



Sieć Kompetencji Otwartej Nauki – Data Steward PL

Grupa Robocza Data Stewardship Competence Center PL

Rekomendacje

W ramach Grupy Roboczej Data Stewardship Competence Center PL pięć podzespołów opracowało rekomendacje w zakresie: infrastruktury danych badawczych (repozytoria danych badawczych); wsparcia dla naukowców (rola data stewarda); polityki udostępniania danych; monitoringu planów zarządzania danymi; kwestii danych badawczych w kontekście ewaluacji.

Należy doprecyzować, iż w styczniu 2024 r. doszło do zmiany nazwy: Grupa Robocza Data Stewardship Competence Center PL przekształciła się w Sieć Kompetencji Otwartej Nauki – Data Steward PL, dlatego też w tytule Rekomendacji ujęto dwie nazwy Grupy.

Grupa Robocza – złożona m.in. z bibliotekarzy szkół wyższych, pracowników naukowych, osób zatrudnionych na stanowiskach data stewardów, uczelnianych/institutowych pełnomocników ds. Otwartej Nauki – kieruje swoje Rekomendacje do: Ministerstwa Nauki i Szkolnictwa Wyższego, Narodowego Centrum Nauki, a także do szeroko pojętych władz szkół wyższych/uczelni oraz instytutów badawczych i placówek naukowych.

Zaproponowane Rekomendacje, opierając się na europejskich i krajowych wytycznych dotyczących otwierania danych badawczych, bazują przede wszystkim na zawodowych doświadczeniach członków Grupy Roboczej, którzy w codziennej praktyce spotykają się z różnorodnymi problemami, jakie nastroczą pracownikom naukowym m.in. wymagania grantodawców, związane z udostępnianiem danych badawczych.

Rekomendacje dotyczące infrastruktury

Rozumiejąc wagę i zasadność tworzenia repozytoriów danych badawczych, Grupa Robocza zwraca uwagę na najważniejsze kwestie związane z prawidłowym funkcjonowaniem tego typu repozytoriów:

- w repozytoriach należy stosować trwałe identyfikatory, czyli PID (Persistent Identifier, np. DOI, handle, URI);
- w kontekście różnorodnych rozwiązań związanych z usuwaniem datasetów (opcja pierwsza: nie można usuwać żadnych danych; opcja druga: wszystko można usunąć; opcja trzecia: można usunąć plik, metadane powinny pozostać), Grupa wskazuje na potrzebę jasnego określenia procedur związanych z usuwaniem danych w repozytorium konkretnej uczelni/institucji;
- instytucjonalne repozytoria danych badawczych powinny być (zgodnie z wymogami m.in. NCN) indeksowane w rekomendowanych serwisach, np. re3data; OpenAIRE;
- zalecamy certyfikowanie repozytoriów danych, jednak wybór konkretnego certyfikatu pozostawiamy administratorom systemu (np. Core Trust Seal; nestor Seal; ISO 16363);
- zalecamy śledzenie postępów w zakresie rozwoju międzynarodowych standardów metadanych.

Podsumowanie Rekomendacji skierowanych do poszczególnych grup odbiorców:

Zagadnienie	NCN	MNiSW	Jednostka naukowa
PID	Rekomendujemy stosowanie PID. Rekomendujemy śledzenie zmian w środowisku, ze szczególnym uwzględnieniem organizacji międzynarodowych.	Rekomendujemy stosowanie PID. Rekomendujemy śledzenie zmian w środowisku, ze szczególnym uwzględnieniem organizacji międzynarodowych. Dofinansowanie PID.	Rekomendujemy stosowanie PID. Rekomendujemy śledzenie zmian w środowisku, ze szczególnym uwzględnieniem organizacji międzynarodowych.
Indeksowanie i certyfikacja repozytoriów			Repozytorium powinno być rejestrowane w indeksach repozytoriów danych badawczych. Istotne jest rozróżnienie pomiędzy serwisami, w których rejestrujemy repozytorium a agregatami/sieciami, w których udostępniane są również dane. Rekomendujemy opracowanie zasad dotyczących wyboru repozytorium przez naukowców.
Usuwanie danych badawczych			Rekomendujemy, aby polityka w zakresie usuwania danych badawczych była jasno określona/rozpisana w danej instytucji.

Wsparcie dla naukowców

Rola data stewarda

Zmiany związane z udostępnianiem danych badawczych, będące przede wszystkim implikacją wymogów grantodawców, spowodowały pojawienie się funkcji data stewarda, czyli opiekuna danych.

W założeniu data steward powinien pełnić rolę doradczą i wspierającą naukowców, np. przy wypełnianiu Planu Zarządzania Danymi lub udostępnianiu danych badawczych (wybór repozytorium). Brak odpowiednich uregulowań prawnych na gruncie polskim spowodował, iż zadania te przejęli w głównej mierze pracownicy bibliotek naukowych, oddelegowani do tych zadań pracownicy naukowcy, niekiedy uczelniani/instytutowi pełnomocnicy ds. Otwartej Nauki, a nawet informatycy.

Mimo heterogeniczności Grupy Roboczej polskich data stewardów, będącej wynikiem różnorodnych doświadczeń zawodowych, udało się wypracować wspólne rekomendacje dotyczące zawodu oraz wykonywanych zadań:

- grupę wsparcia dla naukowców w kwestiach związanych z danymi badawczymi powinni tworzyć, oprócz data stewarda: bibliotekarz, prawnik, informatyk, inspektor ochrony danych osobowych, pracownik odpowiedzialny za bezpieczeństwo teleinformatyczne jednostki, osoba odpowiedzialna za aparaturę badawczą (przykład dobrej praktyki: Uniwersytet Ekonomiczny we Wrocławiu);
- należy wprowadzić uregulowania prawne działalności data stewardów oraz doprecyzować zakres ich kompetencji (np. poprzez zarządzenia rektora/dyrektora);
- rekomenduje się wprowadzenie stałego kontaktu ze środowiskiem data stewardów (np. poprzez newsletter) oraz informowanie na bieżąco o nowych regulacjach i zmianach w obowiązujących zasadach, a także możliwość szybszego kontaktu data stewardów z NCN w celu konsultacji.

Podsumowanie Rekomendacji dotyczących roli i zadań data stewarda:

Zagadnienie	NCN	Ministerstwo	Jednostka naukowa
Data steward	Rekomenduje się wprowadzenie stałego kontaktu z środowiskiem data stewardów (np. poprzez newsletter) oraz informowanie na bieżąco o nowych regulacjach i zmianach w obowiązujących zasadach, a także możliwość szybszego kontaktu data stewardów z NCN w celu konsultacji.	Rekomenduje się ustalenie ścieżki zawodowej data stewardów i data librarian oraz niezbędnych kompetencji na tych stanowiskach.	Rekomenduje się, aby w skład grupy wsparcia naukowców wchodziło: data steward, bibliotekarz, prawnik, informatyk, inspektor ochrony danych, ekspert ds. bezpieczeństwa, osoba z działu projektów, osoba odpowiedzialna za aparaturę badawczą. Rekomenduje się, żeby działalność data stewardów i grup wsparcia była uregulowana w wewnętrznych regulacjach prawnych, np. Zarządzeniach rektorów/dyrektorów. Wprowadzenie na poziomie uczelni ram i norm pracy oraz opracowanie workflow, uwzględniającego podział zadań, procedury, role i terminy nadsyłania PZD. Określenie zakresu wsparcia oraz zakresu zadań data stewardów. Rekomenduje się wyznaczenie dziedzinowych data stewardów.

Polityka udostępniania danych badawczych

Rekomendacje Grupy Roboczej związane z polityką udostępniania danych badawczych dotyczą następujących zagadnień:

- udostępnianie danych badawczych może mieć pozytywny wpływ na ocenę pracownika naukowego; jednakże stosowanie odpowiednich bodźców może zachęcić naukowców do tego typu działań; dlatego też rekomenduje się np. przyznawanie dodatkowych punktów za otwieranie publikacji/danych przy

ocenie pracowniczej, nagrodach, konkursach, bonus przy awansie, a w przypadku braku możliwości udostępnienia samych danych, otwieranie przynajmniej metadanych;

- obecnie otwierane i udostępniane są przede wszystkim dane badawcze powiązane z publikacją - efektem realizacji projektu NCN oraz grantami EU, w przyszłości rekomenduje się udostępnianie wszystkich danych finansowanych ze środków publicznych;

- w kontekście działalności data stewarda, należy rozpatrzyć możliwość zatrudnienia data stewarda przy projektach NCN w ramach kosztów bezpośrednich; ponadto należy rozważyć utworzenie nowej ścieżki zawodowej data stewarda oraz określić zakres zadań, które ma wykonywać;

- zauważalna jest konieczność pracy u podstaw i nauczania na etapie szkół doktorskich na temat otwartości nauki (w tym otwierania danych badawczych);

- należy dokonać zmiany w wymogach NCN i umożliwić udostępnienie zbioru danych na licencji CC BY, bez proponowanego przez NCN okresu przejściowego do 31. 12. 2025 r.;

- postuluje się zwiększenie limitu znaków w wymaganym przez NCN Planie Zarządzania Danymi.

Monitoring planów zarządzania danymi

Zespół opracowujący Rekomendacje dotyczące monitoringu Planów Zarządzania Danymi (PZD) zwrócił uwagę na różnorodność praktyk i standardów stosowanych nie tylko na poszczególnych uczelniach, ale także na wydziałach czy w ramach dyscyplin. W kontekście monitoringu PZD przez bibliotekę podkreślono różne zakresy zadań bibliotek oraz fakt, iż nie wszystkie biblioteki uczelniane są odpowiedzialne za pomoc naukowcom w wypełnianiu PZD.

Rekomendacje Grupy Roboczej:

1. Wsparciem dla badacza przygotowującego Plan Zarządzania Danymi powinna być grupa osób zaznajomionych z tematyką i bieżącymi zagadnieniami z zakresu Open Science, RODO, rozwiązań informatycznych, baz danych, licencji i innych, pokrewnych tematów. Grupę taką powinna koordynować osoba wyznaczona przez uczelnię/institut badawczy i zajmująca się wymienioną tematyką na co dzień. Członkowie takiej grupy mogą w jednostce pełnić niżej wymienione funkcje: data steward; prawnik; specjalista z branży IT; inspektor ochrony danych osobowych; ekspert dziedzinowy; bibliotekarz; rzecznik patentowy; pracownik z obsługi działu projektów; pracownik centrum obsługi/wsparcia nauki.

2. PZD powinien być monitorowany w trakcie realizacji projektu. Zalecamy przygotowanie w każdej jednostce naukowej instytucjonalnych wzorów PZD, dostosowanych do możliwości danej jednostki oraz wymogów najczęściej wybieranych przez badaczy grantodawców. Powinny one być sporządzone tak, aby badacz mógł je intuicyjnie dostosować do potrzeb wynikających z przygotowywanego projektu.

3. Przygotowanie narzędzi, które ułatwią monitorowanie PZD, np. lista Excel lub baza, gdzie można wpisywać takie dane jak:

wnioskodawca; e-mail wnioskodawcy; wydział; dziedzina; grantodawca; konkurs; edycja konkursu; czas trwania konkursu; data złożenia PZD do konsultacji; status PZD; konsultant; ponowna konsultacja; uzyskane finansowanie; repozytorium, w którym dane zostały zdeponowane; raport; liczba danych zgromadzonych w projekcie.

4. W programie ramowym Unii Europejskiej Horyzont Europa zaleca się, aby przysyłać aktualizacje PZD, jeśli zaszły w nim zmiany.

5. PZD powinny być gromadzone w instytucjonalnym zbiorze, aby osoby biorące udział w tworzeniu PZD mogły służyć profesjonalnym wsparciem badaczom i w miarę możliwości doradzać tam, gdzie wystąpi taka potrzeba.

Podsumowanie Rekomendacji dotyczących monitoringu planów zarządzania danymi:

Zagadnienie	NCN	Ministerstwo	Jednostka naukowa
Monitoring			Monitoringiem PZD powinna zajmować się wyznaczona jednostka
Zespół wsparcia dla badacza	Data steward zatrudniony na uczelni w ramach projektu/projektów oraz zespół ekspertów np. prawnik, specjalista z branży IT, inspektor ochrony danych osobowych, ekspert dziedzinowy, bibliotekarz, rzecznik patentowy, pracownik z obsługi działu projektów, pracownik centrum obsługi/wsparcia nauki.	Data steward zatrudniony na uczelni w ramach projektu/projektów oraz zespół ekspertów np. prawnik, specjalista z branży IT, inspektor ochrony danych osobowych, ekspert dziedzinowy, bibliotekarz, rzecznik patentowy, pracownik z obsługi działu projektów, pracownik centrum obsługi/wsparcia nauki.	Data steward zatrudniony na uczelni w ramach projektu/projektów oraz zespół ekspertów np. prawnik, specjalista z branży IT, inspektor ochrony danych osobowych, ekspert dziedzinowy, bibliotekarz, rzecznik patentowy, pracownik z obsługi działu projektów, pracownik centrum obsługi/wsparcia nauki.
Narzędzia			Baza/lista, gdzie można wpisywać dane dotyczące prowadzonych PZD.

Dane badawcze w kontekście ewaluacji

Rekomendacje Grupy Roboczej DSCC w kontekście ewaluacji:

Grupa Robocza postuluje otwarty model uprawiania nauki w Polsce. Tak, by proces otwierania danych badawczych (ale też publikacji i innych osiągnięć naukowych) był istotnym, jeśli nie podstawowym, atrybutem działalności naukowej. Grupa Robocza proponuje wprowadzenie rozwiązań systemowych, które mają w tym pomóc:

1. Udostępnianie danych badawczych, podobnie jak otwarte publikowanie, powinno być jak najszybciej włączone w proces ewaluacji nauki oraz oceny osiągnięć naukowych poszczególnych pracowników. Postulujemy przede wszystkim uwzględnienie udostępniania danych w procesie oceny wewnętrznej pracownika. Wysiłek, jaki pracownicy naukowcy wkładają w proces opracowania oraz udostępniania

danych powinien zostać zauważony. Rekomendujemy przyznawanie dodatkowych punktów za otwieranie publikacji/danych przy ocenie pracowniczej, w konkursach, nagrodach rektora/dyrektora. Udostępnianie danych powinno również procentować w procesie awansu zawodowego lub naukowego.

2. Udostępnianie danych powinno być również elementem procesu ewaluacji nauki, ponieważ dostrzegamy trudności z wyborem „wskaźnika” podlegającego ocenie, postulujemy dalsze prace nad wskaźnikiem oceny udostępniania danych. Należy monitorować tendencje światowe i w miarę możliwości wprowadzać zaobserwowane rozwiązania. Analizując sytuację, warto zwrócić uwagę na kwestie problematyczne, trudności w procesie udostępniania danych i wdrażać/postulować rozwiązania, które te trudności zminimalizują. Ocena udostępniania danych powinna służyć rozwojowi badawczemu, nie karaniu. Powinno się doceniać fakt dzielenia się danymi, nie karać za brak takich działań.

3. Wymóg NCN dotyczący przygotowania PZD jest niewystarczający, należy kłaść nacisk na szersze udostępnianie danych. Wymóg udostępniania danych powinien dotyczyć wszystkich badań, nie tylko tych finansowanych z grantów.

4. Przygotowanie Planu Zarządzania Danymi oraz ich udostępnianie powinno być ważnym elementem procesu kształcenia w Szkołach Doktorskich. Proponujemy włączenie edukacji w zakresie zarządzania danymi w plan kształcenia oraz obowiązek przygotowania minimum jednego Planu Zarządzania Danymi w trakcie, np. poprzez złożenie wniosku projektowego przez doktoranta w celu zapoznania się z OS, wypełnienia PZD. Rekomendujemy, by PZD i udostępnianie danych w repozytorium instytucjonalnym był obowiązkowym elementem dysertacji.

Opracowanie Rekomendacji: maj-czerwiec 2023 r.; kontynuacja: maj-czerwiec 2024 r.

Podzespoły zadaniowe SKON – Data Steward PL w składzie:

Agnieszka Adamiec (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie)

Maja Bogajczyk (Uniwersytet Warszawski)

Anna Cedro (Uniwersytet Szczeciński)

Katarzyna Dudek (Politechnika Krakowska)

Monika Gajewska (Politechnika Warszawska)

Małgorzata Galik (Uniwersytet Jagielloński w Krakowie)

Helena Gałek (Politechnika Krakowska)

Krzysztof Gmerek (Uniwersytet Przyrodniczy w Poznaniu)

Agnieszka Goszczyńska (Uniwersytet Medyczny w Łodzi).

Barbara Grzelczak (Uniwersytet Ekonomiczny we Wrocławiu)

Michał Janiak (Instytut Rozrodu Zwierząt i Badań Żywności PAN w Olsztynie)

Anna Anetta Janowska (Szkoła Główna Handlowa w Warszawie)

Joanna Konik (Uniwersytet Jagielloński w Krakowie)

Szymon Kubik (Uniwersytet Jagielloński Collegium Medicum w Krakowie)

Maria Kuczkowska (Uniwersytet im. Adama Mickiewicza w Poznaniu)
Magdalena Nagieć (Uniwersytet Ekonomiczny w Krakowie)
Rafał Narwojsz (Politechnika Gdańska)
Tomasz Nowocień (Pomorski Uniwersytet Medyczny)
Dorota Olejnik (Politechnika Bydgoska)
Kamila Perlik (Uniwersytet im. Mikołaja Kopernika w Toruniu)
Jakub Rusakow (Gdański Uniwersytet Medyczny)
Zbigniew Ruszczyk (Uniwersytet Gdański)
Dorota Siwecka (Uniwersytet Wrocławski)
Anna Starek (Uniwersytet Przyrodniczy w Lublinie)
Leszek Szafrński (Uniwersytet Jagielloński w Krakowie)
Magdalena Szuflita-Żurawska (Politechnika Gdańska)
Iwona Świerblewska (Instytut Fizyki Jądrowej im. Henryka Niewodniczańskiego PAN w Krakowie)
Marta Urbaniec (Akademia Górniczo-Hutnicza w Krakowie)
Agnieszka Wasilewska (Uniwersytet Gdański)