



Oprac. mgr Agnieszka Adamiec
Samodzielne Stanowisko ds. Otwartych Zasobów Nauki
Biblioteka przy Instytucie Medycyny Weterynaryjnej
e-mail: agnieszka_adamiec@sggw.edu.pl
tel. 59 35 782

Otwarty dostęp w projektach na przykładzie NCN

Spis treści

1. Definicja danych badawczych i otwartych danych badawczych.... 1
2. Struktura planu zarządzania danymi badawczymi stanowiącego część wniosku o finansowanie projektów przez NCN..... 4
3. Najczęstsze błędy w przygotowywaniu planu zarządzania danymi 14

1. Definicja danych badawczych i otwartych danych badawczych

Dane badawcze:

„Dokumenty w formie cyfrowej, inne niż publikacje naukowe, które są gromadzone lub produkowane w ramach działalności badawczo-naukowej i są wykorzystywane jako dowody w procesie badawczym, bądź też są powszechnie akceptowane w środowisku naukowym jako konieczne do weryfikacji poprawności ustaleń i wyników badań”.



Źródło: [Dyrektywa Parlamentu Europejskiego i Rady z 20 czerwca 2019 r. w sprawie otwartych danych i ponownego wykorzystania informacji sektora publicznego \(wersja przekształcona\)](#)

Dane badawcze:

„Wszelkiego rodzaju informacje (ilościowe i jakościowe) wytwarzane, zbierane, przetwarzane, poddawane analizie w ramach projektu badawczego. Definicja ta obejmuje wszystkie możliwe formaty, zarówno cyfrowe, jak i niecyfrowe (np. próbki, wypełnione papierowe kwestionariusze, nagrania dźwiękowe na dyktafonie, papierowe dokumenty archiwalne itd.)”.

Źródło: [Narodowe Centrum Nauki](#)

Przykłady danych badawczych: dane liczbowe, dokumenty tekstowe, notatki, kwestionariusze, wyniki ankiet, nagrania audio i wideo, fotografie, bazy danych, oprogramowanie, wyniki symulacji komputerowych, protokoły metodologiczne, obserwacje laboratoryjne.

Otwarte dane badawcze (ang. open research data) to dane, uzyskane w procesie badawczym i użyte w pracy naukowej, do których każdy ma powszechny, zdalny dostęp z możliwością ich bezpłatnego ponownego wykorzystania z poszanowaniem praw autorskich.

FAIR data

Zgodnie z zaleceniami instytucji europejskich wyniki badań naukowych, które pochodzą z badań finansowanych ze środków publicznych powinny spełniać **zasady FAIR**. Akronim ten oznacza dane: możliwe do znalezienia (**F**indable) - dane opisane metadanymi odczytywanymi przez ludzi i maszyny, przypisany stały identyfikator (PID), PID powiązany z metadanymi, indeksacja metadanych;



dostępne (**Accessible**) - dostęp do (meta)danych z wykorzystaniem standardowych protokołów komunikacji (otwarte, bezpłatne i powszechnie stosowane), protokoły pozwalające na autentyfikację i autoryzowanie dostępu, metadane są przechowywane dłużej niż dane (jest to tańsze i łatwiejsze);

interoperacyjne (**Interoperable**) - (meta)dane odczytywane i przetwarzane przez ludzi i maszyny bez konieczności wykorzystywania tłumaczeń, specjalistycznych algorytmów; format (zapewniający odczyt maszynowy i przez ludzi), powszechnie stosowane słowniki i ontologie, dobrze udokumentowane słowniki (PID), odnośniki do powiązanych zbiorów;

nadające się do ponownego wykorzystania (**Reusable**) - metadane i dane udostępnione zgodnie z wyraźnie wskazaną licencją określającą jednoznacznie warunki ich ponownego wykorzystania i przetwarzania, np. Creative Commons, które są odczytywane zarówno przez ludzi, jak i maszyny; precyzyjne informacje dotyczące cytowania; dostosowanie do standardów obowiązujących w danej dziedzinie.

Ważne! Spełnianie zasad FAIR nie oznacza jeszcze ich otwierania. Przed otwarciem danych należy wziąć pod uwagę kwestie związane m.in. z ochroną danych osobowych, bezpieczeństwem, tajemnicą handlową, komercjalizacją lub prawami własności intelektualnej.

Zgodnie z podstawową zasadą udostępniania danych badawczych powinny być one 'tak otwarte, jak jest to możliwe, tak zamknięte, jak jest to niezbędne' (as open as possible, as closed as necessary).

Dane, które ze względu na ochronę prywatności nie mogą zostać opublikowane, mogą spełniać wszystkie zasady FAIR. Spełnianie tych zasad świadczy o ich prawidłowym generowaniu, przechowywaniu i wykorzystaniu.



2. Struktura planu zarządzania danymi badawczymi stanowiącego część wniosku o finansowanie projektów przez NCN

Zalecenia instytucji europejskich i krajowych:

Polityka powszechnej, ogólnodostępnej nauki wynika z celów Unii Europejskiej i zobowiązuje państwa członkowskie do przyjmowania strategii otwierania dostępu zarówno do publikacji naukowych, jak i danych uzyskanych z badań finansowanych ze środków publicznych.

20 czerwca 2019 r. Parlament Europejski i Rada przyjęli [dyrektywę o otwartych danych i ponownym wykorzystaniu informacji sektora publicznego](#), zgodnie z którą państwa członkowskie są zobligowane do wdrożenia polityk otwartego dostępu do danych badawczych finansowanych ze środków publicznych.

W kwietniu 2021 r. [Minister Edukacji i Nauki powołał zarządzeniem zespół doradczy do spraw otwartych danych naukowych](#), do którego zadań należy m.in. opracowanie polityki otwartego dostępu do danych badawczych finansowanych ze środków publicznych.

W sierpniu 2021 r. w Polsce wprowadzono [Ustawę o otwartych danych i ponownym wykorzystaniu informacji sektora publicznego](#), która wdraża europejską dyrektywę o otwartych danych i ponownym wykorzystaniu informacji sektora publicznego. W Ustawie zastrzeżono, że minister właściwy do spraw szkolnictwa wyższego i nauki opracuje Politykę otwartego dostępu do danych badawczych finansowanych ze środków publicznych.

Następstwem rekomendacji wydawanych na szczeblach europejskim i krajowych są wymagania podmiotów finansujących naukę stawiane swoim beneficjentom.

Coraz więcej programów finansowania wymaga, aby wyniki uzyskane w projektach realizowanych ze środków publicznych, udostępniane



były w sposób otwarty oraz aby cały proces badawczy był właściwie zarządzany.

Przykładami są zalecenia programu Horyzont Europa oraz zalecenia naszego polskiego podmiotu finansującego badania - Narodowego Centrum Nauki.

We wrześniu 2018 r. Narodowe Centrum Nauki (NCN) stało się jednym z sygnatariuszy porozumienia cOAlition S. Celem przyjętego przez koalicjantów [Planu S](#) jest wprowadzenie otwartego dostępu do publikacji i wyników badań naukowych finansowanych ze środków publicznych.

Narodowe Centrum Nauki zarządzeniem Dyrektora NCN z 2020 r. przyjęło [politykę otwartego dostępu](#) dotyczącą otwartego dostępu do publikacji.

Zapisy tego dokumentu regulujące kwestie danych badawczych:

- dane powinny być udostępnione w otwartym repozytorium, tam gdzie to możliwe, zgodnie z warunkami licencji Creative Commons, najlepiej Public Domain [CC0](#) (NCN wymaga, żeby dane, które stanowią podstawę publikacji zawierającej wyniki będące efektem realizacji projektu, były udostępniane w zaufanym repozytorium na licencji CC0),
- metadane powinny spełniać wytyczne podane przez [OpenAIRE](#), w tym powinny zawierać adnotację o finansowaniu ze środków projektu (Narodowe Centrum Nauki, numer projektu),
- standardy cytowania danych powinny być zgodne z wytycznymi [Declaration of Data Citation Principles](#) by FORCE 11 oraz [TOP Guidelines](#) (przykład: autor(zy), rok, tytuł danych, repozytorium, wersja, stały identyfikator).

W 2019 r. NCN zobowiązało autorów wniosków o dofinansowanie projektów badawczych do dołączania do zgłoszeń planu zarządzania



danymi badawczymi (Data Management Plan, DMP). DMP musi być opracowany w języku angielskim (z wyjątkiem DMP dla wniosków składanych w konkursie MINIATURA¹).

DMP jest stosowany we wszystkich konkursach i naborach, dotyczy to również konkursów międzynarodowych.

W konkursach międzynarodowych, które są organizowane zgodnie z procedurą lead agency (np. Opus Lab – NCN jest agencją wiodącą), tylko jedna z agencji zaangażowanych we współpracę przeprowadza proces oceny merytorycznej, z kolei druga agencja uznaje wynik tego procesu i przyznaje finansowanie swojemu zespołowi w oparciu o te rekomendacje. Formularz wniosku zawsze bazujena formularzu agencji wiodącej, dlatego może się zdarzyć, że w formularzu nie będzie DMP. W takich przypadkach DMP dotyczy polskiej części projektu i będzie odnosić się do działań realizowanych przez polski zespół badawczy. Natomiast w przypadku, gdy NCN jest agencją wiodącą, DMP należy przygotować dla całego zespołu badawczego. Przedstawia wtedy całość zadań realizowanych przez wszystkie zespoły, jednak koszty są kwalifikowane tylko dla polskiego zespołu badawczego.

[Wytyczne NCN dla wnioskodawców do uzupełnienia PLANU ZARZĄDZANIA DANYMI](#) w projekcie badawczym

NCN dopuszcza, że w ramach niektórych projektów nie będą wytwarzane, na nowo wykorzystywane, ani poddawane analizie żadne dane badawcze ani inne podobne materiały. **W takich przypadkach potrzebne jest uzasadnienie.**

Kwestionariusz planu zarządzania danymi NCN został oparty w niemal 95% na [Practical guide to the international alignment of research data management](#), który został opracowany przez Science Europe. NCN chce udostępniać dane, które będą wykorzystywane nie tylko

¹ Informacja aktualna na 09.02.2022.



lokalnie, ale przede wszystkim międzynarodowo. Dla NCN ważne jest, żeby opracowywać standardy, które będą miały ogólne zastosowanie.

Elementy planu zarządzania danymi badawczymi:

1. Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych

1.1. W jaki sposób będą pozyskiwane lub wytwarzane nowe dane lub ponownie wykorzystywane dane już istniejące?

- nowe dane: standardy, metody i oprogramowanie, które posłużą do pozyskiwania lub wytwarzania nowych danych
- dane istniejące: opis rodzaju danych (własne lub będące w posiadaniu strony trzeciej)
- sposób udokumentowania pochodzenia danych
- sposób kontroli i dokumentowania ich jakości i spójności (np. procesy kalibracji, powtórne pobieranie próbek lub ponowne wykonywanie pomiarów, standardy zapisywania/nagrywania i utrwalania danych, wykorzystanie narzuconego słownictwa, walidacja wprowadzanych danych, specjalistyczne recenzje, itd.)
- sposób organizacji plików i zarządzanie ich różnymi wersjami (odpowiednie nazewnictwo plików)
- informacja czy konieczna będzie digitalizacja danych analogowych lub wydanych w formie papierowej; jeśli tak, to jakiego rodzaju urządzenie będzie potrzebne do digitalizacji (np. mapy, fotografie, teksty)

1.2. Jakie dane (tj. rodzaje, formaty, objętości) będą pozyskiwane lub wytwarzane w projekcie?



- rodzaj, format (przechowywanie; należy udzielić pierwszeństwa formatom otwartym i szeroko stosowanym) i szacunkowa objętość danych wytworzonych, pozyskanych, ponownie wykorzystanych

2. Dokumentacja i jakość danych

2.1. Jakie metadane i dokumentacja (np. metodologia oraz sposoby pozyskiwania i organizacji danych) będą towarzyszyć danym w projekcie?

- wskazanie, jakie informacje należy uzupełnić, aby potencjalni użytkownicy byli w stanie w przyszłości odczytać i zinterpretować zebrane dane
- sposób utworzenia odpowiedniej dokumentacji (wskazanie sposobu organizacji danych w trakcie projektu poprzez podanie np. przyjętej konwencji, wersji i struktury folderów; należy określić w jaki sposób i gdzie informacje takie zostaną utrwalone, tj. w bazie danych z łączem prowadzącym do poszczególnych elementów, plikach README, książce kodowej itd.)
- standardy środowiskowe, które posłużą do objaśnienia (meta)danych
- międzynarodowe standardy lub schematy służące do organizacji metadanych (Dublin Core, DDI, DataCite)

2.2. Jakie planują Państwo zastosować środki kontroli jakości?

- wpływ metod pozyskiwania, analizy i przetwarzania danych na ich jakość
- eliminacja błędów pomiarowych i stroniczości (np. kalibracja, powtórne pomiary)



- minimalizacja ryzyka dotyczącego poprawności danych (wskazanie przyjętych w instytucji mechanizmów mających na celu zabezpieczenie danych przed nieuprawnioną modyfikacją i określenie, w jaki sposób/na jakim etapie zostanie wdrożona ocena jakości)

3. Przechowywanie i tworzenie kopii zapasowych podczas badań.

3.1. W jaki sposób w trakcie projektu będą przechowywane dane i metadane? W jaki sposób będą tworzone ich kopie zapasowe?

- pojemność pamięci i miejsce przechowywania danych
- sposób tworzenia kopii zapasowych (automatyczny czy manualny)
- procedury tworzenia kopii zapasowych (częstotliwość aktualizacji, podział obowiązków, proces automatyczny/ręcznie kontrolowany, środki bezpieczeństwa, itd.; zasada 3-2-1: trzy kopie, na dwóch nośnikach, przynajmniej jedna kopia w lokalizacji innej niż miejsce przetwarzania danych)
- informacja czy będą potrzebne specjalne metody przenoszenia danych z urządzeń mobilnych, stanowisk terenowych lub sprzętu domowego na główny serwer w miejscu pracy

3.2. W jaki sposób zostanie zapewnione bezpieczeństwo i ochrona danych wrażliwych w okresie trwania projektu?

- w jaki sposób można będzie odzyskać dane utracone w wyniku incydentu?
- informacja, kto będzie miał dostęp do danych w czasie trwania projektu i jak będzie przebiegać kontrola dostępu do danych (zwłaszcza w przypadku współpracy między kilkoma partnerami)
- polityka ochrony danych w instytucji (aktualne na dzień 31.01.2022 [Zarządzenie Nr 120 Rektora Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie z dnia 4 listopada 2021 r.](#))



[w sprawie wprowadzenia Polityki ochrony danych osobowych](#)
w Szkole Głównej Gospodarstwa Wiejskiego w Warszawie; BARDZO
WAŻNE: przedstawienie polityki ochrony danych instytucji nie
może polegać tylko na wymienieniu nazwy dokumentu, ale należy
również pokrótce omówić najważniejszą kwestię, która będzie
się odnosić do wybranego aspektu organizacji danych; trzeba
pamiętać o tym, że **eksperti nie mają dostępu do informacji
spoza wniosku - eksperci nie mogą korzystać z żadnych
informacji, które są poza wnioskiem, nawet jeżeli w planie
zamieszczony jest link)**

4. Wymogi prawne, kodeksy postępowania.

4.1. Jeżeli będzie miało miejsce przetwarzanie danych osobowych,
w jaki sposób zostanie zapewniona zgodność z przepisami dotyczącymi
danych osobowych oraz ich ochrony?

- informacja czy w procesie pozyskiwania danych niezbędna będzie ich anonimizacja (zupełne wyeliminowanie informacji o charakterze osobowym)
- informacja czy należy usunąć informacje umożliwiające identyfikację osób lub zataić tożsamość uczestników (np. za pomocą pseudonimizacji (przetworzenie danych osobowych w taki sposób, by nie można ich było już przypisać konkretnej osobie, której dane dotyczą, bez użycia dodatkowych informacji; równoległym warunkiem jest, by takie dodatkowe informacje były przechowywane osobno i zostały objęte środkami technicznymi i organizacyjnymi uniemożliwiającymi ich przypisanie zidentyfikowanej lub możliwej do zidentyfikowania osobie fizycznej - wówczas trzeba mieć na uwadze, że księga kodów będzie musiała być przechowywana w innym miejscu, w innej lokalizacji niż same dane, tak, żeby nie było możliwości



rozszyfrowania informacji, które są podane przy samych danych)) przed udostępnieniem danych?

4.2. W jaki sposób planują Państwo zapewnić zgodność z innymi przepisami, takimi jak prawa własności intelektualnej i prawa własności? Jakie przepisy znajdują w tym przypadku zastosowanie?

- wskazanie właściciela danych
- wskazanie zastosowanych licencji
- ograniczenia ponownego wykorzystania danych pochodzących od osób trzecich
- informacja czy przed udostępnieniem danych konieczne będzie uzyskanie odpowiedniego zezwolenia w zakresie praw autorskich

5. Udostępnianie i długotrwałe przechowywanie danych

5.1. Kiedy i w jaki sposób będą udostępniane dane z projektu? Czy istnieją ewentualne ograniczenia i zakazy dotyczące ich udostępniania?

- sposób przekazywania informacji o danych potencjalnym użytkownikom
- kiedy dane zostaną udostępnione (możliwie szybko! Najpóźniej w momencie publikacji wyników badań; NCN wymaga, żeby dane, które stanowią podstawę publikacji zawierającej wyniki będące efektem realizacji projektu, były udostępniane w zaufanym repozytorium na licencji CC0)
- okres przechowywania danych (dane minimum 10 lat, metadane bezterminowo, ale na pewno jeszcze dłużej)
- ograniczenia i przeszkody uniemożliwiające ich pełne lub częściowe udostępnienie (opóźnienie lub ograniczenie np. w celu publikacji, ochrony własności intelektualnej, czy ubiegania się o patenty)



- informacja czy wydawcy czasopism będą wymagać składowania danych na poparcie ustaleń publikacji
- informacja czy udostępnianie danych wymaga zgody uczestników badania

5.2. Jak będzie wyglądać selekcja danych przeznaczonych do utrwalenia i gdzie będą one długoterminowo przechowywane (np. w repozytorium danych, archiwum)?

- Informacja, które dane trzeba zachować, a które zniszczyć z przyczyn wynikających z umów, przepisów prawnych lub regulacji
- sposób podjęcia decyzji o tym, które dane zachować
- procedura selekcji danych przeznaczonych do utrwalenia
- jak i w którym repozytorium zostaną umieszczone (wykazy polecane przez NCN: [CoreTrustSeal](#), [Registry of Research Data Repositories](#))
- koszty umieszczenia w repozytorium

5.3. Jakie metody lub oprogramowanie umożliwiają dostęp do danych i korzystanie z danych?

- metody udostępniania danych
- informacja czy należy je przekształcić do formatu standardowego lub otwartego, aby dane mogły być przechowywane przez dłuższy okres i zachowały długi okres ważności (należy korzystać z formatów otwartych lub przynajmniej szeroko używanych)
- informacja czy do skanowania lub konwersji niezbędny będzie dodatkowy sprzęt lub oprogramowanie
- mechanizm, jaki posłuży do udostępniania danych (np. odpowiedzi na żądanie, repozytorium)



5.4. W jaki sposób zagwarantują Państwo stosowanie unikalnego i trwale przypisanego identyfikatora (np. takiego jak cyfrowy identyfikator dokumentu elektronicznego (DOI)) dla każdego zbioru danych?

- krótka informacja czy zostanie uzyskany trwale przypisany identyfikator (PID), jeśli tak, jaki? (np. cyfrowe identyfikatory dokumentu elektronicznego, numery dostępowe; nie jest to sekcja, która wymaga rozbudowania)

6. Zadania związane z zarządzaniem danymi oraz zasoby

6.1. Kto będzie odpowiadał za zarządzanie danymi (tj. kto będzie ich opiekunem)?

- czy instytucja posiada personel odpowiedzialny za zarządzanie danymi w trakcie trwania projektu (należy wyraźnie rozróżnić właściciela danych, formalnie ponosi odpowiedzialność za dane, a ich opiekuna (data stewarda); ten ostatni to specjalista ds. jakości danych odpowiedzialny za ich ocenę (środki zaradcze), ale nie bezpośrednio korektę)
- jeśli NIE: kto jest odpowiedzialny za archiwizację i długoterminowe zarządzanie danymi w jednostce
- jeśli TAK: jakie zajmuje stanowisko w obrębie jednostki
- podział ról (kto odpowiada za weryfikację, korektę, kto wdraża plan)

6.2. Jakie zasoby zostaną przeznaczone na cele zarządzania danymi i zagwarantowanie przestrzegania zasad FAIR?

- dodatkowe zasoby do zarządzania danymi (osoby, czas, sprzęt lub oprogramowanie; może się okazać, że projekt ze względu na swoją specyfikę będzie wymagać wsparcia nie tylko data



stewardów, ale również informatyków, specjalistów w zakresie oprogramowania czy infrastruktury informatycznej w jednostce, specjalistów w zakresie RODO)

- koszty związane z zapewnieniem standardów FAIR w projekcie
- sposób opłacenia kosztów

3. Najczęstsze błędy w przygotowywaniu planu zarządzania danymi

- plan zarządzania danymi w języku polskim (błąd formalny; wyklucza projekt z dalszego postępowania – nie może być nawet wtrąceń w języku polskim²)
- informacja, że dane będą opracowywane papierowo (NCN przyjmuje, że dane papierowe są danymi, ale ważne jest to, żeby je digitalizować)
- ignoruje się metadane
- schemat planu powielany przez naukowców z tej samej uczelni – tymczasem jest on zawsze oceniany w kontekście danego projektu badawczego
- deklaracja, że projekt nie generuje żadnych danych, ani nie wykorzystuje istniejących danych (a projekt przewiduje np. badania empiryczne, analizę tekstów)
- brak informacji o tym, że dane będą udostępniane – brak uzasadnienia, dlaczego nie (udostępnianie danych, które stanowią podstawę publikacji, jest warunkiem umownym)
- pojawiają się ogólnikowe stwierdzenia – brak precyzyjnych informacji (np. „zostanie zastosowany najbardziej odpowiedni format dla tego rodzaju danych”)

² Z wyjątkiem DMP dla wniosków składanych w konkursie MINIATURA (stan na 09.02.2022).

Błędy szczegółowe

Pozyskiwanie / wykorzystywanie pozyskanych danych

- nie sprawdza się, czy są dostępne dane (na odpowiednich licencjach) odnoszące się do planowanych badań - kładzie się nacisk na dane, które zostaną wytworzone w projekcie
- w przypadku danych będących w posiadaniu strony trzeciej nie sprawdza się, czy są na pewno otwarte
- brak informacji w jaki sposób będzie dokumentowane pochodzenie danych
- brak informacji na temat formatu i rodzaju danych, ich objętości
- brak uzasadnienia: dlaczego wybrano dany format (zwłaszcza jeśli stosuje się format, który nie jest wykorzystywany szeroko w danej dziedzinie, lub nie jest formatem otwartym)
- nie daje się pierwszeństwa standardowym otwartym formatom
- stosuje się formaty, które mogą wymagać dodatkowych wtyczek
- w DMP nie są uwzględnione wszystkie rodzaje danych i formatów planowanych we wniosku (np. we wniosku - Opisie szczegółowym projektu - są planowane analizy filmów, a wśród rodzaju danych ich nie ma)
- brak informacji, w jaki sposób zostanie zapewniona wysoka jakość danych (np. podczas digitalizowania manuskryptów, czy podczas transkrypcji wywiadów wykonywanej przez kilka osób)

Dokumentacja i jakość

- brak informacji o zakresie / standardzie metadanych, jaki zostanie zastosowany
- brak informacji, w jaki sposób dane będą organizowane (foldery i subfoldery, nazwy plików, wersje)



- brak informacji, w którym miejscu znajdzie się opis dokumentacji danych, by był czytelny
- dane nie są wyszukiwalne / odczytywalne maszynowo
- brak informacji na temat, w jaki sposób dane zostaną zabezpieczone przed nieuprawnioną modyfikacją
- w jaki sposób będzie prowadzona kontrola jakości danych (np. czyszczenia danych) - tak aby mogły być udostępniane i wykorzystywane przez innych
- brak informacji o dodatkowej dokumentacji wskazującej na stosowaną metodologię - co umożliwi replikowalność badań (jak zebrano dane, procedury dotyczące analiz, definicje zmiennych, jednostki pomiaru, założenia)

Przechowywanie i kopie zapasowe

- brak informacji, gdzie dane będą przechowywane
- dane nie są przechowywane zgodnie z zasadą 3-2-1 (trzy kopie, na dwóch nośnikach, jedna kopia w lokalizacji innej, niż ta, gdzie dane są przetwarzane)
- brak informacji o kopiach zapasowych - czy będą tworzone - w jaki sposób - zwłaszcza jeśli dane są zbierane na laptopie w czasie wyjazdu badawczego
- przechowywanie kopii na CD i DVD

Wymogi prawne, kodeksy postępowania

- podawanie bardzo ogólnikowych informacji („będzie przestrzegane RODO”, „będą przestrzegane regulacje uczelniane”)
- brak informacji o tym, jak będą (pseudo) anonimizowane dane (tam gdzie występują wywiady, ankiety)



- informacje nie są spójne z tymi podawanymi w „Kwestiach etycznych” – eksperci oceniają cały wniosek

Udostępnianie i długotrwałe przechowywanie

- brak deklaracji dotyczącej udostępniania danych lub wyjaśnienia dlaczego dane nie będą udostępniane (puste pola)
- dane mają być udostępniane w repozytoriach, które nie spełniają warunków otwartości zgodnie z zasadami FAIR
- deklaracja, że dane będą udostępniane wyłącznie w formie publikacji (nie wiadomo, co to oznacza)
- ogólne stwierdzenie, że dane będą publikowane natychmiast po opublikowaniu pracy bez ograniczeń (ale nie wiadomo, czy wszystkie?)
- brak informacji na temat licencji, które zostaną zastosowane (Creative Commons)

Zadania związane z zarządzaniem danymi oraz zasoby

- kierownik podawany jako jedyna osoba odpowiedzialna za wszystkie działania związane z zarządzaniem danymi
- brak informacji o tym, czy będzie zaangażowany inny personel – np. kto będzie odpowiedzialny za zapewnienie jakości, za przechowywanie danych, za tworzenie kopii zapasowych, za archiwizowanie danych i ich udostępnianie
- brak informacji na temat kosztochłonności – czy opracowywanie, zapewnienie kontroli jakości, udostępnianie, przechowywanie generuje koszty? Czy działania te są bezkosztowe? Jeśli kosztowe – jak zostaną zapewnione środki? Z kosztów pośrednich projektu? Uczelnia? Z zewnątrz?



Podsumowanie

- DMP musi być napisane w języku angielskim
- (meta)dane odczytywane i przetwarzane maszynowo
- digitalizacja (meta)danych
- metadane mają znaczenie!
- DMP - należy konstruować indywidualnie do projektu
- dane (duplikaty) trzeba przechowywać w różnych lokalizacjach
- DMP nie muszą być obszerne - ważne żeby były konkretne
- DMP nie powinny odnosić do dokumentacji spoza wniosku!
- informacje w DMP muszą być spójne z tymi, które zostały podane w innych miejscach wniosku
- udostępnianie danych!
- dane udostępnianie w repozytoriach, które posiadają certyfikat [CoreTrustSeal](#), lub są umieszczone w [Registry of Research Data Repositories](#)
- bardzo ważna jest kontrola jakości danych (zbieranych, przechowywanych, jak również udostępnianych)

Odpowiedzi na kilka wybranych pytań

Czy w ramach projektu można zatrudnić osobę z uczelni do zarządzania danymi, byłyby to koszty kwalifikowalne?

Odp. Do projektu można zaangażować osobę z uczelni do zarządzania danymi (np. jako opiekuna danych - data stewarda) - koszt



wynagrodzenia może być opłacany wyłącznie z kosztów pośrednich. Wynika to z tego, że opiekunowie danych (data stewards) odpowiadają na uczelni nie tylko za dany projekt finansowany z NCN, ale również za inne projekty realizowane na uczelni (w tym inne projekty finansowane przez NCN).

Czy każdorazowo poprawione DMP powinno się dostarczać do NCN z raportem rocznym? Czy dotyczy to tylko raportu końcowego?

Odp. Nie trzeba dostarczać do NCN poprawionych DMP w raportach rocznych. Zmiany wprowadzone w stosunku do planu przedstawionego we wniosku należy wykazać w raporcie końcowym.

Czy koszty publikacji objętych Polityką NCN dotyczące otwartego dostępu można zaplanować zarówno w kategorii open access, jak i w kategorii koszty bezpośrednie?

Odp. NIE. Środki na publikowanie w ramach otwartego dostępu można zaplanować **wyłącznie w kategorii kosztów pośrednich** open access. W ramach kosztów pośrednich można zaplanować koszty na open access związane z udostępnieniem publikacji lub danych badawczych w otwartym dostępie. Koszty te mogą stanowić maksymalnie 2% kosztów bezpośrednich. Ewentualnie, dodatkowe środki na open access mogą zostać przeznaczone z kosztów pośrednich (20% kosztów bezpośrednich).

We wniosku należy planować je w zakładce open access (należy wprowadzić kwotę kosztów OA, która jest planowana na cały okres realizacji projektu). W sekcji Koszty zestawienia zostanie zaprezentowana procentowa wysokość planowanych kosztów.



Koszty w ramach kategorii Koszty pośrednie OA są niewystarczające na sfinansowanie wyników badań w ramach otwartego dostępu. Czy NCN ma jakieś sugestie, jak w takiej sytuacji należy zaplanować budżet projektu i publikowanie wyników badań?

Odp. Open Access zgodny z Planem S obejmuje trzy ścieżki. OA to nie tylko płatny "full Open Access".

Stąd też, zgodne z Polityką NCN dotyczącą otwartego dostępu do publikacji są również te, które zostały opublikowane w czasopismach zamkniętych lub hybrydowych (czyli opłacanych przez subskrybentów, czytelników). Jest to opcja bezkosztowa dla autora (w przypadku czasopism hybrydowych należy się upewnić, że praca będzie publikowana w formie niewymagającym pokrywania kosztu APC). Trzeba jednak pamiętać, że w tym przypadku muszą być spełnione następujące warunki:

- manuskrypt (wersja AAM lub VoR) **musi zostać umieszczony w repozytorium zarejestrowanym w openDOAR** w tym samym czasie co wersja pracy opublikowana przez wydawnictwo (NCN nie pozwala na embargo, czyli na odroczoną zgodę wydawnictwa na opublikowanie AAM, np. kilka tygodni lub miesięcy po opublikowaniu pracy przez wydawnictwo)
- praca musi zostać opublikowana na licencji CC-BY
- wersja AAM musi posiadać poniższą adnotację dotyczącą Strategii Zachowania Praw Autorskich (Rights Retention Strategy):

"This research was funded in whole or in part by [Funder] [Grant number]. For the purpose of Open Access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission".

Poprzez zastosowanie powyższego zapisu, będą chronione prawa autorskie. Adnotację tę należy umieścić już w **zgłoszeniu pierwotnej**



wersji manuskryptu składanej do wydawnictwa (w treści pracy w sekcji „Acknowledgements”, a także w cover letter składanym do wydawnictwa).

Dzięki temu zapisowi Strategii Zachowania Praw Autorskich, na podstawie licencji CC-BY, autorzy będą mieli możliwość dalszego dystrybuowania wersji AAM (lub VoR jeśli zgodzi się wydawnictwo). Strategia ta wspiera prawa autorskie naukowców.

Zgodność czasopisma z POD NCN można zweryfikować w narzędziu Journal Checker Tool (<https://journalcheckertool.org/>). Narzędzie to nie wskazuje jednak na kwalifikowalność wydatków przez NCN.

Czy jeżeli dane są kupowane, a umowa zabrania udostępniania/otwierania tych danych, to czy jest to wystarczające uzasadnienie, aby nie otwierać tych danych po publikacji artykułu (opcja opt-out)?

Odp. Raczej tak (każdorazowo taka sytuacja będzie podlegała ocenie ekspertów). Kwestię tę należy jasno opisać we wniosku lub w raporcie, by eksperci oceniający DMP nie mieli wątpliwości, że jest to uzasadniona sytuacja.

Czy można podać np. zespół repozytorium odpowiedzialny za zarządzanie danymi? Czy to musi być osoba z imienia i nazwiska, czy może to być zespół pracowników (w punkcie 6. DMP)?

Odp. Nie trzeba podawać osoby z imienia i nazwiska – najważniejsze żeby podać stanowisko i zakres przypisanych kluczowych obowiązków.



Źródła:

1. NCN (b.r.) Open access - najczęściej zadawane pytania [online]. Dostęp (2022.01.28):
<http://www.ncn.gov.pl/finansowanie-nauki/otwarta-nauka/faq>
2. NCN (2020) Wytyczne dla wnioskodawców do uzupełnienia PLANU ZARZĄDZANIA DANymi w projekcie badawczym [online]. Dostęp (2022.01.28):
https://www.ncn.gov.pl/sites/default/files/pliki/regulaminy/wytyczne_zarzadzanie_danymi_16_03_2020.pdf
3. Pazik-Aybar, A. (2021) Narodowe Centrum Nauki. Plan zarządzania danymi (prezentacja szkoleniowa)

**Szkoła Główna Gospodarstwa
Wiejskiego w Warszawie**

**Biblioteka Główna
im. Władysława Grabskiego**

ul. Nowoursynowska 161
02-787 Warszawa
+48 22 593 57 00
bg@sggw.edu.pl
www.bg.sggw.edu.pl