



Wskazówki dla wypełniających Plan Zarządzania Danymi (NCN)

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?

W punkcie tym należy opisać jakie dane będą wykorzystywane:

- Czy będą to dane pierwotne (nowe) czy wtórne (istniejące wcześniej)?
- Jeśli będą to dane pierwotne należy umieścić tu informację o metodach pozyskiwania, analizowania danych i oprogramowaniu służącym ich wytwarzaniu.
- Jeśli będą wykorzystywane dane wtórne, należy podać ich pochodzenie oraz metodę kontroli ich jakości i spójności.
- W przypadku niewykorzystania żadnych danych wtórnych, należy uwzględnić tu informację: „Ze względu na innowacyjność badań, w projekcie naukowym nie będą wykorzystywane żadne dane już istniejące”.
- Należy wymienić tu również rodzaje wytwarzanych/gromadzonych danych (np. dane eksperymentalne, wyniki badań ankietowych, pomiary, nagrania audio/wideo, dane obliczeniowe itp.).
- Czy w projekcie będą występowały dane osobowe lub wrażliwe (pochodzenie rasowe, poglądy polityczne, przekonania religijne, dane o zdrowiu, dane genetyczne, dane biometryczne, informacje o orientacji seksualnej i przynależności do związków zawodowych)?
- Czy konieczna będzie digitalizacja danych analogowych lub wydanych w formie papierowej (mapy, fotografie, teksty)?
- Informacja o ewentualnej współpracy z innymi ośrodkami naukowymi lub jej braku.

*Przykład**:

W związku z realizacją projektu (...) dane zostaną pozyskane w wyniku przeprowadzonych pomiarów z wykorzystaniem specjalistycznej aparatury oraz odpowiedniego oprogramowania. Przyjmą one postać danych numerycznych oraz obrazowych. Całość materiału badawczego zostanie następnie przetworzona zgodnie ze standardowymi procedurami oraz poddana analizie. Ze względu na nowatorską tematykę badawczą, nie

*Przykłady użyte w instrukcji nie mają uniwersalnego charakteru. W poszczególnych dyscyplinach naukowych dane są gromadzone, opracowywane i archiwizowane w odmienny sposób, dlatego w przedstawionych tutaj przykładach nie uwzględniono wszystkich możliwych uwarunkowań, które mogą pojawić się w indywidualnych przypadkach.



planuje się korzystać z żadnych istniejących danych. Nie jest przewidziana współpraca z innymi jednostkami badawczymi.

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

Należy opisać jaki rodzaj, format i objętość danych pozyskano, wytworzono lub ponownie wykorzystano. Jeśli planowane jest użycie formatu zastrzeżonego, należy wyjaśnić ten wybór (np. kiedy konwersja danych do innego formatu znacznie obniży jakość danych).

- Dane ilościowe: CSV, XLSX, TXT (strukturalne dane tabelaryczne).
- Dane jakościowe: DOCX, PDF, RTF (transkrypcje, opisy).
- Dane multimedialne: JPG, MP4, WAV.
- Kody analityczne: R, PY, JSON (np. pliki R, Python, Jupyter Notebook).
- Szacowana objętość danych.

Przykład:

W ramach projektu powstaną dane eksperymentalne w postaci plików .CSV zawierających pomiary spektroskopowe (ok. 15 GB). Dodatkowo powstaną wyniki analiz numerycznych w formacie .txt i .xlsx. Nie jest przewidziane korzystanie z danych osobowych.

2. Dokumentacja i jakość danych

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

- Jakie standardy opisu danych zostaną zastosowane (np. Dublin Core, DataCite, DDI)?
- Należy opisać dokumentację umożliwiającą ponowne wykorzystanie danych, np. księga kodów, zeszyty laboratoryjne, nagłówki plików czy dokument tekstowy Readme, który zawiera informacje na temat organizacji i zawartości danych badawczych.
- Jaki będzie sposób organizacji danych podczas trwania projektu?

Przykład:

Wszystkie dane badawcze w projekcie będą opisywane za pomocą metadanych, zgodnie ze standardem Dublin Core. Każdy zbiór otrzyma metadane zawierające m.in. tytuł, autorów, opis, słowa kluczowe, format, wersję, licencję oraz informację o sposobie pozyskania danych. Dane będą uzupełnione dokumentacją metodologiczną opisującą sposób gromadzenia, narzędzia badawcze, przetwarzanie oraz ograniczenia danych. Każdy zestaw będzie posiadał słownik danych (opis zmiennych) oraz plik Readme wyjaśniający strukturę, nazewnictwo i organizację plików. Całość dokumentacji i metadanych zostanie udostępniona w repozytorium wraz z danymi.



2.2. Jakie planują Państwo zastosować środki kontroli jakości?

W punkcie tym należy opisać co obejmuje kontrolę jakości danych aby wykazać ich wysoką jakość, uwzględniając personel (zespół badawczy, projektowy, kierownik projektu) odpowiedzialny za te działania.

- Standaryzacja procedur zbierania danych.
- Automatyczne i ręczne sprawdzanie poprawności (wykrywanie braków i wartości odstających).
- Walidacja danych.
- Stosowanie wersjonowania plików przez regularny przegląd kompletności i spójności danych.
- Kalibracja sprzętu.

Przykład:

Jakość danych będzie zapewniona poprzez wieloetapową kontrolę realizowaną przez wyznaczonych członków zespołu. W celu zapewnienia jakości danych planuje się weryfikację poprawności danych podczas wprowadzania (kontrola błędów, spójności i zakresów wartości), stosowanie zdefiniowanych procedur czyszczenia danych, podwójną kontrolę losowej próbki danych przez drugiego członka zespołu, automatyczne logi i wersjonowanie plików zapewniające możliwość śledzenia zmian. Przestrzegane będą kalibracje sprzętu zgodnie z wymaganiami producenta.

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?

- Należy wskazać gdzie dane będą przechowywane (np. serwery instytucji, dysk sieciowy, chmura uczelniana).
- Zaleca się aby kopie zapasowe były tworzone zgodnie z zasadą 3:2:1 (polega na posiadaniu 3 kopii danych, przechowywanych na 2 różnych nośnikach, z których 1 kopia znajduje się w innej lokalizacji).
- Jak często wykonywane będą kopie zapasowe?
- Jak będą zabezpieczone (hasła, szyfrowanie, dostęp ograniczony, programy antywirusowe)?
- Bezpieczeństwo przepływu danych między członkami zespołu – dostęp autoryzowany.
- Kto będzie odpowiedzialny (kierownik projektu PI - Principal Investigator)?

Przykład:

Dane będą przechowywane zgodnie z zasadą 3:2:1 na dyskach komputerów lokalnych oraz z wykorzystaniem platformy e-science.pl w przestrzeni chmurowej NextCloud, która umożliwi przesyłanie plików i udostępnianie ich innym użytkownikom oraz wspólną edycję dokumentów. Pliki robocze będą przechowywane w uporządkowanej strukturze katalogów z zachowaniem zasad wersjonowania. Kopie zapasowe będą tworzone automatycznie raz na



miesiąc i przechowywane na oddzielnym serwerze backupowym. Dostęp do danych będą mieli wyłącznie członkowie zespołu. Za prawidłowe przechowywanie i kontrolę dostępu odpowiada kierownik projektu.

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

Punkt odnosi się do danych wrażliwych, jeśli są one wytwarzane/gromadzone podczas projektu. Należy podać tu informacje na temat przechowywania i zastosowanych zabezpieczeniach, zgodnie z instytucjonalnymi i/lub krajowymi politykami. Określić, kto będzie miał do nich dostęp w czasie trwania projektu oraz strategię działania na wypadek ich utraty. W przypadku współpracy między kilkoma partnerami należy wskazać jak będzie wyglądał przebieg kontroli dostępu do danych.

Przykład 1:

W ramach projektu będzie tworzona aplikacja gromadząca dane medyczne uczestników. Dane te zostaną zapisane w zaszyfrowanej bazie danych i przechowywane na komputerze zabezpieczonym hasłem, bez dostępu do Internetu. Dostęp do nich będą mieli wyłącznie członkowie zespołu upoważnieni do ich przetwarzania. Wszelkie operacje na danych będą rejestrowane w dzienniku audytowym, a serwery będą zabezpieczone i regularnie aktualizowane. Przed rozpoczęciem prac każdy członek zespołu przejdzie szkolenie RODO oraz podpisze oświadczenie o zachowaniu poufności. Wszystkie dane wrażliwe będą anonimizowane na etapie pozyskiwania, a ich udostępnianie poza zespół będzie zabronione. Przesyłanie danych odbywać się będzie wyłącznie za pomocą szyfrowanych kanałów komunikacji. Kierownik projektu będzie odpowiedzialny za wersjonowanie kopii.

Przykład 2:

W projekcie nie będą wykorzystywane żadne dane wrażliwe.

4. Wymogi prawne i 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?

- Jeśli w projekcie nie będą gromadzone żadne dane osobowe, należy to napisać.
- Jeśli dane osobowe będą wytwarzane powinien pojawić się zapis: „Przetwarzanie danych osobowych będzie przebiegało zgodnie z przyjętą na uczelni Polityką Ochrony Danych Osobowych (Załącznik do PO 43/2020) a wszelkie kwestie wymagające wyjaśnienia będą konsultowane z Inspektorem Ochrony Danych (IOD) na Politechnice Wroclawskiej”.
- Należy podać informacje o przyjętych procedurach bezpieczeństwa podczas przetwarzania danych osobowych, np. uzyskanie świadomej zgody, anonimizacja, pseudonimizacja, szyfrowanie danych, ograniczenia dostępu tylko do osób



upoważnionych czy sposób przechowywania danych osobowych – najlepiej na dysku bez podłączenia do Internetu.

- Jeśli istnieją ograniczenia wynikające z wykorzystania danych osobowych, również należy je podać.
- Powinny znaleźć się tu informacje na temat zgód osób biorących udział w badaniach. W tym miejscu powinna znaleźć się wzmianka o przewidywanym ubieganiu się o aprobatę Komisji Bioetycznej, o ile badania będą takowej wymagały.

Przydatne linki:

[Przepisy o przetwarzaniu danych i wzory dokumentów](#)

[Przetwarzanie danych osobowych przez NCN](#)

Przykład 1:

Przetwarzanie danych osobowych będzie przebiegało zgodnie z przyjętą na uczelni Polityką Ochrony Danych Osobowych (Załącznik do PO 43/2020) a wszelkie kwestie wymagające wyjaśnienia będą konsultowane z Inspektorem Ochrony Danych (IOD) na Politechnice Wroclawskiej. Przed rozpoczęciem prac każdy członek zespołu przejdzie szkolenie RODO. W celu zapewnienia bezpieczeństwa, podczas przetwarzania danych osobowych zastosowane zostanie ich szyfrowanie a dostęp do nich będzie ograniczony (wyłącznie uprawnieni członkowie zespołu badawczego).

Przykład 2:

W ramach projektu nie będą gromadzone żadne dane osobowe.

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?

- W punkcie tym, w odniesieniu do danych nowo wytworzonych, powołując się na Regulamin zarządzania własnością intelektualną (Załącznik do ZW 16/2024) należy podać, iż prawa własności do wytworzonych w projekcie danych badawczych przysługują Uczelni.
- W przypadku wykorzystania danych już istniejących, należy zaznaczyć, iż będą one udostępniane zgodnie z postanowieniami zawartymi w umowach licencyjnych i obowiązującymi regulacjami prawnymi.
- Jeśli w projekcie współpracuje kilka ośrodków naukowych, należy wspomnieć o umowie między tymi jednostkami, która będzie regulowała kwestie własności.

Przykład:

Właścicielem danych badawczych wytworzonych podczas realizacji projektu jest Politechnika Wroclawska. Wszelkie prawa własności intelektualnej będą przestrzegane i realizowane zgodnie z obowiązującym Regulaminem zarządzania własnością intelektualną (Załącznik do ZW 16/2024). Wszelkie kwestie dotyczące praw własności reguluje dodatkowo umowa między konsorcjantami.



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?

- Powinna pojawić się tu informacja, iż dane zostaną udostępnione w repozytorium najszybciej jak to możliwe, po ukazaniu się artykułu naukowego lub najpóźniej 12 miesięcy po zakończeniu projektu (w zależności od tego co nastąpi później).
- W przypadku konkursu MINIATURA, który nie musi kończyć się opublikowaniem wyników – po zakończeniu realizacji działania naukowego.
- Jeśli istnieją ograniczenia w udostępnianiu danych badawczych (np. potencjał komercjalizacyjny lub dysponowanie danymi wrażliwymi) w tym miejscu Planu należy uwzględnić taką informację, podać jej przyczynę oraz potencjalne rozwiązanie, np. w postaci zastosowania embarga oraz określić kto będzie miał dostęp do danych i na jakich warunkach.
- Należy uwzględnić tu również nazwę repozytorium, w jakim zamierza się zdeponować dane badawcze i zaznaczyć, iż będzie ono spełniało standardy FAIR. Rekomendujemy repozytorium dziedzinowe bądź repozytorium ogólne – [RepOD](#), w którym Politechnika Wroclawska posiada swoją kolekcję.
- Należy podać licencję, na której zostaną udostępnione dane (CC0 lub CC-BY lub innej o tym samym stopniu otwartości) oraz zadeklarować udostępnianie danych przez co najmniej 10 lat.

Przykład:

Dane zostaną udostępnione najszybciej jak to możliwe, po ukazaniu się artykułu naukowego. Raporty końcowe zawierające wyniki badań terenowych oraz laboratoryjnych zostaną zdeponowane w repozytorium RepOD oraz udostępnione na otwartej licencji (CC-BY) przez minimum 10 lat. Dane udostępnione będą zgodnie z zachowaniem zasad FAIR.

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

- Należy ująć tu informacje na temat osoby odpowiedzialnej za wybór danych do udostępnienia po zakończeniu projektu (przeważanie jest to kierownik projektu) oraz kryteriów wyboru (np. na podstawie ich wartości badawczej lub konieczności walidacji wyników badań zawartych w artykule).
- Powinna pojawić się tu odpowiedź na pytanie, które dane muszą zostać zachowane lub usunięte ze względów umownych, prawnych lub regulacyjnych.
- Należy wskazać repozytorium w jakim przechowywane będą dane oraz zaznaczyć, iż będą udostępniane zgodnie z zasadami FAIR.



Przykład:

Wyboru danych przeznaczonych do udostępnienia dokona kierownik projektu. Selekcja zostanie przeprowadzona według kryterium wartości badawczej oraz walidacji wyników zaprezentowanych w publikacji. Miejszem długotrwałego przechowywania będzie repozytorium RepOD, które umożliwi udostępnianie danych z zachowaniem zasad FAIR.

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

- Zgodnie z wymogami NCN, dane powinny być udostępniane w jak najszerszym dostępie, dlatego punkt ten powinien zawierać informację o tym, że dane zdeponowane w wybranym repozytorium będą dostępne w otwartych (ewentualnie powszechnych) formatach, a korzystanie z nich będzie możliwe bez konieczności użycia specjalistycznego oprogramowania lub narzędzi.
- Jeśli wymóg otwartości nie jest zachowany, należy wskazać tu konkretne narzędzie czy oprogramowanie, które umożliwi potencjalnym użytkownikom dostęp do danych.

Przykład:

Dane przeznaczone do długotrwałego przechowywania i udostępniania poprzez repozytorium zapisane będą w otwartych lub standardowych formatach, a korzystanie z nich nie będzie wymagało użycia specjalistycznych narzędzi.

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

W miejscu tym rekomenduje się sporządzenie uwagi, iż repozytorium, które zostało wybrane do długotrwałego przechowywania i udostępniania danych, zgodnie z zasadami FAIR, każdemu zestawowi danych badawczych nada trwały identyfikator (Persistent Identifiers – PID).

Przykład:

Dane zostaną zdeponowane w repozytorium RepOD, które każdemu zestawowi danych nadaje unikalny cyfrowy identyfikator – DOI.

6. Zadania związane z zarządzaniem danymi oraz zasoby

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

Najczęściej wskazuje się tu kierownika projektu (bez podawania imienia oraz nazwiska, z podaniem numeru ORCID). Osobami mogącymi udzielić wsparcia w zakresie zarządzania danymi badawczymi są pracownicy Biblioteki Politechniki Wroclawskiej/ Sekcji Otwartych Danych Badawczych i można w tym punkcie ująć taką informację.



Przykład:

Osobą odpowiedzialną za zarządzanie danymi będzie kierownik projektu (ORCID: 0000-0000-0000-0000).

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

Najczęściej koszty zarządzania danymi i zapewnienia możliwości ich odnalezienia, dostępu, interoperacyjności i ponownego wykorzystania (FAIR) mieszczą się 2% przyznanych kosztów bezpośrednich i wtedy taką informację należy umieścić w tym punkcie.

Jeśli potrzebne będą inne środki (przechowywanie danych, sprzęt, czas personelu czy opłaty repozytoryjne) powinno się jasno wskazać szacunki kosztu oraz sposób ich pokrycia.

Przykład:

Koszty zarządzania danymi badawczymi oraz przeznaczone na ich długotrwałe udostępnianie zgodne z zasadami FAIR, przez wzgląd na wybór bezpłatnego repozytorium, nie przekroczą limitu 2% środków bezpośrednich grantu.