|  |  |
| --- | --- |
| 1 | **Opis danych oraz pozyskiwanie lub ponowne wykorzystanie dostępnych danych** |
| 1.1 | Sposób pozyskiwania i opracowania nowych danych i/lub ponownego wykorzystania dostępnych danych |
|  | Prowadzone badania opierać będą się w głównej mierze na wykorzystaniu istniejących danych badawczych. Będą to dane w postaci tekstów aktów prawnych, powszechnie dostępnych publikacji naukowych oraz archiwaliów. Dane te pozyskiwane będą z zasobów archiwów, bibliotek i muzeów jak również za pośrednictwem Internetu, m.in. dzięki wykorzystaniu Internetowego Systemu Aktów Prawnych. Dane pozyskane z instytucji kultury będą kopiowane w formie zdjęć. Zebrane dane stanowiły będą zatem w przeważającej ilości materiały dostępne w domenie publicznej. W materiałach, które powstaną w oparciu o pozyskane dane, każdorazowo będą one wskazane i opisane. |
| 1.2 | Pozyskiwanie lub opracowanie danych |
|  | W toku badań naukowych będą wykorzystywane dane tekstowe, czyli zwykły tekst, standardowe dokumenty biurowe, tekst ustrukturyzowany itp. Dane badawcze na potrzeby przechowywania będą gromadzone w powszechnie używanych, standardowych formatach takich jak: doc, docx, pdf, odt. (dla dokumentów tekstowych). Reprodukcje archiwaliów w postaci zdjęć gromadzone będą w formacie jpg. lub png. Ostatecznie przewidziana jest natomiast konwersja zebranych danych do otwartych formatów, niezależnych od platformy i oprogramowania, takich jak txt. Dokładna ilość i wielkość plików nie jest możliwa do oszacowania na etapie tworzenia wniosku, nie powinna ona przekraczać jednak 1TB. |
| 2. | **Dokumentacja i jakość danych** |
| 2.1 | Metadane i dokumenty (np. metodologia lub pozyskiwanie danych oraz sposób porządkowania danych) towarzyszące danym |
|  | Dane zebrane w trakcie badań uporządkowane zostaną na dysku komputera osobistego kierownika projektu. Dane pozyskiwane w trakcie projektu porządkowane będą w określonej strukturze folderów, która dla danych archiwalnych wyglądała będzie następująco (od poziomu najwyższego do najniższego): folder archiwum/muzeum/biblioteki (nazwa)folder zespół (numer i nazwa)folder jednostka (numer i nazwa)poszczególne zdjęcia. Nazwy zdjęć otrzymają następujący format: nr zespołu/kolekcji-sygnatura jednostki/obiektu (numer kolejny zdjęcia) np. 456-21 (1).W przypadku braku w danej instytucji określonego poziomu opisu, pominięty zostanie dla niej odpowiedni folder. Tam gdzie to możliwe nazwy folderów zostaną sporządzone zarówno w języku polskim jak i angielskim. Rozwiązanie takie ułatwi dostęp do danych zarówno dla badaczy z Polski jak również z innych części świata.  Dane nie będące archiwaliami przechowywane będą w osobnym folderze i otrzymają stosowny opis bibliograficzny, umożliwiających ich łatwe wyszukiwanie.  Dla każdego folderu utworzone zostaną pliki Readme.txt, które będą zawierały informacje dotyczące zawartości, pochodzenia, struktury, etc.  Dla poszczególnych folderów zostaną także wykonane metadane w formacie XML zgodnie ze standardem Encoded Archival Description (EAD). Wybór tego formatu metadanych podyktowany jest jego przeznaczeniem dla źródeł archiwalnych, które będą stanowiły większość gromadzonych danych.  Dane bibliograficzne będą zarządzane w menedżerze bibliografii Zotero, który umożliwia m. in. eksport spisu bibliograficznego do standardowych formatów bibliograficznych. Celem będzie takie oznaczenie i przechowywanie danych, aby możliwe było ich komputerowe oczytanie i łatwe wyszukiwanie. |
| 2.2 | Stosowne środki kontroli jakości |
|  | W zakresie prowadzonych badań ryzyko dotyczące poprawności danych można zminimalizować jedynie poprzez dokładną analizę źródeł z wykorzystaniem ich krytyki wewnętrznej i zewnętrznej oraz analizę porównawczą danych zawartych w poszczególnych źródłach. Z większą ostrożnością należy podchodzić do danych zawartych w egodokumentach, które z samej swojej natury obarczone są znaczną subiektywnością. Ocena jakości danych będzie wdrażana na różnych etapach prowadzenia prac badawczych. Do jej prowadzenia zobowiązani będą wszyscy uczestnicy projektu badawczego.  Techniczna jakoś wykonanych zdjęć będzie kontrolowana bezpośrednio po ich wykonaniu. Pozwoli to na natychmiastowe zidentyfikowanie kopii o niskiej jakości oraz uzupełnienie ewentualnych braków w trakcie kwerendy. |
| 3. | **Przechowywanie i tworzenie kopii zapasowych podczas badań** |
| 3.1 | Przechowywanie i tworzenie kopii zapasowych danych i metadanych podczas badań |
|  | Dane badawcze będą przechowywane z zachowaniem zasad ostrożności i najwyższej staranności. Zastosowana zostanie zasada backupu 3-2-1. Dane przechowywane będą na komputerze osobistym (1 kopia) oraz na dysku zewnętrznym przechowywanym w miejscu pracy (2 kopia). Kopia zapasowa przechowywana będzie w oddzielnym miejscu (PIONIER box) i uaktualniana raz w miesiącu. Ponadto każdy z wykonawców będzie posiadał roboczą kopię danych.  Aby zapewnić bezpieczeństwo przepływu informacji pomiędzy wykonawcami projektu przyjmuje się, że każdy z nich zobowiązany jest do archiwizowania zgromadzonych przez siebie danych na co najmniej dwóch nośnikach (komputer osobisty oraz dysk zewnętrzny-roboczy). Przekazanie danych kierownikowi nastąpić musi najpóźniej w ciągu 7 dni od zakończenia kwerendy. Za przeprowadzenie kopiowania danych oraz ich odzyskiwanie odpowiedzialny jest kierownik grantu. |
| 3.2 | Sposób zapewnienia bezpieczeństwa danych oraz ochrony danych wrażliwych podczas badań |
|  | Przechowywane dane nie są danymi wrażliwymi, stąd wdrażanie skomplikowanych systemów ochrony jest niecelowe. Badania będą opierać się na danych powszechnie dostępnych dla naukowców, a także zanonimizowanych bądź uzyskanych na podstawie przepisów o dostępie do informacji publicznej lub/i procedur obowiązujących w danej instytucji. Warto podkreślić, że polityka ochrony danych na Uniwersytecie Mikołaja Kopernika w Toruniu opiera się na podstawie powszechnie obowiązujących regulacji prawnych. W miarę potrzeby wykorzystane będą udostępnione przez UMK bezpieczne usługi służące przechowywaniu danych on-line.  Dostęp do danych w trakcie trwania projektu uzyskają docelowo kierownik projektu oraz wykonawcy. Dostęp dla innych badaczy będzie możliwy na podstawie zgody kierownika projektu.  W przypadku awarii i zagrożenia utraty danych ich odzyskanie będzie możliwe zarówno z nośników zewnętrznych (kopia matka i kopie robocze), jak również z PIONIER box. Minimalizuje to prawdopodobieństwo utraty danych. |
| 4. | **Wymogi prawne, kodeks postępowania** |
| 4.1 | Sposób zapewnienia zgodności z przepisami dotyczącymi danych osobowych i bezpieczeństwa danych w przypadku przetwarzania danych osobowych |
|  | Postępowanie ze zgromadzonymi danymi badawczymi oparte będzie o przepisy Rozporządzenia Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenia o ochronie danych), ustawy z dnia 10 maja 2018 r. o ochronie danych osobowych oraz ustawy z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystaniu informacji sektora publicznego. Jeżeli w trakcie trwania badań, zaistnieje taka potrzeba, podjęta zostanie decyzja o zastosowaniu stosownych środków ochrony danych. Ze względu na brak przechowywania danych wrażliwych w chwili obecnej nie zachodzi taka konieczność. Ponadto na Uniwersytecie Mikołaja Kopernika działa Inspektor ds. Ochrony Danych. W przypadku wątpliwości, kierownik projektu zwróci się do niego z prośbą o wskazanie właściwego i zgodnego z prawem sposobu postępowania z danymi. |
| 4.2 | Sposób zarządzania innymi kwestiami prawnymi, np. prawami własności intelektualnej lub własnością. Obowiązujące przepisy |
|  | W toku prowadzonych badań naukowych zastosowanie będą miały przepisy prawa powszechnie obowiązującego na terytorium Rzeczypospolitej Polskiej. W zakresie praw własności intelektualnej jest to zwłaszcza Ustawa z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych. W przypadku kolizji pomiędzy przepisami o prawie własności, a prawach własności intelektualnej zastosowanie ma wskazana wyżej ustawa. Prawa autorskie do pozyskanych danych takich jak np. notatki badacza będą własnością kierownika projektu bądź sporządzającego je wykonawcy. Pozostałe (wybrane) dane będą docelowo udostępniane na podstawie licencji Creative Commons BY. Wszystkie wykonane metadane zostaną natomiast udostępnione na podstawie licencji CC0. Oprócz wyżej wskazanych kwestii wyznacznikiem dla możliwości podjęcia decyzji odnośnie innych kwestii prawnych będzie także Kodeks Narodowego Centrum Nauki. |
| 5. | Udostępnianie i długotrwałe przechowywanie danych |
| 5.1 | Sposób i termin udostępniania danych. Ewentualne ograniczenia w udostępnianiu danych lub przyczyny embarga |
|  | Wybrane dane badawcze, w tym spis bibliograficzny oraz fotografie, o ile nie wystąpi konieczność wprowadzenia dodatkowych ograniczeń, zostaną udostępnione natychmiast po ukazaniu się publikacji końcowej. W przypadku danych, które objęte zostaną ograniczeniami, lub nie zostaną udostępnione, udostępnione zostaną ich metadane. Dane oraz metadane udostępnione i wyszukiwane będą przez Repozytorium Otwartych Danych UMK na platformie RepOD, tak długo jak będzie ono działało. Zgodnie z wytycznymi Narodowego Centrum Nauki będzie to okres obejmujący co najmniej 10 lat. |
| 5.2 | Sposób wyboru danych przeznaczonych do przechowania oraz miejsce długotrwałego przechowywania danych (np. repozytorium lub archiwum danych) |
|  | Wybór danych zostanie dokonany na podstawie ich wartości badawczej oraz przydatności dla dalszych badań. Przede wszystkim będą to zgromadzone odwzorowania cyfrowe archiwaliów, których udostępnienie ułatwi innym badaczom korzystanie z zasobów archiwów, bibliotek i muzeów. Dane udostępniane będą w Repozytorium Otwartych Danych UMK na platformie RepOD. Repozytorium to, prowadzone przez Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego, zapewnia nieodpłatne przechowywanie i udostępnianie danych badawczych. Udostępnione dane, będą mogły być wykorzystane przede wszystkim w badaniach z zakresu nauk humanistycznych i społecznych. Będą to zarówno badania klasyczne, np. procesualne, jak również dzięki udostępnieniu egodokumentów – mikrohistoryczne, czy też wpisujące się w nurty antropologizujące w poszczególnych naukach. |
| 5.3 | Metody lub narzędzia programowe umożliwiające dostęp do danych i korzystanie z danych |
|  | Jeśli tylko będzie to możliwe, dane zostaną wyeksportowane do formatów, których otwarcie możliwe będzie przy użyciu standardowych narzędzi i aplikacji biurowych. Dla plików tekstowych będzie to format txt, dla fotografii jpg, natomiast dla ewentualnych danych liczbowych csv. Repozytorium Otwartych Danych UMK może przechować i upowszechnić każdy format. Dostęp do repozytorium ma każdy zainteresowany bez logowania i innych ograniczeń. Wykorzystanie danych wymagało będzie zatem wyłącznie zastosowania standardowych narzędzi dostarczanych przez systemy operacyjne jak przeglądarka internetowa czy podstawowy edytor tekstu. Dzięki udostępnieniu danych na wolnych licencjach, oraz sporządzeniu dla nich odpowiednich i bogatych metadanych, spełnione zostaną również zasady FAIR ponieważ dane będą do znalezienia (findable), dostępne (accessible), interoperacyjne (interoperable) i wielokrotnego użytku (reusable). |
| 5.4 | Sposób zapewniający stosowanie unikalnego i trwałego identyfikatora (np. cyfrowego identyfikatora obiektu (DOI) dla każdego zestawu danych |
|  | Dane udostępniane w Repozytorium Otwartych Danych Badawczych RepOD otrzymają identyfikator DOI. |
| 6. | Zadania związane z zarządzaniem danymi oraz zasoby |
| 6.1 | Osoba (np. funkcja, stanowisko i instytucja) odpowiedzialna za zarządzanie danymi (np. data steward) |
|  | Do kierownika projektu należy sprawowanie funkcji administratora danych oraz zarządzanie stosownymi danymi. Do jego obowiązków należało będzie także wdrożenie planu zarządzania danymi i jego ewentualne weryfikowanie podczas realizacji projektu. Pozostali wykonawcy projektu zobowiązani będą do przestrzegania ustalonych procedur postępowania z danymi oraz zapewnienia ich bezpieczeństwa. |
| 6.2 | Środki (np. finansowe i czasowe) przeznaczone do zarządzania danymi i zapewnienia możliwości odnalezienia, dostępu, interoperacyjności i ponownego wykorzystania danych |
|  | W projekcie nie przewiduje się konieczności zapewnienia dodatkowych środków przeznaczonych na przygotowanie danych badawczych do przechowywania i archiwizacji w długim okresie. Działania te będą wykonywane w ramach zadań realizowanych przez członków zespołu. Ponadto, na chwilę składania wniosku działalności repozytorium otwartych danych badawczych ma charakter bezpłatny, co eliminuje konieczność ubiegania się o dodatkowe środki na przechowywanie i udostępnianie danych badawczych w długim okresie oraz zagwarantowanie podstawowych założeń FAIR.  Podczas trwania projektu konieczne jest wyłącznie zapewnienie środków na zakup trzech dysków zewnętrznych, które posłużą do wykonywania kopii zapasowych danych tzn. kopii matki, przechowywanej w siedzibie Uniwersytetu oraz kopii roboczych przechowywanych przez wykonawców. |