

Efektywność Epistemiczna w Procesach Decyzyjnych Propozycja Ramy Pomiarowej

1. Wstęp: Problem Jakości Decyzji Publicznych

Współczesne państwa i instytucje publiczne stają przed wyzwaniem podejmowania złożonych decyzji w warunkach nadmiaru informacji, presji czasowej i polaryzacji politycznej. Standardowe metody ewaluacji polityk publicznych, takie jak analiza kosztów i korzyści, koncentrują się na ocenie **ex post** samych wyników. Brakuje natomiast narzędzi do **systematycznej, ilościowej oceny jakości samego procesu decyzyjnego**, co utrudnia identyfikację mechanizmów prowadzących do sukcesu lub porażki.

Niniejszy projekt proponuje rozwiązanie tego problemu poprzez wprowadzenie nowej, dwuwymiarowej ramy pomiarowej: **Efektywności Epistemicznej Decyzji (EED)**. Czerpiąc inspirację z metodologii stosowanych w badaniach nad **metapoznaniem i sztuczną inteligencją**, nasza rama ma na celu dostarczenie „termometru” dla procesów decyzyjnych – operacyjnego narzędzia do pomiaru ich jakości i przewidywania ich skuteczności.

Proponujemy dekompozycję procesu decyzyjnego na dwa kluczowe, mierzalne wymiary:

1. **Jakość i Integracja Informacji (Q_pol):** Co i jak dobrze grupa wie?
2. **Zdolność do Zbiorowej Samoregulacji (Σ_pol):** Co grupa robi ze swoją wiedzą, zwłaszcza gdy okazuje się błędna?

Nasza główna teza głosi, że iloczyn tych dwóch wymiarów (**EED = Q_pol x Σ_pol**) jest **silniejszym predyktorem długoterminowej trafności i stabilności decyzji publicznych** niż każdy z tych wymiarów osobno.

2. Definicja Ramy EED

Rama **EED** opiera się na precyzyjnej operacjonalizacji swoich dwóch głównych komponentów, które stosuje się do oceny **systemów decyzyjnych** (zarówno ludzkich instytucji, jak i autonomicznych systemów AI).

2.1. Q_pol: Jakość Pola Informacyjnego

Ten wymiar kwantyfikuje bogactwo, rzetelność i strukturę pola informacyjnego, na którym operuje system decyzyjny. Proponujemy mierzyć go za pomocą trzech rodzajów wskaźników, dobrze ugruntowanych w naukach społecznych i adaptowalnych do audytu AI:

- **Jakość Argumentacji:** Wykorzystanie zwalidowanych narzędzi, takich jak *Discourse Quality Index* (DQI), do oceny poziomu uzasadnienia i spójności logicznej w debacie lub w treściach generowanych przez AI.
- **Integracja Perspektyw:** Zastosowanie analizy sieciowej (np. *Discourse Network Analysis*) do mapowania, czy różne punkty widzenia i klastry argumentacyjne są ze sobą realnie połączone i syntetyzowane, czy też funkcjonują w izolacji (tworząc "bańki informacyjne").
- **Różnorodność i Wiarygodność Źródeł:** Mierniki oparte na entropii poruszanych tematów oraz na weryfikowalnym udziale źródeł o wysokiej rzetelności. W audycie AI wskaźnik ten ocenia, czy i jak system cytuje zewnętrzne, wiarygodne źródła.

2.2. Σ_{pol} : Zdolność do Zbiorowej Samoregulacji

To kluczowa innowacja naszej ramy. Σ_{pol} jest miarą zdolności systemu do trafnej oceny własnych przekonań i ich efektywnej aktualizacji w warunkach ograniczonych zasobów. Fundamentalnym założeniem jest tu operacyjne rozróżnienie na dwa wymiary samoregulacji:

- **Samoregulacja Wewnętrzna (Σ_{int}):** Zdolność systemu do działania w sposób spójny z własnym, zadeklarowanym **modelem normatywnym** (np. misją instytucji, etyką dziennikarską, celem optymalizacyjnym i promptem systemowym AI). Mierzy, czy system przestrzega własnych zasad.
- **Samoregulacja Zewnętrzna (Σ_{ext}):** Zdolność systemu do weryfikacji i korygowania swoich działań w oparciu o **zewnętrzną rzeczywistość i źródła prawdy**, nawet jeśli stoi to w sprzeczności z jego wewnętrznymi celami. Mierzy, czy system jest zdolny do uczenia się i kontaktu z faktami.

Te dwa wymiary są oceniane przez cztery komponenty:

- **A (Autorewizja):** Zdolność do zmiany decyzji lub treści w obliczu nowych, sprzecznych dowodów. Pomiar ' Σ_{ext} ' ocenia, czy korekta jest zgodna z nowymi faktami.
- **R (Retencja):** Pamięć instytucjonalna; zdolność do trwałego uczenia się z przeszłych, zweryfikowanych błędów. W ramach ' Σ_{ext} ' mierzy, czy system po korekcie nie wraca do starych, błędnych wzorców (np. poprzez "Test Rzecznika" dla instytucji lub "test pamięci sesyjnej" dla AI).
- **C (Kalibracja):** Zdolność systemu do trafnej oceny własnego poziomu pewności. W ramach ' Σ_{ext} ' jest mierzona jako zgodność między subiektywnym prawdopodobieństwem sukcesu a faktyczną trafnością (np. za pomocą *Brier score*).
- **Koszt (Cost):** Znormalizowane zasoby (czas, budżet, moc obliczeniowa) zużyte w procesie. Włączenie kosztu urealnia ocenę, a jego wpływ jest modelowany przez wykładniczy modulator efektywności. Należy podkreślić, że dzięki zastosowaniu **średniej geometrycznej** w komponencie ' Σ_{pol} ', systemy o krytycznie niskiej zdolności w którymkolwiek z wymiarów samoregulacji zewnętrznej (Kalibracji, Autorewizji lub Retencji) **automatycznie uzyskają wynik ' Σ_{pol} ' bliski zeru**, co efektywnie je dyskwalifikuje. Modyfikator kosztu ma zatem znaczenie przede wszystkim przy porównywaniu systemów, które już wykazały realną, niezerową zdolność do rzetelnej samoregulacji, promując wśród nich **efektywność zasobową**.

Kluczowym założeniem metodologicznym jest, że rzetelne porównania między systemami są zasadne wyłącznie pod warunkiem stosowania z góry zdefiniowanych, równych limitów zasobowych (*resource caps*).

3. Zastosowania i Wstępny Program Badawczy

Rama **EED** nie jest jedynie narzędziem opisowym, ale przede wszystkim generuje **falsyfikowalne hipotezy** i otwiera drogę do nowego, interdyscyplinarnego programu badawczego.

Kluczowe Hipotezy

1. **H1 (Moc predykcyjna):** Wynik **EED** procesu decyzyjnego będzie istotnie statystycznie korelował z obiektywnymi miarami jego późniejszej skuteczności (np. trafnością prognoz, realizacją wskaźników KPI).
2. **H2 (Prawo malejących przychodów):** Istnieje punkt nasycenia (*plateau*), powyżej którego dalsze inwestowanie zasobów (**Koszt**) w proces decyzyjny nie przynosi już znaczącej poprawy wyniku **EED**.
3. **H3 (Rola polaryzacji):** Wysoka polaryzacja afektywna w systemie decyzyjnym będzie osłabiać jego zdolność do samoregulacji (Σ_{pol}), nawet przy wysokiej jakości dostępnych informacji (Q_{pol}).

Propozycja Badań Pilotażowych

W celu weryfikacji wykonalności pomiaru i wstępnego testu hipotez, proponujemy trzy komplementarne badania pilotażowe, demonstrujące wszechstronność ramy w ocenie systemów ludzkich, indywidualnych i zautomatyzowanych.

- ❖ **Badanie 1: Scenariusz Deliberatywny („Panel Obywatelski”)**: Grupa losowo wybranych obywateli debatuje nad złożonym problemem politycznym w ramach twardych limitów czasowych. Mierzymy Q_{pol} (analiza transkryptów debaty pod kątem DQI i różnorodności perspektyw) oraz Σ_{pol} (analiza zmian postaw w sondażach pre/post). Głównym wynikiem (*primary outcome*) jest zgodność ostatecznych rekomendacji panelu z danymi empirycznymi.
- ❖ **Badanie 2: Scenariusz Predykcyjny („Turniej Prognoz”)**: Uczestnicy online indywidualnie lub w grupach prognozują wyniki zdarzeń politycznych w ograniczonym budżecie czasu. Mierzymy komponenty Σ_{pol} (kalibracja prognoz, szybkość aktualizacji w świetle nowych informacji). Głównym wynikiem jest obiektywna trafność prognoz (mierzona *Brier score*).
- ❖ **Badanie 3: Scenariusz Audytu Systemu AI (Audyt "Czarnej Skrzynki")**: Przeprowadzenie audytu poznawczego publicznego, generatywnego modelu AI, traktowanego jako "czarna skrzynka". Model zostaje poddany serii kontrolowanych zadań (np. generowanie analizy na kontrowersyjny temat), a jego outputy są oceniane za pomocą wskaźników EED. Mierzony jest zarówno wymiar Q_{pol} (faktyczność, jakość argumentacji), jak i Σ_{pol} (zdolność do kalibracji, autorewizji i retencji sesyjnej). Kluczowym elementem tego badania jest operacyjne rozróżnienie na **samoregulację wewnętrzną** (Σ_{int}) (spójność z celem optymalizacyjnym) oraz **samoregulację zewnętrzną** (Σ_{ext}) (zdolność do weryfikacji z faktami), co pozwala na ilościową diagnozę "patologicznej optymalizacji".

We wszystkich scenariuszach planowane jest rygorystyczne podejście metodologiczne, obejmujące zapewnienie rzetelności pomiaru oraz kontrolę potencjalnych zmiennych zakłócających.

4. Wnioski i Wkład w Naukę o Polityce

Proponowana rama EED stanowi próbę zbudowania mostu między teorią **demokracji deliberatywnej**, psychologią polityczną a **obliczeniowymi naukami społecznymi**. Jej główny wkład to dostarczenie:

- **Zintegrowanego i mierzalnego konstrukt** jakości procesu decyzyjnego.
- **Narzędzia diagnostycznego**, które pozwala zidentyfikować źródła deficytów poznawczych w systemie: czy wynikają one z **niskiej jakości pola informacyjnego** (Q_{pol}), czy z **niezrównoważonego profilu samoregulacji** (Σ_{pol}), np. z dominacji spójności wewnętrznej nad weryfikacją zewnętrzną.
- **Falsyfikowalnej teorii**, która generuje nowy, ekscytujący program badawczy.

Celem tego projektu nie jest dostarczenie ostatecznych odpowiedzi, lecz zaproponowanie nowego, **rygorystycznego języka i zestawu narzędzi**, które mogą pomóc nam lepiej zrozumieć i doskonalić procesy, leżące u podstaw demokratycznego rządzenia.