



FAQ – WYTYCZNE DOT. REALIZACJI PROJEKTÓW W RAMACH INFOSTRATEG IV

Realizacja wymogu przekazania (udostępnienia) zbiorów danych:

1. W jaki sposób Beneficjent ma przekazać do Centrum pierwsze zbiory danych z fazy I?
 - o Wykonawcy zgodnie z Umową o dofinansowanie powinni przekazać do Centrum zbiory danych utworzone w ramach pierwszego zadania fazy I po 6 miesiącach realizacji projektu. Zbiory danych utworzone w ramach pierwszego zadania fazy I powinny zostać udostępnione w taki sposób, aby zapewnić ich reużywalność. W tym zakresie oczekiwane jest udostępnienie zarówno zebranych danych, jak również w późniejszym etapie projektów prototypów opracowywanych systemów, zapewniając zdalny dostęp do komputera, na którym przechowywane są zbiory danych lub zainstalowany działający system. W przypadku zaistnienia wyjątkowych okoliczności, które uniemożliwią Beneficjentom udostępnienie zbiorów danych oraz systemów w sposób zdalny, Beneficjenci obowiązani są niezwłocznie poinformować o tym Centrum. Strony umowy będą mogły się porozumieć celem przyjęcia odpowiedniego sposobu udostępnienia efektów wypracowanych w realizowanych projektach.

FAQ w odniesieniu do kryteriów przejścia do kolejnej fazy:

Kryterium: Posiadanie/pozyskanie/zebranie adekwatnych danych

1. **Wolumen** – Przekazanie wolumenu w jakiej formie? – zip? Link do chmury? Fizyczna kopia?
 - o Zbiory danych utworzone w ramach pierwszego zadania fazy I powinny zostać udostępnione w taki sposób, aby zapewnić ich reużywalność. W tym zakresie oczekiwane jest udostępnienie zarówno zebranych danych, jak również w późniejszym etapie projektów opracowywanych systemów, zapewniając zdalny dostęp do komputera, na którym przechowywane są zbiory danych lub zainstalowany działający system.
 - o W przypadku zaistnienia wyjątkowych okoliczności, które uniemożliwią Beneficjentom udostępnienie zbiorów danych oraz systemów w sposób zdalny, Beneficjenci obowiązani są niezwłocznie poinformować o tym Centrum. Strony umowy będą mogły się porozumieć celem przyjęcia odpowiedniego sposobu udostępnienia efektów wypracowanych w realizowanych projektach.
 - o Zbiory danych powinny być szczegółowo opisane z wykorzystaniem metadanych odpowiednich do specyfiki realizowanego zadania. Dodatkowo, wraz ze zbiorem



danych niezbędne jest udostępnienie opisu tego zbioru danych, który w sposób wystarczający do uzyskania jego reużywalności scharakteryzuje opracowany zbiór:

- opis, który pozwala wykorzystać te dane do uzyskanych wyników (opis „how to use it” – co to jest, do czego to służy i jakim narzędziem należy się postawić, żeby to wykorzystać);
- opis zbioru danych wykorzystujący: metadane, interoperacyjność semantyczną każdego elementu danych, politykę bezpieczeństwa itd.;

2. Reprezentatywność – Czy powinniśmy przedstawić statystyczną analizę datasetu? W jakiej formie? Jakie parametry interesują NCBiR?

- Nie narzucamy statystycznych charakterystyk datasetu ani innych mierzalnych charakterystyk zbiorów danych. Oczekiwane jest natomiast przedstawienie charakterystyki statystycznej zbioru danych (podmiot udostępniający dane i powiązane z nimi metody powinien je tak opisać, aby inni zainteresowani wykorzystaniem tych danych mogli uzyskać statystycznie podobne wyniki na tym zbiorze i aby wiedzieli, jak go wykorzystać do swoich celów – odtwarzalność (reproducibility), reużywalność).

3. Anonimizacja danych – jeśli dataset nie zawiera osób i znaków zastrzeżonych, czy jest wymagana anonimizacja przez NCBiR? Jeśli tak, to w jakiej formie?

- Obowiązek identyfikacji danych, które wymagają anonimizacji, a także ich anonimizacji leży po stronie dysponenta danych – dane wymagające anonimizacji należy poddać anonimizacji do takiego stopnia, by ponowne ich wykorzystanie nie powodowało szkód dla osób fizycznych lub prawnych z udziałem których zgromadzono dane, ani dla właścicieli obiektów, gdzie zgromadzono dane. Anonimizacja danych powinna być jednocześnie przeprowadzona w taki sposób, by dane mogły być ponownie wykorzystane (reużywalność).

4. Standard/format – jaki standard adnotacji jest wymagany przez NCBiR? Jaki format zdjęć i plików? Jaka struktura drzewa plików?

- Nie narzucamy jednego standardu czy formatu danych. Podstawowym wymogiem jest przedstawienie precyzyjnego opisu zbioru danych, który pozwoli na ponowne wykorzystanie danych (reużywalność).

Kryterium: Ocena niezawodności, sprawności, efektywności i dokładności działania rozwiązania.

1. W jaki sposób powinniśmy przedstawić (zgodnie z wnioskiem), że nasze rozwiązanie realizuje zakładaną dokładność? – film, wykres, wyniki uczenia, testy na arbitralnym zbiorze danych dostarczonych przez NCBiR?



- *W raporcie z realizacji danej fazy powinna znaleźć się co najmniej deklaracja Beneficjenta o wynikach przeprowadzonych testów na opracowanym zbiorze danych i osiągniętej dokładności rozwiązania.*
 - *Przyjęta miara dokładności powinna być spójna z miarą przedstawioną we wniosku o dofinansowanie. Nie narzucamy w tym zakresie żadnych specyficznych wymogów.*
 - *Dodatkowo Beneficjent powinien udostępnić szczegółowy i precyzyjny opis eksperymentu pozwalający na odtwarzalność (reproducibility) i (reżywalność) badania, tj. jak przeprowadzono eksperyment, protokół badawczy eksperymentu pozwalający na ocenę trafności przeprowadzonego badania, tj. eksperyment mierzył to, co miał mierzyć, itd.*
2. W jaki sposób będzie ewaluowana elastyczność rozwiązania? – czy mamy przygotować własny program który będzie zaciągał dane czy tylko dostarczyć model z parametrami?
- *Elastyczność sensu stricto nie podlega ocenie w kryterium oceny. Beneficjent powinien przedstawić do oceny projektu po danej fazie program/system, który będzie wczytywał dane. Dane natomiast mogą być rzeczywiście takimi danymi, jakie Beneficjent stosuje – czy to będą formaty danych, czy pliki danych. Natomiast inny podmiot, dysponujący szczegółowo i precyzyjnie opisanymi danymi stosowanymi przez Beneficjenta, może podjąć próbę wygenerowania danych kompatybilnych z danymi Beneficjenta i przetworzenia ich tym programem. Program/system Beneficjenta powinien być na to gotowy. Program powinien być udostępniany zdalnie (np. przez stronę internetową lub w sposób zaproponowany przez Beneficjenta) z interfejsem pozwalającym wykazać funkcjonalność rozwiązania.*
 - *W przypadku, gdy udostępnienie rozwiązania w sposób zdalny nie będzie możliwe – Beneficjent zobowiązany jest dostarczyć model z parametrami. Dodatkowo Beneficjent powinien udostępnić szczegółowy i precyzyjny opis eksperymentu pozwalający na odtwarzalność (reproducibility) i (reżywalność) badania, tj. jak przeprowadzono eksperyment, protokół badawczy eksperymentu pozwalający na ocenę trafności przeprowadzonego badania, tj. eksperyment mierzył to, co miał mierzyć, itd.*
3. Co w sytuacji kiedy osoba uruchamiająca model/program po stronie NCBiR popełni błąd i nieprawidłowo uruchomi program/model? Czy jest możliwość weryfikacji po naszej stronie tego procesu? Czy Weryfikacja ma możliwość podjęcia drugiej próby w przypadku wyżej zaistniałej sytuacji? Czy weryfikacja odbywa się w sposób całkowicie zdalny czy jednak jest możliwa asysta w formie panelu w siedzibie NCBiR?
- *W trakcie okresu przejściowego i oceny raportów z realizacji projektów w danej fazie będzie istniała możliwość konsultacji z Beneficjentami oraz ponownej weryfikacji*



programów/systemów przez Centrum. Wszystkie przypadki będą rozpatrywane indywidualnie.

4. Jak przebiega proces weryfikacji od strony technicznej? Jaki sprzęt jest wykorzystywany do przeprowadzenia weryfikacji?
- *Niemożliwe jest wskazanie wymogów odnośnie sprzętu wykorzystywanego w procesie weryfikacji opracowanych przez Beneficjentów rozwiązań.*
 - *Testy programów/systemów będą prowadzone z wykorzystaniem podstawowego sprzętu biurowego w oparciu o udostępnione przez Beneficjentów programy/systemy oraz dostarczone instrukcje ich użytkowania. Program powinien być udostępniany zdalnie (np. przez stronę internetową lub w sposób zaproponowany przez Beneficjenta) z interfejsem pozwalającym wykazać funkcjonalność rozwiązania. Rolą Beneficjenta jest zapewnienie zdalnego dostępu do programu/systemu lub w uzasadnionych przypadkach i ustalonych z Centrum bezpośredniego dostępu do swojego środowiska.*