

Michał Rupniewski*

Robotyczni urzędnicy a wolność jako nie-dominacja. Republikańska diagnoza zautomatyzowanego podejmowania decyzji w administracji publicznej

Robotic Officials and Freedom as Non-Domination: A Republican Diagnosis of Automated Decision-Making in Public Administration

Słowa kluczowe: zautomatyzowane podejmowanie decyzji, wolność jako nie-dominacja, administracja publiczna, dominacja algorytmiczna, republikanizm

Key words: Automated Decision Making, freedom as non-domination, public administration, algorithmic domination, republicanism

Abstrakt: Artykuł analizuje zautomatyzowane podejmowanie decyzji (ADM) w administracji publicznej z perspektywy republikańskiej koncepcji wolności jako nie-dominacji. Autor bada, w jaki sposób wykorzystanie algorytmów i systemów sztucznej inteligencji wpływa na relacje władztwa pomiędzy obywatelem a urzędnikiem. Wprowadzając pojęcie dominacji algorytmicznej, wskazuje na ryzyko powstania nowych, strukturalnych form zależności, wynikających z nieprzejrzystości, asymetrii władzy oraz ograniczenia możliwości kontestacji systemów podejmowania decyzji administracyjnych.

Abstract: This article analyzes automated decision-making (ADM) in public administration from the perspective of the republican concept of freedom as non-domination. The author examines how the use of algorithms and artificial intelligence systems influences relations between citizens and public officials. By introducing the concept of algorithmic domination, the author points to the risk of new forms of structural dependence resulting from opacity, power asymmetry, and limited opportunities for contestation of ADM systems.

* ORCID ID: <https://orcid.org/0000-0002-1701-5204>; dr hab. nauk prawnych, profesor uczelni w Uniwersytecie Łódzkim, Wydział Prawa i Administracji Uniwersytetu Łódzkiego. E-mail: mrupniewski@wpia.uni.lodz.pl.

Wprowadzenie

W artykule niniejszym podejmuję refleksję nad zautomatyzowanym podejmowaniem decyzji (ang. Automated Decision Making, dalej jako ADM¹) w administracji publicznej, ze szczególnym uwzględnieniem systemów korzystających ze sztucznej inteligencji (ang. Artificial Intelligence, dalej jako AI). Terminy te zdefiniuję w odpowiednim momencie, w ramach wprowadzenia wystarczy powiedzieć, że w dzisiejszej dobie programy komputerowe, często wykorzystujące modele sztucznej inteligencji, coraz silniej wspomagają lub nawet współkształtują rozstrzygnięcia w sprawach migracyjnych, podatkowych, socjalnych, planistycznych, czy policyjnych². Co prawda są one stosowane wciąż ostrożnie i w ograniczonym stopniu, ale już dziś możemy metaforycznie mówić o „robotycznych urzędnikach”. Są to te wszystkie sytuacje, w których człowiek jest w procesach podejmowania decyzji wspomagany, odciażany, czy nawet, w pewnym sensie lub w pewnym zakresie, zastępowany przez systemy komputerowe. Podstawowym celem artykułu jest diagnoza tego fenomenu z perspektywy republikańskiej koncepcji wolności jako nie-dominacji. Chodzi o zbadanie, czy i w jaki sposób ADM przebudowuje tradycyjne relacje władztwa i co z tego wynika dla wolności jako zasady naczelnej. Jeżeli bowiem między jednostką a władzą publiczną mediuje coś o nieokreślonym statusie ontologicznym i politycznym, sytuacja wymaga głębokiego namysłu, a dotychczasowe kategorie czy zasady mogą się bez nowej interpretacji okazać niewystarczające. Nawet, jeśli systemy ADM są tylko narzędziem, są narzędziem tak doskonałym i skomplikowanym, że naturalnie rodzą się pytania o to, czym w rzeczywistości jest rządzenie za ich pomocą.

Przyjmowana perspektywa normatywna i filozoficzno-polityczna bez wątpienia wpływa na wagę, którą przykładamy do danych problemów, a nawet na sam fakt ich dostrzeżenia, dlatego takie badania warto prowadzić z konkretnej perspektywy, nie zaś „z lotu ptaka”³. W niniejszym artykule tą wyróżnioną perspektywą będzie właśnie republikańskie ujęcie wolności jako nie-dominacji, w najbardziej rozpowszechnionej wersji autorstwa Philippa Pettita. Usta-

¹ W literaturze i dyskursie funkcjonują alternatywne nazwy, takie jak Algorithmic Decision Making lub Algorithmic Governance. W tym artykule traktuję je jako bliskoznaczne, nie ma bowiem potrzeby wyodrębniania innych kategorii. „Automated Decision Making” traktuję jako termin zbiorczy na oznaczenie podobnych praktyk podejmowania decyzji (definicję podam w odpowiednim czasie).

² Por. np. F. Chiusi i in. (red.), *Automating Society Report 2020*, Berlin 2020, <https://automatingsociety.algorithmwatch.org/wp-content/uploads/2020/12/Automating-Society-Report-2020.pdf> (2.06.2026).

³ Próbe ogólnej syntezy zagadnień AI z perspektywy filozofii politycznej oferuje Mark Coeckelbergh w książce *The Political Philosophy of AI: An Introduction*, Cambridge 2022.

lenia uzyskane z tej konkretnej perspektywy nie roszczą sobie wyłączności; będzie to raczej próba naświetlenia problemu z danej strony, a wyniki można i powinno się włączyć do szerszego oglądu ADM i AI w administracji publicznej. Wypracowanie tegoż oglądu jest palącą koniecznością, wobec niezwykle szybkiego rozwoju technologii oraz dezorientacji prawnej i moralnej, którą ów rozwój powoduje. Jak dotąd, refleksja nad ADM w ramach doktryny wolności jako nie-dominacji jest stosunkowo słabo rozwinięta, do powstałej już literatury na ten temat odniosę się we właściwym czasie. Artykuł niniejszy ma zatem za zadanie przyczynić się do rozwinięcia republikańskiej refleksji nad ADM. W szczególności, przeprowadzone analizy pozwolą zdefiniować i ocenić ryzyko powstania strukturalnej dominacji nowego typu, którą można nazwać dominacją algorytmiczną.

Zautomatyzowane podejmowanie decyzji i sztuczna inteligencja w administracji publicznej

Aby uniknąć nieporozumień, należy w pierwszej kolejności wprowadzić definicje czy przynajmniej dookreślić podstawowe pojęcia oraz podać przykłady zjawisk, o których artykuł ten traktuje. Zaczniemy od zautomatyzowanego podejmowania decyzji (ADM). Jak dotąd nie osiągnięto powszechnej zgody co do znaczenia tego pojęcia, lecz mamy dobrą kandydaturę do kanonicznej definicji. Przede wszystkim, należy powiedzieć, że będziemy się tutaj zajmować automatyzacją z użyciem technologii informacyjnych. Tym samym przykłady „automatyzacji analogowej”, która polega na zastępowaniu ludzkiego uznania i osądu poprzez narzucone odgórnie schematy i kryteria metryczne, nie będą przedmiotem namysłu. Warto jednak o nich pamiętać. Najnowsze badania pokazują, jak tego rodzaju procesy analogowe przygotowały grunt pod automatyzację cyfrową (na przykładzie Finlandii)⁴. W tym sensie także na przykład zastąpienie swobodnej oceny esejów maturalnych z języka polskiego sztywnymi i scentralizowanymi arkuszami, ma już w sobie coś z ADM. Z naszego punktu widzenia interesujące są jednak sytuacje, w których pomiędzy stanem faktycznym danej sprawy a finalną decyzją w jakiś sposób pośredniczą (lub całkowicie ją „wydają”) technologie cyfrowego przetwarzania informacji. To tutaj bowiem w sposób najbardziej wyraźny pojawia się ryzyko dominacji algorytmicznej, o której wspomniałem we wstępie, a pojęcie której rozwinę we właściwym czasie.

⁴ A. Heikkinen, M. van Gerven, R. Koulou, *From Discretion to Calculation: How Analog Automation Shaped Digitalization of Finnish Social Assistance*, «Public Administration» 2026 (early view, <https://onlinelibrary.wiley.com/doi/10.1111/padm.70041>).

Gdy chodzi o cyfrowe ADM, definicje istniejące są często zbyt wąskie. Na przykład organizacja Informatics Europe definiuje ADM jako "podejmowanie decyzji poprzez środki czysto technologiczne, bez udziału człowieka"⁵. Takiego rodzaju praktyk nie znajdziemy zbyt wiele. Po pierwsze, wszędzie na świecie systemy ADM funkcjonują jako część złożonych struktur, gdzie decyzja człowieka jest na pewnym etapie nieodzowna. Po drugie, w kontekście europejskim, takie „czyste” ADM jest co do zasady zabronione, m.in. przez art. 22 RODO⁶, choć zakaz ten (zdaniem wielu badaczy) doznaje osłabienia przez art. 14 AI Act⁷. Czasem jednak mamy do czynienia z pełną automatyzacją. Na przykład w Estonii funkcjonuje system informatyczny, w ramach którego niektóre decyzje administracyjne z zakresu prawa podatkowego, takie jak wyznaczanie korekt, zapadają w pierwszej instancji w sposób całkowicie zautomatyzowany. Decyzja wydawana jest bez bezpośredniej interwencji urzędnika, przez komputer, program opiera się na algorytmach przetwarzających dane z deklaracji podatkowych i baz danych⁸. Podobne możliwości stwarza prawo w Niemczech czy Kanadzie⁹. W tym wąskim sensie ADM oznaczałoby sytuacje, w których zarówno zbieranie danych, jak i pełny cykl ich przetwarzania aż do momentu wydania decyzji, opiera się w pełni na działaniu programu komputerowego. Jednakże, sytuacje pełnej automatyzacji ograniczają się obecnie do spraw prostych, niewymagających wielokryterialnej oceny sytuacji

⁵ J. Larus i in., *When Computers Decide: European Recommendations on Machine-Learned Automated Decision Making*, Informatics Europe & EUACM 2018, s. 1. W przypadku cytowania tekstów obcojęzycznych bez wskazania tłumacza każdorazowo tłumaczenie autora tekstu.

⁶ Nie jest to jednak zakaz absolutny. Por. F. Geburczyk, *Kryteria dopuszczalności decyzji zautomatyzowanych w świetle brzmienia art. 22 ust. 1 RODO*, «Business Law Journal» 2022, vol. 75, nr 5, s. 34–41.

⁷ Por. np. C. Sarra, *Artificial Intelligence in Decision-making: A Test of Consistency between the "EU AI Act" and the "General Data Protection Regulation"*, «Athens Journal of Law» 2024, vol. 11, nr 1, s. 45–62. Piszac o „osłabieniu” mam na myśli fakt, iż artykuł 22 RODO wprost zabrania decyzji opartych wyłącznie na zautomatyzowanym przetwarzaniu, natomiast art. 14 AI Act, ograniczając się do systemów wysokiego ryzyka, wprowadza jedynie wymóg nadzoru człowieka, który ma być „proporcjonalny”. Wiele tekstów, tak jak przywołany przed chwilą, zwraca uwagę na to, iż art. 14 AI Act pozwala na dalej idącą automatyzację.

⁸ Por. Art. 46(1) i 46(2) Estońskiego Taxation Act (<https://www.riigiteataja.ee/en/eli/508072020001/consolide>). Szersze umówienie patrz I. Pilving, *Guidance-based Algorithms for Automated Decision-Making in Public Administration: the Estonian Perspective*, «Rivista Interdisciplinare Sul Diritto Delle Amministrazioni Pubbliche» 2023, vol. 1, s. 51–74. Jak podkreśla autor, należy pamiętać, że dyskutowane dopuszczenie automatycznych decyzji jest w rzeczywistości bardzo ostrożne, dotyczy spraw nieskomplikowanych i nie deleguje uznania administracyjnego na systemy komputerowe.

⁹ E. Szewczyk, *Algorytmizacja jurysdykcyjnego postępowania administracyjnego i jej skutki społeczne*, Warszawa 2025.

i pozbawionych elementu uznania administracyjnego. Jak już jednak wspomniano, jeżeli przyjęlibyśmy dyskutowaną, wąską definicję, nie obejmowałaby ona „współpracy” człowieka z systemami rządzonymi przez algorytmy, a takich sytuacji jest wiele i są one równie, a może nawet szczególnie, warte namysłu z perspektywy wolności jako nie-dominacji.

Jako przykład definicji uwzględniającej współpracę człowieka z programami komputerowymi można podać propozycję Anny Kathariny Boos, wedle której ADM to

praktyka stosowania algorytmów do rozwiązywania problemów decyzyjnych, gdzie owe algorytmy mogą pełnić rolę sugerowania, odciążania lub zastępowania ludzi, przy czym decyzje są definiowane jako rozstrzygnięcia skutkujące działaniem.¹⁰

Gdy chodzi o ostatni wskazany element, autorka ma na myśli fakt, że interesujące nas „decyzje” i problemy decyzyjne w ADM mają charakter praktyczny, to znaczy dotyczą rozstrzygnięcia co do konkretnego działania, nie tylko co do faktów – chodzi zatem o decyzje praktyczne. Definicja Boos pozwala na pominięcie trudnego problemu ontologicznego – czy i w jakim sensie program komputerowy może rzeczywiście (w sensie ścisłym) podjąć decyzję praktyczną. Zdrowy rozsądek podpowiada, że potrzebne są do tego procesy umysłowe (jakiś rodzaj wewnętrznej deliberacji)¹¹, a natura tych ostatnich jest przedmiotem ożywionych sporów filozoficznych i nie wiadomo, czy działanie umysłu człowieka (autentyczne rozumowanie) może być w ogóle skutecznie emulowane przez komputer¹². Podejście naszej autorki pozwala uniknąć tego problemu (bez udzielania odpowiedzi na pytanie ontologiczne). Wystarczy bowiem traktować ADM jako instytucjonalny i organizacyjny sposób rozwiązywania problemów decyzyjnych. A zatem, na przykład, gdy program komputerowy sugeruje jakieś rozwiązanie, a człowiek podąża lub nie za daną rekomendacją, sprawdzając tylko poprawność działania algorytmu i dane wejściowe, możemy mówić o problemie decyzyjnym rozwiązany przez użycie programu komputerowego – niezależnie od tego, czy jakaś decyzja w tym momencie

¹⁰ A.K. Boos, *Conceptualizing Automated Decision-Making in Organizational Contexts*, «Philosophy & Technology» 2024, vol. 37, nr 3, s. 2.

¹¹ Przykładowo, stanowisko sądu australijskiego z 1999 roku zakładało, że aby mówić o decyzji, konieczny jest jakiś proces umysłowy (w domyśle – aktywność umysłu człowieka), a przez to wynik działania systemu zautomatyzowanego nie kwalifikuje się jako decyzja. Por. *Semunigus v. Minister for Immigration and Multicultural Affairs* [1999] FCA, s. 422.

¹² Najnowszy stan wiedzy na ten temat prezentuje internetowy artykuł T. Wei i in., *Agentic Reasoning for Large Language Models*, 2026, <https://arxiv.org/pdf/2601.12538> (2.06.2026). Tekst ten ukazuje zarówno dalece zaawansowane możliwości obecnych dużych modeli językowych, wskazując jednocześnie, iż nie można jeszcze mówić o rzeczywistych „autonomicznych podmiotach” kreowanych czy konstytuowanych przez AI.

rzeczywiście zapada (w sensie ontologicznym), czy też nie. Autorka oferuje typologię ról pełnionych przez algorytmy, nie musimy jednak się w nie zagłębiać. Wystarczy powiedzieć, iż warunkiem koniecznym ADM, zdaniem Boos, jest to, by zinstytucjonalizowany proces podejmowania decyzji był powierzony algorytmowi w stopniu przekraczającym przygotowanie informacji lub prognoz. Musimy mieć do czynienia przynajmniej z jakąś rekomendacją działania (na przykład – „przyznać” lub „nie przyznawać” danego świadczenia socjalnego)¹³. Możliwe są także sytuacje, gdzie program wydaje decyzję, a człowiek stale monitoruje działanie systemu i może zmienić decyzję algorytmu (to autorka nazywa „odciążaniem”), lub też sytuacje pełnej automatyzacji, czyli zupełnego zastępowania człowieka. Powyższą definicję możemy na potrzeby tego artykułu przyjąć bez modyfikacji i każdy z tych scenariuszy możemy określić mianem tytułowych „robotycznych urzędników”. Konkretnie przykłady tak rozumianej ADM pojawiają się we właściwym czasie.

Systemy ADM mogą funkcjonować bez AI, lecz warto pochylić się także nad tym terminem. Został on zaproponowany przez Johna McCarthy’ego wraz z kolegami w roku 1955, przy okazji poszukiwania funduszy na (odbywające się w roku następnym) seminarium letnie w Dartmouth College¹⁴. Termin ten stosowano następnie na oznaczenie całkowicie odmiennych od siebie podejść i rozwiązań technicznych, czasem go też dyskredytowano, jednakże zawsze chodziło w nim o poszukiwanie sposobów na odwzorowanie inteligencji człowieka przez działanie maszyny cyfrowej¹⁵. Z punktu widzenia tego artykułu możemy pozostawić jednak kwestie konkretnych technologii na boku rozważań, a to dzięki unijnemu AI Act, który wprowadza definicję legalną, koncentrującą się na aspekcie funkcjonalnym i społecznym. Artykuł 3 definiuje termin „system AI”, jako

system maszynowy, który został zaprojektowany do działania z różnym poziomem autonomii po jego wdrożeniu oraz który może wykazywać zdolność adaptacji po jego wdrożeniu, a także który – na potrzeby wyraźnych lub dorozumianych celów – wnioskuje, jak generować na podstawie otrzymanych danych wejściowych wyniki, takie jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne lub wirtualne.

¹³ Z takimi sytuacjami mam do czynienia także w Polsce. Szeroko komentowanym i krytykowanym przykładem jest system profilowania bezrobotnych funkcjonujący w latach 2014–2019. Por. np. K. Sztandar-Sztanderska i in., *Czy algorytmy wprowadzają w błąd? Metaanaliza algorytmu profilowania bezrobotnych stosowanego w Polsce*, «Studia Socjologiczne» 2021, vol. 250, nr 1, s. 89–115.

¹⁴ Potwierdza to choćby notka na stronie tej instytucji edukacyjnej: <https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth>.

¹⁵ Przestępne wprowadzenie oferuje Przemysław Pałka w tekście *Sztuczna inteligencja*, [w:] B. Szlachta, W. Pasierbek (red.), *Słowniki Społeczne*, vol. 18, Kraków 2025, s. 351–369.

Dla zrozumienia powyższej definicji najważniejsze jest pojęcie „autonomii”. Programy komputerowe nie są (przynajmniej jak dotąd) autonomiczne w mocnym sensie, o którym mówimy w przypadku osoby ludzkiej jako podmiotu rozumnego i moralnego¹⁶, czy też posiadającego swój własny plan życia, który ma prawo realizować¹⁷. „Autonomia”, o którą chodzi w przypadku AI, to po prostu zdolność programów komputerowych do generowania na wyjściu czegoś, co nie zostało dokładnie przewidziane i zaplanowane przez programistę. Chociaż wciąż mamy do czynienia z realizacją przez program celu zadanego z góry, w procesie programowania, to różnorodne technologie, a w szczególności różne odmiany uczenia maszynowego¹⁸, pozwalają na to, by programy AI generowały swoje rezultaty coraz bardziej „samodzielnie”, w oparciu o korelacje statystyczne wykrywane przez program w dużych zbiorach danych. Programy z AI działają zatem w ten sposób, że użytkownik deleguje pewne zadania wymagające zdolności kognitywnych bezpośrednio na program i zakłada, że samodzielnie dobierze on odpowiednie (najlepsze, optymalne) środki do realizacji tych zadań¹⁹. Przejście od ustaleń faktów (interpretację czy analizę danych) bezpośrednio do decyzji jest już kwestią konwencji i konkretnych rozwiązań organizacyjnych i prawnych. W kontekście „robotycznych urzędników” może chodzić o sformułowanie rekomendacji rozwiązania problemu, generowanie pism z gotowymi propozycjami rozstrzygnięć, wyszukiwanie informacji kluczowych dla danej decyzji. AI nie jest konstytutywna dla ADM, niemniej jednak to właśnie ona sprawia, że w ramach systemów ADM komputery są w stanie wykonywać coraz więcej zadań, czyniąc to w sposób autonomiczny w wyłuszczonej wyżej sensie.

Wolność jako nie-dominacja

Po charakterystyce fenomenu ADM możemy przejść do próby jego filozoficzno-politycznej oceny. Kluczowym dla tego zadania pojęciem jest „wolność jako nie-dominacja”. Termin ów jest dziś przede wszystkim kojarzony z Philipem Pettitem²⁰, choć autor ten wprost nawiązuje do rzymskiej tradycji repu-

¹⁶ Por. K. Korsgaard, *The Sources of Normativity*, Cambridge 2010. Jest to najbardziej rozpowszechniona współczesna wersja etycznego kantyzmu, opierającego moralność na autonomicznym podmiocie osobowym czy ludzkim.

¹⁷ G. Dworkin, *The Theory and Practice of Autonomy*, Cambridge 2012.

¹⁸ Por. S. Russel, P. Norwig, *Artificial Intelligence: A Modern Approach*, Londyn 2021, s. 651–823.

¹⁹ Por. G. Sartor, *Cognitive Automata and the Law: Electronic Contracting and the Intentionality of Software Agents*, «Artificial Intelligence and Law» 2009, vol. 17, s. 253–290.

²⁰ Najważniejsze pozycje dla tej teorii to P. Pettit, *Republicanism: A Theory of Freedom and Government*, Oxford 1997 oraz P. Pettit, *On The People's Terms: A Republican Theory and Model of Democracy*, Cambridge 2012.

blikanizmu. Najbardziej znanym obrońcą rzymskiego ustroju republikańskiego jest z pewnością Ciceron, ważne źródło inspiracji Pettita²¹, dlatego spójrzmy najpierw w tę stronę. Istotą ustroju republiki rzymskiej jest dla Cicerona oczywiście jego mieszany charakter – element demokratyczny, arystokratyczny i monarchiczny tworzą w nim harmonijną całość. Co bardzo istotne, u Cicerona racją, dla której to właśnie ustrój mieszany jest optymalny, jest wolność (*libertas*). Dobre państwo to państwo wolne, przy czym wolność rozpatrywana jest na niwie politycznej, i „nie polega na tym, by służyć dobremu panu, lecz by nie mieć nad sobą żadnego”²².

Warto zwrócić uwagę na dwa istotne szczegóły przywołanej wyżej myśli Cicerona. Po pierwsze, pojawia się ona w kontekście wyższości republiki nad monarchią. Rzymski jurysta wskazuje, iż poddanym monarchy brakuje wolności, która – powtórzmy cytowane już zdanie w oryginalnym brzmieniu – „*non in eo est ut iusto utamur domino, sed ut nullo*”²³. W tekście występuje zatem słowo *dominus*, które w pogańskim Rzymie oznacza przede wszystkim „pana domu”, czy nawet „właściciela”²⁴. Pan posiada władzę szczególnego typu, którą egzemplifikuje prawna relacja pomiędzy nim a będącymi jego własnością niewolnikami. Jest on przede wszystkim władny wydawać niewolnikom polecenia wedle własnego osądu. I to właśnie w tym sensie wolność polega na tym, by nie mieć nad sobą żadnego pana. Chodzi o to, by nikt, w ramach zrzeszenia obywatelskiego, nie rozporządzał ludźmi wolnymi w sposób arbitralny, na podobieństwo pana domu wydającego polecenia niewolnikom. Zabezpieczeniu przed takim stanem rzeczy ma służyć właśnie ustrój mieszany, monarchia zaś nieuchronnie grozi jego powstaniem.

Po drugie, warto zatrzymać się także na chwilę nad słowem „iustus”, które w cytowanym polskim wydaniu oddano nieściśle jako „dobry”. W pierwotnym znaczeniu słowo to oznacza oczywiście tyle co „sprawiedliwy”. Użycie słowa „iustus” czyni Ciceronijskie stwierdzenie bardzo mocnym i kategorycznym. Pan „sprawiedliwy” zapewne nie będzie bez przyczyny nastawał na dobra czy interesy obywatela, będzie dotrzymywał swoich zobowiązań i troszczył się o poddanych; słowem, jego władza nie będzie tyraniczna. Jednakże sprawiedliwość jako cnota władcy nie legitymizuje jeszcze jego władzy. Niezależnie od

²¹ Szerszą panoramę współczesnego republikanizmu w świecie anglosaskim oraz jego zakorzenienie w myśli starożytnej ukazuje D. Pietrzyk-Reeves, *Neorepublikańska filozofia publiczna i model demokracji*, «Etyka» 2014, vol. 48, s. 120–138.

²² Ciceron, *O państwie. O prawach*, przeł. I. Żółtowska, Kęty 1999, s. 54.

²³ Skorzystano z wydania internetowego dostępnego pod linkiem: https://www.loebclassics.com/view/marcus_tullius_cicero-de_re_publica/1928/pb_LCL213.153.xml (2.06.2026).

²⁴ Skorzystano ze słownika internetowego, hasło „dominus” dostępne pod adresem: <https://www.latin-dictionary.net/search/latin/dominus> (2.06.2026).

tego, jak pozbawiona instytucjonalnych hamulców władza monarchy jest przez jej piastuna wykonywana, koliduje z wolnością ze względów strukturalnych. Władza nad wolnym obywatelem jest bowiem zupełnie innego typu – nie występuje w niej możliwość arbitralnego rozkazywania na podobieństwo pana objawiającego swoją wolę wobec niewolników, czy decydującego o przeznaczeniu swojej własności. Nie przypadkiem u Cycerona na określenie posiadającego odpowiednią legitymację władcy występują słowa takie jak „rector”, „gubernator”, lub „princeps”, nigdy zaś „dominus”. Jedynie władza, w działaniu której wpisane są pewne ustrojowe hamulce, godzi się być wykonywaną nad ludźmi wolnymi. Przykładami instytucji ustrojowych konstytuujących tak rozumianą wolność w ustroju mieszanym są choćby *ius provocationis* (prawo odwołania się od wyroku wysokiego urzędnika do zgromadzenia ludowego), czy też szereg uprawnień trybuna ludowego, dającego plebejuszom realne narzędzia oporu wobec ewentualnie arbitralnej władzy nobilów²⁵.

Właśnie te wątki myśli Cycerona rozwija Pettit, proponując „neorzymską” koncepcję wolności. Dodajmy na marginesie, że jest to wizja zdecydowanie jednostronna i nie oddaje w pełni ducha rzymskiego republikanizmu. Przede wszystkim u Pettita traci na znaczeniu wątek dla Rzymian równie obok wolności istotny, a mianowicie dobro wspólne, u Cycerona często określane mianem *utilitas rei publicae*²⁶. Cynceron – czy szerzej, cała tradycja klasyczna począwszy od Platona – przyjmuje, że w państwie musi istnieć jakieś umiłowanie rzeczy wspólnych, łączące obywateli na poziomie fundamentalnym, zespalające ich wolę na poziomie wspólnotowym (stąd np. angielskie określenie *body politic* na określenie „państwa”, w sensie bliskim pojęciu republiki czy łacińskiej *civitas*). Dobro wspólne częstokroć wymaga poświęceń i wyrzeczenia się własnego wąsko rozumianego interesu²⁷. Tymczasem u Pettita objawia się ono przede wszystkim w formie demokratycznej sprawiedliwości społecznej²⁸, czy też standardów demokratycznych wiążących

²⁵ Por. T. Banach, *Res publica est res populi. Myśl polityczno-prawna Marka Tulliusza Cyncerona*, Łódź 2023, s. 99–106.

²⁶ Co istotne, łaciński termin „utilitas” jest ściśle powiązany z terminem „honestas”, a zatem nabiera charakteru moralnego, nie tylko związanego z pożytkiem (na przykład w sensie ekonomicznym). Por. J. Kulawiak-Cyrankowska, *Utilitas in Roman Jurists' Legal Interpretation*, Stuttgart 2025, s. 39.

²⁷ Dlatego u Cyncerona wolność (*libertas*) na płaszczyźnie politycznej była wolnością państwa, a nie wolnością indywiduum, jak jest u Pettita. Koncepcja wolności politycznej była natomiast integralnie powiązana z wolnością indywidualną, lecz ta ostatnia nie wiązała się z nie-dominacją, lecz z panowaniem nad sobą dzięki nabytym cnotom. Por. R. Piekarski, *Mówca-myśliciel-pisarz humanista, czyli Cynceron o wolności*, [w:] O. Górecki (red.), *Wolność człowieka i jej granice*, Łódź 2019, s. 99–136.

²⁸ Por. P. Pettit, *On the People's Terms...*, s. 110–121.

władzę²⁹, jednak nie stanowi motywu przewodniego ani głównego terminu technicznego. Jak zobaczymy we właściwym czasie, ten nadmierny indywidualizm Pettita wyznacza istotne ograniczenia jego teorii także w refleksji nad ADM. Na razie jednak powróćmy do zagadnienia wolności.

Aby zrozumieć, czym jest dla Pettita wolność, najlepiej rozpocząć od jej przeciwieństwa. Jest nim dominacja, ujmowana jako „bezbronna podatność na ingerencję innego podmiotu”³⁰. Podmiot dominujący zajmuje pozycję „pana” w sensie Cycerońskim, co automatycznie czyni drugiego podobnym do niewolnika. Pettit dookreśla, iż władza tego rodzaju występuje wtedy, gdy jeden wobec drugiego „(1) ma zdolność do ingerowania, (2) bezkarnie i wedle uznania (3) w pewne wybory, których drugi może dokonać”³¹. Pettit szczegółowo wyjaśnia owe przesłanki, wskazując, że do zaistnienia dominacji konieczna i wystarczająca jest realna groźba, iż dominujący zdecyduje się pogorszyć sytuację drugiego w zakresie jego wyborów, obejmujących jego żywotne interesy. Możliwość arbitralnej ingerencji konstituuje trwałą i strukturalną asymetrię między podmiotami, gdzie jedna strona relacji dysponuje realną możliwością działania wbrew woli i interesom drugiej. Warunek bezkarności oznacza brak mechanizmów instytucjonalnych lub społecznych, nadających stronie potencjalnie zdominowanej możliwości kontestacji postępowania strony potencjalnie dominującej, pociągnięcia drugiego do odpowiedzialności lub też zinstytucjonalizowanych form wyrażenia sprzeciwu. Co ważne, u Pettita relacja dominacji musi mieć charakter intencjonalny i świadomy – obie strony muszą wiedzieć o tym, że zachodzi taka realna możliwość i zdawać sobie sprawę ze swojego statusu.

Per opposita cognoscitur wolność oznacza brak dominacji, czyli sytuację, w której wszyscy obywatele cieszą się równym statusem – o nikim nie można

²⁹ P. Pettit, *The general Will, the Common Good, and a Democracy of Standards*, [w:] Y. Elazar, G. Rousselière (red.), *Republicanism and the Future of Democracy*, Cambridge 2019, s. 13–40. Pettit proponuje inter-subiektywną koncepcję dobra wspólnego, co sprowadza się do wymogów proceduralnych oraz możliwości kontestacji decyzji i polityk publicznych (Por. R. Wierchosławski, *Ontologia społeczeństwa, dobro wspólne i kontestacja. Wokół pewnych kategorii państwa republikańskiego w ujęciu Philipa Pettita*, «Roczniki Filozoficzne» 2023, vol. 71, nr 1, s. 183–212). Tym samym odrzucona jest substancjalna wizja dobra wspólnego. Pettit sytuuje się zatem daleko od klasycznego rozumienia dobra wspólnego, które zawsze wiązało się z integralnym rozwojem człowieka i jego podstawowym dobrem o charakterze substancjalnym i społecznym (Idee te znajdują swoje odzwierciedlenie w Konstytucji RP z 1997 r. oraz w nauczaniu kościoła Katolickiego. Por. M. Piechowiak, *Dobro wspólne jako fundament porządku konstytucyjnego*, Warszawa 2012; M. Sadowski, *Godność człowieka i dobro wspólne w nauczaniu papieskim nauczaniu społecznym (1878–2005)*, Wrocław 2010).

³⁰ P. Pettit, *Freedom as Antipower*, «Ethics» 1996, vol. 106, nr 3, s. 576.

³¹ Tamże, s. 578, powtórzone w tegoż, *Republicanism...*, s. 52.

powiedzieć, że zajmuje pozycję bezbronnego niewolnika, którego interes i sfera wyborów zależą od kaprysu pana; kaprysu, którego nie można w żaden sposób kontestować. Wolność jako nie-dominacja wymaga istnienia odpowiednich zabezpieczeń instytucjonalnych i normatywnych, które czynią ewentualne ingerencje przejrzystymi, uzasadnionymi, poddanymi odpowiedniej kontroli instytucji publicznych czy kontestacji obywateli. Uzasadnia to aktywną, rozbudowaną rolę państwa, choć i samo państwo powinno podlegać obywatelskiej kontroli i odpowiednim mechanizmom ustrojowym. Tak rozumianą wolność można rozpatrywać na dwóch płaszczyznach, a mianowicie na płaszczyźnie relacji prywatnych, pomiędzy równorzędnymi podmiotami, oraz na płaszczyźnie prawa i polityk publicznych, czyli relacji pomiędzy jednostką a państwem. Na płaszczyźnie pierwszej Pettit proponuje tzw. „eyeball test” (test spojrzenia), który wymaga, by każdy członek republikańskiej społeczności mógł spojrzeć innym w oczy jako równym sobie, „bez powodu do strachu czy uległości, które może wzbudzać władza ingerencji”³². Spełnienie warunków testu oznacza zatem sytuację, w której każdy, nawet najgorzej uposażony obywatel, posiada na tyle silne zabezpieczenie w systemie republikańskich instytucji, że nie ma powodu by „płaszczyc się” przed nikim, nawet najpotężniejszym współ-obywatelem³³. Natomiast w relacjach pomiędzy jednostką a państwem, analogiczną rolę pełni „tough luck test” (test „braku szczęścia”). Wolność rozpatrywana pod tym względem wymaga istnienia właściwego systemu kontroli nad decyzjami dotyczącymi polityk publicznych oraz prawnych ograniczeń jednostkowych wyborów; systemu, który daje obywatelom „na tyle duży poziom bezpieczeństwa wobec rządu, że mogą postrzegać osobiście źle widziane inicjatywy [tegoż rządu] jako zwykły brak szczęścia”³⁴. Pettit realistycznie uznaje, że zawsze znajdą się jednostki niezadowolone z działań danego rządu, czy uchwalonych praw. Jednakże osoby, których interesy doznały jakiegoś uszczerbku, nie powinny mieć powodów, by uznawać owo niekorzystne dla nich działanie rządu za wynik „złej lub obojętnej woli”³⁵. Wówczas, gdy sytuacja danej jednostki pogarsza się (na przykład, powiedzmy, w wyniku budowy przez podmiot publiczny tunelu kolejowego pod prywatnym budynkiem i wynikłych z tego komplikacji), to zasadne jest w tym przypadku postrzegać ten fakt jako zwyczajnie pechową okoliczność, nie zaś działanie obcej woli zaszkodzenia

³² P. Pettit, *On the People's Terms...*, s. 84.

³³ Tak rozumiana wolność jest bardzo blisko związana z równością. Jest to cecha każdej niemal myśli republikańskiej, widoczna bardzo wyraźnie także w polsko-litewskim republikanizmie szlacheckim. Por. T. Tulejski, *Wstęp*, [w:] J. Sachs, *De Scopo*, Warszawa 2024, s. 68–70.

³⁴ P. Pettit, *The State*, Princeton 2023, s. 10.

³⁵ P. Pettit, *The General will...*, s. 31.

interesowi prywatnemu³⁶. Jak widać, test „braku szczęścia” podobny jest do testu „spojrzenia w oczy”, gdyż obydwie koncentrują się na tym samym – na jednostce, która nie ma dobrych powodów, by uznawać samą siebie za podległą arbitralnej władzy, jeśli dany test wypadnie pomyślnie.

Dominacja algorytmiczna jako zagrożenie związane z ADM

Po zaznajomieniu się zarówno z problematyką ADM, jak i wolności jako nie-dominacji, możemy teraz przystąpić do rozważenia zagrożenia powstania dominacji nowego typu, związanej ze społeczno-technologicznymi przemianami. Choć będę argumentował, że zagrożenie dominacją algorytmiczną rzeczywiście zachodzi, to, po pierwsze, nie jest to na gruncie teorii Pettita wcale oczywiste, po drugie, sama ta teoria wymagać będzie dostosowania do nowej sytuacji związanej z ADM. W szczególności przedmiotem naszego zainteresowania są sytuacje, w których systemy ADM korzystają z AI, a zatem działają z jakimś zakresem autonomii w sensie wyżej wyłożonym. Paweł Księżak (w kontekście podstawowych gwarancji prawa konstytucyjnego) zwraca uwagę, iż właśnie ze względu na autonomię „AI staje się zjawiskiem jakościowo nowym, którego nie można zredukować do wcześniej znanych pojęć, opisanych w istniejących hipotezach norm prawnych.”³⁷ Wypada się zgodzić z tezą tego autora, że zasady ustrojowe będą wymagały dostosowania do nowej sytuacji. Wolność jako nie-dominacja służyć może jako jedno z narzędzi krytycznych w toku prac w tym zakresie, lecz także i samo to pojęcie trzeba do nowej sytuacji dostosować.

Naszą refleksję nad potencjalną dominacją algorytmiczną rozpoczniemy od stanowiska umiarkowanie optymistycznego. Jacob Sparks i Athmeya Jayaram przekonują mianowicie, że ADM może posłużyć raczej wyeliminowaniu dominacji niż jej wzmocnieniu. Autorzy ci słusznie zwracają uwagę, że według teorii Pettita w relacji dominacji strona dominująca „zawsze będzie podmiotem – nie może być to po prostu system czy sieć”³⁸. Na tym opiera się argumentacja przywołanych autorów. Decyzje zautomatyzowane są z założenia zdepersonalizowane, co eliminuje możliwość dominacji na poziomie strukturalnym. Dzieje się tak, zdaniem Sparksa i Jayaram, z trzech powodów:

Po pierwsze, systemy zautomatyzowane nie są podmiotami, nie mają ani woli ani intencji, więc wedle koncepcji Pettita nie mogą dominować. (...) Po drugie, zautomatyzowany

³⁶ Tamże, s. 32.

³⁷ P. Księżak, *My, Naród? Konstytucjonalizacja sztucznej inteligencji, czyli o potrzebie przemodelowania założeń ustrojowych*, «Przegląd Sejmowy» 2021, vol. 165, nr 4, s. 65–87.

³⁸ P. Pettit, *Freedom as Antipower...*, s. 578.

system decyzji stosowałby się do reguł bez pozostawienia miejsca dla dyskrecjonalności. To nie tylko zredukowałoby zależność od woli innych, ale poprawiłoby spójność i przewidywalność (...). Po trzecie, decyzje zautomatyzowane będą lepiej służyły celom testu spojrzenia w oczy. Nie ma możliwości, by twój los zależał od nastroju systemu, albo od tego, czy okazujemy mu wystarczający szacunek.³⁹

Przy założeniu, że systemy ADM funkcjonują poprawnie oraz na podstawie i w granicach prawa, można rzeczywiście argumentować, iż są one po prostu konsekwencją, czy może nawet ostateczną realizacją, idei rządów prawa. Choć komentowany tu artykuł tego nie problematyzuje, być może nawet w większym stopniu ADM realizuje ideę demokratycznego państwa prawnego, znanego z niemieckiej kultury prawno-politycznej jako *Rechtstaatlichkeit*⁴⁰. Ideą regulatywną od zawsze było tu przecież postępowanie powtarzalne, przewidywalne, sprofesjonalizowane i podlegające jednolitym standardom. Co prawda tradycja ta nie jest związana ściśle z republikanizmem, możemy jednak przyjąć, że współcześnie nie-dominacja wymaga tego rodzaju zaawansowanych mechanizmów, a ADM, zdaniem niektórych autorów, może owe mechanizmy wzmocnić.

Wydaje się, iż tego rodzaju optymizm opiera się na pewnym nieporozumieniu co do istoty idei nie-dominacji. Sparks i Jayaram dostrzegają naturalną czy intuicyjną niechęć wobec ADM, lecz w moim przekonaniu mylą się co do jej źródła, a w każdym razie nie uwzględniają wszystkich kluczowych źródeł. Twierdzą oni mianowicie, że sprzeciw wobec ADM wynika przede wszystkim z braku „zamienności ról” (*role-reversibility*), czyli na tym, że systemy ADM nie będą nigdy podlegały własnym decyzjom, czy też decyzjom o charakterze podobnym do decyzji, które wydają. W tym sensie nie jest możliwe stworzenie „demokratycznej wspólnoty” pomiędzy obywatelem a systemem⁴¹. Ludzcy urzędnicy natomiast, będący także obywatelami, nie tylko wydają decyzje, ale

³⁹ J. Sparks, A. Jayaram, *Rule by Automation: How Automated Decision Systems Promote Freedom and Equality*, «Moral Philosophy and Politics» 2022, vol. 9, nr 2, s. 6.

⁴⁰ Relacje i napięcia pomiędzy ideami *rule of law* i *Rechtstaat* analizują Karolina M. Cern i Barbara Janusz-Pohl w tekście „*Tam i z powrotem*” – *the Rule of Law ujęte od źródeł aż po zastosowanie w orzecznictwie TSUE*, «Archiwum Filozofii Prawa i Filozofii Społecznej» 2024, vol. 41, nr 4, s. 36–52.

⁴¹ Autorzy opierają się w dużej mierze na tekście K. Brennan-Marquez, S. Henderson, *Artificial Intelligence and Role-Reversible Judgment*, «Journal of Criminal Law & Criminology» 2019, vol. 109, nr 2, s. 137–164. Ci ostatni autorzy podkreślają, że pewne osądy (na warsztat biorą wyroki karne ferowane przez sędziów, a zatem nie interesujące nas decyzje urzędnicze, lecz można ich myśl przez analogię rozszerzyć), są specyficznej natury – polegają one na tym, że „podejmujący decyzję mogą wyobrazić sobie siebie podlegających tej decyzji”, co ma miejsce we wspólnotach demokratycznych (tamże, s. 149).

także, przy innych okazjach, im podlegają⁴². Podlegają także ocenie moralnej – kogoś, kto ze złej woli wydaje krzywdzące decyzje, może spotkać słuszne moralne oskarżenie, pozbawione tego samego sensu w przypadku systemu⁴³. W moim przekonaniu, choć problem zamienności ról stanowi istotny składnik problemu, istnieją także inne strukturalne zagrożenia związane z ADM, łączące statuujące potencjał dominacji algorytmicznej. Jest to, obok zamienności ról, przede wszystkim niesymetryczna dynamika władzy oraz związane z nim kwestie statusu politycznego. Kolejne dwie sekcje będą poświęcone analizie dominacji algorytmicznej na poziomie wyjaśnionego wyżej „testu spojrzenia”. Tak jak refleksja republikańska od zawsze nakierowana jest na potencjalną degenerację ustroju państwowego, tak i nasza refleksja nad ADM skoncentruje się na potencjalnych patologiach prowadzących w kierunku dominacji algorytmicznej.

ADM jako statuujące asymetryczne relacje władzy

Ze względów objętościowych ograniczę się w moich analizach do najszerzej znanego i komentowanego „testu spojrzenia”. Pytamy: czy zastosowanie ADM w administracji publicznej może budzić wątpliwości z perspektywy tegoż testu? Zaczniemy od ustalenia, czy test spojrzenia w oczy w ogóle można zastosować do relacji z administracją publiczną. Jak pamiętamy, Pettit test ten wprowadził przede wszystkim jako relewantny w relacjach pomiędzy obywatelami, nie zaś w relacjach pomiędzy obywatelem a władzą⁴⁴. Wydaje się jednak, iż w sytuacjach wydawanej przez urzędnika indywidualno-konkretnej decyzji, możemy również mówić o relacjach dominacji, w postaci świadomie szkodliwej (bezprawnej lub nie) ingerencji wydającego decyzję w legitymowane interesy obywatela. W najprostszym przypadku decyzji tradycyjnej (bez użycia ADM) obywatel może być zdominowany w tym sensie, że arbitralna wola urzędnika budzi uzasadnione obawy i uległość, a zatem test spojrzenia w oczy wypada niepomysłnie. Należy teraz rozważyć, czy zastąpienie człowieka, jego odciążenie lub sugestia przez system ADM (w wyżej wyłożonym sensie Boos) wyłączenie łagodzi to niebezpieczeństwo, czy też pojawia się ono w nowej formie.

Przypomnijmy, że wedle argumentacji Sparks i Jayaram użycie ADM (zwłaszcza w najmocniejszej formule zastępowania człowieka) wyklucza działanie systemu w złej wierze (gdyż nie jest to podmiot posiadający własne cele i emocje), a poprawna konstrukcja całości pozwala oprzeć proces podejmowania decyzji

⁴² J. Sparks, A. Jayaram, *Rule by Automation...*, s. 13.

⁴³ Tamże, s. 16.

⁴⁴ P. Pettit, *On the People's Terms...*, s. 87.

wyłącznie na prawie powszechnie obowiązującym i publicznie podzielanych wartościach (na przykład zasadach konstytucyjnych czy prawach człowieka)⁴⁵. Inni autorzy podkreślają jednak inne, bardziej niepokojące aspekty. Koncentrują się bowiem nie na „kaprysie” podejmującego indywidualno-konkretną decyzję (kaprys, a zatem tradycyjnie pojęta uległość są rzeczywiście wykluczone lub złagodzone przez ADM), lecz na generowanych przez technologię nierównościach statusu oraz niesymetrycznej dynamice władzy. Wątki te eksploruje na przykład Jonne Maas⁴⁶. Koncentruje się ona co prawda na relacjach z przedsiębiorstwami prywatnymi, oraz na konkretnej technologii AI, jaką jest uczenie maszynowe, jednak refleksję tę można uogólnić. Maas przekonuje, że systemy uczenia maszynowego powstają w procesie, który na wielu etapach – selekcji danych, trenowania modelu oraz projektowania jego implementacji – zawiera w sobie szereg ludzkich decyzji, które nie są widoczne dla końcowego użytkownika aplikacji AI, ale w sposób istotny wpływają na jego sytuację. A zatem, o ile produkt końcowy działa w sposób zdepersonalizowany i nie jest niczym więcej niż algorytmem, w rzeczywistości jego działanie „dziedziczy” trudną do prześledzenia serię decyzji człowieka. Decyzje owe są bardzo trudne do kontestowania i kontroli⁴⁷.

Refleksja Maas (w podobnym tonie wypowiadają się też inni autorzy⁴⁸) zastosowana do systemów ADM prowadzi do przekonania, że w przypadku wspomaganej lub zastąpionej przez te systemy administracji publicznej brak arbitralności czy asymetrii relacji jest jedynie pozorny. Co prawda systemy ADM muszą działać na podstawie i w granicach prawa, lecz, zawężając dyskrecjonalność urzędnika (lub ją eliminując), usztywniają cały proces. Samo w sobie

⁴⁵ Tamże, s. 11.

⁴⁶ J. Maas, *A Neo-Republican Critique of AI Ethics*, «Journal of Responsible Technology» 2022, vol. 9, s. 1–8; tegoż, *Freedom in the Digital Age: Designing for Non-Domination*. [praca doktorska, Delft University of Technology] 2025.

⁴⁷ Por. zwr. J. Maas, *Freedom in the Digital Age...*, s. 113–121.

⁴⁸ Por. np. M.D. Hamilton, *Digital Domination: A Case for Republican Liberty in Artificial Intelligence*, online, bez daty, <https://academic.oup.com/edited-volume/59762/chapter-abstract/508607957?redirectedFrom=fulltext>; B.L. Hoeksema, *A Republican Perspective on Digital Domination [Dissertation]*, Nijmegen 2025; J. Susskind, *The Digital Republic: On Freedom and Democracy in the 21st Century*, Londyn 2022. Warto zwrócić uwagę, iż prace te koncentrują się przede wszystkim na niesymetrycznej władzy korporacji nad użytkownikami (konsumentami), wpisują się zatem w większym stopniu w krytykę cyfrowego kapitalizmu niżli w uprawianą w tym artykule krytyczną filozofię państwa i administracji publicznej. Tego ostatniego tematu dotyczy N. Frost w tekście *The Impoverished Publicness of Algorithmic Decision Making*, «Oxford Journal of Legal Studies» 2024, vol. 44, nr 4, s. 780–807. Frost koncentruje się jednak w mniejszym stopniu na wspólnocie uznającej równy status każdego obywatela; motyw przewodni jego tekstu stanowi „wspólnota działania” organów administracji publicznej powiązanych publiczno-prawnymi regułami i zasadami odpowiedzialności.

nie jest to problemem – problem pojawia się dlatego, że owo „uszytnienie” odbywa się poprzez szereg decyzji statuujących niesymetryczne relacje władzy, a ludzie podlegli tej władzy nie mogą kontrolować procesu powstawania algorytmu, i jest to przeszkoda zasadnicza, strukturalna. Nie są przy tym istotne szczegóły techniczne, choć z pewnością użycie nietransparentnych technik uczenia maszynowego nie poprawia sytuacji⁴⁹. Status obywateli staje się zatem niepewny, ponieważ nie wiedzą oni, czyje decyzje realnie wpływają na ich żywotne interesy.

Biorąc powyższe pod uwagę, zastosowanie testu spojrzenia w oczy jest niemożliwe, ponieważ dany obywatel często po prostu nie wie, gdzie patrzeć, aby ten test wykonać. Po to, by w jakiś sposób zbliżyć się do możliwości wykonania testu spojrzenia, proponuję dwa terminy na określenie sytuacji niesymetrycznej relacji władzy generowanej przez dominację algorytmiczną. Są to „podległość zaszytej woli” i „podległość obojętnej woli”. Aby wyjaśnić znaczenie tych pojęć, wyobraźmy sobie sytuację pana B., który nie otrzymuje wizy uprawniającej do wjazdu na teren państwa X, a jego wniosek był procedowany przez człowieka, podążającego w tym przypadku za rekomendacją algorytmu. W tej sytuacji nie mamy do czynienia z klasyczną dominacją, ponieważ urzędnik stał się *de facto* „robotycznym” operatorem systemu rozwiązywania problemów decyzyjnych za pomocą algorytmu. Zarazem jednak, pan B. potencjalnie podlega dwójakiego rodzaju arbitralnej władzy. Po pierwsze, możemy mieć do czynienia z sytuacją, w której podczas powstawania systemu ADM ktoś arbitralnie uznał, że na przykład osoby o danej historii przebytych chorób nie będą wpuszczane na terytorium państwa X. Pan B. nie musi o tym wiedzieć, nie musi o tym wiedzieć nawet urzędnik U, teoretycznie „podejmujący decyzję” w świetle prawa. Dlatego arbitralną wolę, która jest tu ukryta, można nazwać „zaszytą”, to znaczy wbudowaną w system bez możliwości republikańskiej kontroli. Po drugie, system mógł być skonstruowany niedbale, a rząd, który dopuścił jego użycie, zrobił to zwyczajnie naprędce, na przykład kierując się względami populistycznymi. W tej sytuacji wniosek pana B. może się spotkać odmową tylko dlatego, że na przykład algorytm kojarzy ze sobą dwie neutralne cechy wnioskodawcy jako konstytuujące niebezpieczeństwo, które w rzeczywistości nie zachodzi. Urzędnik U natomiast nie ma możliwości tego sprawdzić, z uwagi na źle skonstruowane mechanizmy transparentności i kontroli. W takim przypadku mamy do czynienia z „obojętną wolą”, gdzie rozwiązywanie problemów decyzyjnych projektuje się przede wszystkim z myślą o tym, by usprawnić proces, bez należytej troski o zabezpieczenie legitymowanych interesów osób podległych tym

⁴⁹ Por. U. Franke, *First- and Second-Level Bias in Automated Decision-making*, «Philosophy & Technology» 2022, vol. 35, s. 21–41.

decyzjom. Obie te wady są przykładami tego, co nazywam „dominacją algorytmiczną”. Test spojrzenia w oczy okazuje się możliwy dopiero po prześledzeniu procesu konstrukcji systemu podejmowania decyzji oraz poznania szczegółów każdego przypadku.

Konkuzje

Nasze dociekania prowadzą do paradoksalnych wniosków. Cycerońskie przekonanie, że wolność ostatecznie polega na tym, by nie podlegać żadnemu panu, okazuje się trudne do interpretacji w warunkach nowych technologii zaprzęgniętych w struktury władzy. Okazuje się także, że przekonanie Pettita dotyczące tego, że dominacja musi mieć charakter świadomy (obydwie strony muszą o niej wiedzieć), a przez to, że dominującym może być tylko konkretny podmiot ludzki, nie wytrzymuje konfrontacji z nową rzeczywistością. Okazuje się bowiem, że „brak pana” w postaci podległości zdepersonalizowanym systemom ADM może (choć oczywiście nie musi) prowadzić do nowego rodzaju dominacji. Robotyczni urzędnicy nie będą panami w pierwotnym sensie Cicerona i Pettita, natomiast dominacja w systemach ADM może objawiać się w formach bardziej zawołowanych i pośrednich, na przykład tych, które określiłem „zaszytą wolą” oraz „obojętną wolą”. Warunkiem koniecznym obydwu są niesymetryczne relacje władzy. Ocena praktyki administracyjnej pod względem dominacji staje się w tych warunkach bardziej skomplikowana niż dotąd – wymaga zindywidualizowanego podejścia i dokładnego śledzenia procesu powstawania i stosowania struktur ADM w każdym z przypadków.

Refleksja nad systemami zdepersonalizowanymi ujawnia także istotny aspekt pozytywny prawidłowo funkcjonującej sfery publicznej oraz wartość moralną państwa-republiki. Nie wystarczy bowiem usunąć przestrzeni dla kaprysu i arbitralności jednostkowej decyzji, by mówić o wolności jako nie-dominacji. W kontekście dominacji algorytmicznej widać jeszcze wyraźniej, jaką wartość ma wzajemne uznanie równego statusu obywatelskiego w ramach dobrze funkcjonującej Rzeczypospolitej. Takie wątki znajdziemy także u Pettita, choćby piszącego, że być człowiekiem cieszącym się wolnością jako nie-dominacją, oznaczać bycie kimś, „komu należy się odpowiednia pozycja pomiędzy jego rodakami”⁵⁰. Wydaje się jednak, iż idea wolności jako nie-dominacji winna być w stopniu większym niż u Pettita otwarta na perspektywę dobra wspólnego. Nasze rozważania prowadzą bowiem do przekonania, że wolność

⁵⁰ P. Pettit, *On the People's Terms...*, s. 91.

to uczestnictwo w Rzeczypospolitej, oznacza nie tylko możliwość niezdominowanego wyboru (na co kładzie akcent Pettit), lecz także, jeśli nie przede wszystkim, wspólne bytowanie w ramach wartościowego związku społecznego, szanującego każdego człowieka. Patrząc na rzecz pod tym względem, obywatel podlegający ADM będzie wolny wtedy, gdy będzie miał świadomość, czyje decyzje mają wpływ na jego żywotne interesy, oraz gdy system ADM nie zniweczy możliwości zawiązania wspólnoty opartej na równym statusie obywatelskim.

Bibliografia

- Banach T., *Res publica est res populi. Myśl polityczno-prawna Marka Tulliusza Cicerona*, Łódź 2023.
- Boos A.K., *Conceptualizing Automated Decision-Making in Organizational Contexts*, «Philosophy & Technology» 2024, vol. 37, nr 3.
- Brennan-Marquez K., Henderson S., *Artificial Intelligence and Role-Reversible Judgment*, «Journal of Criminal Law & Criminology» 2019, vol. 109, nr 2.
- Cern K.M., Janusz-Pohl B., „Tam i z powrotem” – the Rule of Law ujęte od źródeł aż po zastosowanie w orzecznictwie TSUE, «Archiwum Filozofii Prawa i Filozofii Społecznej» 2024, vol. 41, nr 4.
- Chiusi F., Fischer S., Kayser-Bril N., Spielkamp M. (red.), *Automating Society Report 2020*, Berlin 2020, <https://automatingsociety.algorithwatch.org/wp-content/uploads/2020/12/Automating-Society-Report-2020.pdf> (2.06.2026).
- Coeckelbergh M., *The Political Philosophy of AI: An Introduction*, Cambridge 2022.
- Cycon M.R., *O państwie. O prawach*, przeł. I. Żółtowska, Kęty 1999.
- Dworkin G., *The Theory and Practice of Autonomy*, Cambridge 2012.
- Franke U., *First- and Second-Level Bias in Automated Decision-making*, «Philosophy & Technology» 2022, vol. 35.
- Frost N., *The Impoverished Publicness of Algorithmic Decision Making*, «Oxford Journal of Legal Studies» 2024, vol. 44, nr 4.
- Geburczyk F., *Kryteria dopuszczalności decyzji zautomatyzowanych w świetle brzmienia art. 22 ust. 1 RODO*, «Business Law Journal» 2022, vol. 75.
- Hamilton M.D., *Digital Domination: A Case for Republican Liberty in Artificial Intelligence*, online <https://academic.oup.com/edited-volume/59762/chapter-abstract/508607957?redirectedFrom=fulltext> (bez daty).
- Heikkinen A., van Gerven M., Koulu R., *From Discretion to Calculation: How Analog Automation Shaped Digitalization of Finnish Social Assistance*, «Public Administration» 2026, early view, <https://onlinelibrary.wiley.com/doi/10.1111/padm.70041>.
- Hoeksema B.L., *A Republican Perspective on Digital Domination* [praca doktorska], Nijmegen 2025.
- Korsgaard K., *The Sources of Normativity*, Cambridge 2010.
- Księżak P., *My, Naród? Konstytucjonalizacja sztucznej inteligencji, czyli o potrzebie przemodelowania założeń ustrojowych*, «Przegląd Sejmowy» 2021, vol. 165, nr 4.
- Kulawiak-Cyrankowska J., *Utilitas in Roman Jurists' Legal Interpretation*, Stuttgart 2025.
- Larus J. i in., *When Computers Decide: European Recommendations on Machine-Learned Automated Decision Making*, Informatics Europe & EUACM 2018.

- Maas J., *A Neo-Republican Critique of AI Ethics*, «Journal of Responsible Technology» 2022, vol. 9.
- Maas J., *Freedom in the Digital Age: Designing for Non-Domination* [praca doktorska, Delft University of Technology] 2025.
- Pałka P., *Sztuczna inteligencja*, [w:] B. Szlachta, W. Pasierbek (red.), *Słowniki Społeczne*, vol. 18, Kraków 2025.
- Pettit P., *Freedom as Antipower*, «Ethics» 1996, vol. 106, nr 3.
- Pettit P., *On The People's Terms: A Republican Theory and Model of Democracy*, Cambridge 2012.
- Pettit P., *Republicanism: A Theory of Freedom and Government*, Oxford 1997.
- Pettit P., *The general Will, the Common Good, and a Democracy of Standards*, [w:] Y. Elazar, G. Rousselière (red.), *Republicanism and the Future of Democracy*, Cambridge 2019.
- Piechowiak M., *Dobro wspólne jako fundament porządku konstytucyjnego*, Warszawa 2012.
- Piekarski R., *Mówca-mysliciel-pisarz humanista, czyli Ciceron o wolności*, [w:] O. Górecki (red.), *Wolność człowieka i jej granice*, Łódź 2019.
- Pietrzyk-Reeves D., *Neorepublikańska filozofia publiczna i model demokracji*, «Etyka» 2014, vol. 48.
- Pilving I., *Guidance-based Algorithms for Automated Decision-Making in Public Administration: the Estonian Perspective*, «Rivista Interdisciplinare Sul Diritto Delle Amministrazioni Pubbliche» 2023, vol. 1.
- Russel S., Norwig P., *Artificial Intelligence: A Modern Approach*, Londyn 2021.
- Sadowski M., *Godność człowieka i dobro wspólne w nauczaniu papieskim nauczaniu społecznym (1878–2005)*, Wrocław 2010.
- Sarra C., *Artificial Intelligence in Decision-making: A Test of Consistency between the "EU AI Act" and the "General Data Protection Regulation"*, «Athens Journal of Law» 2024, vol. 11, nr 1.
- Sartor G., *Cognitive Automata and the Law: Electronic Contracting and the Intentionality of Software Agents*, «Artificial Intelligence and Law» 2009, vol. 17.
- Sparks J., Jayaram A., *Rule by Automation: How Automated Decision Systems Promote Freedom and Equality*, «Moral Philosophy and Politics» 2022, vol. 9, nr 2.
- Susskind J., *The Digital Republic: On Freedom and Democracy in the 21st Century*, Londyn 2022.
- Szewczyk E., *Algorytmizacja jurysdykcyjnego postępowania administracyjnego i jej skutki społeczne*, Warszawa 2025.
- Sztandar-Sztanderska K., Kotnarowski M., Zielińska M., *Czy algorytmy wprowadzają w błąd? Metaanaliza algorytmu profilowania bezrobotnych stosowanego w Polsce*, «Studia Socjologiczne» 2021, vol. 250, nr 1.
- Tulejski T., *Wstęp*, [w:] J. Sachs, *De Scopo*, Warszawa 2024.
- Wei T. i in., *Agentic Reasoning for Large Language Models*, 2026, <https://arxiv.org/pdf/2601.12538/> (2.06.2026).
- Wierchosławski R., *Ontologia społeczeństwa, dobro wspólne i kontestacja. Wokół pewnych kategorii państwa republikańskiego w ujęciu Philipa Pettita*, «Roczniki Filozoficzne» 2023, vol. 71, nr 1.