POLITECHNIKA GDAŃSKA

Technologia Studyjna

Wykład

Dźwięk w produkcjach multimedialnych

dr inż. Bartłomiej Mróz (EA726, bartlomiej.mroz@pg.edu.pl)

Wprowadzenie

- Przekaz multimedialny – informacja za pośrednictwem kilku współgrających mediów: tekst, grafika, dźwięk, animacja, wideo
- Interaktywność – coraz częstszy element
- Dźwięk – jeden z podstawowych nośników informacji
- Współczesne produkcje multimedialne (zwłaszcza gry AAA) przewyższają jakością dźwięku wiele produkcji filmowych
- Spatial audio i immersja stały się standardem, nie luksusem

Produkcje multimedialne – Przykłady

- Gry komputerowe i wideo (również VR/AR)
- Gry edukacyjne
- Encyklopedie multimedialne
- Interaktywne reklamy komercyjne
- Wizualizacje i symulacje
- Prezentacje multimedialne
- Materiały szkoleniowe
- Immersyjne doświadczenia muzyczne (360 3D Audio)
- Aplikacje metaverse
- Podcasty przestrzenne

Produkcje multimedialne – Przykłady

https://youtu.be/dkoU-4bRTW8

https://youtu.be/D9ZSwKYROVo

Rodzaje dźwięku

1. **Narracja i dialogi**
   - Głos lektora / narratora
   - Dialogi postaci (w grach – voice acting)
2. **Efekty dźwiękowe (SFX)**
   - Realistyczne (field recording)
   - Wymyślone / syntetyczne
   - Interfejsu (UI sounds)
3. **Muzyka**
   - Tła / atmosferyczna
   - Dźwięki diegetyczne (elementy świata fabulanego)
   - Nie-diegetyczne (tylko dla gracza)
4. **Dźwięk proceduralny**
   - Generowany algorytmicznie w runtime
5. **Audio immersyjne**
   - Spatial, binaural, 3D

Głos w multimediach

- W produktach multimedialnych dość często wykorzystuje się narrację lektora
- Współczesne gry komputerowe używają głosów podobnie jak filmy
- Każda postać ma swój głos, postaci mogą prowadzić dialogi
- Głosy mogą być syntetyzowane AI (ElevenLabs, StyleTTS) z pełną kontrolą emocji
- Voice cloning dla spójności między sesjami

Głos w multimediach

https://www.reddit.com/r/artificial/comments/1bhyi33/

Przygotowanie nagrań głosowych

- **Tradycyjne podejście:**
  - Nagranie w warunkach studyjnych
  - Odpowiednia postprodukcja
  - Kryteria doboru głosu (jak w produkcji filmowej)
- **Synteza mowy (TTS + AI):**
  - ElevenLabs, Google Cloud TTS, Azure Speech Services
  - Naturalna prozodia i kontrola emocji
  - Wiele języków w czasie rzeczywistym
  - Voice cloning – możliwość reprodukowania głosu
- **Najważniejsze: Ciągle wymaga oceny ludzkiej, ale jest to szybsze niż tradycyjny dubbing**
- **Współczesna praktyka**
  - Hybrydowe podejście: mowa syntetyzowana dla NPC, nagrania dla głównych postaci
  - Ulepszenia AI istniejących nagrań (usuwanie szmeru, wyrównywanie dynamiki)

Narracja – wybór głosu

- **Kryteria obiektywne do doboru głosu:**
  - Zrozumiałość – wyraźna dykcja, brak szumów
  - Pasowanie do grupy docelowej – dzieci, dorośli, senior itp.
  - Zgodność z marką – warm/cold, energetyczny/spokojny
  - Rejestr i barwa – zależne od kontekstu, nie od stereotypu płci
  - Spójność – głos musi być stabilny przez wiele sesji

Narracja – wybór głosu

- **Pomocne techniki:**
  - AI voice profiling (analiza parametrów głosu)
  - A/B testing z groupą docelową

Narracja – rejestracja głosu

- **Warunki studyjne – bez pogłosu, zakłóceń, szumów**
- **Jakość studyjna:**
  - Częstotliwość próbkowania: 48 kHz (standard), opcjonalnie 96 kHz
  - Bitowa głębia: 24-bit (współcześnie), minimum 16-bit, opcjonalnie 32 bit
  - Format: WAV PCM (bezstratny)
- **Wiele sesji – każda musi mieć identyczne warunki (mikrofon, pozycja, poziom)**
- **Monitoring w real-time – nie „na słuch", ale na miernikach**
- **Ulepszenia AI jako etap postprodukcji – usuwanie artefaktów, wyrównywanie**

Narracja – wyrównanie dynamiki

- **Problem: Głos „nieszkolony" (amatorski) ma zmienną dynamikę – „pływa"**
- **Rozwiązanie – kompresor:**
  - Obniża głośne części, unosi cichsze
  - Ustawienie: ratio 4:1, attack 10 ms, release 100 ms (start)
  - Cel: LUFS -18 do -16 (standard dla gier)
- **Narzędzia AI:**
  - iZotope RX Voice Leveler
  - Adobe Podcast (auto-enhance)
- **De-esser – redukcja głośnych sybilantów (s, z, sz)**
- **Waves DeBreath / iZotope RX Breath Control – usuwanie / redukcja oddychania między słowami**

Nagrania aktorskie

- **Wyrażenie emocji – aktor może celowo „przesterować" (zbyt głośno)**
- **Monitoring poziomu – w real-time, nie retrospektywnie**
- **Limiter – safety net, aby nic nie „klipiło"**
- **Interakcja aktor-inżynier: aktor musi wiedzieć, co jest za głośne**
- **Bonus: Aktorzy mogą imitować różne odgłosy (SFX na żywo)**

- **W kontekście gier:**
  - Motion capture + voice capture razem (synchro)
  - Real-time monitoring w silniku gry
  - Multitrack recording – osobne warstwy dla efektów/emocji
  - Voice direction w kontekście interaktywnych narracji (gra reaguje na głos)

Specjalne efekty dźwiękowe (SFX)

- **Źródła rzeczywistych efektów:**
  - Używane wprost (realistyczne)
  - Baza do nowego efektu (obróbka, transformacja)
- **Dwa sposoby pozyskania:**
  - Samodzielnie zarejestrować (field recording / nagrania terenowe)
  - Z bibliotek komercyjnych
- **Przykładowe biblioteki:**
  - Sonniss (darmowe, wysokiej jakości)
  - Boom Library (zaawansowane, Atmos support)
  - Pro Sound Effects (streaming)
  - Freesound (crowdsourced, creative commons)
- **Uwaga: Licencjonowanie – zawsze sprawdzaj warunki użytku**

Rejestracja efektów dźwiękowych – sprzęt

- **Rejestrator cyfrowy – Sound Devices MixPre, Zoom F, Tascam Portacapture**
- **Mikrofony kierunkowe – „shotgun"**
  - Selektywnie rejestruje źródło (bez tła)
  - Idealne dla precyzyjnych efektów
- **Mikrofony stereofoniczne**
  - Oddają przestrzeń dźwiękową
  - Kierunek poruszania się źródła
  - Idealne dla tła, ambience
- **Nowe technologie:**
  - Ambisonic microphones (Rode Soundfield) – 360° capture dla VR/AR
  - Binaural recording – stereo dummy head
  - Hydrophone – rejestracja podwodna

Rejestracja efektów dźwiękowych – sprzęt

Efekty – Rejestracja – Praktyka

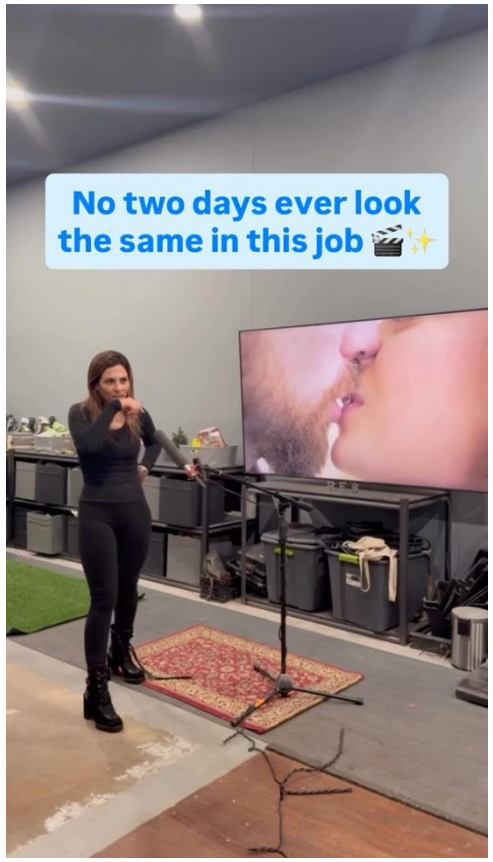
- **W warunkach studyjnych (jeśli możliwe)**
  - Kontrolowana akustyka
  - Bez nieprzewidzianych zakłóceń
- ***Location scouting* – warto rozpoznać miejsce przed sesją**
- ***Multi-track recording* – osobne kanały dla różnych źródeł**
- ***Metadata tagging* – każdy plik dokumentuj (co, gdzie, kiedy, ustawienia)**
- **Backup na miejscu – zawsze 2 kopie (redundancja)**

Imitacja dźwięku – Foley Studio

- **Potrzeba: Generowanie dźwięków, które nie istnieją w rzeczywistości, lub rejestracja jest trudna/niemożliwa**
- **Rozwiązanie: Sztuczna imitacja dźwięków – Foley studio (technika filmowa)**
- **Proces: Aktor/inżynier/realizator tworzy dźwięki na żywo, synchronizując z animacją/wideo**
- **Historia:**
  - Jack Foley (1891-1967) – pionier w Hollywood
  - Podstawowa technika: niezmieniona przez dekady
  - Współcześnie: hybrydowe (Foley + proceduralny audio + sample)

Imitacja dźwięku – Foley Studio

https://youtu.be/Va0syfetL4Y



https://instagram.com/reel/DNKIbqpxgkN/

imitacja dźwięków – co się tworzy w foley studio?

- **Klasyczne Foley:**
  - Odgłosy kroków (różne tereny – beton, las, śnieg)
  - Dźwięki wzmacniające akcję (walka, zamykanie drzwi, zerwane tkaniny)
  - Odgłosy codzienne (oddychanie, ruchy ubrania)
- **Dźwięki wykreowane (fantazja/sci-fi):**
  - Fantasy sounds (magiczne efekty, potwory)
  - Sci-fi sounds (lasery, pojazdy nieznanego pochodzenia)
  - Fantazyjne środowiska (obce planety)

Udźwiękowienie animacji – synchronizacja

- **Wymóg: Precyzyjne dopasowanie dźwięku do każdej sytuacji w video**
- **Tradycyjne: Foley artist synchronizował ręcznie z klatką**
- **Nowoczesne:**
  - *Animation curves*
  - *Event markers* –triggery w osi czasu
  - *Motion capture sync* – zapis głosu oraz ruchu całego ciała

- **Praktyka:**
  - Procedural footsteps – animacja nóg → dźwięk dynamiczny
  - Symulacja tkaniny – silnik fizyczny → dźwięki tkaniny/materiału
  - Wykrywanie uderzeń – pole kolizji → dźwięk zderzenia
  - Dźwięk oparty na fizyce – wysokość odbicia → dźwięk

Efekty interfejsu użytkownika (UI Sounds)

- **Dźwięki towarzyszące interakcji użytkownika z interfejsem**
- **Funkcja:**
  - Potwierdzenie dla użytkownika (coś się stało)
  - Orientacja w interfejsie
  - Feedback loop (audio != visual)
- **Przykłady:**
  - Pojawianie się komunikatów (zapytania, ostrzeżenia, błędy)
  - Wybór opcji (kliknięcia, swipes)
  - Najechanie na opcję / podświetlenie (hover sounds)
  - Inne interakcje (submit, load, error)

Efekty interfejsu użytkownika (UI Sounds)

https://youtu.be/b3X3Fl5LTwg

UI Sounds – Parametry i dostępność

- Charakter dźwięku zależy od:
- Specyfiki użytkownika (dzieci, młodzież, dorośli, senior itp.)
- Częstości interakcji (↑ częstość = ↓ intensywność dźwięku)
- Kontekstu (tyla, gra, edukacja)
- Możliwość dostosowania głośności (interfejs użytkownika oddzielony od muzyki i efektów dźwiękowych)
- Włączanie/wyłączanie każdej kategorii
- Wibracje zamiast dźwięku (dla osób niesłyszących)
- Wskazówki wizualne oprócz dźwięku
- Regulowana częstotliwość (dla osób z ADHD/szumami usznymi)
- Napisy dla osób niedosłyszących
- Obowiązek: WCAG 2.1 AA – wszystkie funkcje nie mogą opierać się wyłącznie na audio

Przetwarzanie efektów dźwiękowych – DSP

- **Prawie każdy efekt wymaga postprodukcji**
- **Cel: adaptacja efektu do danego produktu**
- **Ograniczenia: wyobraźnia i doświadczenie inżyniera**
- **AI-powered tools:**
  - iZotope RX – spectral repair, de-noising
  - Ultimate Vocal Remover – separacja źródeł (open-source, głos z muzyki)
  - Spleeter (Open-source) – dekompozycja muzyki
  - LANDR – auto-mastering
- **Typowe zabiegi DSP:**
  - EQ – korekcja barwy (boost/cut określonych pasm)
  - Kompresja – kontrola dynamiki
  - Reverb – dodanie przestrzeni
  - Modulation – chorus, flanger, phaser
  - Distortion – destrukcja, agresja
  - Time-stretching – zmiana tempa bez zmiany wysokości

Muzyka w produkcji multimedialnej

- **Znaczenie: nie mniejsze niż w filmie**
- **Źródła:**
  - Specjalnie skomponowana dla produktu
  - Już istniejąca muzyka (z zachowaniem praw autorskich)
- **Funkcje muzyki:**
  - Ustanawianie nastroju/atmosfery
  - Sygnalizacja emocji (bez narażania dialogu)
  - Dynamika (akcja vs. spokój)
  - Tożsamość marki (jingle, leitmotiv)
- **Wyzwania w grach:**
  - Interaktywność – muzyka nie jest statyczna
  - Synchronizacja z wydarzeniami (walka, eksploracja, dialog)
  - Pętle / adaptacja

Muzyka w multimediach – starszy model: pętle

- **Specyfika gier (zwłaszcza retro):**
  - Komponowane utwory musiały być łatwe do zapętlenia
  - Wymagania pamięciowe
- **Jak działały:**
  - Utwór 2-3 minutowy zapętlany w nieskończoność
  - Podczas gdy gracz eksploruje / gra
- **Problemy:**
  - ❌ Po 2-3 minutach słuchacz zmęczy się powtórzeniami
  - ❌ Trudne do synchronizacji z eventami (boss pojawiła się, ale muzyka jakoś musi się zakończyć)
  - ❌ Brak emocjonalnej dynamiki
  - ⚠️ Ciągle używane w indie games i mobilnych grach
  - ✅ W AAA – zastąpione systemami adaptacyjnymi

Muzyka w multimediach – systemy dynamiczne

- **Muzyka zmienia się w czasie rzeczywistym w odpowiedzi na zdarzenia gry**
- **Zamiast statycznych loopów – dynamiczny system**
- **Dwa główne podejścia:**

1. **Vertical Re-mixing (Layering)**
   - Warstwy (perkusja, bas, harmonia, linia prowadząca) włączają się / wyłączają
   - Spokojnie: tylko harmonia
   - Bitwa: perkusja + bas + dęte blaszane = pełny mix
   - Płynne przejście w 0.5-2 sekundy
2. **Horizontal Re-sequencing**
   - Instrument przechodzi między sekcjami (A → B → C → D)
   - Gra rozpoznaje: boss health = 50% → zmiana na sekcję B
   - Interaktywne triggery

- **Triggery adaptacyjności:**
  - Stan gry (Combat, exploration, dialogue, cutscene)
  - Parametry numeryczne (intensywność, zdrowie gracza, liczba przeciwników w pobliżu)
  - Lokacja (biom, budynek, podziemia)
  - Postęp fabuły (rozdział, punkt zwrotny, odkryte sekrety)
  - Środowisko (pogoda, pora dnia)

Audio generowane proceduralnie (Procedural Audio)

- **Co to jest dźwięk proceduralny?**
  - Dźwięk tworzony w czasie rzeczywistym (podczas gry, aplikacji)
  - Zamiast odtwarzania nagranych plików – algorytm oblicza falę dźwiękową na podstawie parametrów
  - Wynik: fala sinusoidalna (lub bardziej złożona) generowana żywo
- **Dlaczego to rewolucja w grach?**
  - Problem tradycyjny (sampli nagranych):
  - Wielki plik audio = dużo pamięci
  - Pętle nudzą się po 2-3 minutach
  - Zawsze ten sam dźwięk = mniej immersji

Audio generowane proceduralnie (Procedural Audio)

- **Rozwiązanie proceduralne:**
  - Gigabajty sampli → kilkabajtowe obliczenia
  - Każde odtworzenie inne (np. każdy krok ma inny dźwięk)
  - Dźwięk reaguje dynamicznie na parametry gry (prędkość gracza, teren, siła zderzenia)
  - Brak opóźnień – natychmiastowy feedback
- **Porównanie pamięci:**
  - Biblioteka 100 kroków na różnych materiałach: ~500 MB
  - Proceduralny generator kroków: ~50 KB kodu + ~ 200 KB danych akustycznych
  - Oszczędność: 2500x mniej pamięci

Audio generowane proceduralnie (Procedural Audio) – techniki

1. **Synteza modalna (Modal Synthesis)**
   - Symuluje drżenie i rezonans obiektu (jak struny w gitarze)
   - Każdy materiał ma inne częstotliwości rezonansowe
   - Wynik: realistyczne, "naturalne" brzmienie
   - Przykład: pukanie w różne materiały (drewno, metal, szkło)
2. **Synteza ziarnista (Granular Synthesis)**
   - Dzieli dźwięk rzeczywisty na maleńkie fragmenty (ms)
   - Łączone w nowe sekwencje – powstaje zupełnie nowy dźwięk
   - Wynik: cyfrowe, artystyczne dźwięki (sci-fi, magia)
   - Przykład: efekty magiczne w grach fantasy
3. **Modelowanie fizyczne (Physical Modeling)**
   - Opisuje akustykę za pomocą równań fizyki
   - Symuluje rzeczywistość: jak dźwięk rozchodzi się w materii
   - Wynik: bardzo realistyczne, ale złożone obliczeniowo
   - Przykład: echo w jaskini, propagacja fal w wodzie
4. **Synteza FM (Frequency Modulation)**
   - Moduluje częstotliwość dźwięku drugą częstotliwością
   - Stara technika, ale bardzo efektywna na CPU
   - Wynik: bogata, zmieniająca się barwa
   - Przykład: dźwięki silników, dzwony, brzmienia metalowe
5. **Synteza konkatenatywna (Concatenative Synthesis)**
   - Łączy fragmenty rzeczywistych nagrań w nowe kombinacje
   - Kombinuje naturalne brzmienie nagrań z nową sekwencją
   - Wynik: pół-naturalne, pół-syntetyczne dźwięki
   - Przykład: wariacje głosu (różne wymowy tego samego słowa)

Audio generowane proceduralnie (Procedural Audio) – narzędzia

- **GameSynth (Tsugi)** – edykowana platforma do dźwięków proceduralnych w grach; intuicyjny interfejs graficzny (GUI), eksport VST i middleware
- **Pure Data / Max/MSP** – otwarte, bardzo elastyczne środowiska do syntezy i przetwarzania audio; umożliwiają prototypowanie rozwiązań proceduralnych (wymagają jednak nauki niskopoziomowych pojęć)
- **Wwise (AudioKinetic)** – middleware audio z gotowymi generatorami (noise, tone), obsługą parametrów gry i możliwością tworzenia własnych modułów DSP
- **FMOD Studio** – middleware audio ze wsparciem dla proceduralnego audio i własnych efektów DSP
- **Custom DSP (C++)** – najczęstsze rozwiązanie w projektach AAA wymagających pełnej kontroli; wymaga zaawansowanej wiedzy i własnych implementacji algorytmów

Audio Middleware – Wwise vs FMOD

	Wwise (Audiokinetic)	FMOD (Firelight)
Interfejs	Zaawansowany, więcej opcji, bardziej złożony	Intuicyjny, zbliżony do DAW, szybka nauka
Krzywa uczenia	Stroma (wymaga szkolenia)	Łagodniejsza (szybko osiągnąć rezultaty)
Społeczność/Wsparcie	Droższe, ale profesjonalne szkolenia i certyfikaty	Społeczność open-source, mniej oficjalnego wsparcia
Cena	Darmowa wersja z limitami, droższe licencje dla dużych projektów	Darmowa wersja z limitami przychodów, bardziej dostępne cenowo
Najczęściej używane w	AAA studia (EA, Ubisoft, Sony, Activision)	Indie studia, gry mid-budget (Celeste, Hades)

Dźwięk przestrzenny (Spatial Audio) – fundamenty

- **Co to jest dźwięk przestrzenny/immersyjny?**
  - Iluzja lokalizacji źródła dźwięku w przestrzeni
  - Przód/tył, góra/dół, lewo/prawo
  - Różne od stereo (L/R) – prawdziwa 3D
- **Jak mózg lokalizuje dźwięk?**
  - ITD (*Interaural Time Difference*) – różnica czasowa między uszami
  - ILD (*Interaural Level Difference*) – różnica amplitudy
  - HRTF (*Head-Related Transfer Function*) – filtr specyficzny dla każdej osoby
  - Modulacja przez kształt ucha (określa pion)
  - Wskazówki ruchu głowy, odruch szyi
- **HRTF – klucz do immersji:**
  - Charakterystyka filtracyjna między przestrzenią a uszami
  - Zależy od: kształtu głowy, uszu, ramion, gęstości włosów(!), kształtu ciała
  - Standardowe HRTF (Kemar, ARI) nie pasują każdemu
  - Personalizacja HRTF – ML dla indywidualnych filtrów
  - Nowe technologie: HRTF ze zdjęcia twarzy / skanu 3D

Dźwięk przestrzenny (Spatial Audio) – zastosowania

- **Gry VR i AR**
  - Gracz nosi słuchawki + headset ze śledzeniem głowy
  - Silnik audio binauralizuje każdy dźwięk na podstawie kąta jego głowy
  - Przykład: gra akcji – strzelec podchodzi od tyłu, słyszysz kroki za sobą
- **360° Video (YouTube, FB, DeoVR)**
  - Video zawiera ścieżkę ambisoniczną (FOA, 4 kanały)
  - Player dekoduje binauralnie, gdy patrzysz w określonym kierunku
  - Jak obracasz widok – audio "rotuje" razem z obrazem
- **Instalacje Immersive (Kina, Muzea)**
  - Duża sala ze sferycznym zestawem głośników (kopuła, kostka)
  - Dźwięk ambisoniczny dekodowany do wszystkich głośników
  - Słuchacz siedzi w środku i doświadcza pełnego otoczenia 360°

Dźwięk przestrzenny (Spatial Audio) – zastosowania

- **FOA (4 kanały) to *de facto* standard w 2025**
- **HOA to wciąż nisza (badania, specjalne instalacje)**
- **Binauralizacja to „magiczne” połączenie akustyki i psychoakustyki**
- **Dźwięk 3D bez głośników wokół – to wynik lat badań nad tym, jak mózg lokalizuje dźwięk**
- **Ambisonics ≠ object-based audio (Atmos)**
  - Atmos: "Przesuń trąbkę o 2m w lewo"
  - Ambisonics: "Reprezentuję całą sferę pola dźwiękowego, zdekoduj jak chcesz"
- **VR jest game-changer'em**
  - Ambisonics było niszowe, dopóki VR nie stało się popularne
  - Teraz: każdy VR game powinien mieć spatial audio

Dostępność i inclusive design w audio

- **Dlaczego dostępność audio to biznes, a nie dobroczynność?**

- **Rynkowa perspektywa:**
  - ~1.3 miliarda osób na świecie żyje z niepełnosprawnościami (słuch, wzrok, mobilność, poznawcze)
  - Z tego ~430 milionów doświadcza znaczącej utraty słuchu wymagającej wsparcia
  - Do 2050: prognozuje się, że 2.5 miliarda osób będzie mieć problemy ze słuchem
  - Rynek: każdy gracz/użytkownik z niepełnosprawnościami to konsument, który kupi grę/aplikację, jeśli będzie dostępna

- **Przepisy – nie opcja:**
  - USA: Section 508, ADA (Americans with Disabilities Act) – obowiązkowe dla usług publicznych
  - UE: European Accessibility Act (EAA) – weszła w życie 28 czerwca 2025; dotyczy transportu, handlu, bankowości, mediów
  - Polska: Ustawa o dostępności cyfrowej – zgodnie z dyrektywą UE
  - Sankcje: Kary pieniężne, pozwy, utraty reputacji w mediach społecznych

Dostępność i inclusive design w audio

- **Dlaczego dostępność audio to biznes, a nie dobroczynność?**
- **Korzyści biznesowe:**
  - Poszerzenie rynku (5% populacji to znaczący segment)
  - Pozytywne opinie i rating (accessibility features = dobra opinia)
  - Prestiż (studia z dobrą dostępnością są bardziej atrakcyjne dla talentów)
  - Długożyciowość produktu (gra dostępna pozostaje atrakcyjna dłużej)
- **Etyka:**
  - Włączenie marginalizowanych grup w design procesie od początku
  - Accessibility to nie "dodatkowa funkcja" – to standard

Dostępność i inclusive design w audio – standardy i regulacje

- **WCAG 2.1/2.2 (*Web Content Accessibility Guidelines*)**
  - Międzynarodowy standard dla dostępności cyfrowej
  - Poziom A: Minimum – podstawowe wsparcie
  - Poziom AA: Rekomendowany – prawie wszystkie przypadki pokryte (Standard branżowy)
  - Poziom AAA: Aspiracyjny – wszechstronny (trudny, nie zawsze możliwy)
  - Dla audio: napisy, opisy, transkrypcje, alternatywy
- **Section 508 (USA):**
  - Wymaga dostępu dla osób z niepełnosprawnościami do technologii federalnej
  - Obowiązkowy w kontraktach rządowych
- **EN 301 549 (Europa):**
  - Harmonizacja standardów dostępności w UE
  - Konkretne wymogi dla produktów i usług

Dostępność audio dla osób głuchych, niedosłyszących

- **Główne wyzwania - osoby głuche i niedosłyszące:**
  - Ostrzeżenia dźwiękowe, dialogi, lokalizacja dźwięku, komunikacja głosowa
- **Wyzwania techniczne:**
  - Różne urządzenia (słuchawki vs głośniki)
  - Trudności w VR (czytelność)
  - Wyzwania w komunikacji multiplayer (zamienniki audio)
- **Praktyczne rozwiązania:**
  - Napisy (captions) z identyfikacją mówiącego
  - Wizualne substytuty dźwięków (ikony, animacje)
  - Haptic feedback (wibracje kontrolera)
  - Transkrypcja scen video
  - Możliwość wyłączenia/zmiany głośności powiadomień UI

Dostępność audio dla osób niewidomych, niedowidzących

- **Główne wyzwania - osoby niewidome i słabowidzące:**
  - Brak / utrudniony dostęp do elementów wizualnych i orientacji w przestrzeni
- **Praktyczne rozwiązania:**
  - *Audio description* – opis wzrokowych wydarzeń narratora
  - Sonifikacja – przedstawianie informacji (np. pozycji obiektów, interakcji, zagrożeń) za pomocą wyraźnych sygnałów dźwiękowych
  - Tryby wysokiego kontrastu, powiększania tekstu
  - Kompatybilność z czytnikami ekranu

Dostępność audio dla osób z zaburzeniami słuchu, nadwrażliwością

- **Dla osób z zaburzeniami przetwarzania słuchowego:**
  - Problemy z interpretacją dźwięku w hałasie lub szybkim tempie dialogów
- **Rozwiązania:**
  - Regulacja tempa dialogów (0.5x-2x)
  - Zmiana wysokości dźwięków (pitch)
  - Usuwanie pogłosu (echo reduction)
  - Uproszczenie do mono audio
  - Oddzielna kontrola głośności dialogu, muzyki, efektów
- **Dla osób z ADHD lub nadwrażliwością na dźwięk:**
  - Przeszkadzające, agresywne powiadomienia
  - Nadmiar dźwięków, nadmierna stymulacja
- **Rozwiązania:**
  - Regulacja krzywej głośności
  - Presety audio (Quiet, Normal, Cinematic)
  - Możliwość wyłączenia dźwięków powiadomień
  - Filtry częstotliwości (wysokie, niskie tonacje)