

1. Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych

1.1. Sposób pozyskiwania i opracowywania nowych danych i/lub ponownego wykorzystania dostępnych danych

- informacje na temat standardów, metod i oprogramowania (nazwa oprogramowania), które posłużą do pozyskiwania lub wytwarzania nowych danych
- podanie rodzaju danych, np. teksty, obrazy, zdjęcia, pomiary, statystyki, próbki fizyczne, kody
- wymienienie procesów kontroli jakości i spójności, np. procesy kalibracji, powtarne pobieranie próbek lub ponowne wykonywanie pomiarów, wykorzystanie narzuconego słownictwa, walidacja wprowadzanych danych, specjalistyczne recenzje
- informacja, czy jakieś dane już istniejące (własne lub będące w posiadaniu strony trzeciej) zostaną ponownie wykorzystane, a jeżeli tak, zapewnienie, że przestrzegane będą wszelkie ograniczenia związane z ponownym wykorzystaniem danych
- informacja, czy konieczne będzie przekształcenie do postaci cyfrowej danych analogowych lub digitalizacja danych wydanych w formie papierowej (np. mapy, fotografie, teksty)
- opisanie sposobu organizacji plików i zarządzania ich różnymi wersjami – posługiwanie się odpowiednim nazewnictwem plików, tj. nadanie nazwy umożliwiającej szybką identyfikację typu danych, oznaczenie każdej wersji kolejnym numerem, podanie dokładnej dat

1.2. Pozyskiwane lub opracowywane dane (np. rodzaj, format, ilość)

- wymienienie formatów danych (najlepiej formaty otwarte lub przynajmniej szeroko stosowane wśród naukowców z danej dyscypliny): pliki tekstowe: .txt, .odt, .xml; ankiety online: html; zdjęcia, obrazy: .tiff, .jpg, .png; audio: .wav; wideo: .mkv, .ogg, .ogv; dane tabelaryczne: .csv, .tab, .ods, .rdata, .sav/spv
- podanie szacunkowej objętości danych wytworzonych, pozyskanych, ponownie wykorzystanych (kB, MB, GB, TB)

2. Dokumentacja i jakość danych

2.1. Metadane i dokumenty (np. metodologia lub pozyskiwanie danych oraz sposób porządkowania danych) towarzyszące danym

- opisanie sposobu organizacji plików i zarządzania ich różnymi wersjami (np. struktura folderów, wersjonowanie danych)
- przedstawienie sposobu i miejsca utworzenia dokumentacji towarzyszącej danym badawczym, w tym informacji o metodologii i organizacji danych w trakcie projektu (np. baza danych z łączem prowadzącym do poszczególnych elementów, pliki README, książka kodowa itd.)
- podanie informacji o standardach metadanych, które będą opisywać dane badawcze, w tym dane zdeponowane w wybranym (zaufanym) repozytorium (międzynarodowe standardy lub

schematy służące do organizacji metadanych, np. Dublin Core, DDI, DataCite – [Digital Curation Centre \(DCC\)](#) prowadzi listę szeroko stosowanych standardów metadanych)

2.2. Stosowane środki kontroli jakości danych

- opisanie procesów kontroli jakości i spójności, np. badania laboratoryjne – procesy kalibracji, powtarzane pobieranie próbek lub ponowne wykonywanie pomiarów, walidacja wprowadzanych danych; badania ankietowe – nieterenowe metody kontroli (np. merytoryczna analiza wypełnionych kwestionariuszy), terenowe metody kontroli (np. powtórny wywiad kontrolny)
- informacja o minimalizacji ryzyka dotyczącego poprawności danych (wskazanie przyjętych w instytucji mechanizmów mających na celu zabezpieczenie danych przed nieuprawnioną modyfikacją), np. pliki zabezpieczone silnym hasłem

3. Przechowywanie i tworzenie kopii zapasowych podczas badań

3.1. Przechowywanie i tworzenie kopii zapasowych danych i metadanych podczas badań

- opisanie procedur tworzenia kopii zapasowych (zasada 3-2-1: trzy kopie, na dwóch nośnikach (np. serwer uczelni, dysk sieciowy), przynajmniej jedna kopia w lokalizacji innej niż miejsce przetwarzania danych; sposób tworzenia kopii zapasowych – automatyczny czy manualny – jeśli manualny, trzeba dopisać, kto będzie odpowiedzialny za tworzenie kopii i odzyskiwanie danych; częstotliwość tworzenia kopii zapasowych – najlepiej nie rzadziej niż raz na tydzień)
- informacja, czy będą potrzebne specjalne metody przenoszenia danych z urządzeń mobilnych, stanowisk terenowych lub sprzętu domowego na główny serwer w miejscu pracy

3.2. Sposób zapewnienia bezpieczeństwa danych oraz ochrony danych wrażliwych podczas badań

- jeśli projekt nie zakłada przetwarzania **danych wrażliwych**, to trzeba napisać np.: „Dane wrażliwe nie będą przetwarzane w tym projekcie” lub „W badaniu nie przewiduje się konieczności zbierania, gromadzenia oraz przechowywania danych wrażliwych.”
- informacja, czy podczas realizacji projektu / działania będą przetwarzane **dane wrażliwe** (szczególny typ danych osobowych: dotyczące pochodzenia rasowego lub etnicznego danej osoby, ujawniające przekonania religijne lub światopoglądowe, uwidaczniające poglądy polityczne, obnażające przynależność do związków zawodowych, dane biometryczne, dane genetyczne (przede wszystkim o kodzie DNA i RNA) i te dotyczące zdrowia (stan zdrowia, przebyte choroby), dane dotyczące seksualności w tym orientacji seksualnej), **dane wrażliwe politycznie** lub **dane podwójnego zastosowania**
- informacja o polityce ochrony danych obowiązującej w instytucji [Zarządzenie nr 11 Rektora Uniwersytetu w Białymstoku z dnia 16 kwietnia 2019 r. w sprawie wprowadzenia Polityki bezpieczeństwa danych osobowych oraz Instrukcji zarządzania systemem informatycznym w Uniwersytecie w Białymstoku](#)) zabezpieczanie danych zgodnych z tą polityką.

- 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)
- informacja o sposobie odzyskania danych utraconych w wyniku incydentu, np. jedna z trzech kopii pozwoli odzyskać utracone dane bez straty ich jakości.

4. Wymogi prawne, kodeks postępowania

4.1. Sposób zapewnienia zgodności z przepisami dotyczącymi danych osobowych i bezpieczeństwa danych w przypadku przetwarzania danych osobowych

- informacja, czy w trakcie realizacji projektu / działania będzie miało miejsce przetwarzanie danych osobowych
- jeśli projekt nie zakłada przetwarzania **danych osobowych**, to wystarczy napisać np.: „Dane osobowe nie będą przetwarzane w tym projekcie” lub „W badaniu nie przewiduje się konieczności zbierania, gromadzenia oraz przechowywania danych osobowych.”
- jeżeli będzie miało miejsce przetwarzanie danych osobowych (odpowieź **TAK** na pytanie: „**Czy planowane badania wiążą się z przetwarzaniem danych osobowych**” w formularzu dotyczącym kwestii etycznych) – informacja o sposobach ochrony pozyskanych danych, np. posłużenie się anonimizacją lub pseudonimizacją:
 - anonimizacja – zupełne wyeliminowanie informacji o charakterze osobowym
 - pseudonimizacja – 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
- informacja o oświadczeniu uczestnika badań potwierdzającym jego świadomą zgodę na udział
- powołanie się na odpowiednie regulacje: [Rozporządzenie Parlamentu Europejskiego i Rady \(UE\) 2016/679 z dnia 27 kwietnia 2016 r. \(RODO\)](#) i [Ustawa z dnia 10 maja 2018 r. o ochronie danych osobowych](#)
- informacja o polityce ochrony danych osobowych w instytucji ([Zarządzenie nr 11 Rektora Uniwersytetu w Białymstoku z dnia 16 kwietnia 2019 r. w sprawie wprowadzenia Polityki bezpieczeństwa danych osobowych oraz Instrukcji zarządzania systemem informatycznym w Uniwersytecie w Białymstoku](#)) ochrona danych osobowych zgodnych z tą polityką

4.2. Sposób zarządzania innymi kwestiami prawnymi, np. prawami własności intelektualnej lub własnością. Obowiązujące przepisy

- informacja o polityce dotyczącej praw własności intelektualnej w instytucji ([Uchwała nr 3107 Senatu Uniwersytetu w Białymstoku z dnia 28 września 2022 r. w sprawie Regulaminu](#)

[zarządzania prawami autorskimi, prawami pokrewnymi i prawami własności przemysłowej oraz zasad komercjalizacji w Uniwersytecie w Białymstoku](#)) podanie właściciela praw własności intelektualnej do danych w świetle tej polityki

- informacja, czy w projekcie będą wykorzystywane jakieś dane już istniejące (własne lub będące w posiadaniu strony trzeciej) lub dane, których wytworzenie zostanie zlecone innej jednostce, i czy w związku z tym przed udostępnieniem danych (także dostępnych komercyjnie) wskazane będzie uzyskanie (jeśli to możliwe) odpowiedniego zezwolenia w zakresie praw autorskich ze względu na ewentualne ograniczenia ponownego ich wykorzystania
- informacja o licencji, którą oznaczone będą udostępniane dane (NCN rekomenduje licencję CCO lub CC BY).

5. Udostępnianie i długotrwałe przechowywanie danych

5.1. Sposób i termin udostępnienia danych. Ewentualne ograniczenia w udostępnianiu danych lub przyczyny embarga

- wskazanie planowanego terminu udostępnienia danych (najpóźniej w momencie ukazania się publikacji, bez embarga)
- wskazanie okresu przechowywania danych (co najmniej 10 lat zgodnie z wytycznymi NCN)
- informacja o ograniczeniach i przeszkodach uniemożliwiających pełne lub częściowe udostępnienie danych
- informacja, czy udostępnienie danych wymaga zgody uczestników badania
- informacja, czy wydawcy czasopism będą wymagać dostępu do danych na poparcie ustaleń publikacji (np. w formie tabeli w treści artykułu, jako załącznik, suplement dołączony do artykułu, czy prywatny link do danych, które znajdują się we wskazanych przez wydawnictwo repozytoriach)

5.2. Sposób wyboru danych przeznaczonych do przechowania oraz miejsce długotrwałego przechowywania danych (np. repozytorium lub archiwum danych)

- wskazanie procedury selekcji danych (informacja, które dane będą zachowane, a które zostaną zniszczone z przyczyn wynikających z umów, przepisów prawnych lub regulacji)
- wskazanie repozytoriów, w których zostaną umieszczone dane
- informacja o ewentualnych kosztach umieszczenia danych w repozytorium

5.3. Metody lub narzędzia programowe umożliwiające dostęp do danych i korzystanie z danych

- Informacja, czy dane są zapisane w formacie otwartym lub co najmniej standardowym, czy należy je przekształcić do takich formatów, aby mogły być przechowywane przez dłuższy okres i zachowały długi okres ważności
- informacja, czy do skanowania lub konwersji niezbędny będzie dodatkowy sprzęt lub oprogramowanie
- wskazanie mechanizmu, jaki posłuży do udostępniania danych (np. repozytorium lub odpowiedzi na żądanie – z podaniem wyjaśnienia, dlaczego dane mogą być udostępnione tylko

w tak ograniczonym zakresie)

5.4. Sposób zapewniający stosowanie unikalnego i trwałego identyfikatora (np. cyfrowego identyfikatora obiektu (DOI)) dla każdego zestawu danych

- Informacja, czy zbiór danych uzyska trwale przypisany identyfikator (PID) i jeśli tak, jaki (np. cyfrowe identyfikatory dokumentu elektronicznego (np. DOI, URN, ARK, handle, Purl lub inne numery dostępne). Identyfikator jest najczęściej automatycznie nadawany przez repozytoria danych badawczych. Zwykle jest to identyfikator DOI (digital object identifier).

6. Zadania związane z zarządzaniem danymi oraz zasoby

6.1. Osoba (np. funkcja, stanowisko i instytucja) odpowiedzialna za zarządzanie danymi (np. data steward)

- dokonanie podziału ról i obowiązków, czyli informacja, kto odpowiada za weryfikację, korektę, kto wdraża plan; czy instytucja posiada personel odpowiedzialny za zarządzanie danymi w trakcie trwania projektu (**właściciel danych** – osoba, która formalnie ponosi odpowiedzialność za dane; **opiekun (data steward)** – osoba, która jest specjalistą ds. jakości danych i jest odpowiedzialna za ich ocenę (środki zaradcze), ale nie bezpośrednią korektę).

6.2. Środki (np. finansowe i czasowe) przeznaczone do zarządzania danymi i zapewnienia możliwości odnalezienia, dostępu, interoperacyjności i ponownego wykorzystania danych

- informacja o zasobach, które zostaną przeznaczone na cele zarządzania danymi i zagwarantowanie przestrzegania zasad FAIR, dodatkowe zasoby do zarządzania danymi (osoby, czas, sprzęt lub oprogramowanie), koszty związane z zapewnieniem standardów FAIR w projekcie oraz sposób ich opłacenia.