

X kadencja



KANCELARIA SEJMU

Biuro Komisji Sejmowych

PEŁNY ZAPIS PRZEBIEGU POSIEDZENIA

**Komisja
Kultury
i Środków
Przekazu**

■ **PODKOMISJI STAŁEJ DO SPRAW
CZYTELNICTWA I PRAWA AUTORSKIEGO
(NR 11)
z dnia 21 lutego 2025 r.**

Pełny zapis przebiegu posiedzenia Komisji Kultury i Środków Przekazu

– podkomisji stałej do spraw czytelnictwa i prawa autorskiego (nr 11)

21 lutego 2025 r.

Podkomisja do spraw czytelnictwa i prawa autorskiego, obradująca pod przewodnictwem poseł **Pauliny Matysiak (Razem)**, przewodniczącej podkomisji, rozpatrzyła:

- **informację na temat otwartego dostępu do ponownego wykorzystania informacji i danych sektora administracji publicznej oraz dostępu obywateli do wiedzy o państwie.**

W posiedzeniu udział wzięli: **Michał Gramatyka** sekretarz stanu w Ministerstwie Cyfryzacji wraz ze współpracownikami, **prof. dr hab. Paweł Łuków** przewodniczący Komitetu Etyki w Nauce Polskiej Akademii Nauk, **Maria Drabczyk** prezes zarządu, kierownik działu polityk i rzecznictwa Fundacji Centrum Cyfrowe, **Małgorzata Kanownik** dyrektor generalna Polskiej Izby Książki, **Radosław Łukasik** członek Polskiej Izby Książki, przedstawiciel Wydawnictwa Wolters Kluwer, **Maciej Artur Nadzikiewicz** członek zarządu Wikimedia Europe wraz ze współpracownikami.

W posiedzeniu udział wzięł pracownik Kancelarii Sejmu: **Julia Popławska** – z sekretariatu Komisji w Biurze Komisji Sejmowych.

Przewodnicząca poseł Paulina Matysiak (Razem):

Dzień dobry. Otwieram posiedzenie podkomisji stałej do spraw czytelnictwa i prawa autorskiego. Stwierdzam kworum.

Porządek dzisiejszego posiedzenia przewiduje informację na temat otwartego dostępu do ponownego wykorzystania informacji i danych sektora administracji publicznej oraz dostępu obywateli do wiedzy o państwie.

Na sali są z nami goście. Witam bardzo serdecznie pana Michała Gramatykę, sekretarza stanu w Ministerstwie Cyfryzacji. Witam także przedstawicieli ministerstwa, panią Annę Gos, dyrektor Departamentu Zarządzania Danymi w Ministerstwie Cyfryzacji i panią Iwonę Szelenbaum, naczelnik Wydziału Polityki Otwartości Danych w Departamencie Zarządzania Danymi. Są także inni goście, pan profesor Paweł Łuków, przewodniczący Komitetu Etyki w Nauce PAN, pani Maria Drabczyk, prezes zarządu, kierowniczką działu polityki rzecznictwa Fundacji Centrum Cyfrowe, pani Małgorzata Kanownik, dyrektor generalna Polskiej Izby Książki, pan Radosław Łukasik, członek Polskiej Izby Książki, przedstawiciel wydawnictwa Wolters Kluwer oraz panowie Łukasz Lipiński i Maciej Nadzikiewicz, członkowie społeczności Wikimedia, którzy przedstawiają nam swoje spojrzenie na ten temat. Ale najpierw oddam głos pani Annie Gos. Pani Anna ma upoważnienie pana ministra, żeby przedstawić informację. Następnie będzie prezentacja i później dyskusja, mam nadzieję, że ciekawa. Bardzo proszę, pani Anno.

Dyrektor Departamentu Zarządzania Danymi Ministerstwa Cyfryzacji Anna Gos:

Dzień dobry. Bardzo dziękuję. Pani przewodnicząca, szanowni państwo, z dużą przyjemnością przedstawię informację ministra cyfryzacji na temat otwartego dostępu do ponownego wykorzystywania informacji i danych sektora administracji publicznej z tego względu, że Polska ma się czym chwalić na arenie międzynarodowej. Chronologicznie zacznę od tyłu.

W ostatnim rankingu Open Data Maturity, badaniu, które Komisja Europejska przeprowadza co roku już od dziesięciu lat, Polska przez ostatnie dwa lata, w ostatnim roku zajmuje drugie miejsca po Francji. Badane są cztery obszary. Jest to portal otwartych danych, otoczenie prawne, otoczenie techniczne, czyli portale otwartych danych, ale też wpływ na społeczeństwo

i gospodarkę. W większości tych obszarów jesteśmy na pierwszym miejscu. W związku z tym dojrzałość w otwieraniu danych w Polsce jest na bardzo wysokim poziomie. Teraz przedstawię dlaczego. Otóż dlatego że przez ostatnie lata konsekwentnie budowaliśmy i budujemy ekosystem otwierania danych. Skonstruowaliśmy go wokół trzech filarów.

Pierwszy filar to filar regulacyjny. Najważniejszym aktem prawnym regulującym otwarte dane i ponowne wykorzystywanie informacji sektora publicznego jest ustawa o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego. Kompleksowo reguluje cały system otwartych danych, implementuje dyrektywę o podobnym tytule, ale wyszliśmy poza ramy dyrektywy. Dosyć unikalnym rozwiązaniem w Europie jest wprowadzenie sieci pełnomocników do spraw otwartości danych w każdym ministerstwie, w Kancelarii Prezesa Rady Ministrów i w Głównym Urzędzie Statystycznym. Koordynatorem sieci pełnomocników jest minister właściwy do spraw informatyzacji, czyli minister cyfryzacji. Regularnie realizujemy te zadania poprzez spotkania indywidualne, grupowe, opracowanie harmonogramów otwierania danych, zasilanie portalu dane.gov.pl nowymi danymi. Jesteśmy w stałym kontakcie. Wzrost zakresu danych na portalu dane.gov.pl jest zauważalny.

To, co wprowadziła ustawa, to również nowe kategorie danych, których wykorzystanie ma być w sposób proaktywny, czyli przez API, interfejs programowania aplikacji. Są to dane wysokiej wartości. Dane te zostały określone, również wymienione w akcie wykonawczym do dyrektywy. Są skategoryzowane w pięć obszarów. Jest to obserwacja ziemi i środowisko. Są to dane statystyczne, dane o mobilności. W każdym razie te wszystkie dane mają być udostępnione przez API bezpłatnie i do ponownego wykorzystywania. Inwentaryzację tych wszystkich danych przeprowadziliśmy z właściwymi resortami. Bądź w projektach, bądź już dzisiaj dane sukcesywnie mają dobudowywane API i są udostępnione w sposób proaktywny. Ponadto w ustawie wprowadziliśmy, ułatwiliśmy dostęp do danych naukowych, ale również dynamicznych, czyli takich danych, które są dostępne w czasie rzeczywistym. Ustawa dała również podstawę do opracowania programu do spraw otwierania danych publicznych. O ile sama ustawa jest aktem prawnym, o tyle program jest dokumentem operacyjnym, który wymienia bazy, do których w latach 2021–2027 powinno być dobudowane API, ale też cały proces, harmonogramy udostępniania danych, sprawozdań, sprawozdawania się Radzie Ministrów. Jeżeli chodzi o filar prawny, jest to najważniejszy akt prawny. Poza tym, jak wspomniałam, jest program otwierania danych, czyli dokument operacyjny.

Drugim filarem, który prowadzimy w Ministerstwie Cyfryzacji – współpracujemy także z innymi instytucjami – to filar techniczny. Filar techniczny to prowadzenie portalu dane.gov.pl. Jest to krajowy punkt dostępu do otwartych danych w trybie bezwnioskowym. Dane, które są udostępnione na portalu, mogą być wykorzystywane do dowolnych celów komercyjnych bądź niekomercyjnych, chyba że użytkownik chciałby wykorzystać je na innych warunkach niż te, które są określone w portalu. Proszę mi wierzyć, że chodzi o minimalne warunki określone w ustawie, jak podanie czasu wytworzenia czy miejsca tak, żeby potencjalny użytkownik wiedział, że na przykład jest to materiał, który będzie przerobiony. Dzisiaj w portalu jest udostępnionych ponad 40 tys. informacji przez 500 dostawców. Około 580 danych jest udostępnionych przez API do automatycznego wybierania.

W ministerstwie prowadzimy również drugi portal z otwartymi zasobami. Jest to portal kronika. Znajdują się tam fotografie, książki, czasopisma. Głównie są to metadane, które prowadzą do źródła, ale są również takie zasoby, które mogą być dowolnie wykorzystywane.

Wreszcie trzeci filar, który również jest niezwykle istotny, to filar organizacyjny. O pełnomocnikach do spraw otwartości już powiedziałam. Odgrywają oni niebagatelną rolę tak naprawdę w postępie otwierania danych. Ich rola polega również na inwentaryzacji danych w poszczególnych resortach i podnoszeniu jakości danych, dlatego że portal dane.gov.pl zajmuje pierwsze miejsce w Europie ze względu na funkcjonalności, które zostały wdrożone w tym portalu. Funkcjonalności te wymagają od instytucji publicznych, które dodają swoje zasoby, również dostosowania się do wyższego poziomu otwartości niż, powiedzmy sobie, PDF i ustrukturyzowania danych w taki sposób, żeby możliwe było ich łatwe przetwarzanie do maszynowego odczytu.

Jeżeli chodzi o filar organizacyjny, to wspieramy instytucje już nie tylko pełnomocników, ale też samorząd czy sektor prywatny, który dzisiaj także może zasilać portal swoimi zasobami

również w warstwie edukacyjnej. Przygotowaliśmy i poradniki, i multimedialne szkolenia. Całą bazę danych i wszystkie materiały w zakresie otwartości danych, jak sprawozdania czy harmonogramy, udostępniamy na portalu. Jawność dokumentów, transparentność w tym zakresie jest bardzo duża. Poza tym organizowaliśmy i w planach mamy organizowanie hackathonów, czyli maratonów programowania. Uczyliśmy się od programistów, jakie dane są potrzebne do wytwarzania aplikacji produktów, ale też w jaki sposób dobrze je ustrukturyzować. Były to dwudniowe maratony. Mamy w planach i przeprowadzaliśmy już konferencje, również międzynarodowe, „The future is data”. Najbliższa będzie w kwietniu. Tym razem jest to konferencja w ramach prezydencji. Edukujemy również samorządy w postaci warsztatów w różnych aspektach zarówno prawnych, jak i technicznych. W tym roku odbędzie się 15 warsztatów danych. Dzisiaj również jednostki samorządów są już zauważalnym partnerem na portalu, udostępniają swoje zasoby.

Cały system oparty na trzech filarach traktujemy jako koherentny, dający nie tylko wysoką pozycję Polsce, gdyż nasze przykłady działań są przytaczane także na arenie międzynarodowej, ale także jako system dobrze współpracujący z innymi instytucjami, innymi podmiotami. Tak naprawdę w planach mamy rozbudowę poszczególnych komponentów systemu, jak chociażby portalu dane.gov.pl tak, żeby był przyjazny użytkownikom oraz przyjazny instytucjom, które zasilają portal danymi. W planach mamy dalsze warsztaty i edukowanie samorządów, jak również edukowanie na poziomie administracji centralnej, a także przygotowanie raportów eksperckich związanych z rozwijającym się dynamicznie obszarem otwierania danych.

Są to kluczowe elementy systemu i naszych działań. Bardzo dziękuję. Tak jak powiedziałam, z dużą przyjemnością przedstawiłam informację.

Przewodnicząca poseł Paulina Matysiak (Razem):

Bardzo dziękuję. Pewnie będziemy wracać do państwa z Ministerstwa Cyfryzacji z pytaniami. Teraz poproszę o przedstawienie prezentacji przez pana Macieja Nadzikiewicza tudzież pana Łuksza Lipińskiego. To już panowie podzielą się czasem, tylko jeszcze prośba na początek, żeby jakoś nam zarysować czas, ile będzie trwała prezentacja tak, żebyśmy może jeszcze zdążyli poprosić o jakąś wypowiedź, zapytać o zdanie pana ministra, który musi wyjść trochę wcześniej. Bardzo proszę.

Członek zarządu Wikimedia Europe Maciej Artur Nadzikiewicz:

Dzień dobry. Nazywam się Maciej Nadzikiewicz. Jestem członkiem społeczności Wikimedia. W latach 2022–2024 byłem administratorem polskiej Wikipedii.

Prezentacja, którą dzisiaj przedstawię, zajmie nie więcej niż pięć minut. Raczej jest naszym spojrzeniem na sprawę i wstępem do dyskusji.

Przede wszystkim chciałem zarysować, dlaczego jestem dzisiaj osobiście. Do sierpnia 2022 r. polskie portale ministerstw oraz administracji rządowej udostępniały swoje treści i multimedia na wolnej licencji, to jest z minimalnymi warunkami podania autorstwa owych treści. Po 18 sierpnia 2022 r. została podjęta nigdzie publicznie nieopublikowana decyzja, żeby strony te udostępniać na o wiele bardziej restrykcyjnej, najbardziej restrykcyjnej licencji Creative Commons, która wymaga zakazu utworów zależnych oraz zakazuje utworów komercyjnych, co sprawia, że nie jest otwartą licencją. Sprawa była poruszana przez nas w zeszłym roku właśnie na posiedzeniu podkomisji oraz później 11 czerwca na posiedzeniu Komisji Kultury i Środków Przekazu w ramach dyskusji o nowelizacji ustawy o prawie autorskim. Pod koniec zeszłego roku wysłaliśmy pismo do ministerstwa w tej sprawie. W październiku dostaliśmy odpowiedź w sprawie otwartych danych. Dzisiaj jesteśmy, żeby kontynuować dyskusję.

Chciałem pokazać dane, które udało nam się wyjąć z naszej strony przez naszego kolegę Marcina Szwarca. 20 czerwca ubiegłego roku w serwisie gov.pl ponad 330 tys. stron internetowych było dostępnych na licencji zamkniętej, a niecałe 6,5 tys. na licencji otwartej. Tutaj chciałbym lekko zarysować, czym to się różni. Mogą to być przerażające skróty, ale w dużym uproszczeniu pokażę screen z Wikipedii dostępny na wolnej licencji wyjaśniający w formie tabelkowej, jakie są różne licencje Creative Commons. Zaczniemy od definicji tego, czym są licencje otwarte. Według Open Knowledge Foundation są to licencje, które umożliwiają przekazywanie dalej treści, również w formie lekkiego remiksu, pozwalają na dalsze użycie komercyjne oraz są zgodne z polską ustawą o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego. Art. 15 ust. 3 stanowi, że podmiot zobowiązany może

określić warunki ponownego wykorzystywania informacji sektora publicznego, stosując standardowe otwarte licencje, czyli jedną z trzech, które tutaj wymieniamy, jeżeli mówimy o licencjach Creative Commons.

Licencja, z której korzystano na stronach gov.pl przez lata, była to właśnie licencja CC BY, czy uznanie autorstwa. Możemy korzystać z treści na tym portalu pod warunkiem powiedzenia, kto jest autorem i dania atrybucji. Od 2022 r. mamy wiele więcej wymagań. Zakazane jest użycie niekomercyjne oraz nie możemy tworzyć utworów zależnych, co znacznie utrudnia korzystanie z danych w tym portalu.

Chciałbym również pokazać, jak działa to w innych krajach, jak są udostępniane dane rządowe. Zakładam, że każdy z państwa widział swego czasu zdjęcia, które dzisiaj pokazuję. Tutaj mamy zdjęcia między innymi z Teleskopu Hubble'a oraz nowego Teleskopu Jamesa Webba. To są portrety poprzedniego prezydenta Stanów Zjednoczonych, wiceprezydenta i speaker of the House Nancy Pelosi. Są to wszystkie zdjęcia, które mogą państwu tutaj dzisiaj pokazać, ponieważ system w Stanach Zjednoczonych działa w taki sposób, że rzeczy wytworzone z pieniędzy publicznych, federalnych przez pracowników rządu federalnego nie podlegają prawu autorskiemu. Sekcja 105 Copyright Act z 1996 r. jasno stanowi, że dzieła, które oni tworzą nie podlegają prawu autorskiemu, więc można z nich korzystać na zasadzie domeny publicznej. Nawet oficjalna strona Białego Domu jasno mówi, że wszystkie ich treści są na domenie publicznej, a wszystkie inne treści, które udostępniają, są właśnie dostępne na licencji CC BY, na której nasz serwis był do 2022 r. Wszystkie strony rządu holenderskiego również udostępniają swoje treści na otwartej licencji, w tym przypadku na domenie publicznej. Podobnie strona prezydenta Bułgarii. Jest to starsza licencja CC BY 2.5, ale również licencja otwarta.

W 2019 r. Komisja Europejska wypuściła swój raport dotyczący zasad reuse, czyli ponownego wykorzystywania danych, w którym zobowiązała się, że wszystkie treści pisemne, które będą publikowały agendy europejskie, będą dostępne na licencji Creative Commons BY, czyli uznania autorstwa, a wszystkie pozostałe, prostsze dane, czyli przetworzone dane, metadane i inne dokumenty w tym charakterze będą dostępne na domenie publicznej.

W portalu gov.pl mamy co najmniej 339 tys. stron, które są odwiedzane miesięcznie 40 mln razy przez 7,5 mln użytkowników. Skąd to wiemy? Dane numeryczne, które teraz pokazuję, są właśnie z naszego pięknego portalu dane.gov.pl, o którym wspomniała pani pełnomocnik. Jest to portal, który właśnie sprawia, że jako Polska zajmujemy wysokie miejsca w rankingu Open Data Maturity. Jest to coś, z czego powinniśmy być dumni. Jak zobaczymy warunki wykorzystywania na przykład zbioru danych z portalu gov.pl, jest to licencja domeny publicznej, jest brak praw autorskich. Jest to coś pięknego. Więcej treści rządowych powinno być udostępnianych w ten sposób.

Jednym z problemów, które chciałbym zaadresować, stanowi fakt, że administracja rządowa oraz ministerstwa nie są konsekwentne w podejmowaniu decyzji dotyczących licencji, na jakiej udostępniają swoje treści. Tutaj mamy zdjęcia fotografów różnych ministerstw na portalu flickr. Ministerstwo Cyfryzacji udostępnia swoje zdjęcia na najbardziej restrykcyjnej licencji. Jest to all rights reserved. Ministerstwo Rodziny, Pracy i Polityki Społecznej udostępnia swoje zdjęcia na licencji CC BY-NC-ND, Ministerstwo Spraw Zagranicznych na licencji CC BY-NC, Ministerstwo Kultury i Dziedzictwa Narodowego w domenie publicznej. Nawet jeżeli korzystamy z polskich treści rządowych, jeżeli zgodnie z prawem chcemy dalej podzielić się multimediami, to bardzo często chcąc działać w dobrej wierze, musimy trzy razy sprawdzać, czy ministerstwo, na stronach którego działamy, wybrało licencję otwartą czy nie.

Jesteśmy tutaj dzisiaj, żeby porozmawiać, co dalej, jak można rozwiązać tę sytuację, żebyśmy w zakresie polskich otwartych danych radzili sobie jeszcze lepiej. Jesteśmy tutaj jako przedstawiciele sektora społecznego, żeby porozmawiać o współpracy i o tym, jak możemy uczynić polski internet bardziej dostępnym. Dziękuję bardzo.

Przewodnicząca poseł Paulina Matysiak (Razem):

Bardzo dziękuję. W tym miejscu otwieram dyskusję. Może korzystając z okazji, że jest z nami sekretarz stanu w Ministerstwie Cyfryzacji, pan Michał Gramatyka, poproszę też o komentarz, o pana opinię. Jeżeli pan minister ma wiedzę, co się zadziało w 2022 r., z czego wynika zmiana, to może poprosimy o wyjaśnienie. Jeżeli teraz jest to niemożliwe, to może poprosimy

o informację na piśmie do Sekretariatu Komisji, żeby przekazać ją posłom, którzy pracują w Komisji. Oddaję już głos panu ministrowi.

Sekretarz stanu w Ministerstwie Cyfryzacji Michał Gramatyka:

Bardzo dziękuję. Bardzo dziękuję za zaproszenie. Bardzo dziękuję za tę wypowiedź. W mojej ocenie jest oczywiste, że rządowe treści powinny być udostępniane na licencjach otwartych, wszystko to, co jest wytworzone za wspólne pieniądze, powinno być przeznaczone do wspólnego wykorzystania. Jestem zwolennikiem licencji otwartych. W zupełnie innym obszarze mojej działalności, czyli działalności muzycznej, jestem członkiem pierwszego zespołu, który wydał swoją płytę – było to 12 lat temu – właśnie na licencji Creative Commons w sposób umożliwiający jej pobieranie i odtwarzanie bez ponoszenia żadnych opłat, ale to oczywiście jest tylko jakaś tam informacja poboczna.

Niestety, w Polsce jest tak, że mamy do czynienia z dużą silosowością, w tym również z silosowością myślenia o licencjach. Jest to dla mnie nowa informacja, że zdjęcia, które są publikowane na flickerach rządowych, mają różne licencje. Bardzo dziękuję za przekazanie tej informacji. Jest to bardzo potrzebne, żeby się nad tym pochylić i zobaczyć, jakie są przyczyny. Nie znam przyczyn, które spowodowały, że w 2022 r. zamknięto dostęp, a w zasadzie nie tyle zamknięto dostęp, co go ograniczono. Zwracam uwagę państwa, że w dalszym ciągu mamy do czynienia z licencją Creative Commons. To nie jest tak, że jest ona zamknięta, tylko że nie są otwarte czy domyślnie dozwolone niektóre formy ponownego wykorzystania danych. Pochylimy się nad tym razem z Departamentem Zarządzania Danymi. Naszym celem jest otwieranie jak największej liczby danych.

Ostatnio pracujemy nad tym, żeby do serwisu otwartych danych trafiły na przykład informacje o polskiej muzyce rozrywkowej, trafiły informacje, które są udostępniane przez deweloperów na temat cen mieszkań. Wszystko to, co się da i wszystko to, co się nie da, chcemy właśnie udostępniać do ponownego wykorzystania, dlatego że właśnie w taki sposób powstały najlepsze biznesowe pomysły na wykorzystanie big data czy na znajdowanie jakichś nieoczywistych korelacji pomiędzy danymi. Myślę, że jest to przyszłość, że im więcej danych utworzymy, tym będzie lepiej dla wszystkich nas.

Bardzo dziękuję. Prezentację wydobędę sobie z materiałów Komisji. Od poniedziałku zaczynam się zajmować tym, żeby na dzień dobry wszystkie udostępniane zdjęcia były na licencjach otwartych. Będzie to najprostsze do załatwienia. Jest to po prostu kwestia komunikacji z wydziałami i odpowiedzialnymi za PR we wszystkich ministerstwach. Myślę, że całą resztę zrobimy później. Dziękuję.

Przewodnicząca poseł Paulina Matysiak (Razem):

Dziękuję. Dziękuję też za tę deklarację i chęć pracy oraz rozwiązywania problemów. Oczywiście to nie są tylko kwestie zdjęć. Przed rozpoczęciem posiedzenia podkomisji miałam okazję chwilę porozmawiać z paniami z Departamentu Zarządzania Danymi. Pamiętam dyskusję przy okazji wprowadzania ustawy o otwartych danych na posiedzeniu Komisji Cyfryzacji, Innowacyjności i Nowoczesnych Technologii. Wtedy mocno wybrzmiał, ja też zgłaszałam taki postulat i rozwiązanie, żeby na przykład dane, które gromadzi PKP PLK mogły być udostępniane po to, żeby ktoś, kto chce stworzyć aplikację i informować swoją grupę odbiorców o opóźnieniach, mógł to zrobić. W tym momencie dane te są absolutnie zamknięte. Nie można z nich skorzystać, a przecież to też są dane, za które my jako podatnicy płacimy. Jak rozumiem, idea i nasza dyskusja zmierza ku temu, że to, za co zapłaciliśmy, na co zostały wydane środki publiczne, pieniądze podatników, powinno być do wykorzystania także przez obywateli, żeby nie był to jakiś pakiet zamkniętych działań, rozwiązań, zdjęć, publikacji, raportów i wszystkiego, o czym możemy pomyśleć.

Otwieram dyskusję. Zachęcam do zgłaszania się do głosu. Pan minister, bardzo proszę.

Sekretarz stanu w MC Michał Gramatyka:

Dosyć dobrze rozumiemy się z panią przewodniczącą w zakresie potrzeby otwierania danych. Od wielu lat walczymy o to, żeby baza danych kodów pocztowych została otwarta do powtórznego przetwarzania, ale nie za bardzo udaje się nam przekonać Poczcie Polską. Nawet jak wydaje się nam, że sprawa jest już prawie doprowadzona do szczęśliwego finału, nagle pojawia się ktoś, kto nie do końca rozumiejąc wagę wyzwania, stwierdza, że baza danych może mieć wartość komercyjną. Jest to słowo klucz dla każdego prezesa, który bojąc się

odpowiedzialności za działanie na szkodę spółki, blokuje właśnie takie procesy. Im więcej nas będzie o to apelować, tym będzie lepiej.

Nie biorę żadnej odpowiedzialności za aktywa Skarbu Państwa, dlatego że bardzo często funkcjonują one w sobie tylko znany sposób, kierują się sobie tylko znanymi jakimiś wytycznymi. Wielokrotnie mamy problem z niezrozumieniem pewnych podstawowych potrzeb w zakresie przetwarzania danych, w zakresie udostępniania danych. Wielokrotnie trudno jest nam po prostu przekonać spółki prawa handlowego, żeby podchodziły do tego tematu podobnie jak my. To, co mogę zadeklarować, to kwestia uporządkowania sytuacji na płaszczyźnie rządowej, dlatego że to będzie proste. Myślę, że tutaj po prostu wystarczy opisać proces i wprowadzić go. Jest to do zrobienia. Bardziej skomplikowanymi sprawami zajmujemy się w następnej kolejności.

Przewodnicząca poseł Paulina Matysiak (Razem):

Jasne, bardzo dziękuję. Zachęcam do zgłaszania się do głosu. Czy państwo posłowie albo strona społeczna jest chętna? Proszę o podniesienie ręki, żebym mogła zapisać do głosu. Bardzo proszę.

Prezes zarządu, kierownik działu polityk i rzecznictwa Fundacji Centrum Cyfrowe Maria Drabczyk:

Bardzo dziękuję, pani przewodnicząca. Maria Drabczyk. Centrum Cyfrowe. Mamy też przyjemność być Creative Commons Polska, więc licencje są nam bardzo bliskie.

Bardzo dziękuję za wszystkie prezentacje. Bardzo dziękuję za deklarację, płytę ściagnę. Co dla nas jest tutaj kluczowe? Naszym celem jest doprowadzenie do tego, żeby z licencji, z całego wachlarza korzystać mądrze, oczywiście tam, gdzie jest to możliwe, też w maksymalnie otwarty sposób. To, co pan minister mówił o możliwościach biznesowych, też jest szalenie istotne w kontekście właśnie otwartości danych, otwartości danych rządowych. Faktycznie zrozumienie, skąd ta zmiana, byłoby supercenne. Myślę, że jest to też bardzo spójne z tym, co proponują państwo i co również dla nas było bardzo ważne i bardzo na tak w nowej strategii cyfryzacji. Jest tam duży fokus, duży nacisk na otwartość danych również w kontekstach komercyjnych i biznesowych, ich mądrego wykorzystania w tych kontekstach. Myślę, że potencjał dane.gov.pl chociażby w kontekście tworzenia czy wykorzystania danych do tworzenia modeli AI, polskich modeli AI, byłby supercenny.

To, co chciałabym tylko krótko zadeklarować, to, że my jesteśmy i bardzo chętnie wesprzemy ten proces jako centrum, jako Creative Commons Polska. A więc jeżeli możemy się do czegoś przydać, możemy jakoś pomóc, to bardzo chętnie.

Przewodnicząca poseł Paulina Matysiak (Razem):

Bardzo dziękuję. Czy są jakieś pytania?

Poseł Jacek Świat (PiS):

Chciałbym o coś dopytać.

Przewodnicząca poseł Paulina Matysiak (Razem):

Bardzo proszę, panie pośle.

Poseł Jacek Świat (PiS):

Ponieważ pan minister zwrócił uwagę na problemy ze spółkami Skarbu Państwa, akurat jako ekspocztowiec wiem, że baza kodów pocztowych może mieć dużą wartość komercyjną przy konstruowaniu baz adresowych dla roznoszenia, rozprowadzenia reklam bezadresowych. Nie wiem, jak to dzisiaj wygląda, gdyż rynek może jest trochę inny, ale wiem, że baza ta może mieć dużą wartość komercyjną. Poczta, udostępniając takie bazy w sposób zupełnie otwarty, właściwie budowałaby swoją konkurencję. A więc nie jest tak, że na poczcie ktoś zwariował, tylko rzeczywiście ma to swoje realne podstawy. Chciałbym też zapytać, czy w jakichś innych spółkach Skarbu Państwa jakieś obszary też budzą zastrzeżenia. Myślę, że tutaj nie da się tego uregulować jednym aktem. Chyba trzeba by za każdym razem negocjować i próbować oddzielnie wpłynąć na spółki. Czy spotkał się pan z jakimiś innymi takimi sytuacjami?

Sekretarz stanu w MC Michał Gramatyka:

Z przyjemnością odpowiem. Rozważania dotyczące kodów pocztowych zacząłbym od tego, skąd się w ogóle kody pocztowe wzięły. Jeżeli przyjmujemy, że kody pocztowe wzięły się stąd,

że państwo postanowiło, żeby ułatwić adresowanie przesyłek pocztowych i wprowadziło administracyjnie obowiązek przyznawania kodów pocztowych do poszczególnych miejscowości, to dochodzimy do wniosku, że cały system kodów pocztowych został stworzony za publiczne pieniądze. Moim zdaniem jest to baza i podstawa do jakiegokolwiek odpowiedzi w tym zakresie. Jeżeli został stworzony za publiczne pieniądze, to powinien być publicznie i bezpłatnie dostępny. Dla mnie dyskusja w tym momencie się kończy. Mam takie podejście. Pocztowców trudno jest przekonać do czegokolwiek, ale nie składamy broni, będziemy to robić.

Bardzo dziękuję również za zaadresowanie kwestii trenowania modeli językowych. To też jest niezwykle istotne. Jest społecznościowy projekt, który nazywa się Bielik. Jest to LLM, który jest tworzony przez społeczność w ogóle z pominięciem pieniędzy publicznych. Na przykład nie ma on dostępu do wielu danych, do których mógłby mieć dostęp. Dlatego na przykład nie wie, że Irena Santor z Pawłem Kukizem zaśpiewała piosenkę „Już nie ma dzikich plaż”, a powinien to wiedzieć, ponieważ informacja o tym, że taki duet jest gdzieś w cyberprzestrzeni muzycznej, jest informacją publiczną, gromadzoną przez Związek Autorów i Kompozytorów Scenicznych czy Związek Producentów Audio-Video. Informacje te też powinny być udostępniane. Jeżeli model byłby w stanie wytrenować się między innymi na tych informacjach, to taką wiedzę by posiadał, można by go było o to zapytać i odpowiadałby zgodnie z prawdą. Dzisiaj zgodnie z prawdą nie odpowie, a Chat GPT odpowie, dlatego że Chat GPT wprawdzie nie ma dostępu do tych danych, ale ma dostęp do całości internetu, nieco na tak zwany rympał i nieco na zasadzie brute force, ale jednak jest w stanie przeczesać go ze względu na to, że jest dużo bardziej doinwestowany, i udzielić poprawnej odpowiedzi.

Podobny problem w Polsce mają w zasadzie użytkownicy samochodów elektrycznych, którzy nie mają jednej bazy danych wszystkich ładowarek. Ściślej, mają ją, ale jest to baza stworzona przez społeczność na witrynie plugshare.com czy w aplikacji PlugShare. Jeżeli bazy ładowarek samochodowych będą szukać u korporacji energetycznych, które prowadzą swoje systemy ładowarek, to dostaną tylko pewien wycinek mapy. Ani nie ma udostępnionych informacji o ładowarkach, ani nie ma wystawionych danych, dzięki którym będziemy w stanie zrobić pełną mapę. Są to tego typu rzeczy, które w czysto użytkowy sposób ułatwiłyby życie wielu z nas, gdyby dane były otwarte.

Troszkę zmieniają to dyrektywy europejskie Data Act i Data Governance Act, dlatego że w swoich założeniach koncentrują się wokół tego, żeby udostępniać czy umożliwiać powtórne wykorzystywanie danych gromadzonych przez producentów, na przykład producentów samochodów albo producentów sprzętu, tak zwanego internetu rzeczy, inteligentnych opasek, zegarków, jakichś gadżetów, które rejestrują naszą aktywność. Zasada czy idea jest taka, żeby po anonimizacji dane były dostępne do powtórnego przetwarzania, ale nie wszyscy to rozumieją, a nawet jest bardzo niewiele osób, które rozumieją, że za chwilę to po prostu nadejdzie.

Jeżeli chodzi o przykłady stojące mocno nogami na ziemi, są to te, które podałem. Generalnie energetyka też jest dosyć zamkniętym systemem, do którego nie każdy ma dostęp, a każdy powinien mieć taki dostęp. Dzisiaj mamy do czynienia z często kuriozalnymi sytuacjami ze względu na popularność w Polsce prosumentów, produkowanie energii ze źródeł fotowoltaicznych. Dane dotyczące falowników czy sterowników z tych źródeł są wystawione dla producentów, najczęściej chińskich, a nie są wystawione dla korporacji energetycznych ani nie są wystawione publicznie. Może to spowodować jakieś duże zawirowania na rynku. Dzisiaj prosumenci stanowią kilka procent polskiego miksu energetycznego, dobrze, że nie kilkanaście, dlatego że wtedy mogłoby to rodzić jakieś zagrożenia. A więc to wszystko zaczyna i kończy się na filozofii udostępniania danych, eksponowania danych w taki sposób, żeby było możliwe ich powtórne użycie, czyli w zasadzie na temacie dzisiejszego posiedzenia podkomisji. Bardzo dziękuję.

Przewodnicząca poseł Paulina Matysiak (Razem):

Bardzo dziękuję. Jeszcze może trochę odpowiadając na pytania o spółki Skarbu Państwa, podam konkretny przykład, który wybrzmiewał już w mojej wcześniejszej wypowiedzi. Jeżeli chodzi o PKP PLK, kilka lat temu właśnie przez to, że spółka zablokowała dostęp do swoich danych, zniknęła z rynku popularna aplikacja „Mój Pociąg”, z której korzystało wielu podróżnych. Była bezpłatna i informowała o opóźnieniach pociągów. Autor aplikacji,

którzy ją przygotowali, korzystał z danych, które pobierał z portalu spółki PLK, która właśnie publikowała te dane. Spółka po prostu zablokowała dane. Natomiast są też przykłady zakładów komunikacji miejskiej, które na przykład udostępniają dane o położeniu swoich autobusów czy tramwajów. Udostępniają też w sposób otwarty dane na temat rozkładu jazdy, opóźnień i tego, gdzie znajdują się dane pojazdy, co umożliwia stworzenie komercyjnych wyszukiwarek czy innych aplikacji, które będą wygodne z perspektywy pasażera. Po prostu ma on wybór, jeżeli chodzi o sklepy internetowe z aplikacjami w swoim telefonie. PKP PLK od dobrych kilku lat – zamknięcie aplikacji nastąpiło w 2017 r. – stoi na stanowisku, że nie będzie udostępniać danych, nie będzie się nimi dzielić. Jest to przykład działania, które mogłoby zostać zmienione. Na pewno nie byłoby to ze szkodą dla spółki, a byłoby pewnie z korzyścią dla podróżnych, dla pasażerów.

Podobne kwestie są poruszane w kontekście stworzenia chociażby wyszukiwarki połączeń autobusowych. O ile dane, informacje o połączeniach autobusowych, także tych, które stanowią publiczny transport zbiorowy, znajdziemy na danym przystanku autobusowym, może w internecie na stronie przewoźnika, to nie ma ich zagregowanych. Warto tu określić, że Ministerstwo Infrastruktury pracuje nad wytycznymi, żeby dane te, po pierwsze, można było zbierać w określony sposób. A więc mają być wytyczne, w jakiej dokładnie formie przewoźnik, organizator ma je przekazywać. Dzięki temu będzie szansa na to, żeby stworzyć krajową wyszukiwarkę połączeń autobusowych w całym kraju. Wyszukując połączenie pociągiem nie zastanawiamy się, możemy w jednym miejscu wyszukać połączenie każdego przewoźnika, który jeździ na terenie Polski. W przypadku przewoźników autobusowych jest to niemożliwe. A więc jak rozumiem, podaję przykład, gdzie nastąpiłyby realne korzyści z perspektywy obywateli. Otwartość danych może wpływać bardzo korzystnie na jasny i sprawny dostęp do informacji.

Czy ktoś jeszcze chciałby zabrać głos, ma jakieś pytania, chciałby jeszcze podzielić się jakąś refleksją? Bardzo proszę, pan Nadzikiewicz.

Członek zarządu Wikimedia Europe Maciej Artur Nadzikiewicz:

Dziękuję bardzo. Chciałbym podziękować panu ministrowi za deklarację współpracy przy otwieraniu danych. Podobnie jak pani Maria Drabczyk, my również pozostajemy otwarci na współpracę czy z ministerstwem, czy z innymi ministerstwami, żeby pomóc uwolnić dane. Mówimy tutaj o bardzo dużym potencjale, o poprawie dostępu do wiedzy o tym, jak działa państwo oraz o sprawieniu, że w sytuacji, w której ktoś w dobrej wierze chce dalej rozposzechnić informacje państwowe, będzie to robił zgodnie z prawem, a nie, zacytuję, na rympał, jak powiedział pan minister. Jeżeli będzie potrzeba, to jak najbardziej jesteśmy do dyspozycji. Rozumiem, że takim razem we wtorek będziemy już mogli legalnie udostępnianie zdjęcia fotografów polskich ministerstw.

Przewodnicząca poseł Paulina Matysiak (Razem):

Spokojnie dajmy czas panu ministrowi na rozpoczęcie działań, natomiast trzymamy kciuki. Mamy nadzieję, że efekty zobaczymy do końca pierwszego kwartału tego roku. Może tak się umówmy. Jest to chyba realny termin na komunikację z innymi resortami.

Członek zarządu Wikimedia Europe Maciej Artur Nadzikiewicz:

To był żart. Jak najbardziej jestem bardzo cierpliwym człowiekiem. Zdaję sobie sprawę, że nie jest to najważniejsza sprawa, która przed nami stoi, ale mam nadzieję, że w najbliższym czasie wszystko nam pomoże i w dalszym czasie będziemy mogli zajmować się kolejnymi dotyczącymi tego sprawami. Bardzo dziękuję za deklarację. Dziękuję za pytania i za możliwość przedstawienia stanowiska. Dziękuję.

Przewodnicząca poseł Paulina Matysiak (Razem):

Ja też dziękuję za wypowiedzi, za wszystkie głosy, za przedstawioną informację. Zamykam dyskusję. Poszło nam bardzo gładko. W dodatku mamy zapewnienie ze strony ministerstwa, że część postulatów zostanie zrealizowanych. Życzę nam wszystkim, żeby każde posiedzenie podkomisji i Komisji właśnie tak wyglądało.

Dziękuję państwu za przybycie. Dziękuję panu ministrowi.

Na tym wyczerpaliśmy porządek dzienny posiedzenia podkomisji.

Tym samym zamykam posiedzenie podkomisji. Do widzenia.