WYKŁAD 9

Informatyka afektywna w aplikacjach mobilnych

Programowanie Aplikacji Mobilnych

dr inż. Mateusz Pomianek

Emocje · ML On-Device · Prywatność · UX Adaptacja

- Modele emocji i sygnały afektywne
- Rozpoznawanie mimiki i mowy
- Fizjologia mobilna i EDA, PPG, IMU
- ML on-device: TFLite i CoreML
- Etyka, prywatność i zastosowania

Plan wykładu

1. **Czym jest informatyka afektywna**
2. **Modele emocji**
3. **Sygnały afektywne, taksonomia kanałów pomiarowych:**
    - Rozpoznanie wyrazu twarzy
    - Analiza mowy i prozodii
    - Analiza sentymentu tekstu
    - Fizjologia mobilna
    - Ruch i kontekst, IMU jako kanał afektywny
4. **Fuzja multimodalna i architektury ML on-device**
5. **Interfejs afektywny UX**
6. **Zastosowania, Etyka, prywatność i RODO**
7. **LLM + emocje, generatywna empatia**

Czym jest Informatyka Afektywna?

"Affective computing is computing that relates to, arises from, or deliberately influences emotions" — Rosalind W. Picard, MIT Media Lab, 1997

Definicja (Picard, 1997)

Systemy komputerowe zdolne do rozpoznawania, interpretowania symulowania ludzkich emocji. Cel: naturalna, empatyczna interakcja człowiek-maszyna (HCI).

Dlaczego mobile?

Smartfon jest zawsze przy użytkowniku — dostęp do kamery, mikrofonu, sensorów IMU, GPS, danych biometrycznych. Kontekst 24/7 nieosiągalny dla tradycyjnych systemów.

Wyzwanie: subiektywność emocji

Ten sam bodziec → różne reakcje emocjonalne u różnych osób. Brak złotego standardu etykietowania. Inter-rater agreement często < 70% w datasetach.

Kamienie milowe

1997	Picard — książka Affective Computing, MIT Media Lab
2001	Cohn et al. — automatyczna detekcja AU (Action Units) z wideo
2010	Siri (Apple) — pierwszy masowy asystent z intent recognition
2015	Microsoft Azure Emotion API — publiczne cloud API do mimiki
2017	Apple ARKit + TrueDepth — dense face mesh na urządzeniu
2019	Affectiva AFFDEX SDK — real-time 7 emocji w przeglądarce/mobile
2021	Apple AirPods Pro + headphone IMU — head emotion tracking
2023	GPT-4V + Claude — LLM zdolne do opisu emocji z obrazu/tekstu
2024	Meta Llama-3 on-device — LLM afektywny lokalnie na telefonie

Modele emocji: Ekman, VAD i Russell Circumplex

Wybór modelu emocji determinuje architekturę klasyfikatora.
Dyskretny (klasy) vs ciągły (wymiary) to fundamentalna decyzja projektowa

Model Ekmana: 6 emocji podstawowych

Radość	Smutek	Gniew
Strach	Wstręt	Zaskoczenie

Ekman + pogarda = 7 (FACS basic set) | Mikael Kobayashi:
+26 złożonych emocji z kombinacji podstawowych

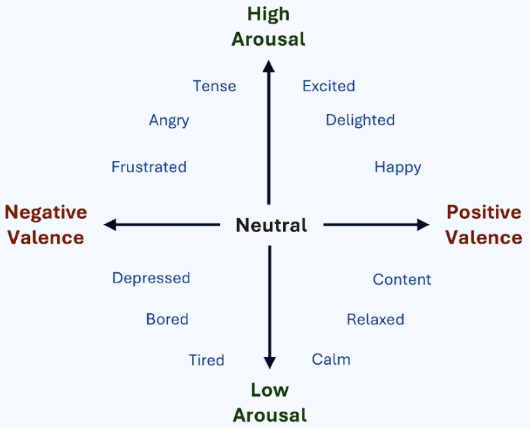
Model VAD (Valence–Arousal–Dominance)

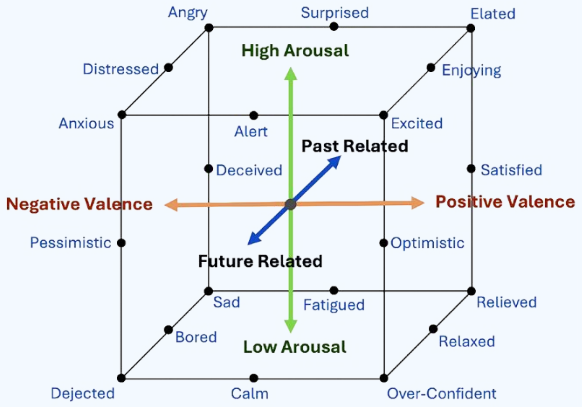
Valence (V)
Przyjemność–nieprzyjemność. Skala −1..+1. Radość: +0.9 | Smutek: −0.8. Koreluje z uśmiechem i kontrastem głosu.

Arousal (A)
Poziom pobudzenia. Spokój (0) – ekscytacja (1). Gniew: 0.7 | Zadowolenie: 0.4. IMU: korelacja z ruchem ciała.

Dominance (D)
Poczucie kontroli. Strach: 0.1 (brak kontroli) | Gniew: 0.8 (dużo kontroli). Rzadziej modelowana w mobilnych systemach.

Russell Circumplex Model of Affect



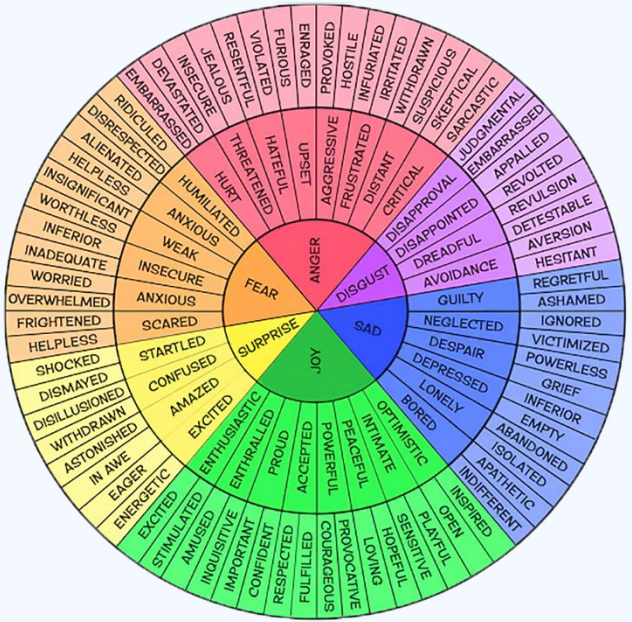


Wykorzystanie VAD w urządzeniach mobilnych: ciągła regresja zamiast klasyfikacji. Model CNN lub LSTM przewiduje wartości (v, a, d) na podstawie sygnału twarzy/głosu, a następnie przyporządkowuje je do etykiety emocji.

Koło emocji Plutchika i emocje złożone

Plutchik (1980): 8 podstawowych emocji bipolarnych + 8 wyższego rzędu.
Model hierarchiczny z intensywnością i kombinacjami diagonalnymi

Emocja podst.	Przeciwnieńs.	Intensywna	Słaba	Kombinacja diagonalna
Radość	Smutek	Ekstaza	Pogoda ducha	Miłość (Joy+Trust)
Zaufanie	Wstręt	Podziw	Akceptacja	Nadzieja (Trust+Anticipation)
Strach	Gniew	Terror	Trwoga	Poddanie (Trust+Fear)
Zaskoczenie	Oczekiwanie	Zdumienie	Rozproszenie	Lęk (Fear+Surprise)
Smutek	Radość	Żal	Zaduma	Wyrzuty sumienia (sadness+disgust)
Wstręt	Zaufanie	Nienawiść	Nuda	Pogarda (Disgust+Anger)
Gniew	Strach	Wściekłość	Irytacja	Agresja (Anger+Anticipation)
Oczekiwanie	Zaskoczenie	Czujność	Zainteresowanie	Optymizm (Anticipation+Joy)



Praktyczne zastosowanie: Affectiva AFFDEX, DeepFace i Microsoft FER+ używają kombinacji Ekmana+Plutchika (7–8 klas). NLP sentiment: często 3 klasy (pos/neg/neutral) lub VAD ciągły z wielowymiarowym BERT fine-tune.

Taksonomia kanałów pomiarowych, czyli skąd odczytujemy emocje?

Bogactwo sygnałów afektywnych w smartfonie. Każdy kanał niesie inne aspekty stanu emocjonalnego, a ich fuzja zwiększa precyzję

Kanał	HW	Sygnał	Dokładność	Latencja	Zastosowanie
Wyraz twarzy	Kamera FFC (selfie)	Ruch mięśni twarzy, AU	60–85%	< 50ms	Reklama, e-learning, gry
Głos i mowa	Mikrofon	Pitch, tempo, barwa, spektrogram	55–80%	100–500ms	Call-center, terapia, nawigacja
Tekst / NLP	Klawiatura, chat	Sentyment, emocja z leksyki i składni	70–90%	< 200ms	Social media, chatbot, HR
EDA / GSR	Wearable elektrody (skóra)	Przewodność skóry [µS], SCR	70–82%	1–5s	Stres, lie detection, gry VR
Tętno PPG	Smartwatch, aparat	HR, HRV, fotopletyzmograf	65–78%	2–10s	Fitness, wellbeing, automotive
IMU / Ruch	Akcelerometr + żyroskop	Tempo chodzenia, gesty, agitacja	50–72%	1–3s	Depresja, agitacja, aktywność

EEG (Muse, Emotiv, OpenBCI) daje najwyższą precyzję (> 85%) ale wymaga zewnętrznego headsetu, poza ekosystemem standardowego smartfona. Potencjał w edukacji i medycynie.

FACS: Facial Action Coding System (Ekman & Friesen, 1978)

FACS koduje ruch twarzy przez 44 Action Units, niezależne aktywacje mięśni.
To złoty standard etykietowania danych treningowych dla algorytmów mimiki.

Kluczowe Action Units

AU	Nazwa	Emocje
AU01	Inner Brow Raise	Smutek, Strach
AU02	Outer Brow Raise	Zaskoczenie
AU04	Brow Lowerer	Gniew, Smutek
AU06	Cheek Raiser	Radość prawdziwa
AU12	Lip Corner Puller	Uśmiech (fałszywy gdy brak AU06)
AU17	Chin Raiser	Wstręt, Smutek
AU23	Lip Tightener	Gniew, dezaprobata
AU43	Eyes Closed	Zmęczenie, intencja nie-uwagi

Google ML Kit Face Mesh: 478 punktów siatki twarzy w czasie rzeczywistym, wykrywa mruganie, Smiling, Left/RightEyeOpen probability bez AU. ARKit ARFaceAnchor: 52 blendshapes ≈ 52 AU-like coefficients.

Pipeline rozpoznawania mimiki na mobile

1. Face Detection
MTCNN / BlazeFace (TFLite) — <5ms. Bounding box + 5 landmarks (oczy, nos, kąciki ust).

2. Face Alignment
Affine warp do 112×112 px na podstawie landmarks. Normalizacja oświetlenia CLAHE.

3. AU Detection
CNN (ResNet-50/MobileNetV3) → AU intensities (0–5 skala lub binarnie). 44 AU wyjście.

4. Expression Class.
Reguły FACS → etykieta emocji LUB regresja VAD. Opcjonalnie mapa uwagi (Grad-CAM).

5. Temporal smooth
Exponential smoothing / Kalman filter — usuwa jitter klatka-po-klatce. 30fps → stable output.

Rozpoznawanie mimiki: praktyczna implementacja na mobile

Android: ML Kit + TFLite Expression Classifier

Kotlin

```
// ML Kit face detection + emotion classifier
val faceDetector = FaceDetection.getClient(
  FaceDetectorOptions.Builder()
    .setPerformanceMode(PERFORMANCE_MODE_FAST)
    .setClassificationMode(CLASSIFICATION_MODE_ALL)
    .build())

faceDetector.process(inputImage)
  .addOnSuccessListener { faces ->
    faces.firstOrNull()?.let { face ->
      val smiling = face.smilingProbability ?: 0f
      val leftEye = face.leftEyeOpenProbability ?: 0f
      // ROI twarzy → TFLite emotion classifier
      val emotion = emotionClassifier.classify(
          cropFaceROI(bitmap, face.boundingBox))
      updateUI(emotion, smiling, leftEye)
    }
  }
```

iOS: ARKit Face Tracking

Swift

```
// ARKit ARFaceAnchor — 52 blendshapes
func renderer(_ renderer: SCNSceneRenderer,
    didUpdate node: SCNNode, for anchor: ARAnchor) {
  guard let face = anchor as? ARFaceAnchor else
{ return }
  let smileL =
face.blendShapes[.mouthSmileLeft]!.floatValue
  let smileR =
face.blendShapes[.mouthSmileRight]!.floatValue
  let browD  =
face.blendShapes[.browDownLeft]!.floatValue
  let blink  =
face.blendShapes[.eyeBlinkLeft]!.floatValue
  // Brak dostępu do personalnych danych poza
urządzeniem
  let emotion = classifyEmotion(smile:
```

Przykład: ML Kit Face Detection + Custom TFLite expression classifier, analiza emocji real-time z kamery selfie na Android/iOS

Praktyczne przykłady aplikacji

Affectiva AFFDEX SDK
B2B SDK — analiza emocji widzów reklam wideo. Używany przez P&G, Unilever. 7 emocji + 20 AU + valence/attention.

Realeyes
Player ogląda wideo → kamera rejestruje reakcje → emocje w czasie rzeczywistym → raport dla marketerów. GDPR compliant: blur twarzy po AU.

iMotions Mobile
Platforma naukowa: face + EDA + eye tracking na iPadzie. Badania UX, consumer research. Akademickie użycie.

Snap Lenses (dynamiczne)
Snap Camera Kit — filtr reaguje na mrugnięcie, uśmiech. Miliony użytkowników, ML Kit BlazeFace < 3ms.

Nura (spersonalizowane słuchawki)
Mierzy ASSR (auditory steady-state response) przez aparat: analiza odpowiedzi emocjonalnej na dźwięk. Personalizacja EQ.

Rozpoznawanie emocji z mowy: Speech Emotion Recognition (SER)

Głos niesie emocje przez treść (co mówimy - NLP), prozodię (jak mówimy - F0, tempo), barwę i rytm. SER operuje głównie na cechach akustycznych.

Ekstrakcja cech akustycznych

MFCC (Mel-Frequency Cepstral Coefficients)
13–40 współczynników na każdą ramkę 25ms. Mimika spektrum percepcji słuchowej. Złoty standard SER — cechy + delta + delta-delta = 39-dim.

F0 / Pitch (Fundamental Frequency)
Częstotliwość podstawowa tonu głosu [Hz]. Gniew: wysoki wzrost. Smutek: niski, opadający. Algorytm CREPE (CNN) lub YIN do estymacji.

Energy & Loudness (RMS)
Energia sygnału → głośność. Gniew/Ekscytacja: wysokie. Smutek/Spokój: niskie. Prostota wydobycia — używane jako baseline.

Mel Spectrogram (2D image)
STFT → mel-scale → log-scale → obraz 2D. Wejście do CNN (ImageNet pretrain). Najlepsza reprezentacja dla głębokich sieci end-to-end.

OpenSMILE (eGeMAPS feature set)
88 cech (GeMAPS) lub 6373 (IS13). Standard w INTERSPEECH challenges. Dostępne w mobile przez JNI/wrapper.

Architektura modelu SER

CNN na Mel Spectrogram — Acc: 78%
ResNet-34 na 128×128 px mel. Pretrain ImageNet → fine-tune IEMOCAP. 15MB TFLite.

LSTM na MFCC sequences — Acc: 74%
2×LSTM 256 units → CRF output. Modeluje temporal dynamics. Trochę słabszy od CNN.

Transformer (wav2vec2) — Acc: 87%
facebook/wav2vec2-base fine-tuned → SOTA. 400MB — wymaga kompresji (quantization int8) do ~50MB na mobile.

TinySER (MobileNet v3) — Acc: 71%
Kompaktowy model dedykowany mobile. < 2MB, < 5ms latencja. Dobry baseline dla produkcji.

Kotlin

```
// Android — SER z TFLite + AudioRecord
val audioData = ShortArray(BUFFER_SIZE)
audioRecord.read(audioData, 0, BUFFER_SIZE)
// Konwertuj do mel-spektrogramu (Librosa-compatible)
val melSpec = MelSpectrogram.compute(audioData,
sampleRate=16000)
val input = TensorBuffer.createFixedSize(
    intArrayOf(1, 128, 128, 1), DataType.FLOAT32)
input.loadArray(melSpec.normalized())
emotionInterpreter.run(input.buffer, outputBuffer.buffer)
// output: [anger, disgust, fear, happy, neutral, sad, surprise]
val emotionIndex = outputBuffer.floatArray.indexOfMax()
```

Analiza sentymentu i emocji w tekście: NLP on Mobile

Tekst jest najłatwiej dostępnym kanałem afektywnym w aplikacjach mobilnych: chat, dziennik, komentarze, e-mail. Klucz: on-device NLP dla prywatności.

Pipeline NLP afektywnego

Tokenizacja i preprocessing

WordPiece/SentencePiece tokenizer (BERT-like) lub Word2Vec. Usuwanie stopwords, lowercase, stemming. Emoji → sentiment token (→ [HAPPY]).

Leksykalne podejście (szybkie)

VADER: lexicon 7500 słów z ocenami sentymentu. Reguły dla negacji (not happy). SentiWordNet. Doskonały baseline, 0.05ms na zdanie, offline.

BERT-based (dokładne)

DistilBERT (66M→40MB) lub MobileBERT (25MB TFLite) fine-tuned na GoEmotions dataset (27 emocji, Reddit). State of the art 0.7+ F1.

GoEmotions Dataset (Google)

58k komentarzy Reddit, 27 klas emocji + neutral. Hierarchia: 27 → 7 (Ekman) → 3 (pos/neg/ambig). Publiczny dataset do fine-tuningu.

```kotlin
// Android — TFLite Text Classifier (MediaPipe)
val clf = TextClassifier.createFromFile(ctx, "sentiment.tflite")
// → [{label:"joy", score:0.82}, {label:"sadness", score:0.12}]
```

Praktyczne przykłady NLP afektywnego

Woebot (chatbot terapeutyczny)

Analiza sentymentu wiadomości CBT chatbota — personalizacja kolejnych promptów. 1.5M+ użytkowników. NLP on-device dla prywatności.

Wysa (mental health app)

Konwersacyjny AI z emotion tracking z tekstu i samooceny PHQ-9. Trigger dla human escalation gdy sentyment bardzo negatywny > 3 dni.

Day One / Reflectly (dzienniki)

Analiza emocji wpisów dziennika → moodtracking wykres. Reflectly używa BERT lite do tygodniowych raportów nastroju.

Gmail Smart Reply / Keyboard

Gboard Emoji Prediction — BERT analizuje napisany tekst → sugestia emoji odpowiadającego emocji. Działanie w pełni on-device.

Apple Intelligence (iOS 18)

Podsumowanie powiadomień z emotion-aware tone. Priorytet: wiadomości z pilnymi/stresującymi emocjami wyżej w liście.

LLM jako silnik afektywny: GPT, Gemini i modele On-Device

Duże modele językowe rewolucjonizują NLP afektywne. Rozumieją kontekst, ironię, sarkasm i metafory emocjonalne na poziomie wcześniej nieosiągalnym

GPT-4 / Claude jako emotion analyser

Zero-shot: prompt zawiera zapis rozmowy → LLM zwraca JSON z emocjami i intensywnościami. Few-shot: 3–5 przykładów poprawia F1 o 8–12%.

Gemini Nano (on-device, Pixel/Samsung)

Quantized 1.8B model uruchamiany lokalnie przez Android ML API. Wnioskowanie emocji z kontekstu rozmowy bez sieci. Prywatny, szybki (200ms).

Llama 3.2 1B/3B (iOS/Android)

Meta open-source — fine-tune na EmoBank/GoEmotions → affective conversational model. ollama mobile lub llama.cpp binding przez JNI.

Ograniczenia LLM dla emocji

Brak real-time 30fps (latencja 200ms–2s). Halucynacje emocjonalne. Brak dostępu do sygnałów fizjologicznych. Wymaga orchestracji.

Prompt engineering dla analizy emocji

```kotlin
// System prompt dla affective analysis
val systemPrompt = """
Analyze the emotional state of the user's message.
Return JSON: {"primary":"joy|sadness|anger|fear|disgust|surprise|neutral",
"valence":-1.0..1.0, "arousal":0.0..1.0, "crisis":false}
Be concise. If crisis indicators present, set crisis:true
"""
```

Architektura hybrydowa LLM + klasyczny ML

```kotlin
// Orchestration: szybki model + LLM fallback
suspend fun analyzeEmotion(text: String): EmotionResult {
// 1. Szybki on-device BERT (~20ms)
val quickResult = bertClassifier.classify(text)
// Jeśli pewność wysoka → zwróć natychmiast
if (quickResult.confidence > 0.85f) return quickResult
// 2. Fallback do LLM gdy niepewność lub kryzys
val llmResponse = if (quickResult.crisisKeywords) {
geminiNano.generate(systemPrompt + text) // on-device
} else {geminiApi.generate(systemPrompt + text) // cloud}
return parseEmotionJson(llmResponse)}
data class EmotionResult(
val primary: String,
val valence: Float, // -1..1
val arousal: Float, // 0..1
val confidence: Float,
val crisis: Boolean,
val crisisKeywords: Boolean
)
```

Przypadek: Replika AI

Chatbot companion — LLM odpowiada empatycznie uwzględniając emotion state z historii rozmów. Użytkownicy reportują redukcję samotności (Shimokawa 2022)

Etyczny dylemat LLM

LLM może symulować empatię bez jej rozumienia. Ryzyko: user attachment do nieświadomego systemu. IEEE Ethically Aligned Design wymaga transparentności.

EDA / GSR: Elektrodermalna aktywność jako wskaźnik pobudzenia



EDA (Electrodermal Activity) = GSR (Galvanic Skin Response). Zmiana przewodności skóry pod wpływem aktywacji układu współczulnego. Najlepszy wskaźnik pobudzenia.

Fizjologia i sygnał

Mechanizm biologiczny

Gruczoły potowe ekrynowe na dłoniach i palcach aktywowane przez układ współczulny (nie temperaturą). Pot → zwiększona przewodność [µS]. Reakcja: 1–5 s po bodźcu.

Dekompozycja sygnału EDA

Tonic (SCL, Skin Conductance Level): powolny baseline. Phasic (SCR, Skin Conductance Response): szybkie szczyty po bodźcu. Filtr cvxEDA lub Ledalab do separacji.

Sprzęt mobilny

Empatica E4 wristband: EDA + PPG + temp + ACC, BLE → Android/iOS API. Shimmer3: badawczy. Polar: tylko PPG. FitBit nie udostępnia surowego EDA.

Zastosowanie kliniczne: stres

SCR frequency > 6/min → stres. Mean SCL > 10 µS → wysoka aktywacja. Kalibracja indywidualna niezbędna — duże różnice bazowe między osobami.

```kotlin
// Empatica E4 BLE → Android
e4Client.subscribe(EmpaStatus.CONNECTED) { status ->
    e4Client.subscribeToSignal(EmpaDataDelegate.GSR) { gsr,
ts ->
        val conductance = gsr  // [µS]
        edaBuffer.add(Pair(ts, conductance))
        if (edaBuffer.hasSCR()) triggerArousalEvent()
    }
}
```

PPG przez kamerę: rPPG bez wearable

Remote PPG (rPPG)

Kamera selfie rejestruje mikrozmiany koloru skóry twarzy → ekstrakcja PPG bez kontaktu. Dokładność: HR ±5 bpm, HRV gorsza. Google: Pixel Heart Rate z kamerą.

Algorytm rPPG (ICA / POS)

Plane of Skin (POS): 3 kanały RGB twarzy ROI → normalizacja → PCA/ICA → dominantna częstotliwość HR. Wymagana stabilna głowa i oświetlenie.

HRV jako wskaźnik stresu

Heart Rate Variability (HRV): zmienność odstępów RR. Niskie HRV to stres i przemęczenie. RMSSD < 20ms → wysoki stres. Najlepszy nieinwazyjny biomarker

```kotlin
// rPPG przez kamerę — szacowanie HR
// Używa biblioteki rPPGlib lub OpenCV
fun extractRPPG(frames: List<Bitmap>): Float {
    // 1. Wykryj ROI twarzy (czoło/policzki)
    val rois = frames.map { detectFaceROI(it) }
    // 2. Uśrednij kolor R,G,B w ROI
    val traces = rois.map { computeRGBTrace(it) }
    // 3. POS algorithm decomposition
    val ppgSignal = POSAlgorithm.compute(traces)
    // 4. FFT → dominantna częstotliwość
    val fftResult = FastFourierTransform(ppgSignal)
    val hrHz = fftResult.peakFrequency(0.7f, 3.5f)  // 42–210
bpm
    return hrHz * 60f  // [bpm]
}
```

IMU jako kanał afektywny: ruch ciała i stres

Ruch ciała odzwierciedla stan emocjonalny, agitacja, spowolnienie, tremor, wzorzec chodzenia.

Akcelerometr i żyroskop bez uprawnień runtime, czyli zawsze dostępne.

Sygnał IMU	Opis sygnału	Emocja/Stan
Agitacja psychomotoryczna	Wysoka wariancja ACC, nieregularne szarpnięcia > 0.8g	Gniew, lęk, mania
Spowolnienie psychomotoryczne	Niska aktywność ACC < 0.1g std, wolne tempo kroków	Depresja, zmęczenie, smutek
Tremor dłoni (drżenie)	Oscylacje 4–12 Hz w ACC i żyroskopu	Stres, lęk, niedobór snu
Wzorzec chodzenia (gait)	Symetria kroków, kadencja, długość kroku	Nastrój depresyjny
Interakcja z telefonem	Szybkość typowania, pressure, obroty ekranu	Frustracja, pośpiech, odprężenie
Pozycja ciała (posture)	Żyroskop: przedłużone pochylenie do przodu	Skupienie, smutek, zmęczenie

Kotlin

```
// Detekcja agitacji: wariancja ACC w oknie 5s
val variance = accWindow.map{it.magnitude()}.variance()
val agitated = variance > AGITATION_THRESHOLD  // 0.15 m/s²²
```

Kotlin

```
// Wzorzec typowania jako emotion proxy (Android InputMethod)
fun onKeyPress(keyCode: Int, pressTime: Long) {
    val iti = pressTime - lastKeyTime  // inter-tap interval [ms]
    typingSpeedBuffer.add(iti)          // niska ITI → stress/pośpiech
}
```

EEG Wearables: Mose, Emotiv i fale mózgowe



EEG (elektroencefalografia) daje bezpośredni dostęp do aktywności mózgu. Najwyższa precyzja rozpoznawania emocji, kosztem niskiej dostępności w masowym mobile

Pasma częstotliwościowe EEG i emocje

Delta (0.5–4 Hz)	Głęboki sen, brak świadomości — nie używane w emotion recognition
Theta (4–8 Hz)	Senność, medytacja, kreatywność. FRONTAL theta → meditation apps (Muse).
Alpha (8–13 Hz)	Relaks przy zamkniętych oczach. FAA (Frontal Alpha Asymmetry): L>R → approach/pozytywne, R>L → withdrawal/negatywne.
Beta (13–30 Hz)	Aktywne myślenie, koncentracja, stres. Wysoki beta → lęk i stres. Podstawowe pasmo emotion recognition.
Gamma (30–100 Hz)	Skupienie, aktywne przetwarzanie. Wysoki gamma → flow state, efektywna nauka.

Konsumenckie EEG headsets

Muse 2 — 4 kanały
Meditation, Alpha/Theta feedback. iOS/Android SDK. 200 Hz. $299.

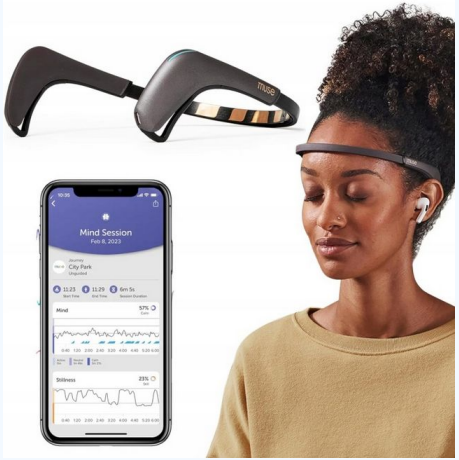
Emotiv Insight — 5 kanałów
Performance tracking, emotion SDK (Affectiv Suite: stress, engagement). $499.

OpenBCI Cyton — 8–16 kan.
Badawczy, open-source. Python/Android API. Najlepsza jakość. $1000+.

InteraXon Muse S — 4+1 ref.
Sleep tracking EEG. Theta/Alpha biofeedback. Aplikacja mobilna.

EEG Wearables: Mose, Emotiv i fale mózgowe





MUSE 2 THE BRAIN SENSING HEADBAND BLACK OPASKA EEG

4,83 (6 ocen)1 osoba kupiła



Firma | poleca 100%

919,99 zł

196,14 zł x 5 rat z PAY sprawdź

Liczba sztuk

1 z 1 sztuki

DODAJ DO KOSZYKA

KUP I ZAPŁAĆ

Allegro Ochrona Kupujących
2 lata na reklamację online i wsparcie w odzyskaniu pieniędzy.

Przewidywana dostawa
pojutrze

Dostawa za 0 zł z aktywnym SMART

Dostawa od 10,49 zł bez Smarta

Fuzja multimodalna: łączenie wielu kanałów afektywnych

Strategie fuzji multimodalnej

Early Fusion (Feature Fusion) — +acc: +5–8%

Konkatenacja wektorów cech wszystkich modalności → jeden klasyfikator. Prosto, ale wymaga synchronizacji czasowej i dostępu do wszystkich modalności jednocześnie.

Late Fusion (Decision Fusion) — +acc: +8–12%

Każda modalność → własny model → probabilistyczne połączenie wyników (weighted average, voting, SVM na scores). Elastyczne, bo działa gdy modalność niedostępna.

Hybrid Fusion — +acc: +12–18%

Early fusion dla modalności o podobnej częstotliwości (face+body), late fusion z fizjologią niskiej częstotliwości. Najlepsza w praktyce.

Attention-based Fusion (Transformer) — +acc: +15–22%

Cross-modal attention: model uczy się które modalności ważniejsze w danym kontekście. MulT (Multimodal Transformer): state-of-the-art.

CMU-MOSI / CMU-MOSEI Datasets

Benchmark multimodalny: 23 485 klipów z audio+wideo+tekst. Adnotacje: sentyment (-3..+3) + 7 emocji Ekmana. Standard oceny modeli fuzji.

Wyzwanie: brakujące modalności

Odporność na brakujące modalności. dropoutowy trening (35% losowe pominięcia podczas treningu). Model musi działać z 1–4 modalnościami.

Żaden pojedynczy kanał nie jest wystarczający. Fuzja twarzy + głosu + tekstu+ fizjologii znacząco poprawia precyzję i odporność na zakłócenia

Synchronizacja temporalna i architektura

Kotlin

```
// Late fusion — weighted decision fusion
data class ModalityResult(
    val emotion: String,
    val confidence: Float,
    val available: Boolean
)
fun fuseEmotions(
    face: ModalityResult, voice: ModalityResult,
    text: ModalityResult, eda: ModalityResult
): EmotionResult {
    val w = mapOf("face"->0.35f,"voice"->0.30f,
                  "text"->0.25f,"eda"->0.10f)
    val avail = listOf(face,voice,text,eda)
        .zip(w.values).filter { (m,_)->m.available }
    val totalW = avail.sumOf { it.second.toDouble() }
    val scores = mutableMapOf<String,Float>()
    avail.forEach { (m,wt) ->
        scores[m.emotion] =

(scores[m.emotion]?:0f)+m.confidence*(wt/totalW).toFloat()
    }
    return EmotionResult(scores.maxBy { it.value }.key)
}
```

Deployment On-Device: TFLite, CoreML i ONNX Runtime

On-device inference to wymóg dla aplikacji afektywnych. Dane emocji użytkownika są wyjątkowo wrażliwe i nie powinny opuszczać urządzenia

Framework	Platforma	Akceleracja	Format	Quantization	Użycie
TensorFlow Lite	Android, iOS	NNAPI, GPU, Hexagon DSP	.tflite	INT8, FP16, dynamic	Szerokie ekosystemy, MediaPipe
CoreML	iOS 11+, macOS	Neural Engine, GPU, CPU	.mlmodel	INT8, FP16, palettized	Natywna iOS, Xcode tools
ONNX Runtime Mobile	Android, iOS	NNAPI, CoreML backend	.onnx	INT8, FP16, QDQ	Cross-platform z PyTorch/HF
MediaPipe	Android, iOS, Web	GPU, Vulkan, Metal	task bundle	INT8	Gotowe solucje (face, pose)
ExecuTorch (Meta)	Android, iOS	XNNPACK, Vulkan, MPS	.pte	INT8, INT4	PyTorch native, Llama mobile

Kwantyzacja modeli afektywnych

```python
# Post-training quantization — TFLite INT8
converter = tf.lite.TFLiteConverter.from_saved_model(model_path)
converter.optimizations = [tf.lite.Optimize.DEFAULT]
# Calibration dataset dla representative inputs
converter.representative_dataset = representative_emotion_data
converter.target_spec.supported_ops =
[tf.lite.OpsSet.TFLITE_BUILTINS_INT8]
tflite_quant = converter.convert()
# FER model: 45MB FP32 → 11.5MB INT8, acc drop < 1.5%
```

```kotlin
// Android — TFLite Interpreter z GPU delegate
val options = Interpreter.Options().apply {
    addDelegate(GpuDelegate())   // lub NnApiDelegate()
    setNumThreads(2)
}
val interpreter = Interpreter(
    FileUtil.loadMappedFile(context, "emotion_model_int8.tflite"),
    options)
// iOS — CoreML z Neural Engine
let config = MLModelConfiguration()
config.computeUnits = .all  // NE + GPU + CPU fallback
let model = try! EmotionClassifier(configuration: config)
let prediction = try! model.prediction(
    melSpectrogram: melInput)
// prediction.emotionLabel, prediction.emotionLabelProbs
```

Affective UX: adaptacja interfejsu do stanu emocjonalnego

Affective UX przekształca rozpoznane emocje w konkretne modyfikacje interfejsu: treść, ton, rytm, kolorystykę i powiadomienia dopasowuje do nastroju użytkownika

Strategie adaptacji i praktyczne przykłady

Stres/lęk
Przykład: Calm, Headspace

Arousal > 0.7, HR > 90 bpm

Zwolnij animacje (duration ×1.5), wyłącz powiadomienia, uruchom breathing exercise, pastelowe kolory

Radość/ekscytacja
Duolingo streak, Fitness rings

Valence > 0.6, Arousal > 0.5

Dodaj confetti efekty, pokaż społeczne udostępnianie, intensywna kolorystyka, szybkie animacje

Smutek/depresja
Przykład: Wysa crisis detection

Valence < −0.4 przez > 10 min

Ciepłe kolory, propozycja rozmowy z przyjacielem, wstrzymaj negatywne newsy, opcja połączenia z terapeutą

Gniew/frustracja
Przykład: DDA w grach mobilnych

Valence < −0.3 + high arousal

Pauza w grze, reset zamiast porażki, haptic feedback uspokajający, propozycja przerwy

Zmęczenie/senność
Samsung wellness, Android auto

Low arousal, blink rate ↑, head drop

Reduce brightness, większe czcionki, mniej czynności na ekranie, remind break/sleep

Skupienie/flow
Przykład: Forest, Focus Mode

Stable low arousal, regular interaction

Minimalistyczny tryb, blokada powiadomień, schowaj niepotrzebne UI, music focus mode

Affective UX: implementacja systemu adaptacji

```kotlin
// Affective State Manager — centralne zarządzanie
class AffectiveStateManager {
    private val stateFlow =
MutableStateFlow(EmotionalState.NEUTRAL)
    // Historia ostatnich 5 minut pomiarów
    private val history =
CircularBuffer<EmotionReading>(300)

    fun updateState(newReading: EmotionReading) {
        history.add(newReading)
        val smoothed =
history.exponentialSmooth(alpha=0.3f)
        val newState = classifyState(smoothed)
        if (newState != stateFlow.value) {
            stateFlow.value = newState
            notifyAdaptationEngine(newState)
        }
    }
    // Stan trwały: nie zmieniaj UI co klatkę
    private fun classifyState(avg: EmotionReading) = when {
        avg.arousal > 0.7f && avg.valence < -0.2f    ->
STRESSED
        avg.valence > 0.5f && avg.arousal > 0.4f     -> JOYFUL
        avg.valence < -0.5f                            -> SAD
        avg.arousal < 0.2f                             -> TIRED
        else                                        -> NEUTRAL
    }
}

// UI Adaptation Engine
class UIAdaptationEngine(val context: Context) {
    fun adapt(state: EmotionalState) {
        when (state) {
            STRESSED -> {
                setAnimationScale(1.5f)             // slower
                NotificationManager.silence(30)     // 30 min
                ColorTheme.apply(PASTEL)
                showBreathingExercise()
            }
            TIRED -> {
                WindowManager.setBrightness(0.4f)
                TextSize.increase(2.sp)
                DND.enable()
            }
            JOYFUL -> ColorTheme.apply(VIBRANT)
            SAD -> showSupportResources()
        }
```

Kotlin

Architektura systemu affective UX: rozpoznawanie → stan emocjonalny → reguły adaptacji → modyfikacja UI — wszystko w < 500ms

Kluczowe zasady affective UX

Smoothing: nie reaguj na każdą klatkę

Temporal smoothing: podejmuj decyzje w oknie 30–60s, a nie natychmiast. Użytkownik denerwuje się gdy UI skacze. Deadband: ±0.1 do VAD zmiany.

Explicability: pokaż dlaczego

Micro-copy: 'Widzę że masz trudny dzień - czy chcesz wyciszyć powiadomienia? Jest transparentny. Ukryty system irytuje. Daj kontrolę.

User Override: zawsze pozwól nadpisać

Przycisk 'To nie mój nastrój' → manual control. Modele afektywne mylą się w 20–30%. Użytkownik musi mieć ostatnie słowo.

Cultural Sensitivity

Emocja radości wyraża się inaczej w Japonii (subdued) vs USA (ekspresyjny). Model trenowany na EMODB (niem.) może źle rozpoznawać np. polskich użytkowników.

Avoid Emotional Manipulation

Adaptujesz UX dla korzyści użytkownika, a nie dla zwiększenia time-in-app. Etyczna zasada: adaptacja musi służyć wellbeing, nie metryce zaangażowania.

Optymalizacja modeli afektywnych: wydajność vs dokładność

Modele afektywne muszą działać w czasie rzeczywistym bez wyczerpania baterii.
Kompromis: accuracy vs latency vs model size vs power consumption

Benchmarki modeli FER na Pixel 7 / iPhone 14

Model	Accuracy FER+	Rozmiar	Latencja (CPU)	Latencja (GPU)	Bateria/h
ResNet-50 FP32	87.3%	97 MB	285 ms	38 ms	~35%
MobileNetV3-L INT8	83.1%	8.2 MB	28 ms	11 ms	~8%
EfficientNet-B0	85.8%	20 MB	52 ms	16 ms	~12%
MediaPipe FER task	81.4%	3.1 MB	12 ms	6 ms	~5%
Custom TinySER	79.2%	1.8 MB	8 ms	4 ms	~3%

Techniki kompresji modeli

Knowledge Distillation

Student (mały) uczy się od Teacher (duży). EfficientFER jako student dla ResNet-50. Zachowuje 97% accuracy przy 10× mniejszym modelu.

Pruning (przycinanie)

Usuwanie wag < threshold. Structured pruning: całe filtry. Unstructured: indywidualne wagi. Sparse model → 60–80% wag zero.

Neural Architecture Search

AutoML: NAS optymalizuje architekturę dla mobile constraints (latency budget). MNasNet, EfficientNet-Lite projektowane pod mobile.

Digital Phenotyping: pasywne monitorowanie zdrowia psychicznego

Digital phenotyping *(fenotypowanie cyfrowe)* to proces gromadzenia i analizy obserwowalnych, mierzalnych cech fizycznych lub behawioralnych organizmu przy użyciu technologii cyfrowych

Cyfrowe biomarkery

Aktywność fizyczna (ACC)
Depresja: < 2000 kroków/dzień. Mania: > 15 000 kroków.
Niepokój: czas nocny agitacja.

Wzorce snu (IMU + light)
Depresja: → > 9h lub < 5h snu, fragmentacja. Mania: > 5-dniowy deficyt snu.
Nieregularność rytmu dobowego.

Wzorce komunikacji (metadane)
Izolacja → < 3 outgoing call/tyg. Mania → > 50 messages/h.
NLP sentymentu wiadomości (za zgodą).

Lokalizacja GPS
Home dwell time > 20h/dzień → depresja, agorafobia.
Nocne zmiany lokalizacji → mania, używanie substancji.

Użycie ekranu i telefonu
Phone pick-ups > 200/dzień → lęk.
App usage patterns: media vs social → isolation biomarker.

Klawiatura (keystroke dynamics)
Typing speed variance → brak równowagi emocjonalnej
Deletion rate wzrost → ruminacja, niepewność.

Badania i wdrożenia

Mindstrong Health (Dror Ben-Zeev, Dartmouth)
BiAffect keyboard study: 500+ uczestników, schizophrenia + bipolar. Typing dynamics przewidują epizody maniakalne z 3-tygodniowym wyprzedzeniem.

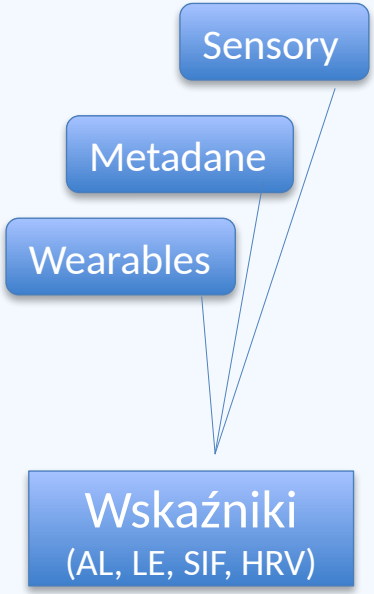
AWARE Framework (Cambridge)
Open-source library do digital phenotyping. Zbiera: accelerometer, location, screen, battery, calls (metadata), notifications. Używane w 200+ badaniach.

CrossCheck (Cornell + Columbia)
Schizophrenia monitoring: GPS + calls + screen + surveys → ML classifier dla relapse detection. 75% sensitivity 2 tygodnie przed nawrotem.

NetHealth (Notre Dame)
5-year longitudinal study 700+ students. Wearable + phone → wellbeing, socialization, depression prediction. Najdłuższe badanie cyfrowej fenotipizacji.

Wyzwanie etyczne: kto ma dostęp? Uczelnia, pracodawca, ubezpieczyciel? Digital phenotyping wymaga rygorystycznych zasad zarządzania oraz wyraźnej zgody użytkownika w odniesieniu do każdego markera.

Digital Phenotyping: aktywność psychomotoryczna



Activity Level (AL)

Wzór:

$$AL = (1/T) * \Sigma ||a_t||$$

Źródła danych:

- akcelerometr (wektor przyspieszenia XYZ)
- częstotliwość próbkowania (np. 10–100 Hz)

Dane wejściowe:

- wartości przyspieszeń w czasie
- liczba próbek w oknie czasowym

Activity Fragmentation (AF)

Wzór:

$$AF = liczba\ przejść\ (aktywny\ nieaktywny)\ /\ czas$$

Źródła danych:

- akcelerometr

Dane wejściowe:

- binarna klasyfikacja aktywności (threshold | |a_t| |)
- liczba zmian stanu

Activity Variability Index (AVI)

Wzór:

$$AVI = Var(||a_t||)$$

Źródła danych:

- akcelerometr

Dane wejściowe:

- rozkład wartości przyspieszeń
- segmentacja czasowa (np. okna 1–5 min)

Inactivity Duration (ID)

Wzór:

$$ID = średnia\ długość\ epizodów\ bezruchu$$

Źródła danych:

akcelerometr

Dane wejściowe:

segmenty poniżej progu aktywności

Digital Phenotyping: rytm dobowy

Circadian Regularity Index (CRI)

Wzór:

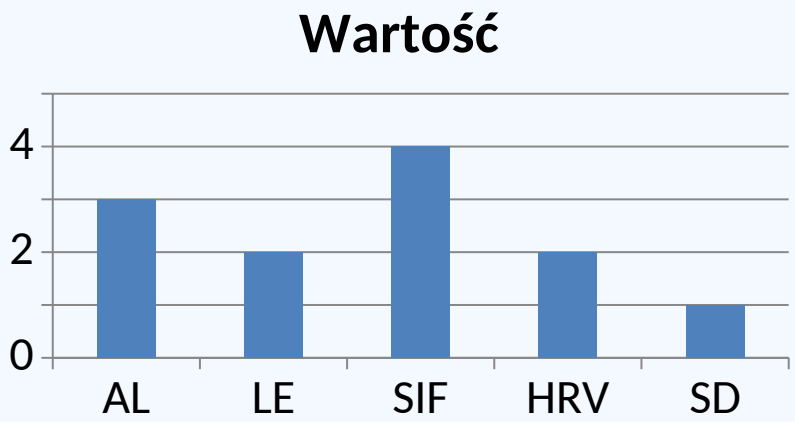
$$CRI = autocorr(activity(t), activity(t - 24h))$$

Źródła danych:

- akcelerometr
- czas systemowy

Dane wejściowe:

- sygnał aktywności w czasie
- znaczniki czasu



Interdaily Stability (IS)

Wzór:

$$IS = wariancja\ międzydniowa\ /\ wariancja\ całkowita$$

Źródła danych:

- akcelerometr

Dane wejściowe:

- profile aktywności dla kolejnych dni

Intradaily Variability (IV)

Wzór:

$$IV = \Sigma(x_t - x_\{t-1\})^2 \ / \ \Sigma(x_t - mean)^2$$

Źródła danych:

- akcelerometr

Dane wejściowe:

- kolejne próbki aktywności

Digital Phenotyping: sen

Sleep Duration (SD)

Wzór: $SD = czas\ między\ onset\ a\ wake$
Źródła danych:
•akcelerometr
•czujnik światła
•logi użycia telefonu
Dane wejściowe:
•okresy bezruchu
•brak aktywności telefonu

Sleep Onset Latency (SOL)

Wzór:
$SOL = czas\ od\ ostatniej\ aktywności\ do\ snu$
Źródła danych:
•akcelerometr
•screen events
Dane wejściowe:
•czas ostatniego użycia telefonu
•początek bezruchu

Wake After Sleep Onset (WASO)

Wzór: $WASO = suma\ wybudzeń$
Źródła danych:
•akcelerometr
•ekran (on/off)
Dane wejściowe:
•epizody aktywności w nocy

Nocturnal Phone Usage (NPU)

Wzór: $NPU = usage_night\ /\ usage_total$
Źródła danych:
•logi systemowe
Dane wejściowe:
•czas aktywności ekranu w godzinach nocnych

Etyka informatyki afektywnej: bias, prawa i manipulacja

"The idea that we can reliably read emotions from faces is not supported by the science" *— Lisa Feldman Barrett, Northeastern University, Science 2019*

Algorytmiczny bias (systemic)

FER modele trenowane na EMODB (German actors) / AffectNet (westerners) → gorsza precyzja dla ciemniejszych karnacji, kobiet starszych, kultur azjatyckich.

Dyskryminacja w rekrutacji, samochodach, systemach edukacyjnych

Zgoda na specjalne przetwarzanie danych

Użytkownicy nie wiedzą że są analizowani emocjonalnie. GDPR wymaga explicit consent dla danych biometrycznych. TikTok FaceFilter, gdzie idą dane?

Naruszenie prywatności, brak autonomii, sprzedaż profili emocjonalnych

Manipulcja emocjonalna

Systemy mogą używać emocji do manipulacji: kiedy jesteś smutny → pokaż płatne treści. Social media feed optymalizowany pod gniew (Facebook internal docs 2021).

Uzależnienie, polaryzacja, eksploatacja stanów emocjonalnych

Podstawy naukowe

Barrett (2019) w Science: wyrazy twarzy nie mapują się jednoznacznie na emocje wewnętrzne. Ekman 6 emocji krytykowane jako oversimplification aktualnie.

Fałszywe decyzje bazowanie na podstawach pseudonaukowych

Power Imbalance

Pracodawca/rząd/platforma zbiera emocje → asymetria wiedzy. User nie ma dostępu do swoich danych emocjonalnych ani do decyzji na ich podstawie.

Surveillance capitalism, social credit, manipulation at scale

RODO, AI Act i regulacje prawne

Dane afektywne = dane biometryczne + dane zdrowotne = najwyższy poziom ochrony RODO.
AI Act dodatkowo klasyfikuje systemy emocjonalne jako High Risk.

Ramy prawne

RODO Art. 9 - Kategorie szczególne

Dane biometryczne do identyfikacji (twarz, tęczówka) = kategoria szczególna.
Przetwarzanie danych wymaga: wyraźnej zgody LUB istotnego interesu publicznego LUB interesu niezbędnego do zachowania życia.

RODO Art. 22 - Automated decisions

Zakaz decyzji opartych wyłącznie na zautomatyzowanym przetwarzaniu (w tym analizie emocji) bez możliwości odwołania i ingerencji człowieka.

AI Act Art. 50 - Emotion recognition ban

Zakaz systemów AI do rozpoznawania emocji w: miejscu pracy, edukacji.
Wyjątki: bezpieczeństwo (DMS w samochodach), medycyna. Obowiązuje od 2026.

AI Act Art. 6 - High Risk AI

Systemy do ewaluacji: rekrutacja, credit scoring, prawo, edukacja = High Risk.
Wymaga: conformity assessment, transparency report, human oversight.

FTC / US - BIPA Illinois

Illinois BIPA: consent przed zbiorem biometrii + $1000/violation.
FTC: unfair/deceptive practices w emotion tech.

Wymogi techniczne

Privacy by Design, czyli on-device

Przetwarzaj dane emocjonalne lokalnie. Nigdy nie wysyłaj twarzy, ani surowych sygnałów do serwera. Wysyłaj TYLKO wynikową emocję jeśli niezbędne.

Data minimization

Zbieraj tylko emocję wynikową (np. 'engaged'), nie surowe wektory cech.
Czas przechowywania: sesja lub max 24h bez re-consent.

Consent UX - granular

Osobna zgoda na każdy kanał: 'Czy mogę analizować wyraz twarzy? [Tak/Nie]'. 'Głos? [Tak/Nie]'. Łatwe przeformuowanie. Nie łącz z ToS.

Explainability (XAI)

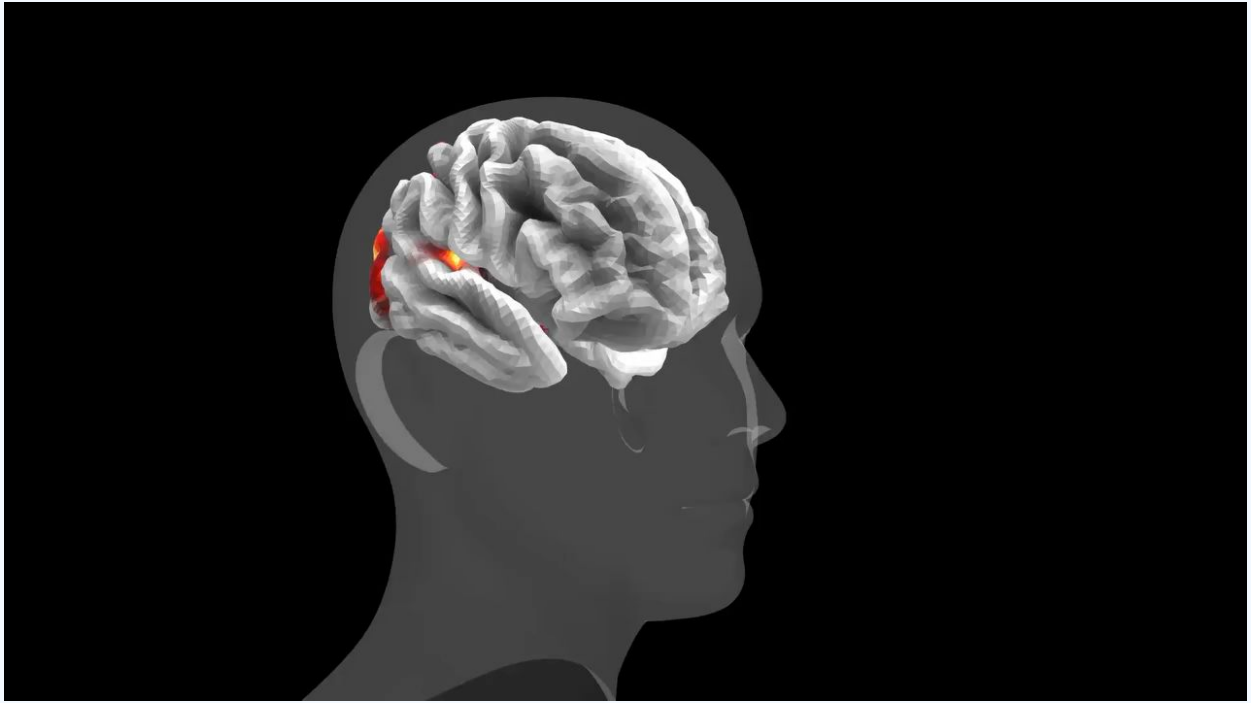
Użytkownik może zapytać: 'Dlaczego dostosowałeś UI?' → wyjaśnienie modelu.
SHAP values lub simpler human-readable explanation.

DPIA (Data Protection Impact)

Przed wdrożeniem systemu emocjonalnego: DPIA obowiązkowe (RODO Art. 35). Ocena ryzyka + środki zaradcze dokumentowane.

TRIBE v2: model predykcyjny reakcji na naturalne bodźce

TRIBE v2 (od Meta) to model predykcyjny przeszkolony w celu zrozumienia, w jaki sposób ludzki mózg przetwarza złożone bodźce



TRIBE v2 to głęboki, multimodalny model kodowania mózgowego, który **przewiduje reakcje mózgu** rejestrowane za pomocą fMRI na **bodźce naturalistyczne** (wideo, audio, tekst). Łączy on najnowocześniejsze ekstraktory cech: LLaMA 3.2 (tekst), V-JEPA2 (wideo) i Wav2Vec-BERT (dźwięk), w ujednoliconą architekturę Transformer, która odwzorowuje reprezentacje multimodalne na powierzchnię kory mózgowej.

https://github.com/facebookresearch/tribev2

Case Study: kompletna implementacja systemu afektywnego

Aplikacja wellness dla studentów. Monitoring nastroju z tekstu
+ HRV ze smartwatcha + IMU pasywny, z adaptowanymi rekomendacjami

Architektura systemu

Data Collection	NLP z dziennika (keyboard) | HRV z Wear OS | IMU pasywny | GPS home dwell
On-Device ML	DistilBERT TFLite 25MB | HRV feature extraction | Activity classifier | Late fusion weighted
Affective State Engine	Smoothed VAD state | Temporal pattern (7-day window) | Crisis detection rule-based
Intervention Logic	CBT micro-intervention library | Breathing exercises | Push notification timing | Escalation to human

Kotlin

```kotlin
// Kompletny flow wellness app
class WellnessEngine(
    private val nlpAnalyzer: NLPEmotionAnalyzer,
    private val hrvTracker: HRVTracker,
    private val activityTracker: ActivityTracker,
    private val interventionLibrary: CBTInterventions
) {
    suspend fun dailyCheck(journalEntry: String):
Intervention {
        val textEmotion = nlpAnalyzer.analyze(journalEntry)
        val hrvScore    = hrvTracker.getLastHour()
        val activity    = activityTracker.getDailySteps()
        // Late fusion
        val moodScore = textEmotion.valence * 0.5f +
                        hrvScore.normalized() * 0.3f +
                        activity.normalized() * 0.2f

        return when {
          moodScore < -0.6f ->
interventionLibrary.crisisProtocol()
           moodScore < -0.3f ->
interventionLibrary.cbtExercise()
           moodScore > 0.5f  ->
interventionLibrary.maintainFlow()
           else              ->
interventionLibrary.minorNudge()
        }
```

Wyniki i wnioski

MAE nastroju (predykcja vs samoocena)	**0.82 (Skala 1–10)**
Czas interwencji od detekcji kryzysu	**< 2 min**
Retencja użytkowników (30 dni)	**71% (vs 34% baseline)**

Kluczowa lekcja: transparentność

Użytkownicy którzy rozumieli jak system działa mieli 2× wyższe zaangażowanie. Wyjaśniaj co mierzysz, po co, co z tym robisz.

Zastosowanie: zdrowie psychiczne, czyli wirtualna terapia i monitoring

Mental health to najbardziej obiecujące zastosowanie informatyki afektywnej,
1 miliard ludzi z zaburzeniami psychicznymi, dostęp do terapii ograniczony

1B+	75%	3× tańszy	68%
osób z zaburzeniami psychicznymi (WHO)	braku dostępu do profesjonalnej pomocy	CBT chatbot vs terapia indywidualna	redukcja objawów (Woebot RCT, 2017)

Rzeczywiste aplikacje i ich architektura afektywna

Woebot (US FDA Authorized)

NLP (tekst), PHQ-9 ankieta

GPT fine-tuned CBT + rule-based crisis detection. Human escalation gdy crisis flag.

Badanie RCT: −28% depresja po 2 tygodniach. 50M+ konwersacji.

Wysa (UK, India)

Tekst NLP + voice + Wysa Score (aktywność)

BERT sentyment, SER z głosu, IMU aktywność → composite mood score. HIPAA.

NHS UK partnership. 5M+ użytkowników. 98% wskazań dostępność bariery.

Mindstrong Health

Smartphone behavioral markers: typing, swipes, app use patterns

Digital phenotyping: NLP+IMU → biomarkers depresji/mani. Dashboard dla lekarza.

Współpraca z CDC. Wykrywa nawrót schizofrenii 3 tygodnie wcześniej.

Calm (meditation)

HRV z Apple Watch + interaction patterns

HRV coherence score → personalizacja sesji medytacji. Tętno po sesji feedback loop.

25M+ użytkowników. Kliniczne wyniki: −27% kortyzol (badanie Stanford).

Zastosowanie: e-learning afektywny, personalizacja nauki

Afektywne systemy e-learningowe wykrywają zaangażowanie, frustację i niepewność, i adaptują te treści w czasie rzeczywistym, poprawiając retencję wiedzy

Stany afektywne istotne w uczeniu się (Graesser 2006)

Zaangażowanie (Engagement)	Uśmiech, forward lean, zwyczajna interakcja, mała wariancja EDA → kontynuuj tempo, nie przeszkadzaj
Niepewność (Confusion)	Zmarszczone brwi (AU04), dłuższy czas na odpowiedź, wsteczne przewijanie → dodaj podpowiedź, powtórz wyjaśnienie
Frustracja (Frustration)	Gniew (AU23+AU04), szybkie dotknięcia ekranu, wiele błędów → upraszcza ćwiczenie, pozytywne wzmocnienie
Znudzenie (Boredom)	Brak mimiki, rzadkie dotknięcia ekranu, odwrócenie wzroku → zmień format (video→quiz), podbij trudność
Flow	Skupienie, regularne odpowiedzi, niska aktywacja EDA → nie przerywaj, kontynuuj tempo, maksymalizuj sesję

Przykłady wdrożeń

Duolingo + face tracking (patent)

Śledzi uśmiech i zmarszczenie podczas ćwiczeń. Personalizuje trudność i dobiera ekstatyczne animacje gdy pozytywny nastrój. Patent US11257382B2.

Case study: Affectiva w kursach online

Harvard edX: emotion-aware MOOC

Analiza facial expressions 80k studentów → identyfikacja najtrudniejszych konceptów przez wzrost konfuzji. Reorganizacja treści podwoiła completion rate.

Zastosowanie: automotive, bezpieczeństwo i DMS

Driver Monitoring Systems (DMS) stały się obowiązkowe w EU (UNECE R79, Euro NCAP 2024). Rozpoznawanie senności i rozproszenia kierowcy ratuje życie

PERCLOS (PERcentage of eyelid CLOSure)

Standard NASA: % czasu gdy powieki zamknięte > 80%. PERCLOS > 0.15 → senność. Kamera podczerwień (NIR) na kierownicy. Działanie w ciemności.

Head Pose i Gaze Estimation

Odchylenie głowy od osi przód > 20° przez 2s → rozproszenie. Gaze off-road > 2s → ostrzeżenie. Modele: FSA-Net, HopeNet dla head pose.

Wykrywanie senności (drowsiness)

MAR (Mouth Aspect Ratio) dla ziewnięcia. PERCLOS + blink rate + head nod → SVM lub LSTM. Alert haptyczny i dźwiękowy. Slow-down suggestion.

Emocje kierowcy i DDA

Gniew (road rage) → adaptacja systemu nawigacji: unikaj korków, sugeruj przerwy. Stres: zmniejsz głośność muzyki, wycisz powiadomienia. BMW affective AI 2024.

```kotlin
// PERCLOS z ML Kit face landmarks
fun computePERCLOS(eyeProbs: List<Float>): Float {
    val closed = eyeProbs.count { it < 0.2f }
    return closed.toFloat() / eyeProbs.size
    // > 0.15 → alert, > 0.25 → emergency stop
}
```

Wdrożenia przemysłowe

Seeing Machines (OEM)

OEM DMS w Volvo, Subaru, GM. NIR 940nm kamera. Real-time 30fps gaze + head + PERCLOS. Certified NCAP 2024.

Smart Eye (aviation→auto)

Aerospace heritage. BMW, Daimler. Multi-camera DMS. Emocje: gniew, stres przez VAD z twarzy. Vehicle speed adaptation.

Tobii DX (eye tracking)

Eye tracking w Subaru Solterra. Gaze heatmapy dla wnętrza. Poziom skupienia na drodze vs na ekranie.

NavInfo (China, Baidu)

5G connected DMS + cloud emotion analytics. Dane flotowe. ADAS integration: pre-brake gdy sen detektowany.

Bosch DMS Gen2 (Android)

Android Automotive OS + Bosch SDK. On-device TFLite INT8. Head + eyes + mouth w jednym modelu 3.2MB.

Zastosowanie: gry mobilne, affective gaming i DDA emocjonalne

Sygnały afektywne dostępne w grach

Kamera selfie
Wyraz twarzy podczas gameplayu: strach, ekscytacja, konfuzja, frustracja. ARKit w iOS.

Mikrofon
Okrzyki, śmiech, przekleństwa → SER w czasie rzeczywistym. Tylko za zgodą użytkownika.

Touchscreen pressure & speed
Gwałtowne tapnięcia → frustracja. Spokojne → zaangażowanie. Android MotionEvent.pressure.

Czas reakcji
Długi czas → trudność/konfuzja. Zbyt krótki → automatyzm/nuda. Optymalne okno: wg flow theory.

IMU telefonu
Trzęsące się ręce podczas napięcia.
Rzucenie telefonem → rage quit (> 5g szarpnięcie).

Resident Evil biofeedback (Capcom)
Prototyp na E3 2019: EDA z wearable → im bardziej przerażony gracz, tym ciemniejsze pomieszczenia i wolniejszy ruch postaci. Strach jako mechanika gry.

Nevermind (2015, PC→Mobile)
Gra horror gdzie strach mierzony HRV i EDA → poziom trudności rosnący gdy zrelaksujesz się. Cel: nauka regulacji emocjonalnej. Research game.

Gry mobilne to naturalne laboratorium dla informatyki afektywnej. Kontroler gry może reagować na strach gracza, jego nudę i ekscytację w czasie rzeczywistym

Emotional DDA (Dynamic Difficulty Adjustment)

Kotlin
```
// Emotional DDA — reaguj na stan emocjonalny gracza
class EmotionalDDA(val game: GameEngine) {
    private var emotionalState = EmotionalState.NEUTRAL
    // Reguły adaptacji Csikszentmihalyi Flow
    fun adapt(state: EmotionalState, level: Level) {
        when (state) {
            BORED -> {
                level.increaseEnemySpeed(1.1f)
                level.addMiniChallenge()
                // Boredom: zbyt łatwe → podnieś trudność
            }
            FRUSTRATED -> {
                level.decreaseEnemyDamage(0.85f)
                level.showHint()
                // Frustration: zbyt trudne → ułatw
            }
            SCARED -> {
                // Strach = zaangażowanie! Nie zmieniaj.
                level.addAmbientSound()  // potęguj tension
            }
            JOYFUL -> {
                // Flow state! Utrzymaj wyzwanie =
umiejętności
                level.matchDifficultyToSkill()
            }
        }
    }
}
```

Zastosowanie: HR i rekrutacja, kontrowersje i etyka

Analiza emocji w rekrutacji to jedno z najbardziej kontrowersyjnych zastosowań.
Potencjalne korzyści vs poważne zagrożenia dyskryminacji i błędów algorytmu

Technologia i firmy

HireVue (NYSE: SHL)

Video interview platform — analiza twarzy, głosu i treści odpowiedzi. 700+ klientów korporacyjnych. AI score → shortlist. Unilever, Goldman Sachs, Delta.

Pymetrics (Harriet AI)

Gry neuronaukowe mierzące poznawcze i emocjonalne reakcje. Analiza: czas reakcji, risk aversion, attention, EDA. Match do job profiles.

Crystal Knows / Humantic AI

NLP z tekstów publicznych (LinkedIn, email) → osobowość DISC. Emotional intelligence score. 'Nie pisz zbyt dużo emocji w CV do tej firmy.'

Cognizin / Uniphore

Call centre analytics — emocje klienta i agenta w czasie rzeczywistym. Supervisor alert gdy frustracja klienta wysoka. Real-time coaching.

AI Act (EU) — zakaz emocji w rekrutacji

AI Act Art. 6: systemy AI do ewaluacji kandydatów = High Risk.
Art. 50: ban emotion recognition w rekrutacji od 2026 w EU.

Krytyka i dokumentowane problemy

Rasowy bias w FER

MIT study (Joy Buolamwini): celność dla ciemniejszych karnacji o 34% niższa. Microsoft, Amazon, IBM FER API — udokumentowane. FaceApp controversy.

Dyskryminacja neuroróżnorodności

Osoby autystyczne nie wyrażają emocji 'normatywnie' — system może je odrzucić. ADHD: zmienna mimika. Parkinson: twarzowe spowolnienie.

Kulturowa różnorodność wyrazu

Uśmiech w Japonii ≠ radość (może oznaczać dyskomfort). Kontakt wzrokowy w kulturach arabskich. Modele trenowane na zachodzie.

Brak validacji naukowej

Science (2019, Barrett et al.): brak naukowych podstaw że wyrazy twarzy = emocje wewnętrzne. HireVue nie opublikował metodologii do 2021.

GDPR i dane biometryczne

Dane twarzy = dane biometryczne (Art. 9 GDPR). Wymagana jest wyraźna zgoda. Przetwarzanie na cel rekrutacji wymaga DPA i DPIA.

Zastosowanie: asystenci głosowi, empatyczne AI i Metaverse

Asystenci głosowi i avatary w metaverse to nowa forpoczta informatyki afektywnej.
System musi rozpoznawać i generować emocje w dialogu

Apple Siri - Emotional Intelligence

iOS 17: Siri analizuje tonację głosu → adaptuje ton odpowiedzi.
Smutek w głosie → bardziej łagodna, empatyczna odpowiedź.
Proaktywne bodźce promujące zdrowy tryb życia.

Google Assistant - Emotions in TTS

System TTS WaveNet / SoundStorm generuje generuje głos o bogatej gamie emocji. „Podekscytowany" ton w przypadku dobrych wiadomości. Tryb „współczujący" po wykryciu smutku.

Amazon Alexa - Emotion Detection

Alexa Voice Service: Frustration Detection — rozpoznaje frustrację z głosu → 'Przepraszam, że to nie zadziałało. Czy chcesz aby spróbowałam inaczej?' O 45% mniejszy wskaźnik rezygnacji.

Inworld AI / Convai (NPC w grach)

LLM + SER + TTS z emocjami = NPC czujące emocje gracza i reagujące empatycznie. Minecraft, Roblox, AAA games. Pionierzy affective NPC.

Empatyczna synteza mowy (Emotional TTS)

Kotlin

```kotlin
// Emotional TTS — ElevenLabs / Google SoundStorm
tts.synthesize("Rozumiem, to musi być trudne",
    style = EmotionalStyle.EMPATHETIC,
    stability = 0.6f,   // ↓ więcej ekspresji
    speakingRate = 0.9f)  // wolniej = empatyczniej
```

Emocjonalny avatar: Affective Metaverse

Kotlin/Swift

```kotlin
// Meta Quest — Face Tracking → Avatar Expression
// Meta Movement SDK (OpenXR extension XR_FB_face_tracking)
val faceTracker = FaceExpressionTracker(session)
faceTracker.subscribe { expressions ->
    // expressions: Map<FaceExpression, Float>
    // > 100 blendshapes na Meta Quest Pro
    avatar.updateBlendShapes(expressions)
    // Też dostępne: eye gaze, tongue, cheeks
}
// Apple Vision Pro
ARView.session.addAnchor(ARBodyAnchor(trackingEnabled: true))
// 52 ARFaceAnchor blendshapes → Reality Kit avatar
```

Oculus/Meta: Empathy Research

Meta Social VR: poczucie połączenia z innymi dzięki expresyjnym avatarom jest 3× wyższe niż bez śledzenia twarzy. Badania CSAIL MIT i Meta Reality Labs.

Digital Humans (Unreal MetaHuman)

Photorealistic avatary z FACS-based animation. Unreal Engine 5 MetaHuman + Apple ARKit → realistic emotional expression. Standard w produkcji filmowej.

Microsoft Mesh: Affective Spaces

Hybrid meetings z emocjonalnymi avataramii. Automatyczna animacja twarzy z kamerki. Zmniejsza zmęczenie video-call poprzez naturalne wyrazy.

Podsumowanie

LLM + Affective Fusion

GPT-4o / Gemini 1.5 rozumieją emocje z tekstu, obrazu i audio jednocześnie. Multimodal LLMs rewolucjonizują SER i FER.

Generatywna empatia

AI generuje empatyczne odpowiedzi dostosowane do emocji, nie tylko rozpoznaje, ale odpowiada emocjonalnie.

Ciągłe pasywne monitorowanie

Smartwatch + ring → 24/7 EDA, HRV, temperatura. Digital phenotyping bez jawnego działania użytkownika.

Neurosymboliczne AI

Połączenie sieci neuronowych z regułami psychologicznymi, umożliwia wnioskowanie o emocjach z kontekstu kulturowego.

EU AI Act → Etyczny design

Zakaz emotion recognition w rekrutacji od 2026. Wymuś re-design systemów afektywnych na Explainable+ Privacy-First.

Spatial computing (Vision Pro)

Meta Quest + Vision Pro → 6DoF + kamera twarzy. Avatar emocjonalny staje się naturalną formą zdalnej komunikacji.

Wnioski

- Emocje to dane wielowymiarowe: używaj modelu VAD lub Plutchika, a nie tylko 6 klas Ekmana
- On-device inference to wymóg, nie opcja: dane afektywne są wrażliwsze od lokalizacji GPS
- Fuzja multimodalna jest kluczowa: żaden pojedynczy sygnał nie osiąga > 80% celności bez kontekstu

Przykładowe pytania zaliczeniowe

Które stwierdzenia najlepiej uzasadniają przewagę ciągłych modeli emocji nad klasyfikacją dyskretną?

A. Nie wymagają etykiet danych
B. Zawsze są dokładniejsze
C. Opisują valence i arousal
D. Lepiej modelują stany pośrednie

Które stwierdzenia o LLM w analizie afektywnej są trafne?

A. A. Nadają się do każdej pętli czasu rzeczywistego
B. Nie potrzebują dodatkowej orkiestracji
C. Dobrze wykorzystują kontekst wypowiedzi
D. Architektura hybrydowa bywa praktyczna

Które stwierdzenia najlepiej opisują cyfrowe fenotypowanie?

A. Zastępuje diagnozę kliniczną
B. Nie rodzi problemów etycznych
C. Opiera się na pasywnym zbieraniu danych
D. Może wspierać wczesne wykrywanie zmian stanu

Które cechy sprawiają, że EDA/GSR jest użytecznym kanałem afektywnym?

A. A. Rozróżnia wszystkie emocje
B. Nie wymaga kalibracji
C. Wiąże się z pobudzeniem
D. Umożliwia analizę tonic i phasic

Które zalety ma late fusion w systemach multimodalnych?

A. Wymaga pełnej synchronizacji ramek
B. Zawsze przewyższa attention Fusion
C. Działa mimo braku części modalności
D. Pozwala łączyć wyniki osobnych modeli.

Które stwierdzenia o wdrożeniach on-device są poprawne?

A. Chmura zawsze lepiej chroni prywatność
B. Akceleracja mobilna nie jest wspierana
C. Przetwarzanie lokalne sprzyja ochronie danych
D. Kwantyzacja zmniejsza rozmiar i opóźnienie

Dlaczego mimika twarzy nie powinna być traktowana jako prosty odczyt emocji?

A. A. Jest całkowicie obiektywna
B. Ma uniwersalne znaczenie
C. Zależy od osoby i kultury
D. Nie zawsze odzwierciedla stan wewnętrzny

Które zasady są kluczowe dla affective UX?

A. A. Interfejs powinien reagować na każdą zmianę emocji
B. Adaptacja powinna być ukryta
C. Użytkownik powinien móc skorygować system
D. System nie powinien reagować zbyt gwałtownie

Które zastosowania są etycznie i regulacyjnie najbardziej problematyczne?

A. Lokalna redukcja powiadomień za zgodą użytkownika
B. Aplikacja wellness działająca offline
C. Analiza emocji w rekrutacji
D. Profilowanie emocjonalne bez zgody

Które decyzje są uzasadnione przy projektowaniu mobilnego systemu SER?

A. A. Należy używać tylko PCM
B. Prozodia i tekst to ten sam kanał
C. Warto używać MFCC lub mel-spektrogramów
D. Warto stosować małe modele mobilne

Które sygnały należą do typowych cyfrowych biomarkerów w psychoinformatyce?

A. Ankiety kliniczne
B. Wyłącznie historia GPS
C. Wzorce snu i rytmu dobowego
D. Metadane komunikacyjne

Które wymagania wynikają z podejścia privacy-first i regulacji dotyczących AI?

A. A. Wystarczy ogólna zgoda w regulaminie
B. Nie trzeba minimalizować danych
C. Należy stosować zgodę granularną (szczegółową)
D. Należy preferować przetwarzanie lokalne