczytałem i w pełni akceptuję Oświadczenie Uczestnika Testów

### Pytania pomocnicze

Odpowiedz na poniższe pytania pozwoli na dokładniejszą analizę rezultatów testu.

Płeć

Mężczyzna ▼

Wiek

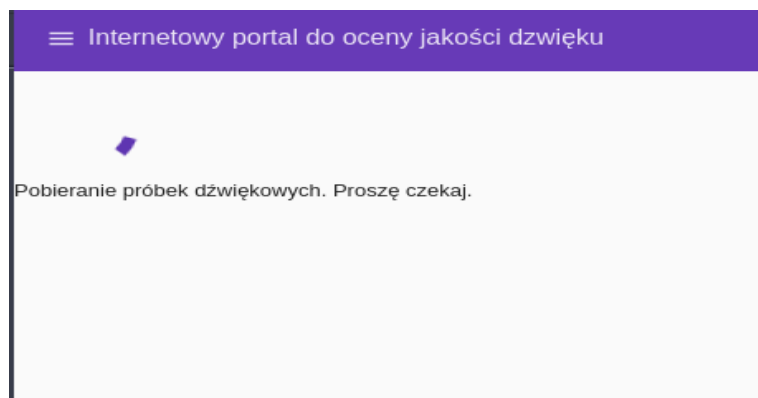
23

Dalej

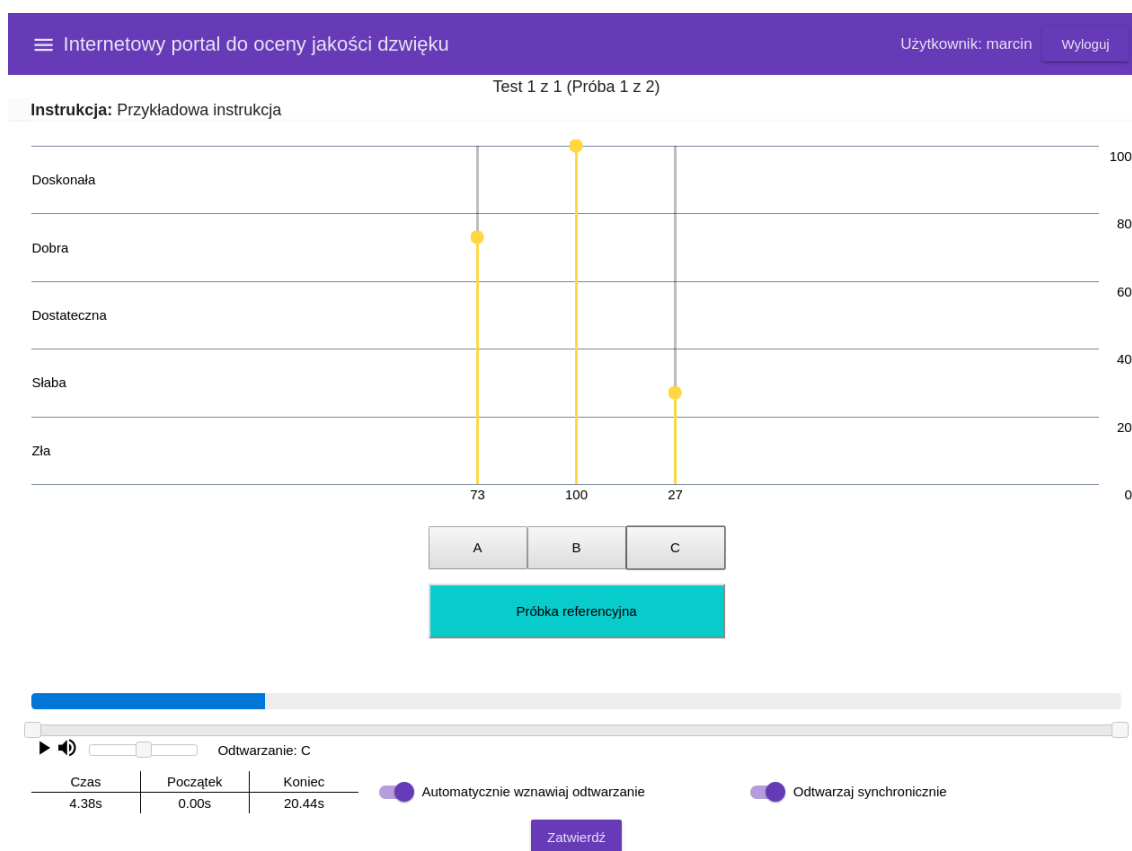
Aby uczestniczyć musisz zaakceptować oświadczenie i odpowiedzieć na pytania pomocnicze, a następnie nacisnąć przycisk 'Dalej'

### Rys. 8. Wstępny formularz przed rozpoczęciem wykonywania testu

Po uzupełnieniu formularza wstępnego i kliknięciu przycisku *Dalej*, słuchacz oczekuje na pobranie próbek (Rys. 9), a następnie wyświetlany jest mu stosowny do danej metody testowej interfejs (Rys. 10).

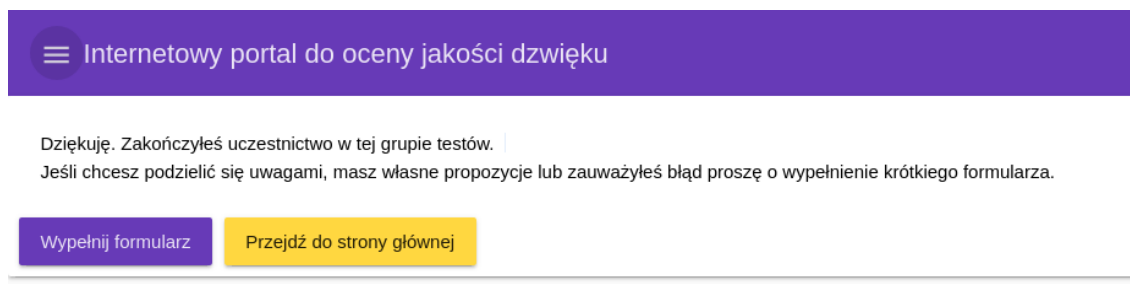


**Rys. 9.** Oczekiwanie na pobranie próbek dźwiękowych



**Rys. 10.** Interfejs podczas wykonywania testu dla metody MUSHRA

Po zakończeniu wykonywania testu, słuchaczowi prezentowane jest krótkie podziękowanie za wzięcie udziału w teście i opcja wypełnienia formularza, w którym może on podzielić się swoimi uwagami, propozycjami, czy zauważonymi błędami (Rys. 11). Pytania zamknięte formularza przedstawione są na Rys. 12.



**Rys. 11.** Strona widoczna po wykonaniu grupy testowej

Internetowy portal do oceny jakości dźwięku   Użytkownik: marcin   Wyloguj

### Uwagi i propozycje użytkowników

Odpowiedzi na poniższe pytania pomogą ulepszyć serwis.

Czy przełączanie pomiędzy dźwiękami powodowało słyszalny efekt chwilowego wyłączenia i włączenia dźwięku?

☐ Tak  
☐ Nie

Jeśli efekt ten występował, to czy utrudniał porównywanie dźwięków?

☐ Nie mam zdania  
☐ Tak  
☐ Nie

Jeśli efekt ten występował, proszę o wskazanie przeglądarki z jakiej Pan/Pani korzystał/a.

☐ Google Chrome  
☐ Mozilla Firefox  
☐ Chromium  
☐ Opera  
☐ Safari  
☐ Nie wiem  
☐ Inne: \_\_\_\_\_

**Rys. 12.** Przykładowe pytania zamknięte formularza z uwagami i propozycjami użytkowników

W celu sprawdzenia statystyk własnych testów należy wybrać zakładkę *Menadżer testów* w menu serwisu. Po wybraniu tej zakładki, użytkownikowi prezentowana jest tabela krótko charakteryzująca posiadane grupy testowe (Rys. 13). Po kliknięciu przycisku *Szczegóły* dla danego zestawu, następuje przejście do zakładki prezentującej szczegóły danej grupy (Rys. 14) i rezultaty (Rys. 15 i Rys. 16). W zakładce tej istnieje też możliwość eksportu rezultatów do plików CSV i JSON (Rys. 17).



Moje testy							
Nazwa grupy	Opis grupy	Cel badań	Szacunkowy czas trwania (w minutach)	Identyfikator	Data utworzenia	Liczba testów w grupie	Akcja
Beta testy serwisu	Dwa testy, każdy po 3 razy, sprawdzające poprawność mieszania próbek i przesyłania rezultatów.	Beta testy serwisu	5	190	27.01.2017 21:02	2	<a href="#">Szczegóły</a>
Beta testy - tylko ABX	Beta testy - tylko ABX	Beta testy - tylko ABX	5	245	27.01.2017 23:45	1	<a href="#">Szczegóły</a>

Rys. 13. Stworzone testy przez użytkownika

Internetowy portal do oceny jakości dźwięku

Użytkownik: marcinWyloguj

Nazwa grupy: Beta testy serwisu

Opis grupy: Dwa testy, każdy po 3 razy, sprawdzające poprawność mieszania próbek i przesyłania rezultatów.

Cel badań: Beta testy serwisu

Szacunkowy czas trwania (w minutach): 5

Identyfikator: 190

Data utworzenia: 27.01.2017

Data ostatniej aktualizacji: 27.01.2017

Numer testu: 1

Nazwa testu: Mushra

Metoda testowa: MUSHRA

Instrukcja: Proszę o zaznaczenie na skali wartości dla usłyszanego głosu: A - 100, B - 80, C - 60, D - 40.

Próbki	Rezultaty	Rezultaty niesformatowane
Próbka: A		
Pełna nazwa pliku: A.ogg		
Próbka: B		
Pełna nazwa pliku: B.ogg		
Próbka: C		
Pełna nazwa pliku: C.ogg		
Próbka: D		
Pełna nazwa pliku: D.ogg		

Numer testu: 2

Nazwa testu: Abx

Metoda testowa: ABX

Instrukcja: Prawidłowo dopasuj próbkę X do A lub B.

Próbki	Rezultaty	Rezultaty niesformatowane
Próbka: A		
Pełna nazwa pliku: A.ogg		
Próbka: B		
Pełna nazwa pliku: B.ogg		

Oznaczyć jako ukończony

Usuń

Rys. 14. Strona pokazująca szczegóły danej grupy testowej

Internetowy portal do oceny jakości dźwięku

Użytkownik: marcin

Wyloguj

Nazwa grupy: Beta testy serwisu

Opis grupy: Dwa testy, każdy po 3 razy, sprawdzające poprawność mieszania próbek i przesyłania rezultatów.

Cel badań: Beta testy serwisu

Szacunkowy czas trwania (w minutach): 5

Identyfikator: 190

Data utworzenia: 27.01.2017

Data ostatniej aktualizacji: 27.01.2017

Numer testu: 1

Nazwa testu: Mushra
Metoda testowa: MUSHRA

Instrukcja: Proszę o zaznaczenie na skali wartości dla usłyszanego głosu: A - 100, B - 80, C - 60, D - 40.

Próbki	Rezultaty	Rezultaty niesformatowane			
Próbka	Odpowiedzi	Średnia	Odchylenie standardowe próby	95% przedział ufności dla średniej	99% przedział ufności dla średniej
A (Hidden reference)	12	100.0	0.0	100.0 < m < 100.0	100.0 < m < 100.0
B	12	79.75	0.62	79.36 < m < 80.14	79.19 < m < 80.31
C	12	60.0	0.43	59.73 < m < 60.27	59.62 < m < 60.38
D	12	40.08	0.29	39.9 < m < 40.27	39.82 < m < 40.34

Numer testu: 2

Nazwa testu: Abx
Metoda testowa: ABX

Instrukcja: Prawidłowo dopasuj próbkę X do A lub B.

Próbki	Rezultaty	Rezultaty niesformatowane		
Odpowiedzi: 12				
Prawidłowe odpowiedzi: 8				
Procent prawidłowych odpowiedzi: 66.67%				
Prawdopodobieństwo uzyskania tylu lub więcej prawidłowych odpowiedzi w rozkładzie dwumianowym B(12, 0.5): 0.19385				
Skumulowana statystyka Pearsona $\chi^2$		$\chi^2$ dla p = 0.05		$\chi^2$ dla p = 0.01
1.333		3.841		6.635
Próbka	Odpowiedzi	Nieprawidłowe odpowiedzi	Prawidłowe odpowiedzi	Procent prawidłowych odpowiedzi
A	8	2	6	75.0%
B	4	2	2	50.0%

Oznacz jako ukończony

Usuń

Rys. 15. Rezultaty dla poszczególnych testów

Z1.10

Internetowy portal do oceny jakości dźwięku

Użytkownik: marcin Wyloguj

Nazwa grupy: Beta testy - tylko ABX

Opis grupy: Beta testy - tylko ABX

Cel badań: Beta testy - tylko ABX

Szacunkowy czas trwania (w minutach): 5

Identyfikator: 245

Data utworzenia: 27.01.2017

Data ostatniej aktualizacji: 27.01.2017

Numer testu: 1

Nazwa testu: ABX

Metoda testowa: ABX

Instrukcja: Dopasuj X do A lub B

Próbki	Rezultaty	Rezultaty niesformatowane
--------	-----------	---------------------------

Odpowiedzi: 9

Prawidłowe odpowiedzi: 9

Procent prawidłowych odpowiedzi: 100.0%

Prawdopodobieństwo uzyskania tylu lub więcej prawidłowych odpowiedzi w rozkładzie dwumianowym  $B(9, 0.5)$ : 0.00195

Skumulowana statystyka Pearsona $\chi^2$	$\chi^2$ dla $p = 0.05$	$\chi^2$ dla $p = 0.01$
9.0	3.841	6.635

Próbka	Odpowiedzi	Nieprawidłowe odpowiedzi	Prawidłowe odpowiedzi	Procent prawidłowych odpowiedzi
A	4	0	4	100.0%
B	5	0	5	100.0%

Oznacz jako ukończony

Usuń

Rys. 16. Rezultaty dla kolejnej grupy testowej

Numer testu: 1

Nazwa testu: Mushra

Metoda testowa: MUSHRA

Instrukcja: Proszę o zaznaczenie na skali wartości dla usłyszanego głosu: A - 100, B - 80, C - 60, D - 40.

Próbki	Rezultaty	Rezultaty niesformatowane
--------	-----------	---------------------------

[Pobierz plik w formacie JSON](#)

[Pobierz plik w formacie CSV](#)

Rys. 17. Zakładka umożliwiająca eksport rezultatów do pliku.

**Załącznik 2**  
**Instrukcja instalacji opracowanego portalu**  
**internetowego**

Aplikacja była poprawnie uruchamiana na systemie operacyjnym *Linux Mint 18.1* z jądrem systemu *Linux Kernel* w wersji 4.6.2-040602-generic (64 bit) i przeglądarką internetową *Chromium* (wersja 55.0.2883.87) oraz *Google Chrome* (wersja 56.0.2924.87).

## 1. Wymagania

1. Java JDK w wersji 1.8 lub wyższej,
2. Apache Maven w wersji 3.2 lub wyższej,
3. Node.js w wersji 7.4.0 lub wyższej,
4. npm w wersji 4.05 lub wyższej,
5. PostgreSQL 9.5,
6. Python (wersja 3.5) - do wykonania pomocnego w uruchomieniu aplikacji skryptu.

## 2. Konfiguracja

Wymagana konfiguracja przed rozpoczęciem procesu instalacji przedstawiona jest w następujących krokach:

1. Utworzyć repozytorium systemu kontroli wersji *Git* dla folderu *sqap-configurations* i umieścić je np. w serwisie repozytoriów takim jak *Github*.
2. Adres zamieszczonego repozytorium należy wstawić dla klucza *spring.cloud.config.server.git.uri* w pliku */sqap/sqap-config/src/main/resources/bootstrap.yml*
3. Należy stworzyć bazę danych, a dane konfiguracyjne umieścić w plikach: */sqap/sqap-api/src/main/resources/bootstrap.yml* i */sqap/sqap-auth/src/main/resources/bootstrap.yml* pod kluczem *spring.datasource*, zgodnie z nazwami kluczy (m.in. *url*, *username*, *password*) widocznych w obu plikach.

## 3. Instalacja

W celu instalacji należy przejść do katalogu */sqap* i wykonać następującą komendę: *mvn clean install -DskipTests -P prod-standalone*.

Po pomyślnej instalacji, utworzone moduły znajdować się będą w folderze wynikowym (ścieżka do folderu podana będzie w terminalu).

## 4. Uruchomienie

W celu uruchomienia aplikacji należy przejść do folderu zawierającego stworzone moduły i uruchomić dostarczony skrypt *start-script.py*. Przed uruchomieniem skryptu należy zmodyfikować w nim ścieżki do folderów zawierających moduły aplikacji (tj. pliki *Jar*).

Interfejs użytkownika portalu domyślnie dostępny będzie pod adresem: *http://localhost:8080*.