

Polityka otwartego dostępu do danych badawczych finansowanych ze środków publicznych



Ministerstwo Nauki
i Szkolnictwa Wyższego

**Polityka otwartego dostępu do danych badawczych
finansowanych ze środków publicznych**

Ministerstwo Nauki i Szkolnictwa Wyższego

Warszawa, lipiec 2025

Spis treści

Wprowadzenie	4
Odbiorcy Polityki i interesariusze	5
Kontekst międzynarodowy i krajowy	5
Dotychczasowe działania ministra właściwego ds. szkolnictwa wyższego i nauki	6
Otwarty dostęp do danych badawczych w Polsce – aktualny stan	7
Mocne strony	8
Słabe strony	8
Szanse	9
Zagrożenia	10
Cele Polityki	10
Cele szczegółowe Polityki	12
1. Zapewnienie środowiska prawnego umożliwiającego udostępnienie i ponowne wykorzystanie danych badawczych, przy zapewnieniu adekwatnej ochrony prawnej uprawnionych do danych	12
2. Wskazanie ram organizacyjnych w zakresie zarządzania otwartymi danymi badawczymi na poziomie krajowym (krajowa Sieć Centrów Kompetencji Otwartych Danych Badawczych)	15
3. Ustalenie, przy uwzględnieniu kontekstu międzynarodowego, wymaganych standardów udostępniania danych badawczych, odpowiadających specyfice poszczególnych obszarów badawczych	17
4. Zapewnienie na poziomie krajowym ekosystemu informatycznego umożliwiającego przechowywanie, odnalezienie, uzyskanie dostępu i ponowne wykorzystywanie rozproszonych zasobów danych w każdym obszarze badawczym	18
Dalsze kroki i ewaluacja	21
Podsumowanie	21
Słowniczek wybranych pojęć i skrótów	24

Wprowadzenie

Dane są fundamentalnym elementem nauki, a ich wartość jest współcześnie znacząco zwiększona dzięki otwartemu dostępowi. Otwarte dane badawcze zwiększają widoczność i cytowalność naukowców, co przekłada się na lepsze możliwości współpracy zarówno na szczeblu krajowym jak i międzynarodowym. Poprawiają one także przejrzystość procesu badawczego, umożliwiając weryfikację rzetelności naukowej realizowanych projektów oraz uwiarygadniając wyniki prezentowane w publikacjach naukowych. Dostęp do danych podnosi efektywność nauki, eliminując konieczność powtarzania badań w ramach realizacji nowych scenariuszy badawczych. W ten sposób otwieranie danych i ich ponowne wykorzystywanie umożliwia bardziej racjonalne wydatkowanie środków publicznych. Ponadto, dane badawcze są również wykorzystywane w procesach decyzyjnych w organizacjach spoza publicznego sektora szkolnictwa wyższego i nauki¹, co zwiększa racjonalność i efektywność podejmowanych decyzji. Z tych względów udostępnianie danych leży w interesie całego środowiska naukowego i jego otoczenia.

Kwestie podejmowane w tym dokumencie są szczególnie istotne w kontekście finansowania publicznego. Należy unikać podwójnego finansowania zbliżonych zadań badawczych, których realizacja wiąże się z generowaniem danych, a jednocześnie utworzyć powszechnie dostępny zbiór danych, z którego mogą korzystać inni naukowcy, innowacyjni przedsiębiorcy oraz zaangażowani naukowo obywatele.

Niniejszy dokument opisuje kontekst rozwoju otwartej nauki w Polsce oraz podsumowuje dotychczasowe działania ministra właściwego do spraw szkolnictwa wyższego i nauki. Zawiera on zwięzłą analizę obecnego stanu otwartej nauki w Polsce oraz przedstawia cel główny Polityki, wraz z celami szczegółowymi, których realizacja ma pomóc w jego osiągnięciu. W podsumowaniu znajduje się informacja o kolejnym dokumencie – planie działań, który ma skonkretyzować sposób realizacji poszczególnych zadań, ze wskazaniem odpowiedniego finansowania. Dokument zawiera również słowniczek istotnych pojęć i skrótów pojawiających się w tekście Polityki w kontekście otwartej nauki.

Dokument przygotowano we współpracy z ekspertami z dwóch zespołów, powołanych przez Ministra Nauki: Zespołu doradczego do spraw otwartych danych naukowych (Dz. Urz. Min. Nauki z 2021, poz. 58) oraz Zespołu doradczego do spraw Polityki otwartego dostępu do danych badawczych finansowanych ze środków publicznych (Dz. Urz. Min. Nauki z 2024, poz. 26).

¹ Art. 7 ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2024 r. poz. 1571 z późn. zm.) określa podmioty tworzące system szkolnictwa wyższego i nauki jako: uczelnie, federacje podmiotów systemu szkolnictwa wyższego i nauki, Polską Akademię Nauk, instytuty naukowe PAN, instytuty badawcze, międzynarodowe instytuty naukowe utworzone na podstawie odrębnych ustaw, Centrum Łukasiewicz, instytuty działające w ramach Sieci Badawczej Łukasiewicz, Polską Akademię Umiejętności, inne podmioty prowadzące głównie działalność naukową w sposób samodzielny i ciągły.

Odbiorcy Polityki i interesariusze

Polityka otwartego dostępu do danych badawczych jest skierowana do podmiotów sektora nauki i szkolnictwa wyższego, które dysponują danymi badawczymi wytworzonymi dzięki finansowaniu ze środków publicznych. Biorąc pod uwagę zakres dokumentu, będzie on miał bezpośredni wpływ także na agencje działające na rzecz sektora szkolnictwa wyższego i nauki, takie jak Narodowe Centrum Nauki (NCN), Narodowe Centrum Badań i Rozwoju (NCBR), Agencja Badań Medycznych (ABM), Narodowa Agencja Wymiany Akademickiej (NAWA) oraz Fundacja na rzecz Nauki Polskiej (FNP), szczególnie w zakresie finansowania badań naukowych, prac rozwojowych i twórczości artystycznej. Polityka będzie miała również wpływ na naukowców, podmioty sektora prywatnego, podmioty sektora edukacji działające w obszarze nauki oraz uczestników tzw. nauki obywatelskiej.

Kontekst międzynarodowy i krajowy

Działania na rzecz otwartej nauki są prowadzone od kilku dekad przez organizacje międzynarodowe oraz instytucje Unii Europejskiej. Niniejsza Polityka dąży do zapewnienia spójności ekosystemu polskiej nauki z tymi inicjatywami uwzględniając rekomendacje Unii Europejskiej², OECD³ i UNESCO⁴ oraz działania takich organizacji jak Science Europe, cOAlition S, EOSC Association czy Europejska Rada ds. Badań Naukowych (ERC, ang. *European Research Council*). Warto także podkreślić, że wiele państw Unii Europejskiej wdrożyło polityki lub strategie dotyczące otwartych danych i/lub otwartej nauki, w tym m.in: Niemcy, Francja, Holandia, Portugalia, Słowacja, Szwajcaria, Finlandia.

Szczególny kontekst dla niniejszej Polityki wyznacza rozwój Europejskiej Chmury Otwartej Nauki (EOSC, ang. *European Open Science Cloud*), opierając się na zasadach FAIR, wspiera wyszukiwalność, dostępność, interoperacyjność i ponowne wykorzystanie danych naukowych.

EOSC jest wdrażany poprzez partnerstwo pomiędzy Komisją Europejską a EOSC Association. Członkami EOSC Association są m.in. Narodowe Centrum Nauki jako tzw. krajowa organizacja przedstawicielska (ang. *mandated organization*) oraz kilka innych polskich instytucji naukowo-badawczych, które oferują zaawansowane usługi w obszarze infrastruktury otwartej nauki. W ramach uczestnictwa w EOSC, koordynowana jest krajowa współpraca pomiędzy

² Commission Recommendation on Access to and Preservation on Scientific Information, 2018; Directive no 2019/1024 of the EP and of the Council (EU) on Open Data and the Re-use of Public Sector Information, 2019; EC Communication on the New Era, 2020; Council Conclusions on the Future Governance of the ERA including the ERA Policy Agenda, 2021; Council Conclusions on Research Assessment and Implementations of Open Science, 2022; European Strategy for Data, 2022; Regulation of the European Parliament and of the Council on European data governance (Data Governance Act), 2022; Regulation of the European Parliament and of the Council on European data governance (Data Governance Act), 2022.

³ OECD Revised Recommendation of the Council on Access to Research Data from Public Funding, 2021, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347>.

⁴ UNESCO Recommendation on Open Science, 2021, <https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en>.

zainteresowanymi instytucjami, m.in. w ramach Sieci EOSC Polska⁵. Polskie podmioty biorą udział w projektach typu INFRAEOSC, finansowanych przez Komisję Europejską w ramach programów Horyzont 2020 i Horyzont Europa, oraz współtworzą systemy i narzędzia EOSC. Na poziomie krajowym, Polityka stanowi realizację obowiązku nałożonego na ministra właściwego do spraw szkolnictwa wyższego i nauki przez art. 23 ust. 1 ustawy z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego (Dz. U. z 2023 poz. 1524). Zgodnie z delegacją zawartą w ww. ustawie, Polityka koncentruje się na otwartym dostępie do danych badawczych. Jest to również pierwszy etap budowy krajowej polityki na rzecz długoterminowego rozwoju otwartej nauki w Polsce.

Otwarta nauka, w tym otwieranie danych badawczych, jest jednym z priorytetów wymienionych w Polityce Naukowej Państwa⁶ (PNP) w kontekście podnoszenia jakości realizowanych badań. PNP zwraca uwagę na potrzebę zachęcania polskich naukowców do publikowania wyników swoich badań w światowym obiegu nauki, oraz na konieczność efektywnego stosowania systemu własności intelektualnej.

Niniejsza Polityka uwzględnia również zapisy zawarte w Programie Otwierania Danych na lata 2021-2027 w zakresie dotyczącym danych badawczych. W szczególności istotne są wyszczególnione w Programie filary otwartości danych określające, że dane powinny być: dostępne, upublicznione w wersji źródłowej, kompletne, aktualne, odczytywalne maszynowo, udostępnione w sposób niedyskryminujący, dostępne bez ograniczeń licencyjnych oraz niezastrzeżone.

Dotychczasowe działania ministra właściwego ds. szkolnictwa wyższego i nauki

Działania ministra właściwego ds. szkolnictwa wyższego i nauki wspierające otwarte modele komunikacji naukowej rozpoczęły się w 2004 roku. Kluczowy dokument, „Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce”⁷, został wdrożony w 2015 roku. Dokument ten zdefiniował główne pojęcia i zalecenia dla interesariuszy, w tym podmiotów finansujących badania, uczelni i wydawców czasopism naukowych. Ustawa Prawo o szkolnictwie wyższym i nauce z 2018 roku określiła ramy prawne dla wdrażania otwartej nauki, wprowadzając trzecie kryterium ewaluacji wpływu działalności naukowej na społeczeństwo i gospodarkę.

Ministerstwo monitoruje efekty działań w zakresie otwartej nauki:

⁵ Więcej informacji: eosc.gov.pl.

⁶ <https://www.gov.pl/attachment/c47d37f9-eaab-4701-adce-84cf7585013e>

⁷ Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce, https://www.gov.pl/documents/1068557/1069061/20180413_Kierunki_rozwoju_OD_wersja_ostateczna.pdf

2017 - przeprowadzono badanie ankietowe uczelni, które sprawdziło poziom realizacji zaleceń zawartych w „Kierunki rozwoju otwartego dostępu do publikacji i wyników badań naukowych w Polsce”;

2018 - sporządzono „Raport nt. realizacji polityki otwartego dostępu do publikacji naukowych w latach 2015-2017”, zawierający przegląd działań, omówienie problemów i rekomendacje;

2020 - przeprowadzono kolejne badanie ankietowe, które obejmowało współpracę międzynarodową i interfejsy API;

2023 - uruchomiono szeroko zakrojoną analizę, obejmującą badanie ankietowe i wywiady pogłębione, oceniającą zaangażowanie podmiotów naukowych w latach 2017-2021 w otwartą naukę.

Badania przeprowadzone przez MNiSW/MEiN w latach 2017 i 2020 wyraźnie wskazały ryzyko prawne oraz brak znajomości ram prawnych jako jedne z głównych barier dla wdrażania instytucjonalnych polityk otwartego dostępu. Zapewnienie jednoznacznej ochrony własności intelektualnej jest kluczowe dla skutecznej realizacji otwartych modeli komunikacji naukowej.

Pomoże to zwiększyć poczucie bezpieczeństwa badaczy i zniwelować obawy przed utratą korzyści płynących z eksploatacji wyników własnych badań. Ważne jest również, aby dane badawcze były otwierane w sposób zgodny z prawem i etyką badań naukowych.

Otwarty dostęp do danych badawczych w Polsce – aktualny stan

Otwarta nauka w Polsce zyskuje na znaczeniu dzięki różnym inicjatywom i projektom. Badania z 2023 roku pokazują, że około 27% ankietowanych instytucji wprowadziło politykę otwartego dostępu do danych badawczych, a połowa pozostałych planuje jej wdrożenie. Finansowanie działań związanych z otwartym dostępem pochodzi głównie ze środków własnych uczelni. Dane badawcze są najczęściej udostępniane w repozytoriach cyfrowych, przy czym około 35% instytucji posiada własne repozytoria, a 61 z 102 repozytoriów umożliwia deponowanie danych badawczych.

Uczelnie coraz częściej zatrudniają specjalistów z obszaru otwartej nauki. Są również udostępniane odpowiednie szkolenia i warsztaty. Do głównych przeszkód w tworzeniu polityk otwartej nauki w Polsce należą z kolei priorytety finansowe oraz brak wiedzy na temat praw autorskich⁸.

W zakresie otwartych danych badawczych w Polsce oferowane są szkolenia online, takie jak webinaria dziedzinowe czy też kursy typu MOOC (ang. *Massive Open Online Courses*) dla naukowców i data stewardów⁹. Rozwijana jest również współpraca między data stewardami

⁸ Ośrodek Przetwarzania Informacji Państwowy Instytut Badawczy, *Analiza stanu otwartej nauki w Polsce*, 2023, <https://radon.nauka.gov.pl/analizy/analiza-stanu-otwartej-nauki-w-Polsce>.

⁹ Szkolenia z zakresu *data stewardship*, poziom podstawowy; Szkolenia z zakresu *data stewardship*, poziom średnio zaawansowany; Szkolenia z zakresu rozwijania wiedzy i kompetencji naukowców i doktorantów

zatrudnionymi w polskich jednostkach naukowo-badawczych. Organizowane są wydarzenia promujące otwartą naukę, oraz dyskusje na temat jej wdrażania i rozwoju w kraju. Wzmacniana i rozszerzana jest także współpraca między krajowymi jednostkami a międzynarodowymi sieciami i organizacjami zaangażowanymi w otwartą naukę.

Aktualny stan otwartej nauki w Polsce wraz z jej mocnymi i słabymi stronami, a także pojawiającymi się szansami i zagrożeniami, został przedstawiony w poniższej analizie SWOT.

Mocne strony

- Istniejące zasoby i zaangażowanie specjalistów zajmujących się zarządzaniem danymi badawczymi;
- Aktywność Polski w międzynarodowych inicjatywach dotyczących otwartej nauki np. EOSC, CoNOSC, Science Europe, cOAlition S, Research Data Alliance (RDA), GO FAIR, OpenAIRE, The Declaration on Research Assessment (DORA), Coalition for Advancing Research Assessment (CoARA);
- Udział w projektach INFRAEOSC finansowanych przez Komisję Europejską;
- Istniejące infrastruktury badawcze, zarówno dziedzinowe jak i ogólnego przeznaczenia, w tym centra Komputerów Dużej Mocy (KDM), Krajowy Magazyn Danych (KMD) oraz sieć PIONIER;
- Istniejące narzędzia operacyjne (infrastruktura PL-GRID, serwisy sieci PIONIER, usługi KMD);
- Projekty na Polskiej Mapie Infrastruktury Badawczej finansowane ze środków polityki spójności UE;
- Rozwój standardów dotyczących planów zarządzania danymi (DMP, ang. *Data Management Plan*) zgodnie z zasadami FAIR;
- Polityki instytucjonalne dotyczące otwartej nauki posiadane przez wybrane agencje finansujące działalność naukową i wiodące jednostki naukowe;
- Kursy MOOC dotyczące zarządzania danymi badawczymi, realizowane na zlecenie MNiSW przez NCN;
- Istniejące repozytoria zarządzane przez podmioty krajowe;
- Zdefiniowanie charakterystyk drugiego stopnia efektów uczenia się dla kwalifikacji na 8. poziomie Polskiej Ramy Kwalifikacji obejmujące zasady otwartego dostępu do wyników badań naukowych.

Słabe strony

- Brak uporządkowania i klaryfikacji przepisów dotyczących udostępniania danych badawczych zgodnie z zasadami FAIR;

w obszarze zarządzania danymi badawczymi, poziom podstawowy; Szkolenia z zakresu rozwijania wiedzy i kompetencji naukowców i doktorantów w obszarze zarządzania danymi badawczymi, poziom średnio zaawansowany.

- Niska świadomość naukowców w zakresie obowiązków dotyczących zarządzania danymi badawczymi i ich znaczenia dla działalności naukowej oraz wynikająca z tego niska aktywność;
- Niski poziom wiedzy i kompetencji w zakresie otwierania danych według zasad FAIR;
- Nierówny poziom zaangażowania uczelni w działania związane z otwartą nauką;
- Zróżnicowane zaangażowanie władz uczelni we wdrażanie odpowiednich strategii;
- Ograniczenia prawne wpływające na proces otwierania danych na poziomie krajowym i międzynarodowym;
- Ograniczenia finansowe i organizacyjne dotyczące infrastruktury do przechowywania i otwierania danych badawczych;
- Niewystarczająca dynamika rozwoju kompetencji osób odpowiedzialnych za zarządzanie danymi;
- Niska dostępność szkoleń w zakresie zarządzania danymi;
- Niski poziom wiedzy i kompetencji środowiska naukowego w zakresie wykorzystania istniejących infrastruktur i narzędzi;
- Obawy naukowców dotyczące skutków udostępniania danych wynikające z braku wiedzy w zakresie praw autorskich;
- Ograniczona liczba certyfikowanych repozytoriów funkcjonujących w oparciu o międzynarodowe standardy;
- Niewystarczająca interoperacyjność i widoczność niektórych krajowych repozytoriów;
- Niski poziom wykorzystania istniejących danych badawczych przez naukowców;
- Brak systemu zachęt do otwierania zasobów danych.

Szanse

- Korzystne otoczenie międzynarodowe umożliwiające zaangażowanie polskich zespołów badawczych w prace prowadzone przez zagraniczne jednostki naukowe oraz prace na zagranicznych infrastrukturach badawczych, jak również udział zagranicznych i międzynarodowych zespołów badawczych w prace prowadzone przez polskie jednostki naukowe oraz prace na polskich infrastrukturach badawczych;
- Udział polskich naukowców w międzynarodowych programach finansowania badań naukowych, wymagających zarządzania danymi badawczymi zgodnie z zasadami FAIR (np. Horyzont Europa);
- Rozwój narzędzi wspierających zarządzanie danymi badawczymi w krajowych i międzynarodowych projektach (np. w ramach EOSC, DMP Creator Argos);
- Prawne rozwiązania związane z wykorzystywaniem danych publicznych oraz ewaluacją jakości działalności naukowej (możliwość wykazania pozytywnego wpływu dostępu do danych badawczych na społeczeństwo i gospodarkę);
- Wymogi czasopism naukowych, światowych konferencji naukowych czy też środowisk naukowych dotyczące udostępniania danych powiązanych z artykułami naukowymi;
- Zaangażowanie polskich zespołów badawczych w działania mające na celu otwieranie danych badawczych, umożliwiające transfer wiedzy oraz rozwój dobrych praktyk i polskich zasobów;

- Spodziewane efekty działań edukacyjnych prowadzonych w kraju (np. szkolenia dziedzinowe, kolejne edycje kursów MOOC na platformie NAVOICA);
- Rozwój federacyjnego środowiska EOSC na poziomie krajowym i europejskim.

Zagrożenia

- Brak nowych zasobów kadrowych i przeciążenie istniejących (w świetle rosnącego zapotrzebowania) oraz ich negatywny wpływ na warunki pracy naukowców (konieczność przejęcia obowiązków związanych z zarządzaniem danymi);
- Nasilenie się wątpliwości dotyczących wpływu otwierania danych na prawa osobiste naukowców, z powodu ograniczonego dostępu do szkoleń;
- Ryzyko pogorszenia sytuacji finansowej jednostek naukowych w wyniku niekorzystnej sytuacji makroekonomicznej w kraju i za granicą, w tym ograniczenie finansowania sektora nauki.

Cele Polityki

Podstawowym celem Polityki jest wskazanie ram zapewniających bezpieczeństwo i dostępność danych badawczych wytworzonych w ramach badań finansowanych lub współfinansowanych ze środków publicznych¹⁰. W szczególności dotyczy to odpowiednich mechanizmów magazynowania danych i ich udostępniania zgodnie z zasadą „otwarte na ile to możliwe i zamknięte na ile to konieczne” (ang. *as open as possible, as closed as necessary*)¹¹.

Polityka ma pomóc doprowadzić do sytuacji, w której otwarty dostęp do danych badawczych finansowanych ze środków publicznych stanie się standardem. Należy w związku z tym pamiętać, że obecnie nie wszystkie badania finansowane ze środków publicznych są realizowane w ramach grantów przyznawanych w trybie konkursów przez agencje finansujące badania. Uczelnie realizują działalność naukową przede wszystkim ze środków publicznych w postaci subwencji oraz innych ministerialnych programów. Programy te, tam gdzie nie ma przesłanek wyłączeniowych, zostaną objęte wymogiem zapewnienia otwartego dostępu do danych badawczych. Proces ten będzie następował stopniowo, w miarę przyjmowania standardów opisanych w celu szczegółowym 3.

Należy jeszcze raz podkreślić, że Ministerstwo uwzględni w niniejszej Polityce uzasadnione sytuacje, które uniemożliwiają otwarte publikowanie danych.

¹⁰ Rekomendowane jest również, aby w przypadkach, w których nie godzi to w interes przedsiębiorstwa, dane były otwierane również przez podmioty prywatne (np. na portalu dane.gov.pl).

¹¹ Istnieją liczne przypadki, w których otwieranie danych nie jest wskazane ani możliwe. Mogą wystąpić sytuacje, w których musimy uznać ograniczenie dostępu (ang. opt-out criteria) lub embargo (opóźniony czas otwarcia danych). Należą do nich ograniczenia jawności danych ze względu na konieczność ochrony prawa własności intelektualnej, w tym ochrony patentowej czy tajemnicy przedsiębiorstwa, względy bezpieczeństwa narodowego i dobra publicznego, brak możliwości anonimizacji danych osobowych, brak zgody uczestników badania, brak zgody uprawnionego do danych (np. w przypadku prywatnych zbiorów) lub embargo zawarte z w umowie z partnerem, a także kwestie etyczne. Ograniczenia takie muszą być rzetelnie uzasadnione.

Cel ten zostanie osiągnięty poprzez realizację następujących celów szczegółowych:

1. Zapewnienie środowiska prawnego umożliwiającego udostępnienie i ponowne wykorzystanie danych badawczych, przy zapewnieniu adekwatnej ochrony prawnej uprawnionych do danych.
2. Wskazanie ram organizacyjnych w zakresie zarządzania otwartymi danymi badawczymi na poziomie krajowym (krajowa sieć kompetencji).
3. Ustalenie, przy uwzględnieniu kontekstu międzynarodowego, wymaganych standardów udostępniania danych, odpowiadających specyfice poszczególnych obszarów badawczych¹².
4. Zapewnienie na poziomie krajowym ekosystemu informatycznego umożliwiającego przechowywanie, odnalezienie, uzyskanie dostępu i ponowne wykorzystywanie rozproszonych zasobów danych badawczych.

Realizacja powyższych celów Polityki przyniesie polskiej nauce, gospodarce i społeczeństwu szereg długoterminowych korzyści:

Korzyści dla rozwoju nauki:

- Wzmacnianie wysokiej jakości badań: podniesienie standardów badawczych, systemu ewaluacji/oceny naukowej oraz promowanie rzetelności naukowej;
- Podniesienie rzetelności i efektywności procesu badawczego: lepsze zarządzanie danymi badawczymi i optymalizacja metod badawczych;
- Zwiększenie transparentności procesu badawczego: ułatwienie dostępu do danych i metodologii badawczej;
- Optymalizacja wykorzystania danych badawczych: lepsze wykorzystanie istniejących danych, zmniejszenie duplikacji badań; możliwość przeprowadzania analiz opartych o unikalne dane, których nie można z różnych przyczyn ponownie zebrać;
- Umożliwienie lepszej replikowalności¹³ i powtarzalności¹⁴ badań: ułatwienie odtwarzania wyników przez innych badaczy;
- Rozwój współpracy między naukowcami, w tym o charakterze interdyscyplinarnym: promowanie współpracy między różnymi dziedzinami nauki oraz nawiązywanie współpracy, nie tylko w ramach środowiska naukowego, ale również i poza nim;
- Wzmocnienie (międzynarodowej) rozpoznawalności polskiej nauki: zwiększenie udziału polskich naukowców w międzynarodowych projektach badawczych;
- Wzmocnienie wpływu badań realizowanych przez naukowców afiliowanych w polskich jednostkach naukowo-badawczych: lepsze wykorzystanie wyników badań w praktyce,

¹² Pojęcie obszaru badawczego rozwinęto w opisie celu szczegółowego 2.

¹³ Replikowalność (ang. *Replicability*) - odnosi się do uzyskania tożsamyh wyników eksperymentu przy użyciu tych samych metod, ale przy użyciu innych danych. Nie wymaga to dostępu do oryginalnych danych, ale wymaga opisu metody badawczej, tak aby inny badacz mógł zastosować te same metody na innym zestawie danych i uzyskać podobne wyniki.

¹⁴ Powtarzalność (ang. *Reproducibility*) - zdolność do dokładnego powtórzenia/odtworzenia wyników eksperymentu przy użyciu tych samych metod, danych i warunków, które zostały użyte pierwotnie.

wzrost liczby cytowań zarówno samych danych badawczych, jak i publikacji z nimi powiązanych;

- Wzrost zaufania społecznego do nauki: budowanie wiarygodności wyników badań i transparentności;
- Rozwój nauki obywatelskiej (citizen science) jako nowy wymiar budowania społeczeństwa obywatelskiego, a zarazem najdojrzalsza forma dialogu między nauką a obywatelami i lokalnymi społecznościami;
- Dostosowanie do wymogów technologii AI: przygotowanie danych badawczych do analizy przy użyciu zaawansowanych narzędzi technologicznych;
- Wzmocnienie wiarygodności wyników badań naukowych: poprawa jakości i przejrzystości publikowanych wyników.

Korzyści dla gospodarki:

- Wzmocnienie przedsiębiorczości i innowacyjności społeczno-gospodarczej: ułatwienie dostępu do danych badawczych dla sektora prywatnego;
- Rozwój regionów w paradygmacie *smart region* (np. *smart city*): wykorzystanie danych badawczych do rozwoju inteligentnych rozwiązań miejskich;
- Tworzenie społeczeństwa i gospodarki opartej na wiedzy: promowanie edukacji i innowacji opartej na rzetelnych danych;
- Zwiększenie zaufania do wyników badań naukowych i polepszenie warunków ich wykorzystania: ułatwienie praktycznego zastosowania wyników badań;
- Zwiększenie dostępności do wyników badań naukowych dla otoczenia społeczno-gospodarczego: poprawa dostępności danych dla wszystkich zainteresowanych stron.

Korzyści dla społeczeństwa:

- Przeciwdziałanie dezinformacji: poprawa dostępu do rzetelnych i sprawdzonych informacji;
- Rozwój nauki obywatelskiej: angażowanie obywateli w procesy badawcze;
- Włączanie społeczeństw w aktywne szukanie rozwiązań współczesnych problemów cywilizacyjnych: zachęcanie do udziału w inicjatywach badawczych;
- Wsparcie dla opracowywania interdyscyplinarnych i multidyscyplinarnych programów mających na celu sprostanie nowym wyzwaniom społecznym i globalnym, promowanie programów badawczych o szerokim zakresie tematycznym.

Cele szczegółowe Polityki

1. Zapewnienie środowiska prawnego umożliwiającego udostępnienie i ponowne wykorzystanie danych badawczych, przy zapewnieniu adekwatnej ochrony prawnej uprawnionych do danych.

Działania na rzecz realizacji celu:

1.a) Przeprowadzenie analizy za pomocą dostępnych narzędzi, w tym zintegrowanej platformy analitycznej¹⁵ i sporządzenie raportu dotyczącego obecnego stanu prawnego w udostępnianiu danych badawczych oraz niezbędnych zmian legislacyjnych wspierających otwieranie i ponowne wykorzystywanie tych danych wraz z propozycją zmian w odpowiednich aktach prawnych.

Polski system prawny nie reguluje w wystarczającym stopniu kwestii zabezpieczenia praw autorskich w obszarze otwierania danych badawczych. Osiągnięcie sytuacji, w której udostępnianie danych badawczych stanie się standardem będzie możliwe pod warunkiem, że naukowcy będą mieli pewność zabezpieczenia ich praw w należyty i przejrzysty sposób w krajowych aktach normatywnych. Wstępna analiza sugeruje, że zmiany muszą być wprowadzone m.in. w:

- Ustawie z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych:
 - doprecyzowanie zakresu uprawnień pracodawcy w odniesieniu do danych oraz zawartego w nim materiału naukowego;
 - uregulowanie kwestii danych badawczych powstających na uczelniach i instytutach poza stosunkiem pracy;
 - wprowadzenie wzorem prawodawstwa np. niemieckiego i francuskiego ustawowego wyjątku pozwalającego na otwarte udostępnianie utworów naukowych niezależnie od postanowień umowy zawartej z wydawcą;
 - wprowadzenie wyjątku pozwalającego na maszynowe przetwarzanie i analizę chronionych utworów w celach badawczych.
- Ustawie z dnia 20 lipca 2018 r. - Prawo o szkolnictwie wyższym i nauce:
 - uwzględnienie w ewaluacji jakości działalności naukowej udostępniania danych badawczych w modelu otwartym oraz ich cytowania i dalszego wykorzystania w badaniach;
 - uwzględnienie zarządzania danymi badawczymi w finansowaniu systemu szkolnictwa wyższego i nauki;
 - wyraźne wymienienie obowiązku otwartego udostępniania danych badawczych jako równoległego do regulacji zawartych w tej ustawie związanych z komercjalizacją oraz doprecyzowanie, że komercjalizacja nie wyklucza zapewnienia otwartego dostępu do danych badawczych (w tym pozyskania koniecznych do tego praw oraz odpowiedniego podziału zasobów na udostępniane w sposób otwarty oraz komercjalizowane);
 - zobowiązanie podmiotów tworzących system szkolnictwa wyższego i nauki do przyjmowania polityk otwartości (przy ustawowym określeniu minimalnego zakresu takich polityk);

¹⁵ Ustawa z dnia 17 lutego 2005 r. o informatyzacji działalności podmiotów realizujących zadania publiczne (Dz. U. 2024 poz. 1557, z późn. zm.)

- uwzględnienie informacji o sposobie upowszechniania wyników badań naukowych, zwłaszcza poprzez otwarty dostęp, we wnioskach stanowiących podstawę do ubiegania się o stopnie i tytuły naukowe;
- wprowadzenie ustawowego zobowiązania podmiotów sektora nauki do wdrożenia albo aktualizacji polityk instytucjonalnych w zakresie otwierania danych naukowych, tak, aby zapewnić otwartość gromadzonych danych badawczych.
- Ustawie z dnia 27 lipca 2001 r. o ochronie baz danych:
 - doprecyzowanie kwestii przysługiwania praw *sui generis* do danych powstających w ramach działalności naukowej finansowanej ze środków publicznych;
 - doprecyzowanie ustawowych wyjątków od praw wyłącznych w kontekście wykorzystywania danych do badań, w szczególności za pomocą ich maszynowego przetwarzania i analizowania.
- Ustawie z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego:
 - wprowadzenie wyraźnego obowiązku otwartego udostępniania danych badawczych, w tym uwzględniania podczas całego procesu badawczego ewentualnych praw osób trzecich aby umożliwić udostępnianie danych w najszerszym możliwym zakresie.

1.b) Przeprowadzenie procesu legislacyjnego w zakresie propozycji legislacyjnych opracowanych na podstawie analizy powstałej w ramach działania 1.a).

Przeprowadzanie analizy wymienionej w działaniu 1.a pozwoli na zidentyfikowanie pełnej listy zmian legislacyjnych niezbędnych dla realizacji celu szczegółowego 1. Na tej podstawie, Ministerstwo Nauki i Szkolnictwa Wyższego rozpocznie prace mające na celu wprowadzenie odpowiednich poprawek do zidentyfikowanych aktów normatywnych.

Oczekuje się, że wprowadzenie ustawowego obowiązku udostępniania danych badawczych pozyskanych w działaniach realizowanych z udziałem środków publicznych nastąpi po zbudowaniu kompleksowego systemu wsparcia w rozumieniu kompetencyjnym i infrastrukturalnym. W dalszej perspektywie warto wdrażać również rekomendację stopniowego otwierania danych badawczych wtórnych (istniejących wcześniej), zgromadzonych przed przyjęciem Polityki.

1.c) Uwzględnienie w procedurach, wszystkich publicznych podmiotów finansujących badania naukowe, zasady udostępniania danych badawczych, pozyskanych w działaniach realizowanych z udziałem środków publicznych zgodnie z zasadami FAIR.

Zakłada się, że wszystkie dane badawcze, zebrane lub wytworzone w ramach działań realizowanych z udziałem środków publicznych, które nie stanowią tajemnicy przedsiębiorstwa albo są niejawne z innych względów, powinny być archiwizowane i udostępniane zgodnie z zasadami FAIR według standardów wspomnianych w celu szczegółowym 3 i z wykorzystaniem zasobów opisanych w celu szczegółowym 4.

Dodatkowo, podmioty prowadzące badania naukowe powinny wdrożyć albo zaktualizować własne polityki instytucjonalne zgodnie z polityką krajową i obowiązującymi aktami prawnymi oraz zapewnić swoim pracownikom wsparcie w zarządzaniu danymi badawczymi, w tym w procesie ich otwierania, poprzez zatrudnienie wykwalifikowanego personelu, przede wszystkim specjalistów w zakresie danych badawczych (m.in. *data steward*, *data curator*, *data librarian*) lub przeszkolenie istniejącego personelu. Ważne jest sformalizowanie wsparcia *data stewarda* w procesie zarządzania danymi badawczymi w jednostce badawczej, w tym podkreślenie kluczowej roli: zespołu IT, rzecznika patentowego, centrum transferu technologii, działów prawnych, naukowych, bibliotek.

Zakłada się, że ewentualne dotacje na rzecz organizacji pozarządowych, które będą związane z pośrednim finansowaniem badań naukowych, docelowo będą warunkowane wdrożeniem podobnych rozwiązań.

2. Wskazanie ram organizacyjnych w zakresie zarządzania otwartymi danymi badawczymi na poziomie krajowym (krajowa Sieć Centrów Kompetencji Otwartych Danych Badawczych).

Biorąc pod uwagę wciąż niewystarczający poziom wiedzy nt. udostępniania danych badawczych w skali kraju, jak również niewystarczającą liczbę specjalistów w zakresie danych badawczych niezbędnym instrumentem wspierającym realizację Polityki jest utworzenie instytucjonalnej sieci składającej się z Krajowego Centrum Kompetencji Otwartych Danych Badawczych (KCKODB) pełniącego funkcje koordynacyjne, oraz Obszarowych Centrów Kompetencji Otwartych Danych Badawczych (OCKODB), których zadaniem będzie rozwój wiedzy i odpowiednich kompetencji w podmiotach sektora nauki. Poszczególne obszary tematyczne powinny być wytyczone z uwzględnieniem już obecnie funkcjonujących w kraju infrastruktur badawczych, wzorem klasyfikacji stosowanej przez Europejskie Forum Strategii ds. Infrastruktur Badawczych (ESFRI, ang. *European Strategy Forum on Research Infrastructures*).

Działania na rzecz realizacji celu:

2.a) Wskazanie i zapewnienie finansowania dla Krajowego Centrum Kompetencji Otwartych Danych Badawczych.

Niezbędne jest wskazanie jednego podmiotu odpowiedzialnego za nadzór merytoryczny i strategiczny nad procesem budowania kompetencji w obszarze zarządzania danymi badawczymi, koordynację działań Obszarowych Centrów Kompetencji Otwartych Danych Badawczych oraz zapewnienie im kompleksowego wsparcia eksperckiego, prawnego i szkoleniowego.

Jest to warunek zapewnienia w tworzonej synergii pomiędzy aktywnościami OCKODB, ale również sprawnego reagowania na ich problemy i potrzeby. W pierwszej fazie wdrażania systemu, Krajowe Centrum Kompetencji Otwartych Danych Badawczych odpowiedzialne będzie również za wyznaczanie ogólnych standardów w zakresie zarządzania metadanymi badawczymi w Polsce. Przeprowadzi

ono również pierwsze szkolenia dla specjalistów wspierających otwieranie danych w podmiotach tworzących system szkolnictwa wyższego i nauki.

W dalszej kolejności, KCKODB, wspólnie z MNiSW wyłoni centra obszarowe. Jego rolą będzie monitorowanie światowych standardów w zakresie otwartej nauki, monitorowanie polskich danych badawczych, łączenie OCKODB z partnerami międzynarodowymi, oraz inne działania o charakterze międzyobszarowym. KCKODB zapewni wsparcie polegające na organizacji kampanii szkoleniowych i informacyjnych oraz doradztwie na poziomie przekraczającym specyfikę danego obszaru.

Zadaniem KCKODB będzie również udział w międzynarodowych gremiach tworzących standardy zarządzania danymi badawczymi, ich monitorowanie i wdrażanie w kraju, w tym m.in. zagwarantowanie spójności z regulacjami UE. Pozwoli to na zwiększenie zdolności do realizacji międzynarodowej współpracy badawczej oraz rozpoznawalności (ang. *visibility*) polskiej nauki. Rolę KCKODB przyjmie wskazany przez MNiSW istniejący już podmiot, dysponujący odpowiednim doświadczeniem i kompetencjami w pracy na rzecz implementacji modelu otwartej nauki w Polsce.

2.b) Wybór i zapewnienie finansowania dla Obszarowych Centrów Kompetencji Otwartych Danych Badawczych.

Obszarowe Centra Kompetencji Otwartych Danych Badawczych zostaną powołane spośród podmiotów dysponujących odpowiednimi kompetencjami w obszarze zarządzania metadanymi badawczymi w danym obszarze naukowym oraz aktywnością na polu otwierania danych. Rolą OCKODB będzie w pierwszej kolejności wybór standardów zarządzania danymi badawczymi dla poszczególnych obszarów oraz ustalenie wraz z KCKODB wytycznych dotyczących metadanych w oparciu o istniejące standardy krajowe i międzynarodowe. OCKODB będą dbać o uwzględnienie specyfiki poszczególnych dyscyplin nauki, które nierzadko posługują się własnymi standardami.

Oprócz tego, OCKODB będą oferować wsparcie merytoryczne i szkoleniowe w ramach swojego obszaru, stanowiąc swoisty punkt kontaktowy. Przewiduje się, że podział zadań OCKODB będzie odpowiadał podziałowi istniejącemu w Polskiej Mapie Infrastruktury Badawczej:

- nauki techniczne i energetyka,
- nauki o ziemi i środowisku,
- nauki biologiczno-medyczne i rolnicze,
- nauki fizyczne i inżynierskie,
- nauki społeczne i humanistyczne,
- cyfrowe infrastruktury badawcze.

W dalszej perspektywie i w miarę potrzeby, w OCKODB mogą powstać zespoły merytoryczne odpowiedzialne np. za zapewnienie wsparcia konkretnej dyscyplinie naukowej lub w ramach określonego tematu badawczego.

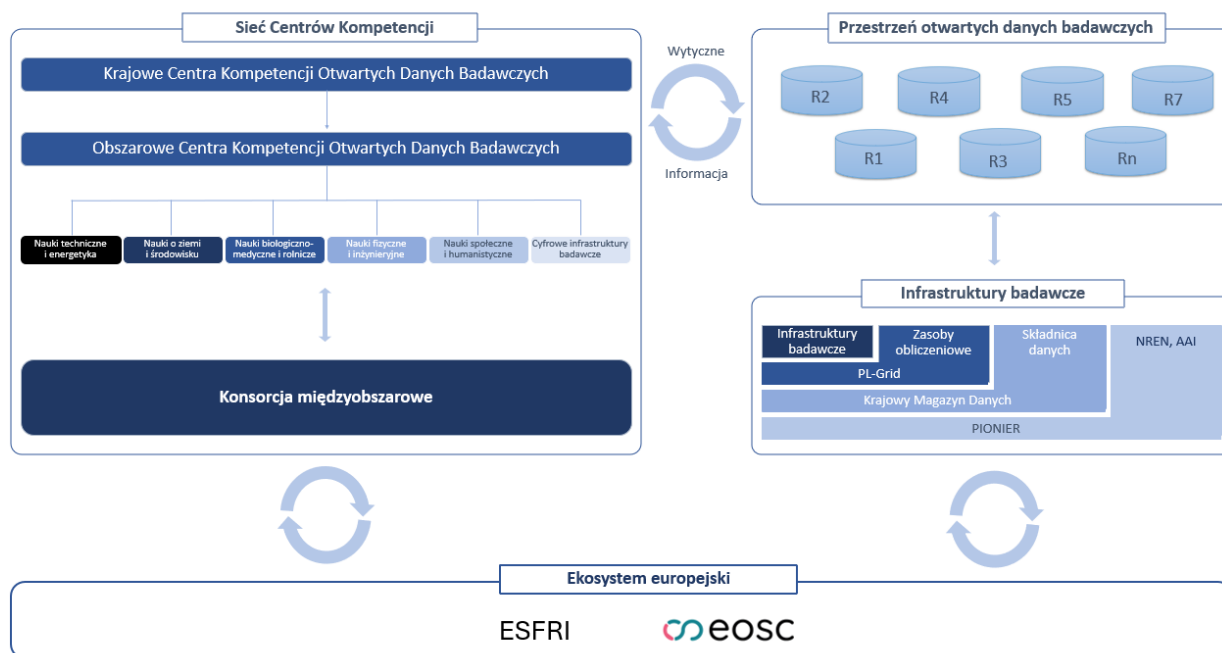


Diagram 1. Zarys krajowego ekosystemu otwartych danych badawczych

3. Ustalenie, przy uwzględnieniu kontekstu międzynarodowego, wymaganych standardów udostępniania danych badawczych, odpowiadających specyfice poszczególnych obszarów badawczych.

Działania na rzecz realizacji celu:

3.a) Wybór standardów zarządzania metadanymi badawczymi, katalogu rekomendowanych formatów otwartych danych, trwałych identyfikatorów obiektów cyfrowych (ang. *Persistent Identifiers*) oraz licencji, na których udostępniane będą dane badawcze dla sześciu pilotażowych obszarów naukowych.

Każde OCKODB będzie odpowiedzialne za wytypowanie jednej pilotażowej dyscypliny naukowej w ramach reprezentowanego przez siebie obszaru oraz za wybór obowiązującego dla niej standardu zarządzania metadanymi badawczymi i katalogu rekomendowanych formatów otwartych danych. Proces ten będzie koordynowany przez KCKODB w uzgodnieniu z MNiSW.

Dodatkowo, zgodnie z zasadą FAIR – ang. *Reusable*, każdy zbiór danych powinien być opatrzony licencją określającą warunki, na jakich można z niego korzystać. Najlepszą praktyką w tym zakresie wydaje się stosowanie gotowych licencji *Creative Commons*. W tym kontekście warto rozważyć również licencje dedykowane dla oprogramowania, np. GNU General Public License Massachusetts Institute of Technology (MIT license) oraz dla baz danych, np. Open Data Commons.

3.b) Ustalenie standardów zarządzania metadanymi badawczymi i katalogu rekomendowanych formatów otwartych danych dla wszystkich obszarów naukowych.

Zakładając powodzenie pilotażu opisanego w działaniu 3.a) zakłada się, że OCKODB rozpoczną prace nad wyborem standardów zarządzania metadanymi badawczymi, opracowaniem katalogu rekomendowanych formatów otwartych danych dla wszystkich dyscyplin. W ramach tych prac, każde OCKODB będzie koncentrować się na jednym obszarze badawczym.

3.c) Aktualizacja katalogów standardów i formatów danych zgodnie z międzynarodowymi trendami.

W miarę rozwoju globalnych standardów i praktyk, przyjęte w Polsce rozwiązania mogą wymagać aktualizacji. OCKODB będą odpowiedzialne za monitorowanie sytuacji w ramach swojego obszaru i w miarę potrzeby aktualizowanie katalogów standardów dotyczących metadanych oraz formatów (w uzgodnieniu z KCKODB).

4. Zapewnienie na poziomie krajowym ekosystemu informatycznego umożliwiającego przechowywanie, odnalezienie, uzyskanie dostępu i ponowne wykorzystywanie rozproszonych zasobów danych w każdym obszarze badawczym.

Utworzenie spójnego, dostępnego krajowego ekosystemu informatycznego, umożliwiającego przechowywanie i udostępnianie danych badawczych zgodnie z zasadami FAIR, jest kluczowym warunkiem realizacji celów Polityki. Uruchomienie takiego ekosystemu pozwoli również na bieżący monitoring w tym zakresie. Za przygotowanie danych do przekazania oraz ich transmisję do systemu odpowiedzialne będą poszczególne podmioty tworzące system szkolnictwa wyższego i nauki. Proces ten wymaga wdrożenia instytucjonalnych polityk dotyczących otwierania danych oraz ewentualnych zmian w istniejących, a także rozbudowy potencjału kadrowego i infrastrukturalnego podmiotów. Wsparcie merytoryczne i szkoleniowe w tym obszarze będzie zapewniane przez krajową Sieć Centrów Kompetencji Otwartych Danych Badawczych.

Działania na rzecz realizacji celu:

4.a) Utworzenie schematu przechowywania i udostępniania danych badawczych w oparciu o federację krajowych repozytoriów oraz o zasoby KDM/MAN/sieć PIONIER, ze szczególnym wskazaniem na zasoby projektów Krajowy Magazyn Danych oraz Narodowa Infrastruktura Chmurowa PLGrid dla EOSC.

Schemat będzie stanowił projekt architektury infrastrukturalnej łączącej istniejące zasoby w ramach krajowej federacji otwartych danych. Schemat będzie bazował na rozproszonej architekturze ekosystemu, będzie również określał kanały i metody transmisji danych z uczelni do infrastruktury krajowej.

Przewiduje się, że krajowy schemat przechowywania i udostępniania danych badawczych w znacznej mierze oprze się na federacji już istniejących krajowych repozytoriów. W tym celu, koniecznym pierwszym krokiem jest zmapowanie

dostępnych obecnie zasobów i zbadanie potrzeb środowiska naukowego. Wsparcie dla podmiotów nieposiadających odpowiednich zasobów zapewni Krajowy Magazyn Danych (KMD) współfinansowany z Programu Operacyjnego Inteligentny Rozwój i wpisany na Krajową Mapę Infrastruktury Badawczej. Celem KMD jest przebudowa istniejącej architektury informatycznej do postaci otwartego, modularnego, rozszerzalnego, zdecentralizowanego i skalowalnego magazynu danych. Magazyn będzie wyposażony w szereg interfejsów dostępowych oraz zintegrowanych usług i aplikacji, w tym mechanizmów wspomagających efektywne magazynowanie i dostęp do danych, długoterminowe zarządzanie danymi, ich eksplorację, analizę i efektywne przetwarzanie. Rozproszona architektura systemu ułatwi magazynowanie, jednak najpierw konieczne jest ustalenie kanałów i metod transmisji uprzednio przygotowanych przez uczelnie danych do infrastruktury krajowej. Wymagania dotyczące udostępniania danych KMD zostaną opracowane w oparciu o filary otwartości danych wymienione w Programie Otwierania Danych na lata 2021-2027.

Uzupełnieniem dla KMD będą zasoby projektu Narodowa Infrastruktura Chmurowa PLGrid dla EOSC, który również został wpisany na Polską Mapę Infrastruktury Badawczej. Pozwoli on na opracowanie, walidację i udostępnienie usług (ogólnych i dedykowanych), a tym samym wykorzystanie nowoczesnych technologii i efektywnych technik zarządzania, przetwarzania i ponownego wykorzystania danych przez środowiska naukowe, podmioty gospodarki i społeczeństwo. W ramach projektu powstanie rozproszona infrastruktura chmurowa, która zostanie wykorzystana m.in. do badań związanych z danymi, infrastrukturami i platformami dedykowanymi do ich przetwarzania, opracowania efektywnych algorytmów i aplikacji do rozwiązania określonych potrzeb. Warto podkreślić, że w oba powyższe projekty zaangażowane są wszystkie polskie centra Komputerów Dużej Mocy (KDM), które w ramach swojej działalności świadczą szereg usług na rzecz podmiotów tworzących system szkolnictwa wyższego i nauki. Pozwoli to wykorzystać istniejące kompetencje oraz dotychczasową współpracę pomiędzy tymi instytucjami.

Technologiczna i funkcjonalna integracja zasobów PLGrid z KMD, przy odpowiednim podziale zadań (PLGrid - usługi horyzontalne i dziedzinowe, KMD - usługi przetwarzania danych) oraz wpisanie zasobów wchodzących w skład centralnej krajowej infrastruktury do krajowych i międzynarodowych rejestrów otwartych repozytoriów pozwolą na zbudowanie efektywnego modelu wsparcia dla otwierania danych badawczych w Polsce.

4.b) Wdrażanie schematów opracowanych w ramach działań 3.a) i 4.a).

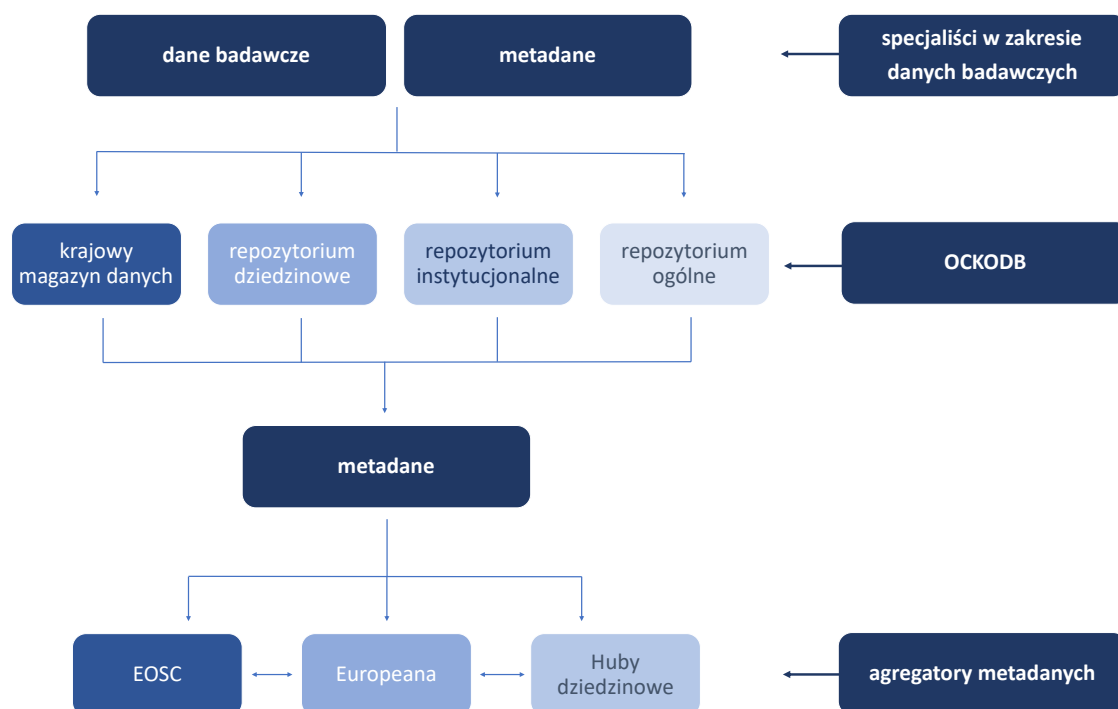
Praktyczne wdrożenie procesów przekazywania danych badawczych do infrastruktury krajowej będzie procesem długotrwałym i w nieunikniony sposób wymagającym od uczelni przeprowadzenia działań przygotowawczych, powiązanych z realizacją celu szczegółowego 1. Rekomenduje się, by przekazując dane, uczelnie o ile to możliwe,

korzystały z dostępnych lokalnych zasobów superkomputerowych, optymalizując ich wykorzystanie.

4.c) Wsparcie repozytoriów we wdrożeniu kryteriów jakościowych (np. w ramach procesu certyfikacji CoreTrustSeal).

Istotnym elementem działań będzie pomoc przy wdrożeniu kryteriów jakościowych, co przełoży się na wyższą jakość danych przekazywanych do magazynu centralnego. W procesie tym rolę koordynacyjną pełnić będzie KCKODB a wsparcia merytorycznego udzielać będzie OCKODB.

Diagram 2. Schemat deponowania danych



Powyższy diagram prezentuje schemat deponowania danych badawczych oraz ich metadanych. W obszarze zależnym od specjalistów w zakresie danych badawczych przedstawiono symbolicznie wytworzone dane oraz metadane które są następnie deponowane w Krajowym Magazynie Danych lub repozytorium (dziedzinowym, instytucjonalnym albo ogólnym). Wskazania dotyczące tego, w jakich repozytoriach dane powinny być deponowane będą zależne od obszaru badawczego i zostaną wskazane przez OCKODB. Metadane zasobów podlegają agregacji/indeksowaniu przez specjalne infrastruktury niezależne od systemu szkolnictwa wyższego i nauki.

Dalsze kroki i ewaluacja

Po przyjęciu niniejszej Polityki, MNiSW opracuje *plan działań*, w którym doprecyzowane zostaną działania służące realizacji celów szczegółowych, ze wskazaniem przewidywanego finansowania i harmonogramu. Co do zasady MNiSW będzie szacować nakłady finansowe na zadania związane z realizacją Polityki, w konsultacji z ekspertami oraz podmiotami realizującymi, jak również uwzględniać je w swoim budżecie na kolejne lata.

MNiSW będzie prowadzić bieżący monitoring efektów Polityki. Po upływie trzech lat od jej przyjęcia, Polityka będzie podlegać ewaluacji, w ramach której oceniane będą przede wszystkim:

- poziom osiągnięcia wskaźników przedstawionych w tabeli 1 poniżej i doprecyzowanych w planie działań,
- poziom wykorzystania infrastruktury krajowej,
- wyniki ewaluacji poziomu zadowolenia użytkowników.

Podsumowanie

Zasadniczym celem opisanych w tym dokumencie działań jest doprowadzenie do sytuacji, w której otwarty dostęp do danych badawczych stanie się w Polsce normą funkcjonującą w praktyce. Polityka nakreśla szereg zadań, które posłużą osiągnięciu tego celu, jednak należy zaznaczyć, że sama w sobie stanowi jedynie swoistą mapę drogową. Do realizacji jej celów potrzebne są długoterminowe działania ukierunkowane z jednej strony na usunięcie barier dla otwartego dostępu (np. w warstwie legislacyjnej i kompetencyjnej), a z drugiej zapewniające badaczom dostęp do środowiska informatycznego, które pozwoli na przechowywanie i udostępnianie danych zgodnie z określonymi standardami. W ten sposób Ministerstwo dąży do wzmocnienia globalnej otwartości polskiej nauki.

Tabela 1. Podmioty odpowiedzialne i wskaźniki realizacji zadań w ramach Polityki

Działanie	Wskaźnik realizacji	Docelowa wartość wskaźnika	Podmioty realizujące zadanie
1.a) Przeprowadzenie analizy i sporządzenie raportu dotyczącego obecnego stanu prawnego w udostępnianiu danych badawczych oraz niezbędnych zmian legislacyjnych wspierających otwieranie i ponowne wykorzystywanie tych danych wraz z propozycją zmian w odpowiednich aktach prawnych	Przyjęcie przez Ministra właściwego ds. szkolnictwa wyższego i nauki raportu z realizacji działania 1.a	1	MNiSW
1.b) Przeprowadzenie procesu legislacyjnego w zakresie propozycji legislacyjnych opracowanych na podstawie analizy powstałej w ramach działania 1.a)	Liczba przyjętych przez Radę Ministrów nowelizacji aktów prawnych	do ustalenia w trakcie realizacji celu szczegółowego 1.a)	MNiSW
1.c) Uwzględnienie w procedurach wszystkich publicznych podmiotów finansujących badania naukowe zasady udostępniania danych badawczych, pozyskanych w działaniach realizowanych z udziałem środków publicznych zgodnie z zasadami FAIR	Liczba przyjętych lub odpowiednio zaktualizowanych polityk agencji finansujących badania naukowe	4	NCBR, NCN, ABM, NAWA
2.a) Wskazanie i zapewnienie finansowania dla Krajowego Centrum Kompetencji Otwartych Danych Badawczych	Powołane Krajowe Centrum Kompetencji Otwartych Danych Badawczych	1	MNiSW
2.b) Wybór i zapewnienie finansowania dla Obszarowych Centrów Kompetencji Otwartych Danych	Liczba utworzonych Obszarowych Centrów Kompetencji Otwartych Danych Badawczych	6	KCKODB wspólnie z MNiSW

3.a) Wybór standardów zarządzania metadanymi badawczymi, katalogu rekomendowanych formatów otwartych danych, trwałych identyfikatorów obiektów cyfrowych (ang. <i>Persistent Identifiers</i>) oraz licencji, na których udostępniane będą dane badawcze dla sześciu pilotażowych obszarów naukowych	Liczba obszarów naukowych, dla których określono standard metadanych i katalog formatów otwartych danych	6	OCKODB
3.b) Ustalenie standardów zarządzania metadanymi badawczymi i katalogu rekomendowanych formatów otwartych danych dla wszystkich obszarów naukowych	n/d	działanie ciągłe	KCKODB, OCKODB
4.a) Utworzenie schematu przechowywania i udostępniania danych w oparciu o federację krajowych repozytoriów oraz o zasoby KDM/MAN/sieć PIONIER, ze szczególnym wskazaniem na zasoby projektów <i>Krajowy Magazyn Danych</i> oraz <i>Narodowa Infrastruktura Chmurowa PLGrid dla EOSC</i>	Akceptacja przez Ministra właściwego ds. szkolnictwa wyższego i nauki dokumentu zawierającego specyfikację schematu przechowywania i udostępniania danych	1	MNiSW
4.b) Wdrażanie schematów opracowanych w ramach działań 3.a) i 4.a)	n/d	działanie ciągłe	MNiSW, podmioty tworzące system szkolnictwa wyższego i nauki
4.c) Wsparcie repozytoriów we wdrożeniu kryteriów jakościowych (np. w ramach procesu certyfikacji CoreTrustSeal)	n/d	działanie ciągłe	KCKODB, OCKODB

Słowniczek wybranych pojęć i skrótów

Centra Komputerów Dużej Mocy (KDM) - specjalistyczne jednostki dysponujące superkomputerami i zaawansowaną infrastrukturą obliczeniową, wspierające badania naukowe, symulacje, analizę dużych zbiorów danych oraz rozwój technologii i innowacji.

Coalition for Advancing Research Assessment (CoARA) - inicjatywa mająca na celu poprawę metod oceny badań naukowych. CoARA została utworzona, aby wspierać i promować zmiany w sposobach, w jakich ocenia się wyniki badań, z naciskiem na wartości takie jak jakość, różnorodność, odpowiedzialność i przejrzystość.

Council for National Open Science Coordination (CoNOSC) - sieć krajowych koordynatorów otwartej nauki.

Creative Commons (CC) - zestaw standardowych licencji prawnych, które umożliwiają twórcom łatwe udostępnianie ich prac na określonych warunkach. Creative Commons oferuje różne typy licencji, pozwalające twórcom kontrolować sposób, w jaki ich prace mogą być używane przez innych, jednocześnie ułatwiając dzielenie się wiedzą i kreatywnością.

Dane badawcze - to wszelkiego rodzaju informacje (ilościowe i jakościowe) cyfrowe i analogowe, zaobserwowane, zbierane, przetwarzane lub wytworzone jako materiał do analizy, w celu uzyskania oryginalnych wyników naukowych.

Data curator - osoba odpowiedzialna za długoterminową archiwizację danych badawczych i ich użyteczność w przyszłości (migracja danych, monitoring trendów technologicznych).

Data librarian - osoba odpowiedzialna za wspieranie społeczności naukowych w udostępnianiu danych badawczych, edukowanie i uświadamianie.

Data steward - osoba odpowiedzialna za tworzenie procedur, zarządzanie danymi badawczymi w organizacji oraz dbanie o utrzymanie wysokiej jakości danych.

Digital Object Identifier (DOI) - cyfrowy identyfikator dowolnego przedmiotu własności intelektualnej, który jest przypisany na stałe do danego obiektu.

Europejska Chmura Otwartej Nauki (EOSC, ang. European Open Science Cloud) - ogólnoeuropejska inicjatywa na rzecz otwartej nauki, która ma zapewnić federacyjne, otwarte i wielodzielnicowe środowisko dostępu i zarządzania szerokim zakresem obiektów cyfrowych (w tym publikacji, danych badawczych, oprogramowania) finansowanych ze środków publicznych.

Przechowywanie, udostępnianie, przetwarzanie i ponowne wykorzystywanie tych obiektów zgodnie z zasadami FAIR jest jednym z kluczowych założeń EOSC. Inicjatywa ta wspiera również rozwój mechanizmów umożliwiających współpracę w obszarze otwartej nauki.

GO FAIR - oddolna, kierowana przez interesariuszy i samorządna inicjatywa, której celem jest wdrożenie zasad FAIR w zakresie danych, dzięki czemu dane będą możliwe do znalezienia, dostępne, interoperacyjne i wielokrotnego użytku.

Krajowy Magazyn Danych (KMD) - otwarty, modularny, rozszerzalny oraz zdecentralizowany i skalowalny magazyn danych, wyposażony w szereg interfejsów dostępowych oraz

zintegrowanych usług i aplikacji, w tym mechanizmów wspomagających efektywne składowanie i dostęp do danych (wielopunktowa obsługa I/O, buforowanie), długoterminowe zarządzanie danymi, ich eksplorację, analizę i efektywne przetwarzanie.

Metadane - ustrukturyzowane informacje służące do opisu publikacji naukowych i danych badawczych, pozwalające na ich jednoznaczną identyfikację i umożliwiające ich wyszukiwanie (dane o danych).

Nauka Obywatelska (ang. *Citizen Science*) - forma działalności badawczej, w której osoby niezawodowe współpracują z naukowcami, uczestnicząc w zbieraniu, analizie i interpretacji danych, wspomagając tym samym proces badawczy i popularyzację nauki.

Otwarta nauka - otwarty dostęp do wyników badań, danych, publikacji i narzędzi badawczych, w celu zwiększenia przejrzystości, współpracy i dostępności w nauce.

Otwarte dane badawcze (ang. *open research data*) - dane z badań naukowych, które są publicznie dostępne bez ograniczeń prawnych, technologicznych ani finansowych, umożliwiając ich swobodne wykorzystywanie, modyfikowanie i udostępnianie z uznaniem autorstwa i zachowaniem integralności danych.

Otwarty dostęp - dostępność treści lub danych nieodpłatnie w Internecie, co pozwala każdemu czytać, ściągać, kopiować, rozprowadzać, drukować, przeszukiwać lub zamieszczać odnośniki do pełnych wersji tekstów lub danych, indeksować, przekazywać jako dane do oprogramowania oraz używać w dowolnym innym zgodnym z prawem celu, bez barier finansowych, prawnych lub technicznych.

Plan S (cOAlition S) - inicjatywa publikowania wyników badań w otwartym dostępie, zainicjowana w 2018 roku przez cOAlition S, międzynarodowe konsorcjum agencji i fundacji finansujących badania naukowe.

Plan zarządzania danymi (ang. *Data Management Plan, DMP*) - dokument opisujący, w jaki sposób dane badawcze będą gromadzone, przechowywane, zarządzane, udostępniane i archiwizowane podczas i po zakończeniu projektu badawczego. DMP obejmuje kwestie dotyczące ochrony danych, prywatności, zgodności z przepisami oraz strategii długoterminowego przechowywania i udostępniania danych.

Polska Mapa Infrastruktury Badawczej - strategiczny dokument opracowywany przez Ministerstwo Nauki i Szkolnictwa Wyższego, zawierający wykaz kluczowych obiektów, zasobów i systemów infrastrukturalnych w Polsce, które wspierają prowadzenie badań naukowych i prac rozwojowych. Celem Mapy jest koordynacja inwestycji w infrastrukturę badawczą, optymalizacja jej wykorzystania oraz wspieranie współpracy międzynarodowej i międzysektorowej w zakresie badań naukowych.

Repozytorium danych badawczych - systematycznie zarządzana cyfrowa platforma, która przechowuje, udostępnia i archiwizuje dane pochodzące z badań naukowych. Repozytoria te umożliwiają długoterminowe przechowywanie danych, zapewniając ich dostępność dla innych badaczy oraz wspierając otwarty dostęp do wyników badań, co promuje transparentność i współpracę w nauce.

Repozytorium publikacji - platforma cyfrowa przeznaczona do przechowywania i udostępniania publikacji naukowych oraz innych dokumentów badawczych. Jej głównym celem jest zapewnienie trwałego dostępu do publikacji, promowanie otwartego dostępu do wyników badań oraz ułatwianie wyszukiwania, przeglądania i cytowania naukowych artykułów, raportów i innych materiałów naukowych.

San Francisco Declaration on Research Assessment (DORA) - deklaracja uznająca potrzebę poprawy sposobów oceny wyników badań naukowych. Deklaracja została opracowana w 2012 roku podczas dorocznego spotkania Amerykańskiego Towarzystwa Biologii Komórkowej w San Francisco.

Trwały identyfikator (PID, PI, ang. *Persistent Identifier*) - długotrwałe odniesienie do obiektu cyfrowego, np. dokumentu, pliku, strony internetowej lub innego obiektu cyfrowego przypisane na cały cykl jego „życia”. PID posiada unikalny identyfikator powiązany z aktualnym adresem metadanych lub treści, co pozwala na jednoznaczną identyfikację obiektu niezależnie od jego lokalizacji. W przeciwieństwie do adresów URL, PID są zarządzane przez usługi umożliwiające aktualizację lokalizacji obiektu, dzięki czemu identyfikator konsekwentnie wskazuje właściwą lokalizację. Dla zapewnienia spójności i przejrzystości danych w infrastrukturze krajowej, należy określić standardowy identyfikator stosowany do wszystkich takich materiałów oraz zbadać możliwość stosowania unikalnej charakterystyki PID polskich danych w celu ułatwienia monitorowania ich udostępniania.

Zasady FAIR - FAIR to akronim angielskich słów *Findable, Accessible, Interoperable, Reusable*. Odnosi się on do wytycznych dotyczących prawidłowego przygotowania danych badawczych do udostępniania. Zgodnie z nimi, dane powinny być możliwe do wyszukania zarówno przez człowieka, jak i maszynowo, dostępne dla wszystkich bez wykorzystania dodatkowych narzędzi, zapisane standardzie odpowiednim do odczytu przez ludzi i komputery i możliwe do łączenia z innymi zbiorami danych, jak również możliwe do wielokrotnego wykorzystania.



Ministerstwo Nauki
i Szkolnictwa Wyższego