



Zarządzanie jakościowymi danymi badawczymi

mgr Maciej Melon

**Instytut Filozofii i Socjologii
PAN**

Zadanie realizowane przez Narodowe Centrum Nauki na podstawie zlecenia Ministra Edukacji i Nauki dot. krajowej koordynacji partnerstwa European Open Science Cloud w latach 2022-2023.



NARODOWE CENTRUM NAUKI



Ministerstwo
Edukacji i Nauki

Dlaczego zarządzanie jakościowymi danymi badawczymi jest ważne

1. Powody formalne – jesteśmy zobowiązani do udostępnienia danych z naszego projektu (prawo unijne, prawo krajowe, wytyczne grantodawców)
2. Mamy kontrolę nad postępem w realizacji projektu (wiemy, które dane już zebraliśmy, co pozostało do zebrania, gdzie napotkaliśmy problemy)
3. Mamy kontrolę nad pracą naszego zespołu badawczego (pod warunkiem, że DMP jest regularnie aktualizowany)
4. Możemy na bieżąco analizować jakość pozyskanego materiału i reagować w razie potrzeby
5. Wiemy, na czym oparte są nasze konkluzje (np. w artykułach)

Charakterystyka jakościowych danych badawczych

1. Są to dane wytworzone przez badacza (i, jak niektórzy twierdzą, współtworzone przez osobę badaną)
2. Dominującą formą danych jakościowych są zapisy wywiadów (pliki wideo, pliki audio, transkrypcje)
3. Dane jakościowe to również zdjęcia (np. dokumentacja towarzysząca badaniom terenowym) czy rysunki
4. Warto tu również wymienić materiały towarzyszące, takie jak scenariusze wywiadu czy dyspozycje do wywiadu, notatki terenowe

Zasady FAIR

Findable: odnajdywalne (dane dobrze opisane, wyposażone w trwałe identyfikatory, indeksowane w bazach)

Accessible: dostępne (dane udostępnione w przeznaczonym do tego celu serwisie, np. w repozytorium a nie na stronie internetowej, bez dodatkowych barier technicznych; czasami możliwy będzie dostęp do samych metadanych, zasady przechowywania i udostępniania muszą być zawsze czytelne)

Interoperable: interoperacyjne (dane i metadane w standardowych formatach, tak, żeby można je było łatwo odczytać i wykorzystać razem z innymi danymi)

Reusable: możliwe do ponownego wykorzystania przez innych badaczy (czyli wyposażone w bogaty opis metadanych, z czytelną strukturą plików, z materiałami dodatkowymi, z jasnymi zasadami dostępu)

Zasady otwierania jakościowych danych badawczych

Dane powinny być tak otwarte, jak to możliwe i na tyle zamknięte, na ile to jest konieczne.

[jęz. ang.: as open as possible, as closed as necessary]

Co to w praktyce oznacza?

1. Nie ma obowiązku automatycznego udostępniania wszystkich danych w wolnym dostępie!
2. Czasami otwarte udostępnienie danych jest niemożliwe (głównie ze względu na dobro osób badanych), warto się wtedy zastanowić nad innymi sposobami dostępu (np. dostęp ograniczony, embargo)
3. Niezależnie od sposobu dostępu do danych, metadane muszą być publicznie dostępne
4. Ewentualne zamknięcie dostępu do danych zawsze wymaga uzasadnienia

Plan Zarządzania Danymi

Ogólne Informacje

Dlaczego Plan Zarządzania Danymi jest ważny w przypadku danych jakościowych?

- Jest to „żywy” dokument – jest to szczególnie ważne w przypadku projektów jakościowych, które są podatne na zmienność
- Umożliwia śledzenie postępów w realizacji projektu – co już mamy, czego jeszcze potrzebujemy a co nam nie wyszło (np. rozmówca odmówił wywiadu)
- Daje wgląd w jakość danych (w jaki sposób i kiedy będziemy reagować np. na zbyt krótkie wywiady czy niskiej jakości transkrypcje?)
- Porządkuje kwestie opisu i organizacji danych (warto od początku projektu się do tego przyłożyć, będziemy mieli mniej pracy przy okazji archiwizowania i udostępniania danych)

Dlaczego Plan Zarządzania Danymi jest ważny w przypadku danych jakościowych?

- Określa dostęp do danych: kto i na jakich zasadach będzie miał dostęp do których danych? Warto tu również uwzględnić zmienność zespołu badawczego w czasie
- Uwzględnia prawa osób trzecich, zgody, kwestie własności danych
- Opisuje przechowywanie, kopie zapasowe, udostępnianie danych
- Reguluje postępowanie w razie utraty danych

Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych

Opis danych

Sposób pozyskiwania lub wytwarzania nowych danych lub ponownego wykorzystywania danych istniejących

Dane pierwotne (nowe), np.:

- Wywiady audio lub wideo (twarzą w twarz lub online) i ich transkrypcje
- Zdjęcia osób badanych, inna dokumentacja foto (np. zdjęcia mieszkań osób badanych)
- Co nam posłuży do wytworzenia tych danych? Np. rejestratory cyfrowe, kamery, oprogramowanie typu Zoom czy Skype?
- Kto i w jaki sposób stworzy transkrypcje wywiadów? Czy np. będzie to firma zewnętrzna czy też członkowie zespołu badawczego?

Opis danych

Sposób pozyskiwania lub wytwarzania nowych danych lub ponownego wykorzystywania danych istniejących

Dane wtórne (istniejące wcześniej):

- Czy będziemy korzystać z danych istniejących wcześniej? Jeśli tak, to do kogo należą? Jakie są zasady dostępu do nich?
- W jaki sposób będziemy w stanie udokumentować pochodzenie danych? (Np. umowa o udostępnienie, linki do udostępnionych już zbiorów danych czy do innych źródeł internetowych)
- W jaki sposób skorzystamy z tych danych? Czy np. zamierzamy digitalizować dane analogowe?
- Przykłady danych wtórnych: istniejące już wywiady na interesujący nas temat, pochodzące z zakończonych projektów, materiały archiwalne, artykuły prasowe

Opis danych

Jakie dane będą pozyskiwane lub wytwarzane w projekcie?

- Np. planujemy 100 wywiadów audio. Nagrania zapiszemy w formacie wav a transkrypcje w docx. Zrobimy również dokumentację foto (300 zdjęć, które zapiszemy w formacie tiff)
- Będziemy musieli uzasadnić wybór tych formatów (w szczególności docx jest formatem zamkniętym, ale bardzo rozpowszechnionym i możliwym do odczytania przez wolne oprogramowanie)
- Jaka będzie szacunkowa objętość danych wytworzonych, pozyskanych, ponownie wykorzystanych w projekcie? Np. wywiady zajmą 60GB a pozostałe materiały 10 GB (te szacunki powinniśmy potem na bieżąco zmieniać w DMP)

Dokumentacja i jakość danych

Dokumentacja i jakość

Metadane dotyczące danych jakościowych

- Metadane to „dane o danych” – nawet, jeśli same dane nie są dostępne w otwarty sposób, metadane są dostępne i mogą (a nawet powinny) być łatwo odczytywane przez ludzi i maszyny
- Do opisu danych jakościowych można używać metadanych w różnych standardach (np. DataCite czy oparte na nim Open-AIRE)
- Nie trzeba stać się specjalistą od schematów metadanych, ale warto wiedzieć, które z nich obsługuje repozytorium wybrane przez nas do zdeponowania w nim danych
- Schematy metadanych różnią się nieco zakresem informacji, które trzeba podać. Poza tym te same informacje mogą być obligatoryjne lub opcjonalne w zależności od systemu metadanych (np. informacja o instytucji finansującej grant)
- Trzeba więc wiedzieć, które „dane o danych” musimy zbierać – z reguły nie będzie tu jednak niespodzianek, bo informacje te to np. tytuł projektu, autor, afiliacja, opis zbioru danych, słowa kluczowe czy właśnie instytucja finansująca i nr projektu. Repozytoria ułatwiają to zadanie – np. Repozytorium Danych Społecznych wskazuje, które pola są obligatoryjne a które nie

Dokumentacja i jakość

Metadane – przykład pól obowiązkowych i fakultatywnych

*Wymagane pola oznaczone są gwiazdką

Metadane cytowania ^

Tytuł* ?

Tłumaczenie tytułu ?

Treść ? **Język tłumaczenia** ?

Autor* ?

Nazwisko i imię autora* ? **System identyfikacji** ?

Identyfikator ?

ROR instytucji afiliującej ?

Osoba do kontaktu* ?

Nazwisko i imię ?

E-mail* ?

Opis* ?

Treść* ?

Wybierz...

- Wybierz...
- ORCID
- ISNI
- LCNA
- VIAF
- GND

Dokumentacja i jakość

Metadane – przykład wypełnionych pól

Trwały identyfikator zbioru danych ⓘ	doi:10.18150/G2KWLV
Data publikacji ⓘ	2022-09-30
Tytuł ⓘ	Między sztuką a badaniem naukowym. Ewaluacja działalności artystycznej w systemach ewaluacji nauki
Tłumaczenie tytułu ⓘ	Between art and research: Evaluation of creative arts in performance-based research funding systems (angielski)
Autor ⓘ	Lewandowska, Kamila 0000-0002-2189-239X – Akademia Teatralna im. A. Zelwerowicza w Warszawie
Osoba do kontaktu ⓘ	Aby nawiązać kontakt, skorzystaj z linku "Kontakt" znajdującego się powyżej. Lewandowska, Kamila (Akademia Teatralna im. A. Zelwerowicza w Warszawie)

Dokumentacja i jakość

Metadane – przykład wypełnionych pól

Język danych ⓘ polski

Producent ⓘ Lewandowska, Kamila (Akademia Teatralna im. A. Zelwerowicza w Warszawie)
Smolarska, Zofia (Akademia Teatralna im. A. Zelwerowicza w Warszawie)

Informacje o grantcie ⓘ Narodowe Centrum Nauki (NCN) – SONATA – 2019/35/D/HS5/00009

Dystrybutor ⓘ Archiwum Danych Jakościowych (ADJ)

Deponujący ⓘ Lewandowska, Kamila

Data zdeponowania ⓘ 2022-09-29

Powiązany zbiór danych ⓘ Lewandowska, Kamila: Between art and research: Evaluation of creative arts in performance-based research funding systems [dane]. Lewandowska, Kamila, Akademia Teatralna im. A. Zelwerowicza w Warszawie [producent]. PADS. Repozytorium Danych Społecznych [wydawca], 2022. <https://doi.org/10.18150/JREG5Y>

Dokumentacja i jakość

Dokumentacja dotycząca danych jakościowych

- Oprócz metadanych, ważna jest również dokumentacja dotycząca pozyskiwania danych: np. opis metodologii badania, dyspozycje do wywiadów, scenariusze, notatki terenowe – wszystko to, co pozwoli potencjalnym użytkownikom, nie znającym kontekstu badania, odczytać i zinterpretować zbiór danych
- Wśród informacji o naszych danych warto również opisać strukturę folderów czy konwencję nazw plików. Można to zrobić np. w pliku `readme.txt`, dołączonym do zbioru danych

Dokumentacja i jakość

Dokumentacja dotycząca danych jakościowych – spis wywiadów

2. Zawartość zbioru danych				
ID wywiadu	ID osoby badanej	Data wywiadu	Dyscyplina naukowa	Uwagi
A_5	A_5	17.02.2021	sztuki filmowe i teatralne	ID osoby badanej oznaczone jako „A” oznacza wywiad przeprowadzony z pracownikiem administracji uczelni; podwójne ID (np. A_13, A_14) oznacza wywiad przeprowadzony z dwiema osobami
A_13-14	A_13, A_14	25.02.2021	sztuki filmowe i teatralne	
A_16	A_16	11.03.2021	sztuki filmowe i teatralne	
A_40	A_40	10.05.2021	sztuki muzyczne	
A_43	A_43	11.05.2021	sztuki plastyczne i konserwacja dzieł sztuki	
A_46	A_46	06.07.2021	sztuki muzyczne	
AR_1	AR_1	18.02.2021	sztuki filmowe i teatralne	ID osoby badanej oznaczone jako „AR” oznacza wywiad przeprowadzony z pracownikiem naukowo-dydaktycznym pełniącym jednocześnie funkcję administracyjną/ kierowniczą na uczelni (np. dziekan, prorektor);
AR_3	AR_3	19.02.2021	sztuki filmowe i teatralne	
AR_15	AR_15	11.03.2021	sztuki filmowe i teatralne	
AR_19	AR_19	04.03.2021	sztuki muzyczne	
AR_20	AR_20	19.03.2021	sztuki filmowe i teatralne	
AR_21	AR_21	24.03.2021	sztuki filmowe i teatralne	
AR_24-25	A_24; AR_25	29.03.2021	sztuki filmowe i teatralne	

Dokumentacja i jakość

Środki kontroli jakości danych

Merytoryczne:

- rozpoczęcie projektu od etapu pilotażowego, żeby sprawdzić, jak narzędzia badawcze wpływają na dane (np. nasz pomysł na scenariusz wywiadu umożliwia uzyskanie zadowolających wywiadów)
- omawianie problematycznych kwestii na spotkaniach zespołu badawczego (szczególnie istotne na początku projektu)

Dokumentacja i jakość

Środki kontroli jakości danych

Techniczne:

- zapewnienie wszystkim członkom zespołu badawczego sprzętu wysokiej jakości, takiego jak rejestratory cyfrowe, kamery czy skanery
- bieżące monitorowanie jakości transkrypcji, jeśli są zaplanowane w projekcie
- sprawdzanie poprawności działania sprzętu przed rozpoczęciem pracy (prostym sposobem jest wykonanie próbnego nagrania, zdjęcia czy skanu, możemy wówczas skorygować parametry)
- ustalenie zasad dostępu do danych tak, żeby nie było ryzyka przypadkowego skasowania czy modyfikacji danych (np. zastąpienia nowszej wersji pliku starszą)

Przechowywanie i tworzenie kopii zapasowych podczas badań

Przechowywanie i kopie zapasowe

Przechowywanie danych i metadanych

- Przechowujemy dane w trzech kopiach, na trzech różnych nośnikach, z których dwa mogą być w miejscu pracy, a jeden gdzie indziej (jest to tzw. reguła 3-2-1)
- Dane powinny być zaszyfrowane a dostęp do nich możliwy po zalogowaniu się
- Nie każdy musi mieć dostęp do wszystkich danych, np. tylko kierownik projektu może mieć takie uprawnienia
- Powinniśmy regularnie tworzyć kopie zapasowe (jak często? to już zależy od specyfiki projektu). Tworzenie kopii minimalizuje ryzyko utraty danych!
- W przypadku badań terenowych / kwerend archiwalnych i zapisywania danych na urządzeniach przenośnych, musimy stworzyć procedurę przenoszenia tych danych do miejsca docelowego (np. codziennie po zakończonej pracy kopiujemy dane do chmury i na dysk przenośny)

Przechowywanie i kopie zapasowe

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

- Dane osobowe uczestników badania przechowujemy oddzielnie - np. danym badawczym możemy nadać oznaczenia słowne. Klucz, który będzie zawierał dane osobowe i odpowiadające im oznaczenia, powinien być przechowywany w innym miejscu, niż dane badawcze
- Członkowie zespołu badawczego, którzy mają do czynienia z danymi osobowymi osób badanych, powinni mieć upoważnienia do przetwarzania danych osobowych, wydane przez instytucję realizującą projekt. Jeżeli przetwarzającym będzie firma zewnętrzna – konieczna jest umowa o powierzeniu przetwarzania
- W przypadku projektu, w którym będzie kilku partnerów, musimy przemyśleć, kto będzie miał dostęp do danych wrażliwych a kto wyłącznie do danych zanonimizowanych – i odpowiednio dostosować nasze dokumenty

Wymogi prawne, kodeksy postępowania

Wymogi prawne, kodeksy postępowania

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

- Dane badawcze a dane osobowe: realizując badanie z udziałem ludzi, będziemy mieli do czynienia z ich danymi osobowymi, czyli wszelkimi danymi umożliwiającymi ich identyfikację (nie tylko imiona i nazwiska!)
- Podstawowe regulacje: RODO i kodeks etyki socjologa (uchwalony przez PTS)
- Wg RODO przetwarzamy dane osobowe osób uczestniczących w badaniu. Osoby te muszą się na to zgodzić i mogą w każdej chwili wycofać zgodę. Muszą też wiedzieć, co się stanie z ich danymi osobowymi - tzn. do kiedy i jak będą przechowywane. Osoby z naszego zespołu badawczego muszą mieć upoważnienia do przetwarzania danych osobowych

Wymogi prawne, kodeksy postępowania

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

- Anonimizacja, czyli usunięcie danych osobowych, oznacza, że zgoda na udostępnienie danych nie jest już formalnie wymagana, ale udostępnienie danych zanonimizowanych bez zgody byłoby wbrew dobrym praktykom badawczym oraz kodeksowi etyki socjologa, który wymaga zgody badanych na późniejsze udostępnienie danych badawczych. **Poza tym całkowita, skuteczna anonimizacja nie zawsze jest możliwa**
- Jeśli udostępniamy dane zanonimizowane, to nie możemy udostępnić zdjęć rozmówców
- Podsumowując: dbamy o pozyskanie od osób badanych zgody, która opisuje, co się stanie z danymi osobowymi i badawczymi (gdzie zostaną zarchiwizowane dane badawcze i na jakich zasadach)
- Przykładowy wzór zgody rozmówcy: http://adj.ifispan.pl/wp-content/uploads/2022/02/formularz_zgody_osoby_badanej_2022-2.docx

Wymogi prawne, kodeksy postępowania

W jaki sposób zostanie zapewniona zgodność z innymi przepisami, takimi jak prawa własności intelektualnej i prawa własności? Jakie przepisy mogą znaleźć zastosowanie w przypadku omawianej dyscyplin/dziedziny?

- Własność danych: dane powstałe w projekcie są własnością instytucji, która z reguły posiada regulamin własności intelektualnej. Czasami trzeba zadbać o odpowiednie zapisy w umowach z współpracownikami
- Dane wtórne są wykorzystywane tak, jak ustalił to ich właściciel, czasami będzie tu potrzebna umowa dwustronna, czasami wykorzystanie tych danych reguluje licencja, na której są udostępnione
- Opinia komisji etycznej, działającej w instytucji

Udostępnianie i długotrwałe przechowywanie danych

Udostępnianie i długotrwałe przechowywanie

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

- Udostępniamy dane pod koniec projektu lub w jego trakcie, jeśli są to tzw. dane powiązane z publikacją, którą raportujemy w sprawozdaniu cząstkowym z realizacji projektu
- Czy zawsze musimy udostępnić dane? Nie, ale musimy to wówczas uzasadnić (powodem może być np. dobro naszych rozmówców). Natomiast metadane zawsze udostępniamy publicznie
- Jw.: udostępnianie zanonimizowanych danych formalnie nie wymaga zgody uczestników, ale etyka badawcza wymaga jednak uzyskania takiej zgody
- Dane powinny być przechowywane przez minimum 10 lat, ale wiele repozytoriów umożliwia przechowywanie bezterminowe

Udostępnianie i długotrwałe przechowywanie

Jak będzie wyglądać selekcja danych przeznaczonych do utrwalenia i gdzie będą one długoterminowo przechowywane?

- Podczas realizacji projektu zebraliśmy bardzo dużo materiału, ale nie musimy udostępniać go w całości. Udostępniamy wszystkie dane, na których oparliśmy wnioski z naszego projektu, czyli w praktyce zapewne wszystkie wywiady czy transkrypcje. Nie musimy jednak udostępniać ich roboczych wersji, nagrań połowy wywiadu, niedokończonych transkrypcji, etc.
- Warto udostępnić materiały towarzyszące, takie jak scenariusze wywiadów, dyspozycje, etc. – to zwiększy dostępność danych dla innych badaczy
- Nadajemy danym czytelną strukturę nazwa i folderów – z myślą o ich przyszłych użytkownikach
- Udostępniamy tylko te dane, do których udostępniania mamy prawo (co oznacza, że niektóre dane, zwłaszcza pozyskane od innych instytucji, musimy zniszczyć)
- Dane udostępniamy w repozytorium ->

Kryteria wyboru repozytorium

Typy repozytoriów:

- Repozytoria instytucjonalne (np. repozytoria uniwersyteckie, często działające przy bibliotekach)
- Kolekcje instytucjonalne w Repozytorium Otwartych Danych (<https://repod.icm.edu.pl/>)
- Repozytorium Danych Społecznych (<https://rds.icm.edu.pl/>)

Wybór repozytorium:

- Repozytoria posiadające certyfikat CoreTrust Seal (w tej chwili jedno polskie repozytorium go posiada: CLARIN-PL)
- Kwestie formalne wyboru (stosowanie FAIR, trwałego identyfikatora (PID), licencja, czas przechowywania, embargo, rozmiar plików, inne)
- Kwestie merytoryczne: „popularność” repozytorium wśród naukowców z dyscypliny, wpływa na widzialność, cytowalność, popularyzacja dorobku naukowego, indeksowanie, inne (np. Zenodo jest bardzo rozpoznawalnym repozytorium)

Rejestry repozytoriów:

- Registry of Research Data Repositories (<https://www.re3data.org/>)

Udostępnianie i długotrwałe przechowywanie

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

- Dostęp do danych może być różny, w zależności od pozyskanych zgód. przykład dostępu otwartego: <https://rds.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/UVNZ3S>
- Przykład dostępu ograniczonego, możliwego po zalogowaniu się w repozytorium i poproszeniu o dostęp pod określonymi warunkami: <https://rds.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/SGPFSL>


[ELITY_SYMBOLICZNE_transkrypcje_wywiadow.zip](#)

Archiwum ZIP - 6.5 MB - 2022-10-05 - 0 Pobrań

MD5: 24e6c2c9b0fa192bc70cc1901648efd2

Warunki dostępu: Ograniczony dostęp - Wyłącznie w celach naukowych, bez możliwości udostępniania go innym osobom i instytucjom

Transkrypcje wywiadów

 Poproś o dostęp

- W obu przypadkach metadane są publicznie dostępne, zgodnie z regulaminem Repozytorium:
5. Dostarczenie Danych Badawczych do RDS jest równoznaczne ze złożeniem przez Użytkownika oświadczenia o treści zgodnej ze wzorcem Creative Commons 0, którego pełne postanowienia dostępne są pod adresem: <https://creativecommons.org/publicdomain/zero/1.0/legalcode> („CC0”). Zakresem przedmiotowym tego oświadczenia objęte są:
- a. metadane dotyczące Danych Badawczych;
 - b. ewentualne prawa związane z Danymi Badawczymi jako zestawieniem plików.

Udostępnianie i długotrwałe przechowywanie

Formaty plików

- Zalecane jest używanie formatów otwartych lub (w drugiej kolejności) takich, które są powszechnie używanym standardem. Obecnie również otwarte oprogramowanie bardzo dobrze radzi sobie z odczytywaniem formatów zamkniętych (takich, jak np. docx), dane jakościowe w zdecydowanej większości mogą być jednak udostępniane w formatach otwartych, takich jak np. odt
- Czasami będzie potrzebna konwersja plików z formatów zamkniętych, np. z wma (format plików dźwiękowych, należący do Microsoft)

Udostępnianie i długotrwałe przechowywanie

W jaki sposób zagwarantować stosowanie unikalnego i trwale przypisanego identyfikatora [ang. Persistent Identifier - PID] dla każdego zbioru danych?

- Najpowszechniejszym trwałym identyfikatorem jest DOI – Digital Object Identifier. Przykładowo - w Repozytorium Danych Społecznych DOI zostaje przypisany w momencie utworzenia roboczej wersji zbioru a staje się aktywny w momencie publikacji zbioru
- Nie musimy dodatkowo się starać o pozyskanie DOI, jest to „usługa” wbudowana w Repozytorium

Udostępnianie i długotrwałe przechowywanie

Wersje zbioru danych

- W momencie publikacji w repozytorium, zbiór danych otrzymuje numer wersji V1
- Drobne poprawki, np. zauważone w opisie błędy, są oczywiście możliwe, ale w wyniku ich wprowadzenia powstanie nowa wersja zbioru: V1.1
- Modyfikacja plików w zbiorze oznacza konieczność opublikowania wersji V2

<https://rds.icm.edu.pl/dataset.xhtml?persistentId=doi:10.18150/C7ABB1>:

Mrozowicki, Adam: PREWORK: Młodzi pracownicy prekaryjni w Polsce i Niemczech. Materiały jakościowe z Polski [dane]. Szkoła Główna Handlowa w Warszawie [producent], Uniwersytet Wrocławski [producent], Warszawa – Wrocław, 2018. Archiwum Danych Jakościowych [dystrybutor], Repozytorium Danych Społecznych [wydawca], 2021. <https://doi.org/10.18150/C7ABB1>, V2

Zadania związane z zarządzaniem danymi oraz zasoby

Zadania związane z zarządzaniem danymi oraz zasoby

- Ustalmy, jaki jest podział zadań w zespole badawczym, związanych z zarządzaniem danymi: kto wytwarza dane, kto je zbiera np. do roboczych folderów, nadaje im sygnatury, tworzy na bieżąco tabelaryczne zestawienie danych - już z myślą o przyszłym udostępnieniu; kto aktualizuje i wersjonuje DMP (pamiętajmy, że to ma być żywy dokument, odpowiadający bieżącym postępom w realizacji projektu)
- Kto zdeponuje dane w repozytorium?
- Jeśli dane będą udostępniane po spełnieniu określonych warunków, to kto będzie udzielał dostępu po zakończeniu projektu?
- W miarę potrzeb korzystajmy ze wsparcia dla procesu zarządzania danymi badawczymi w jednostce badawczej (obsługa prawna, iod, data steward, redaktorzy repozytorium)

Zadania związane z zarządzaniem danymi oraz zasoby

Budżet na cele zarządzania danymi i zagwarantowanie przestrzegania zasad FAIR

- Uwzględnijmy w budżecie zasoby potrzebne do zarządzania danymi (zadania osób odpowiedzialnych za zarządzanie danymi oraz ew. sprzęt, np. dyktafony czy dyski zewnętrzne czy oprogramowanie)
- Czy udostępnienie danych w repozytorium będzie darmowe?
- W projektach NCN: dodatkowe 2% kosztów pośrednich na zadania związane z open access

Wdrażanie i raportowanie Planów Zarządzania Danymi

Wdrażanie i raportowanie Planu Zarządzania Danymi

- Raport z wdrażania – należy opisać stan faktyczny na koniec realizacji projektu: planowano vs. zrealizowano.
- DMP na pewno będzie się zmieniał się w trakcie realizacji projektu – pozyskamy nieco inną ilość danych, niż zakładaliśmy, o innej objętości.

Polecane strony

- Archiwum Danych Jakościowych IFiS PAN: <https://adj.ifispan.pl/>
- Repozytorium Danych Społecznych: <https://rds.icm.edu.pl/dataverse/adj>
- Repozytorium Otwartych Danych: <https://repod.icm.edu.pl/>
- Materiały z projektu „Dziedzinowe Repozytoria Otwartych Danych Badawczych”: <https://drodb.icm.edu.pl/materialy-2/>



Narodowe Centrum Nauki Zespół ds. Otwartej Nauki otwarta.nauka@ncn.gov.pl



GRAMY DLA POLSKIEJ NAUKI

