Konstytucja neurotechnologii: instytucjonalne granice ingerencji w pamięć i prywatność mentalną

AUTOR

Fundacja Dobre Państwo

STRESZCZENIE / ABSTRACT

Artykuł analizuje przejście od teoretycznej etyki do funkcjonalnej architektury instytucjonalnej w obszarze neurotechnologii. Autor podkreśla, że ochrona godności i prywatności mentalnej wymaga konkretnych procedur prawnych i technicznych, a nie tylko deklaracji wartości. Przywołując rekomendacje OECD i UNESCO, tekst wskazuje na konieczność stworzenia wielowarstwowego systemu ochrony, obejmującego prawa podstawowe, regulacje danych oraz audytowalność algorytmów. Kluczowym elementem jest wdrożenie zasady privacy by design oraz ochrona tzw. łańcucha inferencyjnego, gdyż neurotechnologia pozwala na wyciąganie wniosków o stanie psychicznym użytkownika, które wykraczają poza pierwotny cel zbierania danych. Bezpieczeństwo instytucjonalne ma zapobiegać nadużyciom w sferze pracy i zdrowia, chroniąc integralność mentalną jednostki.

The article analyzes the transition from theoretical ethics to a functional institutional architecture in the field of neurotechnology. The author emphasizes that protecting dignity and mental privacy requires concrete legal and technical procedures, not just declarations of values. Citing OECD and UNESCO recommendations, the text points to the necessity of creating a multi-layered protection system encompassing fundamental rights, data regulations, and algorithm auditability. A key element is the implementation of the privacy by design principle and the protection of the so-called inferential chain, as neurotechnology allows for conclusions about the user's mental state that go beyond the original purpose of data collection. Institutional security is intended to prevent abuses in the spheres of work and health, protecting the mental integrity of the individual.

Artykuł stawia tezę, że w obliczu gwałtownego rozwoju neurotechnologii – od interfejsów mózg-komputer po systemy closed-loop modyfikujące pamięć – tradycyjne deklaracje etyczne są niewystarczające. Autor argumentuje, że aby zapobiec redukcji ludzkiej tożsamości do optymalizowanego zasobu, konieczne jest zastąpienie ogólnych wartości funkcjonalną architekturą instytucjonalną. Kluczowym wyzwaniem nie jest już sama ochrona surowych danych neuronalnych, lecz zabezpieczenie całego łańcucha inferencyjnego, czyli procesów wnioskowania o stanach mentalnych użytkownika, które mogą prowadzić do głębokiej asymetrii władzy w relacjach pracowniczych czy edukacyjnych. Proponowana "Konstytucja neurotechnologii" zakłada wielowarstwowy system kontroli: od rygorystycznego audytu algorytmów i zasady privacy by design, przez model zgody dynamicznej, aż po demokratyczny nadzór nad granicami ingerencji w podmiotowość. Celem nie jest stworzenie izolowanej fortecy wokół mózgu, lecz budowa ekosystemu prawnego, który zagwarantuje, że człowiek pozostanie podmiotem własnej historii w świecie, gdzie granica między biologią a technologią staje się płynna.

SŁOWA KLUCZOWE

konstytucja neurotechnologii | prywatność mentalna | łańcuch inferencyjny | neurodata | systemy closed-loop | privacy by design | dynamiczna zgoda | integralność mentalna | audytowalność algorytmiczna | praktyczna odwracalność | inferencja danych neuronalnych | neuroetyka instytucjonalna | ograniczenie celu (purpose limitation) | asymetria relacji władzy w neurotechnologii

Od etyki do funkcjonalnej architektury instytucjonalnej

Konstytucja dla technologii umysłu: prawo, rozliczalność i instytucjonalne granice ingerencji w pamięć

Aksjologia neurotechnologii pozostaje bezsilna, jeśli godność, autonomia, prywatność mentalna, sprawiedliwość i prawo do niedoskonałości nie zostaną przełożone na konkretne procedury ograniczające działania poszczególnych aktorów. Sama deklaracja o konieczności ochrony integralności mentalnej nie rozstrzyga kwestii fundamentalnych: czy producent urządzenia może przechowywać surowe dane EEG w chmurze, czy pracodawca ma prawo uzależnić zatrudnienie od

korzystania z monitoringu koncentracji, kto jest uprawniony do aktualizacji algorytmu sterującego implantem closed-loop i jakie logi decyzji systemu powinny być archiwizowane. Pozostają otwarte pytania o moment ponownego wyrażenia zgody przez pacjenta oraz o to, kto ponosi odpowiedzialność za szkodę wynikającą z interakcji sprzętu, oprogramowania, modelu uczenia maszynowego i fizjologii użytkownika.

Neuroetyka osiąga dojrzałość dopiero w momencie przejścia od katalogu wartości do projektowania instytucji. Taki kierunek przyjmuje współczesne międzynarodowe governance neurotechnologii. Rekomendacja OECD z 2019 roku szeroko definiuje neurotechnologię jako urządzenia i procedury umożliwiające dostęp, monitorowanie, ocenę, manipulację lub emulowanie struktury i funkcji systemu nerwowego. Dokument ten wyraźnie wiąże odpowiedzialną innowację z bezpieczeństwem, inkluzywnością, deliberacją społeczną, kompetencjami organów nadzoru, ochroną danych mózgowych oraz antycypowaniem nadużyć. Co istotne, OECD definiuje również **personal brain data** jako dane dotyczące funkcjonowania lub struktury mózgu zidentyfikowanej lub możliwej do zidentyfikowania osoby, które mogą zawierać wyjątkowo wrażliwe informacje o fizjologii, zdrowiu lub stanach mentalnych. Przyjęta w listopadzie 2025 roku rekomendacja UNESCO idzie jeszcze dalej, wskazując wprost na prywatność mentalną, godność, integralność i autonomię, zwracając uwagę na szczególne ryzyko wobec dzieci oraz niebezpieczeństwo zastosowań pracowniczych służących monitorowaniu produktywności czy profilowaniu pracowników.

Nie oznacza to jednak powstania jednego globalnego „kodeksu mózgu". Rekomendacje UNESCO i OECD są instrumentami normatywnymi typu soft-law, a nie jednolitym ponadnarodowym kodeksem karnym lub cywilnym, który bezpośrednio rozstrzygałby każdy przypadek. Ich znaczenie polega na wyznaczeniu zestawu dóbr, które państwa, regulatorzy, instytucje badawcze i przedsiębiorstwa powinny uwzględniać przy tworzeniu prawa oraz standardów technicznych. Powstaje w ten sposób zalążek konstytucyjnej warstwy neurotechnologii, jednak szczegółowe normy muszą być nadal konstruowane w ramach prawa ochrony danych, prawa medycznego, regulacji wyrobów medycznych, prawa pracy, ochrony konsumenta, regulacji AI, prawa badań naukowych oraz klasycznych systemów odpowiedzialności.

Tę architekturę można porównać do obrony wielowarstwowej. Najwyższą warstwę stanowią prawa podstawowe i zasady godności, autonomii, prywatności, niedyskryminacji oraz integralności. Warstwa następna określa status konkretnego zastosowania: czy mamy do czynienia z badaniem naukowym, wyrobem medycznym, usługą konsumencką, narzędziem pracowniczym, systemem edukacyjnym czy urządzeniem wojskowym. Kolejna warstwa reguluje dane: ich pozyskiwanie, przechowywanie, wnioskowanie, udostępnianie i wtórne użycie. Następna dotyczy algorytmu – jego walidacji, aktualizacji i audytowalności. Jeszcze kolejna obejmuje relację człowieka z urządzeniem:

zgodę, kontrolę, możliwość wyłączenia, wycofania się oraz długoterminowe monitorowanie. Dopiero na styku tych poziomów powstaje rzeczywiste bezpieczeństwo instytucjonalne.

Szczególnego znaczenia nabiera tu zasada **privacy by design**. W przypadku neurodanych próba ochrony prywatności dopiero po zebraniu pełnego strumienia danych może okazać się spóźniona. OECD już w pracach nad odpowiedzialną innowacją wskazywała na konieczność projektowania ochrony danych od samego początku tworzenia systemu. W praktyce należałoby preferować minimalizację zakresu rejestracji, przetwarzanie lokalne (jeśli jest technicznie wystarczające), ograniczenie retencji danych, kryptograficzną ochronę transmisji oraz kontrolę dostępu. Kluczowe jest również rozdzielenie surowych sygnałów neuronalnych od profilów i inferencji wytwarzanych na ich podstawie – to rozróżnienie ma fundamentalne znaczenie.

Ochrona łańcucha inferencyjnego zamiast samych danych

W klasycznym podejściu do ochrony danych intuicja użytkownika obejmuje zazwyczaj informacje, które przekazał świadomie: imię, diagnozę, historię zakupów czy lokalizację. Neurotechnologia pozwala jednak generować nowe dane poprzez inferencję. Użytkownik może udostępniać surowy sygnał EEG w celu sterowania interfejsem, lecz algorytm jest w stanie wykrywać w tym sygnale dodatkowe właściwości, niezwiązane z pierwotnym celem. Europejski Inspektor Ochrony Danych podkreśla, że neurodata mogą służyć nie tylko ocenie stanu fizycznego zdrowia, ale również probabilistycznemu wnioskowaniu o procesach rozumowania, podejmowaniu decyzji, pamięci, percepcji, języku czy emocjach. EDPS zauważa ponadto, że dane neuronalne w Unii Europejskiej często mogą kwalifikować się jako szczególne kategorie danych – na przykład dane dotyczące zdrowia lub biometryczne – choć zależy to od charakteru i celu konkretnego przetwarzania. To zastrzeżenie ma kluczowe znaczenie prawne, gdyż RODO nie definiuje osobnej kategorii zwanej "neurodanymi".

Artykuł 9 chroni między innymi dane dotyczące zdrowia, genetyczne oraz biometryczne przetwarzane w celu jednoznacznej identyfikacji osoby. Nie każdy sygnał neuronalny automatycznie trafia więc do tej samej kategorii prawnej. Właśnie dlatego rozwój neurotechnologii wystawia na próbę technologicznie neutralne konstrukcje ochrony danych: regulacje muszą chronić nie tylko to, czym dane są w momencie ich zebrania, ale również to, co można z nich później wywnioskować. Rada Europy zidentyfikowała podobny problem w raporcie dotyczącym neurotechnologii i Convention 108+. Eksperci analizują neural data jako szczególny przypadek przetwarzania, który może wpływać na prywatność i podstawowe wolności, wymagając ochrony

także w obliczu przyszłych, coraz bardziej zaawansowanych możliwości inferencyjnych. W czerwcu 2025 roku Komitet Konwencji 108 kontynuował prace nad wytycznymi dotyczącymi ochrony danych w neuronauce. Można z tego wyprowadzić istotną zasadę: regulacjom nie powinien podlegać wyłącznie plik zawierający sygnał neuronalny, lecz cały **łańcuch inferencyjny**, dzięki któremu z surowego sygnału powstaje konkretna informacja o osobie.

W przypadku pamięci różnica ta jest szczególnie doniosła. Firma nie musi odczytać kompletnego wspomnienia autobiograficznego, aby naruszyć istotne interesy jednostki. Wystarczy, że potrafi z określonym prawdopodobieństwem sklasyfikować reakcję rozpoznania, poziom pobudzenia lub inne cechy związane z wydobywaniem informacji. W przyszłości wzrost precyzji takich systemów może sprawić, że dane zebrane dziś w banalnym celu staną się jutro znacznie bardziej informacyjne. Tradycyjna zgoda na przetwarzanie nie może zatem zakładać, że znaczenie zbioru danych pozostaje niezmienne. Z tego względu instytucjonalna konstytucja neurotechnologii powinna silnie chronić **purpose limitation**, czyli ograniczenie celu.

Konieczność ograniczenia inferencji danych w relacjach zależności

Dane pozyskane w celach terapeutycznych nie powinny automatycznie trafiać do systemów oceny zatrudnienia, reklamy, ubezpieczeń czy bezpieczeństwa. Zgoda na sterowanie protezą nie może być utożsamiana ze zgodą na budowanie komercyjnego profilu emocjonalnego użytkownika. Szczególnie ryzykowne okazuje się łączenie neurodanych z historią wyszukiwania, lokalizacją, danymi zakupowymi, genomiką czy profilem społecznym; w takim układzie pojedynczo słabe źródła inferencji mogą wspólnie tworzyć znacznie dokładniejszy model osoby.

W tym miejscu ochrona danych spotyka się z regulacjami dotyczącymi sztucznej inteligencji. W Unii Europejskiej AI Act od 2 lutego 2025 roku wprowadza zakazy wybranych praktyk uznanych za niedopuszczalne, a od 2 sierpnia 2026 roku Komisja oraz organy państw członkowskich zyskają szersze kompetencje egzekucyjne wobec obowiązujących części aktu. Dokument ten zakazuje m.in. systemów AI służących do inferowania emocji w miejscu pracy i placówkach edukacyjnych, z wyjątkiem zastosowań medycznych lub związanych z bezpieczeństwem. Choć zakaz ten nie został sformułowany wyłącznie z myślą o neurotechnologiach, oddaje on istotną logikę regulacyjną: nawet jeśli określona inferencja jest technicznie możliwa, kontekst relacji władzy może uzasadniać jej ograniczenie. Dla przyszłych urządzeń neuronalnych ma to znaczenie fundamentalne. System rozpoznający stan poznawczy pracownika nie staje się etycznie neutralny tylko dlatego, że

wykorzystuje sygnał mózgowy zamiast kamery. Wręcz przeciwnie – bezpośredniość pomiaru może zwiększać asymetrię pomiędzy jednostką i instytucją.

UNESCO w rekomendacji z 2025 roku wprost ostrzega przed stosowaniem neurotechnologii w miejscu pracy do monitorowania produktywności oraz tworzenia profili pracowników. Autor książki sugeruje, że instytucjonalna zasada powinna zatem brzmieć: im silniejsza ekonomiczna lub hierarchiczna zależność osoby od podmiotu żądającego neurodanych, tym mniej można polegać na samej formalnej zgodzie. Podobna ostrożność jest konieczna w przypadku dzieci. UNESCO rekomenduje szczególną ochronę rozwijającego się mózgu i odradza nieterapeutyczne zastosowania neurotechnologii wobec dzieci i młodzieży. Wynika to zarówno z biologicznej plastyczności, jak i szczególnej sytuacji normatywnej: dziecko nie tylko nie posiada pełnej zdolności do samodzielnej zgody, ale pozostaje w relacji zależności od rodziców, szkoły i instytucji. Zgoda rodzicielska nie może być automatycznie traktowana jako wystarczające upoważnienie do dowolnego enhancementu, gdyż taka decyzja może ukształtować późniejszą osobę, która nie uczestniczyła jeszcze w wyborze własnych wartości. Prawo powinno zatem rozróżniać terapię przywracającą istotną funkcję od eksperymentalnego zwiększania sprawności zdrowego mózgu. Choć granica ta nie zawsze będzie oczywista, ciężar uzasadnienia powinien wyraźnie rosnąć wraz z oddalaniem się od leczenia ciężkiej dysfunkcji.

Szczególna ochrona powinna objąć również marketing konsumencki. Rodzicielskie aspiracje mogą bowiem zostać łatwo wykorzystane przez rynek oferujący urządzenia z obietnicami „lepszej koncentracji", „szybszego uczenia" czy „optymalnej pamięci", mimo że opierają się one na znacznie słabszych podstawach dowodowych niż te wymagane w przypadku terapii medycznej.

Konieczność regulacji funkcjonalnej i audytowalności systemów closed-loop

Prowadzi to do problemu granicy pomiędzy wyrobem medycznym a produktem konsumenckim. Europejskie MDR traktuje jako wyrób medyczny między innymi oprogramowanie, jeżeli producent nadał mu konkretny cel medyczny; oprogramowanie ogólnego zastosowania lub związane jedynie z lifestyle i well-being nie staje się automatycznie wyrobem medycznym. Intencja producenta i deklarowany cel urządzenia stają się zatem regulacyjnie istotne. Powstaje jednak potencjalna asymetria: produkt wykorzystujący fizjologicznie podobny mechanizm może znaleźć się w odmiennym reżimie prawnym, zależnie od sposobu jego pozycjonowania. Neurotechnologia szczególnie wyraźnie ujawnia słabość takiego podziału. Urządzenie konsumenckie może nie deklarować leczenia żadnej choroby, a mimo to rejestrować wrażliwe neurodane, trenować

algorytmy na zachowaniu użytkownika i wpływać na jego stan poznawczy. Ryzyko dla autonomii i prywatności nie znika tylko dlatego, że producent nazywa produkt wellness. Konstytucja neurotechnologii potrzebuje więc regulacji funkcjonalnej, która patrzy nie tylko na marketingową kategorię produktu, ale na rzeczywistą zdolność do inferowania i modulowania stanów neuronalnych. W przypadku urządzeń medycznych obowiązek nie kończy się natomiast w dniu dopuszczenia technologii do rynku. MDR opiera system również na klinicznej ocenie, nadzorze po wprowadzeniu do obrotu, vigilance oraz ciągłej ocenie bezpieczeństwa i działania.

Dla neurotechnologii powinno to zostać rozszerzone o systematyczne monitorowanie skutków psychologicznych i autobiograficznych. Tradycyjna kategoria adverse event może być zbyt wąska, jeżeli implant nie powoduje bólu ani uszkodzenia neurologicznego, lecz użytkownik doświadcza istotnej zmiany poczucia sprawstwa, afektu lub relacji z własnym urządzeniem.

Badania nad closed-loop neurotechnology ujawniają obecnie lukę pomiędzy formalną zgodnością regulacyjną a pogłębioną refleksją etyczną. Przegląd 66 badań klinicznych opublikowany w 2025 roku wykazał, że tylko jedno zawierało dedykowaną ocenę problemów etycznych, podczas gdy w pozostałych zagadnienia takie jak autonomia, prywatność, doświadczenie pacjenta czy ryzyko algorytmiczne występowały przede wszystkim pośrednio i proceduralnie. Autorzy publikacji podkreślają również problem nieprzejrzystości algorytmów, potencjalnych różnic skuteczności między grupami oraz braku jasnych mechanizmów odpowiedzialności. Jest to ważne ostrzeżenie: przejście przez komisję bioetyczną i spełnienie minimalnych wymogów technicznych nie wyczerpuje etyki technologii, która w czasie rzeczywistym obserwuje i modyfikuje stan mózgu. Closed-loop wymaga dlatego szczególnej zasady audytowalności decyzji urządzenia.

Jeżeli system wykrywa określony stan neuronalny, klasyfikuje go jako wymagający interwencji i automatycznie zmienia parametry stymulacji, powinno być możliwe późniejsze ustalenie, jaki sygnał został wykryty, jaka wersja modelu dokonała klasyfikacji, jakie parametry zastosowano i jakie reguły ograniczające były aktywne. Nie oznacza to, że każdy klinicysta musi rozumieć miliony wag modelu uczenia maszynowego.

Audytowalność i dynamiczna zgoda jako fundamenty odpowiedzialności

Oznacza natomiast potrzebę prowadzenia wystarczającego dziennika zdarzeń i wersjonowania systemu, aby móc rekonstruować ciąg decyzji istotnych dla bezpieczeństwa pacjenta. Audytowalność powinna obejmować również aktualizacje. Model po wprowadzeniu na rynek nie

może zmieniać się w sposób uniemożliwiający stwierdzenie, jaka wersja działała u konkretnego pacjenta w danym czasie. W tradycyjnym leku skład chemiczny nie aktualizuje się nocą przez internet. Adaptacyjny system cyfrowy może natomiast otrzymać nowy model, nowe progi lub zmienioną logikę klasyfikacji, a każda taka zmiana może modyfikować bilans korzyści i ryzyka.

Konieczne stają się zatem kontrola wersji, testy regresji, dokumentowanie zakresu zmian oraz procedura przywracania wcześniejszej, bezpiecznej konfiguracji. Stąd wynika zasada projektowania pod kątem praktycznej odwracalności. Sama możliwość wyłączenia elektrody nie oznacza bowiem pełnej odwracalności interwencji. Współczesna literatura neuroetyczna rozróżnia m.in. odwracalność strukturalną, funkcjonalną, psychologiczną i etyczną, ostrzegając jednocześnie, że termin ten jest często używany nieprecyzyjnie.

Implant można usunąć, a mimo to pozostawić bliznę, zmianę zachowania, nowo wyuczone nawyki albo konsekwencje życiowych decyzji podjętych podczas działania urządzenia. Konstytucja neurotechnologii powinna więc wymagać informowania nie tylko o tym, czy urządzenie „da się wyłączyć", ale które skutki mogą, a które nie mogą zostać realnie cofnięte. Praktyczna odwracalność powinna wpływać na projekt sprzętu i oprogramowania. Tam, gdzie jest to zgodne z celem klinicznym, użytkownik powinien mieć jasno określony mechanizm przerwania działania, klinicysta możliwość bezpiecznego przejęcia kontroli, a system procedurę awaryjną na wypadek błędnej klasyfikacji. Im większa autonomiczność urządzenia, tym ważniejsze staje się projektowanie bezpiecznych stanów awaryjnych i limitów, których algorytm nie może przekroczyć samodzielnie.

Równolegle konieczna jest modyfikacja klasycznej koncepcji świadomej zgody. Jednorazowy dokument podpisany przed implantacją słabo odpowiada technologii używanej przez lata, aktualizowanej programowo i stopniowo włączanej w codzienne życie, co potencjalnie zmienia sposób, w jaki osoba doświadcza własnego ciała oraz sprawstwa. Najnowsze analizy implantowalnych BCI proponują model zgody dynamicznej, w którym początkowa zgoda jest okresowo odnawiana wraz z nabywaniem przez użytkownika doświadczenia życia z urządzeniem. W 2026 roku pojawiają się również propozycje standaryzowania procesów informed consent dla implantowanych BCI, właśnie dlatego, że dotychczasowe praktyki pozostają fragmentaryczne.

Dynamiczna zgoda nie oznacza jednak, że pacjent musi codziennie podpisywać formularz. Jej istotą jest diachroniczność: pierwsza decyzja powinna rozpoczynać proces, a nie definitywnie zamykać wszelkie późniejsze pytania. Ponowna zgoda może być szczególnie konieczna po istotnej aktualizacji algorytmu, pojawieniu się nowych danych o ryzyku, zmianie funkcji urządzenia, rozszerzeniu wykorzystania danych albo zauważalnej zmianie doświadczenia siebie przez użytkownika.

Jeszcze bardziej subtelny problem dotyczy prawa do wycofania się. W klasycznym badaniu człowiek może odwołać zgodę na dalszy udział. W przypadku implantu wycofanie się z projektu nie musi jednak oznaczać chęci usunięcia urządzenia, zwłaszcza jeżeli stało się ono niezbędne do komunikacji, poruszania się albo funkcjonowania. Współczesna dyskusja o BCI podkreśla dlatego potrzebę rozdzielenia prawa do wyjścia z badania od zgody na eksplantację. Jest to zasadnicze dla autonomii: prawo do odrzucenia relacji badawczej nie może automatycznie oznaczać utraty funkcji, którą użytkownik uzyskał dzięki urządzeniu. Z tego wynika instytucjonalny obowiązek planowania opieki po zakończeniu badania. Neurotechnologia implantowana w człowieku nie powinna być projektowana tak, jak krótkotrwały prototyp programistyczny, którego wsparcie kończy się razem z grantem.

Należy wcześniej ustalić, kto pokryje koszty serwisu, wymiany baterii, aktualizacji, leczenia powikłań i ewentualnej eksplantacji. Sprawiedliwość badań wymaga, aby osoba, która podjęła ryzyko dla rozwoju technologii, nie została później pozostawiona z nieobsługiwanym implantem tylko dlatego, że projekt naukowy się zakończył lub przedsiębiorstwo zmieniło strategię biznesową. Tutaj pojawia się problem odpowiedzialności rozproszonej. W klasycznej procedurze medycznej stosunkowo łatwo wskazać operatora i producenta produktu. W systemie closed-loop szkodliwe działanie może być jednak wynikiem błędu elektrody, niewłaściwie ustawionego progu, błędu modelu, danych treningowych słabo reprezentujących daną populację, aktualizacji producenta, niewłaściwej konfiguracji klinicznej albo niezwykłej reakcji biologicznej pacjenta.

Nie można rozwiązać kwestii odpowiedzialności prostym stwierdzeniem, że „algorytm podjął decyzję”, ponieważ algorytm nie jest autonomiczną osobą prawną, odpowiedzialną moralnie w tym samym sensie co człowiek czy organizacja. System odpowiedzialności powinien więc opierać się na identyfikowalności łańcucha projektowania i wdrożenia. Producent odpowiada za określone właściwości sprzętu i oprogramowania, podmiot dostarczający model za jego walidację i dokumentację, klinika za dobór pacjenta i użycie zgodne z przeznaczeniem, a regulator za ustanowienie i egzekwowanie odpowiedniego standardu dowodowego. Granice odpowiedzialności mogą być prawnie różne, ale nie powinno powstać pole, w którym złożoność systemu staje się pretekstem do zaniku odpowiedzialności. Audyt algorytmiczny powinien dodatkowo obejmować zróżnicowanie populacyjne.

Bezpieczeństwo systemowe zamiast deklaracji etycznych

Systemy closed-loop neurotechnology opierają się na klasyfikacji sygnałów, co sprawia, że model trenowany na wąskiej grupie może działać nierównomiernie w szerszej populacji. Przegląd z 2025 roku zwraca uwagę właśnie na ryzyko biasu wynikającego z ograniczonych i homogenicznych zbiorów danych oraz na możliwość wystąpienia nierównej skuteczności terapii. Sprawiedliwość neurotechnologiczna wymaga zatem czegoś więcej niż formalnie równego dostępu do urządzenia; wymaga pewności, że działa ono równie bezpiecznie i skutecznie w różnych grupach użytkowników. W systemach ingerujących w pamięć pojawia się dodatkowo wymóg rozdzielenia celu terapeutycznego od produktu danych. Klinika może potrzebować określonych informacji, aby bezpiecznie dostroić terapię, podczas gdy producent może postrzegać ogromny zbiór neurodanych jako zasób służący ulepszaniu modeli, rozwojowi kolejnych produktów lub zastosowaniom komercyjnym.

Cele te nie powinny być stapiane w jednym formularzu zgody. Zgoda niezbędna do leczenia nie może automatycznie oznaczać blankietowej akceptacji wszystkich przyszłych form komercyjnego wykorzystania danych. Co więcej, możliwość anonimizacji neurodanych nie powinna być przedstawiana jako absolutna gwarancja. Dane wysokowymiarowe mogą posiadać cechy umożliwiające ponowną identyfikację, zwłaszcza gdy są łączone z innymi zbiorami. Problem dotyczy więc nie tylko usunięcia nazwiska, lecz kontroli dostępu, środowiska analitycznego i zakazu określonych form łączenia danych. Ochrona neural data musi być traktowana jako problem bezpieczeństwa systemowego.

Z tego wynika również obowiązek cyberbezpieczeństwa. Implant lub urządzenie zamkniętej pętli nie jest wyłącznie aparaturą medyczną, lecz systemem sieciowym, którego bezpieczeństwo informatyczne staje się bezpośrednim składnikiem bezpieczeństwa biologicznego. O ile naruszenie konta w serwisie internetowym prowadzi do wycieku danych, o tyle naruszenie systemu neuromodulacyjnego może teoretycznie wpływać na parametry fizycznej interwencji. Neurotechnologia zaciera więc tradycyjną granicę między cyberbezpieczeństwem a bezpieczeństwem pacjenta. W kontekście pamięci standardy powinny być wyjątkowo wysokie, ponieważ skutkiem nadużycia może być nie tylko ujawnienie informacji, ale i manipulacja procesami zapamiętywania lub wydobywania danych. W urządzeniach zdolnych do stymulacji funkcje odczytu i zapisu powinny być traktowane jako odrębne powierzchnie ryzyka – prawo dostępu do informacji nie może implikować prawa do wykonania interwencji.

Znakomitym przykładem prawnego znaczenia tego rozróżnienia jest dyskusja o neurorights. Zwolennicy nowych uprawnień wskazują, że neurotechnologie tworzą jakościowo nowe zagrożenia dla prywatności mentalnej, ciągłości psychologicznej i poznawczego samostanowienia. Przeciwnicy argumentują natomiast, że mnożenie praw może prowadzić do **rights inflation** i neuroexceptionalism, podczas gdy większość interesów można chronić poprzez rozwinięcie istniejących praw do prywatności, integralności, wolności myśli, cielesności, autonomii i niedyskryminacji. Najnowsza literatura nie rozstrzyga tego sporu jednoznacznie; raczej precyzuje sposoby konstruowania zakresu już istniejącego prawa do integralności mentalnej. Chile pozostaje w tym względzie ważnym eksperymentem prawnym – w 2021 roku wprowadziło do konstytucyjnej ochrony integralności fizycznej i psychicznej szczególną klauzulę dotyczącą działalności mózgowej i informacji z niej pochodzącej. Znaczenie tego przykładu nie polega jednak na dowiedzeniu, że każdy porządek prawny powinien skopiować rozwiązanie chilijskie.

Literatura z 2025 roku pokazuje wręcz, że w Ameryce Łacińskiej powstaje fragmentaryczny krajobraz różniących się koncepcji neurorights, co utrudnia harmonizację. Przykład Chile pokazuje raczej, że zagadnienie przestało być wyłącznie filozoficzną spekulacją i stało się przedmiotem rzeczywistego prawa konstytucyjnego. Najbardziej pragmatyczne stanowisko powinno unikać rozstrzygania sporu samą etykietą. Jeżeli istniejące prawo do prywatności można wiarygodnie interpretować tak, aby chroniło neuronalne inferencje przed nieuprawnionym przetwarzaniem, nie ma potrzeby tworzyć odrębnego prawa wyłącznie ze względów retorycznych. Jeśli jednak pojawia się konkretna luka – na przykład brak jasnej ochrony przed niedobrowolnym neuronalnym profilingiem – trzeba ją usunąć, niezależnie od tego, czy ustawodawca nazwie rozwiązanie neuright, rozszerzoną ochroną danych czy szczególnym prawem osobistym. Dobra regulacja powinna chronić interes, a nie rywalizować o atrakcyjność terminologiczną.

„Konstytucja neurotechnologii" powinna być funkcjonalna. Jej podstawową jednostką nie jest technologia jako marka ani konkretna metoda, lecz typ ingerencji. Inne normy będą potrzebne dla biernego rejestrowania sygnału, inne dla inferowania stanu mentalnego, jeszcze inne dla modulacji, adaptacyjnego systemu closed-loop czy interwencji mogącej trwale zmienić pamięć. Taki model jest bardziej odporny na zmiany technologiczne niż regulacja oparta na katalogu nazw aktualnie istniejących urządzeń. Podobnie należy projektować stopniowanie wymagań: prosty, nieinwazyjny system fitness rejestrujący ograniczony sygnał i niegenerujący wrażliwych inferencji nie powinien podlegać tym samym procedurom co wszczepialny system modyfikujący aktywność hipokampalną.

Regulacja powinna pozostawać proporcjonalna do ryzyka, które należy rozumieć wielowymiarowo: jako ryzyko somatyczne, psychiczne, informacyjne, społeczne, dyskryminacyjne i dotyczące tożsamości. Z tego wynika potrzeba oceny ex ante i ex post. Przed wdrożeniem należy

przeanalizować spodziewane zagrożenia, możliwe nadużycia, populacje podatne oraz alternatywy. Po wdrożeniu konieczna jest obserwacja skutków, których projektant nie przewidział. OECD już w 2019 roku uwzględniło antycypowanie i monitorowanie niezamierzonego użycia oraz nadużyć wśród swoich podstawowych zasad.

Adaptacyjne zarządzanie i wielowarstwowa kontrola instytucjonalna

Neurotechnologia jest obszarem szczególnie podatnym na zastosowanie adaptive governance, gdyż dynamiczny postęp sprawia, że sztywne, jednorazowo uchwalone normy szybko przestają odpowiadać praktyce. Nie może to jednak oznaczać regulacyjnej dowolności. System ten wymaga trwałego rdzenia praw podstawowych, wokół którego ewoluować będą szczegółowe standardy techniczne. O ile dopuszczalny poziom szyfrowania, procedury certyfikacji czy sposób raportowania incydentów mogą ulegać zmianom, o tyle znacznie trudniej powinno być naruszyć zasadę, że człowiek nie może być zmuszany do pozamedycznej ingerencji neuronalnej bez szczególnego uzasadnienia prawnego. Kluczową rolę powinny tu odgrywać niezależne organy interdyscyplinarne. Choć klasyczna komisja bioetyczna jest niezbędna w badaniach, przyszłe systemy neurotechnologiczne łączą w sobie zagadnienia medyczne, programistyczne, prawne, psychologiczne i społeczne. Przegląd closed-loop neurotechnology z 2025 roku proponuje zatem powołanie interdyscyplinarnych grup obejmujących klinicystów, neuronaukowców, prawników, ekspertów AI, etyków oraz przedstawicieli pacjentów.

Włączenie użytkownika nie jest dodatkiem partycypacyjnym; pozwala ono ujawnić rodzaje szkód niewidoczne w standardowym modelu technicznym. Pacjent może bowiem uznać zmianę, którą badacz sklasyfikował jako pożądaną redukcję objawu, za utratę części własnej osobowości. Może z kolei zaakceptować znaczącą zmianę stanu psychicznego tylko dlatego, że odzyskał zdolność działania. Takie dane nie powinny zastępować biomarkerów, lecz zostać włączone do oceny technologii jako patient-reported outcomes i jakościowe dane o doświadczeniu użytkownika. Szczególną uwagę należy przy tym poświęcić konfliktom interesów.

Neurotechnologia jest dziedziną kapitałochłonną, a rozwój implantów i BCI wymaga ścisłej współpracy nauki, kliniki i biznesu. Sama kooperacja nie stanowi problemu – bez niej wiele technologii nigdy nie opuściłoby laboratoriów. Problem pojawia się w momencie, gdy interesy właściciela urządzenia, badacza, klinicysty i pacjenta uznaje się za automatycznie zbieżne. Niezależna ocena bezpieczeństwa oraz przejrzyste ujawnianie relacji finansowych są więc częścią architektury epistemicznej, nie tylko etykiety zawodowej.

Analogiczny problem dotyczy własności algorytmu. Producent ma uzasadniony interes w ochronie tajemnicy przedsiębiorstwa, jednak regulator, klinika i pacjent mają prawo do wiedzy niezbędnej do oceny bezpieczeństwa. Pełna publikacja kodu nie zawsze jest konieczna ani właściwa, lecz ochrona własności intelektualnej nie może uniemożliwiać niezależnego audytu systemu wpływającego bezpośrednio na mózg. Niezbędne są mechanizmy kontrolowanego audytu: dostęp regulatora lub certyfikowanej jednostki do dokumentacji modelu, danych walidacyjnych, historii wersji, wyników testów fairness oraz informacji o incydentach. W zastosowaniach wysokiego ryzyka argument o „tajemnicy handlowej" nie powinien eliminować możliwości zewnętrznej weryfikacji.

Dalszym wyzwaniem jest długoterminowa zależność technologiczna. Implant może funkcjonować przez wiele lat, podczas gdy firma może zbankrutować, zostać przejęta lub zaprzestać wspierania produktu. W przypadku zwykłej aplikacji oznacza to utratę usługi; dla użytkownika urządzenia zintegrowanego z funkcją neurologiczną może to oznaczać utratę części sprawstwa. Konstytucja neurotechnologii potrzebuje więc mechanizmów continuity of care i planów na wypadek zakończenia działalności producenta. Można tu wprowadzić obowiązkowe fundusze lub gwarancje finansujące minimalne wsparcie urządzeń, procedury przekazywania dokumentacji innemu operatorowi oraz regulowane mechanizmy dostępu do krytycznego oprogramowania w razie upadku producenta. Zagadnienie to dowodzi, że prawo neurotechnologii musi myśleć w skali biograficznej człowieka, a nie w skali cyklu życia produktu konsumenckiego.

Szczególnej regulacji wymagają zastosowania wojskowe, policyjne i sądowe – nie dlatego, że każda aktywność instytucji publicznych jest niedopuszczalna, lecz ze względu na wyjątkowo silną asymetrię władzy. Neuromonitoring czy neurostymulacja żołnierza nie mogą być oceniane wyłącznie przez pryzmat poprawy sprawności operacyjnej.

Neurotechnologia stosowana wobec podejrzanego nie może opierać się na założeniu, że „mózg nie kłamie", gdyż stan neuronalny jest reprezentacją osoby, a nie bezpośrednim zapisem obiektywnej historii. Wymiar sprawiedliwości powinien być szczególnie odporny na neurorealism, czyli tendencję do nadawania obrazowi lub pomiarowi neuronalnemu nadmiernego statusu dowodowego tylko dlatego, że pochodzi on „z mózgu". Neural evidence może być niezwykle użyteczne, lecz jego interpretacja zależy od przyjętego modelu, jakości danych i warunków eksperymentalnych. Autor książki słusznie wskazuje problem niewiarygodności świadectwa pamięciowego, jednak stwierdzenie, że eksperymenty optogenetyczne „podważają autentyczność materiału dowodowego w prawodawstwie", należy formułować ostrożniej: badania te wzmacniają wiedzę o konstruktywnym charakterze pamięci, ale nie tworzą automatycznie nowego testu prawdziwości konkretnego zeznania.

Stopniowanie rygorów kontroli i projektowanie odwracalności systemów

Regulacje muszą ponadto chronić naukę przed przedwczesną penalizacją ryzyka. Gdyby każdy eksperyment neuronalny obciążono wymogami właściwymi dla gotowego systemu komercyjnego, innowacja mogłaby zostać zahamowana, zanim pojawiłaby się szansa na zastosowanie terapeutyczne. Architektura prawna powinna zatem odróżniać eksploracyjne badania laboratoryjne i z udziałem ludzi od eksperymentalnej terapii, dopuszczenia do obrotu czy masowego produktu konsumenckiego. Wraz ze zbliżaniem się technologii do społeczeństwa powinien rosnąć ciężar dowodowy oraz liczba mechanizmów kontroli. Takie stopniowanie odpowiada klasycznej logice translacji biomedycznej: dowód skuteczności w modelu zwierzęcym uzasadnia dalsze badania, a nie bezpośredni marketing; mała próba bezpieczeństwa pozwala na kolejną fazę, lecz nie na masowe wdrożenie; z kolei badanie kliniczne w starannie wyselekcjonowanej populacji nie dowodzi jeszcze skuteczności i bezpieczeństwa we wszystkich grupach użytkowników. Neurotechnologia pamięci wymaga szczególnej cierpliwości translacyjnej, ponieważ jej skutki mogą być subtelniejsze i bardziej odroczone niż klasyczne działania niepożądane leków. Obowiązek monitorowania nie powinien zatem kończyć się wraz z osiągnięciem statystycznie istotnej poprawy pamięci; konieczne są obserwacje wpływu na codzienne funkcjonowanie, afekt, sen, zachowania społeczne, poczucie tożsamości, zdolność do pracy oraz satysfakcję użytkownika.

Technologia oddziałująca na pamięć może osiągnąć cel laboratoryjny, jednocześnie generując nieoczekiwany koszt na innym poziomie funkcjonowania. Tu pojawia się znaczenie **reversibility-by-design**, choć pojęcie to należy rozumieć jako postulat projektowy, a nie istniejący uniwersalny standard prawny. Jeśli dwie technologie osiągają podobny efekt, lecz jedna oferuje większą możliwość bezpiecznego wyłączenia, cofnięcia parametrów i etapowego zwiększania dawki interwencji, jej profil aksjologiczny jest co do zasady korzystniejszy. Nie chodzi o to, że każda zmiana musi być odwracalna, lecz o to, że odwracalność zachowuje dodatkową przestrzeń dla późniejszej autonomii człowieka. Zasadę tę można rozszerzyć na **contestability-by-design**. Użytkownik systemu powinien mieć możliwość zakwestionowania klasyfikacji lub działania algorytmu, zwłaszcza gdy interwencja wpływa na jego autonomię. W systemie medycznym może to oznaczać prawo do konsultacji i zmiany parametrów; w systemie administracyjnym lub pracowniczym – prawo do informacji o wykorzystaniu neurodanych oraz możliwość zakwestionowania wynikającej z nich decyzji. Nie każda decyzja musi być jednak w pełni wyjaśnialna na poziomie pojedynczej cechy neuronalnej. Explainability nie powinna stać się rytualnym żądaniem stworzenia prostej narracji tam, gdzie model jest złożony. Ważniejsze jest, aby

system był weryfikowalny: byśmy znali jego przeznaczenie, dane walidacyjne, ograniczenia, zakres błędu i warunki, w których nie powinien być używany.

Transparentność bez zrozumiałości może być pozorna, jednak uproszczona „wyjaśnialność" również bywa myląca. Dlatego obok audytowalności niezbędna jest dokumentacja epistemicznych granic urządzenia. Producent systemu dekodującego określone klasy stanów nie powinien sugerować, że urządzenie odczytuje myśli w zakresie, którego nie wykazano. System klasyfikujący reakcję rozpoznania nie staje się wykrywaczem prawdy, a model poprawiający kodowanie listy słów nie jest jeszcze protezą autobiograficznej tożsamości. Normatywne bezpieczeństwo zaczyna się od semantycznej uczciwości opisu możliwości technologii. Można zatem mówić o obowiązku anti-hype by design w komunikacji klinicznej i marketingowej.

Ograniczenie inferencji i zakaz technologizacji problemów systemowych

Kwestia ta nabiera szczególnego znaczenia w przypadku neurotechnologii, ponieważ sam przedmiot analizy – mózg – posiada ogromną siłę retoryczną. Użytkownik może bowiem przypisywać wynikowi neuronalnemu większy autorytet, niż ten w rzeczywistości zasługuje. Informacje o urządzeniu nie powinny zatem ograniczać się do wyliczania korzyści, lecz jasno określać, czego dana technologia nie potrafi.

Instytucjonalna konstytucja musi objąć również rynek ubezpieczeń, gdzie neurodata lub predykcje z nich wyprowadzane mogą potencjalnie tworzyć nowe kategorie ryzyka. Z perspektywy aktuarialnej wykorzystanie bardziej informacyjnego predyktora może wydawać się racjonalne, jednak z punktu widzenia sprawiedliwości może oznaczać karanie osoby za cechę neuronalną, której nie wybrała i która nie musi doprowadzić do przewidywanego stanu. Należy zatem rozważać nie tylko poufność danych, ale przede wszystkim ograniczenia określonych zastosowań inferencji, nawet jeśli sam dostęp do danych uzyskano zgodnie z prawem.

Analogicznie w sferze zatrudnienia nie wystarczy stwierdzenie, że pracownik „zgodził się" na EEG – asymetria relacji ekonomicznej może znacząco ograniczać realną woluntarność takiej zgody. Unijne zakazy stosowania AI do inferowania emocji w pracy pokazują, że ustawodawca uznaje niektóre konteksty za zbyt podatne na nadużycia, by sama zgoda mogła rozwiązać problem. W przypadku neurotechnologii należałoby podobnie przyjąć silne domniemanie zakazu pozamedycznego monitorowania stanu mózgu pracowników, z wyjątkiem bardzo wąskich sytuacji dotyczących realnego bezpieczeństwa. System ochronny powinien być jeszcze bardziej restrykcyjny wobec prób

wpływania na stan neuronalny pracownika. Istnieje bowiem zasadnicza różnica między ochronnikiem słuchu, który chroni organizm przed uszkodzeniem, a neuromodulacją mającą zwiększyć odporność na zmęczenie, stres czy lęk. Pierwsze narzędzie chroni osobę przed środowiskiem; drugie może dostosowywać osobę do środowiska, które samo w sobie być może wymaga zmiany. Etyka pracy musi zatem chronić przed technologizowaniem problemów organizacyjnych.

Ta sama zasada dotyczy edukacji. Jeśli uczniowie nie potrafią utrzymać uwagi w źle zaprojektowanym otoczeniu, pierwszą odpowiedzią nie powinno być automatyczne zwiększanie ich wydajności neurokognitywnej. Choć neurotechnologia może być cenna w przypadku konkretnych zaburzeń, instytucja nie powinna przesuwać odpowiedzialności za niedostosowany system edukacyjny na biologiczny substrat ucznia. W tym punkcie regulacja neurotechnologii spotyka się z klasyczną teorią sprawiedliwości społecznej.

Nie każda dysfunkcja relacji człowiek-środowisko tkwi „w człowieku". Czasem najbardziej proporcjonalną interwencją jest zmiana samego środowiska. Neuroredukcjonizm staje się instytucjonalnie niebezpieczny w momencie, gdy wszystkie problemy społeczne zaczynają być reprezentowane jako defekty mózgu użytkownika wymagające korekty.

Wielowarstwowa architektura odpowiedzialności zamiast ogólnych deklaracji

Stąd konieczność przyjęcia zasady **least intrusive effective alternative**. Jeśli cel można osiągnąć metodą mniej inwazyjną, a o porównywalnej skuteczności, głębsza ingerencja w mechanizmy neuronalne wymaga dodatkowego uzasadnienia. W terapii może to oznaczać wybór farmakoterapii lub psychoterapii zamiast implantu; w pracy – zastąpienie neuromonitoringu lepszą organizacją czasu i ergonomią; w edukacji zaś modyfikację warunków nauczania zamiast technologicznej kontroli uwagi.

Zasada ta nie czyni psychoterapii czy interwencji środowiskowej moralnie "czystszymi" z definicji. Każda metoda może okazać się nieskuteczna, uciążliwa lub manipulacyjna. Kluczowa jest proporcjonalność: im bliżej procesów konstytuujących pamięć, afekt i sprawstwo znajduje się interwencja, tym silniejsze muszą być argumenty za jej wyborem zamiast rozwiązania mniej ingerującego.

Niezbędny jest tutaj również mechanizm demokratycznej deliberacji. OECD od początku uwzględnia społeczną deliberację w zasadach odpowiedzialnej innowacji neurotechnologicznej. Nie

chodzi o to, by obywatele głosowali nad parametrami stymulacji hipokampa, lecz by uczestniczyli w definiowaniu granic dopuszczalnych zastosowań technologii. Jej skutki wykraczają bowiem poza relację pacjent-lekarz i mogą trwale zmienić normy pracy, edukacji, prywatności oraz odpowiedzialności. Eksperci są niezbędni, by określić możliwości techniczne narzędzi, lecz nie posiadają mandatu do samodzielnego rozstrzygania, czy pracodawcy powinni monitorować mózgi pracowników lub czy dopuszczalne jest neurotechnologiczne ulepszanie dzieci.

Są to pytania ustrojowe i aksjologiczne. Nauka dostarcza warunków brzegowych dla decyzji politycznej, ale jej nie zastępuje. Tym samym właściwa Konstytucja neurotechnologii powinna cechować się pluralizmem epistemicznym z jednoznaczną odpowiedzialnością. Neuronaukowiec wnosi wiedzę o mechanizmie, klinicysta o chorobie i leczeniu, inżynier o systemie, prawnik o prawach, etyk o konflikcie wartości, pacjent o doświadczeniu własnego życia, a społeczeństwo o akceptowalnych granicach instytucjonalnej władzy.

Żaden z tych aktorów nie może zastąpić pozostałych, a jednocześnie odpowiedzialność nie może zostać rozproszona do stopnia, w którym nikt nie odpowiada za całość. Każdy system musi mieć wyraźnie wskazanego podmiotu odpowiedzialnego za bezpieczeństwo, kontrolę danych i nadzór, a także przejrzystą procedurę zgłaszania szkód i ścieżkę odwoławczą dla użytkownika. Technologiczna złożoność nie może produkować normatywnej bezpańskości.

Najważniejszym osiągnięciem regulacyjnym prawdopodobnie nie będzie stworzenie jednego, abstrakcyjnego "prawa do mózgu", lecz budowa spójnego ekosystemu. W takim układzie klasyczne prawa człowieka zostają rozwinięte o nowe formy danych i interwencji: prawo medyczne chroni pacjenta, prawo danych zabezpiecza informacje neuronalne i inferencje, regulacje AI ograniczają niedopuszczalną automatyzację, prawo pracy chroni przed strukturalnym przymusem, a instytucje badawcze stosują rygorystyczne standardy wobec technologii wpływających na tożsamość i sprawstwo. Spór o "neurorights kontra stare prawa" może być zatem fałszywą alternatywą. Nowe konfiguracje problemów mogą wymagać nowych nazw, lecz ich normatywna legitymizacja powinna opierać się na ugruntowanych prawach do prywatności, wolności, integralności, godności i niedyskryminacji. Ważniejsza od nazewnictwa jest realna ochrona konkretnej osoby w konkretnej sytuacji, dlatego kryterium skuteczności prawa powinno być operacyjne.

Czy użytkownik wie, jakie neural data są zbierane? Czy rozumie powstające z nich inferencje? Czy może odmówić zgody bez ponoszenia niedopuszczalnej szkody? Czy może zażądać usunięcia danych zgodnie z prawem lub zakwestionować decyzję algorytmu? Czy system posiada dziennik ingerencji i podlega niezależnemu audytowi? Czy producent gwarantuje utrzymanie urządzenia, by pacjent po zakończeniu badania nie został sam z implantem? Czy aktualizacja algorytmu wymaga ponownej oceny?

To właśnie takie pytania przekształcają godność z deklaracji w instytucję. W przypadku pamięci należy dodać jeszcze jedno: czy osoba wie, która funkcja pamięci ma zostać zmieniona? Jak pokazano wcześniej, kodowanie, konsolidacja, wydobycie, walencja, dostępność i integracja autobiograficzna to odrębne wymiary. Zgoda na ogólną "terapię pamięci" jest zbyt nieprecyzyjna. Pacjent powinien wiedzieć, czy procedura ma zwiększyć zdolność wydobycia, osłabić reakcję afektywną, wspierać konsolidację czy zmniejszyć określoną reaktywność. Ta precyzja musi zostać wpisana w regulacyjną klasyfikację ryzyka. Technologia poprawiająca jakość snu (a tym samym pośrednio pamięć) należy do innej kategorii niż implant zamkniętej pętli reagujący na stan kodowania. Obie wpływają na pamięć, lecz różnią się głębokością oddziaływania na podmiotowość.

W ten sposób powstaje nie pojedyncza reguła, lecz instytucjonalna architektura odpowiedzialności, której rdzeniem jest godność i autonomia. Jej warstwa informacyjna to ochrona neural data i ograniczenie inferencji; techniczna – bezpieczeństwo, audytowalność, cyberbezpieczeństwo, wersjonowanie i praktyczna odwracalność; kliniczna – proporcjonalność, dowody skuteczności, dynamiczna zgoda i długoterminowa opieka; społeczna – sprawiedliwość oraz ochrona dzieci, pracowników i osób zależnych; polityczna zaś – demokratyczna kontrola granic pozamedycznego zastosowania technologii.

Odpowiedzialność instytucjonalna jako gwarant podmiotowości człowieka

Taka konstrukcja nie eliminuje ryzyka, lecz zmienia jego dystrybucję. Zamiast obarczać pojedynczego użytkownika obowiązkiem obrony przed potężną instytucją i technologią, przenosi część odpowiedzialności na projektanta, producenta, klinikę oraz regulatora. Jest to klasyczna funkcja dobrego prawa: korygowanie asymetrii, której sama zgoda jednostki nie jest w stanie zrównoważyć. Ostatecznie problem nie sprowadza się do zbudowania fortecy wokół mózgu.

Człowiek nie potrzebuje bowiem prawa do całkowitej izolacji od zewnętrznych wpływów – byłoby to niemożliwe i sprzeczne z jego społeczną naturą. Potrzebuje natomiast ochrony przed nieuprawnioną, nieprzejrzystą i nieproporcjonalną ingerencją w procesy, dzięki którym interpretuje świat i własną przeszłość. Instytucjonalna konstytucja neurotechnologii musi chronić warunki podmiotowości, a nie metafizyczną czystość mózgu. Pamięć ulega zmianom; człowiek może pragnąć tych zmian, a medycyna mieć ku temu silne przesłanki. Granica zostaje jednak przekroczona w momencie, gdy możliwość modyfikacji staje się cudzym uprawnieniem do projektowania osoby, bez wystarczającej kontroli samego podmiotu.

Wywód ten prowadził nas od Semona i materialnego śladu, przez Hebba, engramy Tonegawa Lab, Xu Liu i Ramireza, optogenetykę, fałszywą pamięć, walencję, ciche engramy, prospekcję, replay, sen, żałobę, tożsamość, władzę, enhancement, closed-loop, AI i uzależnienie, aż po socjologię nauki, godność i prawo. Każdy z tych etapów poszerzał odpowiedź na pytanie „co człowiek może zrobić z pamięcią?", a jednocześnie komplikował kwestię tego, „co dzieje się wtedy z człowiekiem?". Czas dokonać syntezy problemu na poziomie antropologii naukowej. Człowiek po raz pierwszy w historii uzyskuje coraz bardziej selektywny dostęp do biologicznych mechanizmów współtworzących jego przeszłość, wyobraźnię i tożsamość. Nie odnalazł jednak w mózgu prostego centrum „Ja", podobnie jak nie znalazł pojedynczego pliku pamięci. Znalazł procesy dynamiczne, rekonstrukcyjne, relacyjne i podatne na ingerencję.

Ostateczną granicą nie jest czaszka. Granicą jest odpowiedzialność.

Pytanie finałowe nie brzmi już, czy można wejść do mózgu. Nauka dowodzi, że w określonym znaczeniu jest to możliwe i że będziemy potrafili robić to z coraz większą precyzją. Pytanie brzmi: czy rozwój wiedzy o mechanizmach pamięci będzie nadążał za rozwojem wiedzy o tym, jak żyć z władzą nad tymi mechanizmami?

Odpowiedź musi być pozytywna, gdyż technologia nie może otrzymać prawa do samodzielnego definiowania człowieka. Nauka dostarcza wiedzy o możliwościach, technika daje sprawczość, a prawo wyznacza granice instytucjonalne – lecz dopiero aksjologia pozwala określić, czemu ta sprawczość ma służyć. Tak skonstruowana Konstytucja neurotechnologii nie byłaby hamulcem postępu. Byłaby warunkiem tego, by postęp zachował człowieka jako podmiot własnej historii, zamiast przekształcać tę historię w kolejny zasób dostępny do optymalizacji.

> Ostateczną granicą ochrony człowieka nie jest już fizyczna bariera czaszki, lecz zakres naszej odpowiedzialności instytucjonalnej. Pytanie o przyszłość neurotechnologii nie brzmi zatem, czy potrafimy wejść do mózgu, lecz czy nasza mądrość w zarządzaniu władzą nad tymi mechanizmami zdoła nadążyć za precyzją narzędzi, które je otwierają. W przeciwnym razie ryzykujemy, że postęp techniczny zamiast wyzwalać potencjał ludzki, przekształci naszą tożsamość w kolejny zasób podlegający rynkowej optymalizacji.

Inspiracje

[1] 

"How to Change a Memory" Steve Ramirez

Princeton University Press | ISBN: 9780691266688

Steve Ramirez (ur. 1988) jest amerykańskim neurobiologiem i profesorem nadzwyczajnym psychologii i nauk o mózgu na Uniwersytecie Bostońskim. Uzyskał tytuł licencjata na Uniwersytecie Bostońskim, a następnie doktorat z neurobiologii w Massachusetts Institute of Technology w laboratorium Susumu Tonegawy, a następnie odbył stypendium na Uniwersytecie Harvarda. Ramirez jest znany na całym świecie ze swojej pionierskiej pracy w dziedzinie biologii engramów pamięci i optogenetyki. Wraz ze swoim zmarłym partnerem badawczym, Xu Liu, przeprowadził przełomowe eksperymenty, które wykazały, że indywidualne wspomnienia u myszy można identyfikować, reaktywować, modyfikować i sztucznie wszczepiać za pomocą światła. Jego laboratorium bada mechanizmy obwodów neuronalnych związane z formowaniem i odzyskiwaniem pamięci, poszukując nowatorskich interwencji terapeutycznych mających na celu tłumienie traumatycznych wspomnień i reaktywację pozytywnych w leczeniu zaburzeń psychicznych, takich jak depresja i zespół stresu pourazowego.

Steve Ramirez *(born 1988) is an American neuroscientist and associate professor of psychological and brain sciences at Boston University. He earned his bachelor's degree from Boston University and completed his Ph.D. in neuroscience at the Massachusetts Institute of Technology in the laboratory of Susumu Tonegawa, followed by a fellowship at Harvard University. Ramirez is internationally recognized for his pioneering work in memory engram biology and optogenetics. Alongside his late research partner Xu Liu, he conducted landmark experiments demonstrating that individual memories in mice can be identified, reactivated, altered, and artificially implanted using light. His laboratory investigates the neural circuit mechanisms of memory formation and retrieval, seeking novel therapeutic interventions to suppress traumatic memories and reactivate positive ones to treat psychiatric conditions such as depression and post-traumatic stress disorder.*

Boston University

Literatura uzupełniająca

Książki

Książki autora i dzieła przywołane w tekście:



[1] **"Project Total Recall"**

Steve Ramirez

2025 | Robinson Press | ISBN: 9781472141743

Książki autorów przywoływanych i literatura pokrewna:



[2] **"Errors and Biases in Modern Healthcare: Public Health, Medico-legal and Risk Management Aspects"**

Simone Grassi, Davide Ferorelli, Francesco De Micco , Francesco Ausania

2025 | Frontiers Media SA | ISBN: 9782832568408

Artykuły naukowe

Artykuły naukowe autora:

[1] **"Memory Engram Cells Have Come of Age"**

Susumu Tonegawa, Xu Liu, Steve Ramirez, Roger L. Redondo

2015 | Neuron | DOI: 10.1016/j.neuron.2015.08.002

Źródło

[2] **"Optogenetic stimulation of a hippocampal engram activates fear memory recall"**

Xu Liu, Steve Ramirez, Petti T. Pang, Corey B. Puryear, Arvind Govindarajan

2012 | Nature | DOI: 10.1038/nature11028

Źródło

[3] **"Creating a False Memory in the Hippocampus"**

Steve Ramirez, Xu Liu, Pei-Ann Lin, Junghyup Suh, Michele Pignatelli

2013 | Science | DOI: 10.1126/science.1239073

Źródło

[4] **"Bidirectional switch of the valence associated with a hippocampal contextual memory engram"**

Roger L. Redondo, Joshua Kim, Autumn L. Arons, Steve Ramirez, Xu Liu

2014 | Nature | DOI: 10.1038/nature13725

Źródło

[5] "Activating positive memory engrams suppresses depression-like behaviour"

Steve Ramirez, Xu Liu, Christopher J. MacDonald, Anthony Moffa, Joanne Y. Zhou

2015 | Nature | DOI: 10.1038/nature14514

Źródło

[6] "Assessments of dentate gyrus function: discoveries and debates"

Mia Borzello, Steve Ramirez, Alessandro Treves, Inah Lee, Helen E. Scharfman

2023 | Nature reviews. Neuroscience | DOI: 10.1038/s41583-023-00710-z

Źródło

[7] "Inception of a false memory by optogenetic manipulation of a hippocampal memory engram"

Xu Liu, Steve Ramirez, Susumu Tonegawa

2013 | Philosophical Transactions of the Royal Society B Biological Sciences | DOI: 10.1098/rstb.2013.0142

Źródło

[8] "Artificially Enhancing and Suppressing Hippocampus-Mediated Memories"

Briana K. Chen, Nathen J. Murawski, Christine Cincotta, Olivia P. McKissick, Abby Basya Finkelstein, Steve Ramirez

2019 | Current Biology | DOI: 10.1016/j.cub.2019.04.065

Źródło

[9] "Identification and optogenetic manipulation of memory engrams in the hippocampus"

Steve Ramirez, Susumu Tonegawa, Xu Liu

2014 | Frontiers in Behavioral Neuroscience | DOI: 10.3389/fnbeh.2013.00226

Źródło

[10] "An implantable piezoelectric ultrasound stimulator (ImPULS) for deep brain activation"

Jason F. Hou, Md Osman Goni Nayeem, Kian A. Caplan, Evan A. Ruesch, Albit J. Caban, Steve Ramirez

2024 | Nature Communications | DOI: 10.1038/s41467-024-48748-6

Źródło

Treść tworzy Fundacja Dobre Państwo.
Redaguje i publikuje APA ONE, autorski system redakcyjny Fundacji oparty na AI.
APA ONE Publishing System
Fundacja Dobre Państwo © 2026
Article ID: af76cb70-870c-415d-9443-c0f52398e19e