

## Załącznik do KARTY OCENY ZAŁOŻEŃ PROJEKTU INFORMATYCZNEGO NR P508

<b>Nazwa dokumentu:</b> Opis założeń projektu informatycznego pn. „FBC.AI: Cyfrowa dostępność i ponowne wykorzystanie zasobów nauki wspierane sztuczną inteligencją” - wnioskodawca: Minister Nauki i Szkolnictwa Wyższego, beneficjent: Instytut Chemii Bioorganicznej Polskiej Akademii Nauk — Poznańskie Centrum Superkomputerowo-Sieciowe						
Lp.	Organ wnoszący uwagi	Jednostka redakcyjna, do której wnoszone są uwagi	Treść uwagi	Propozycja zmian zapisu	Odniesienie do uwagi	
1.	RA IT	Uwaga ogólna	Jako Beneficjenta wskazano Instytut Chemii Bioorganicznej Polskiej Akademii Nauk – Poznańskie Centrum Superkomputerowo-Sieciowe, podczas gdy Poznańskie Centrum Superkomputerowo-Sieciowe stanowi jedynie jednostkę organizacyjną Instytut Chemii Bioorganicznej Polskiej Akademii Nauk i nie posiada odrębnej osobowości prawnej.	Proszę o analizę i korektę opisu założeń.		
2.	RA IT	Uwaga ogólna	Federacja Bibliotek Cyfrowych agreguje dane z wielu bibliotek cyfrowych i repozytoriów, w tym licznych instytucji wykorzystujących system dLibra oraz innych podmiotów udostępniających zasoby w FBC. Proszę o wyjaśnienie, dlaczego współpracę partnerską w projekcie ograniczono akurat do Uniwersytetu Jagiellońskiego i Uniwersytetu Wrocławskiego oraz jaki będzie wpływ budowanego rozwiązania na pozostałe podmioty korzystające z systemu dLibra i/lub udostępniające zasoby za pośrednictwem FBC.	Proszę o analizę i korektę opisu założeń.		
3.	RA IT	Uwaga ogólna	W związku z faktem, że podmioty będące Beneficjentami lub Partnerami projektu w obecnej chwili planują lub realizują projekty o podobnym charakterze, należy zweryfikować konieczność uspołnienia opisu gromadzonych i przetwarzanych zasobów informacyjnych oraz wzajemnego wykorzystania danych i elementów infrastruktury, w celu zwiększenia ich przydatności i efektywności realizowanych przedsięwzięć.	Proszę o analizę i korektę opisu założeń.		
4.	RA IT	Uwaga ogólna	W dokumencie brakuje szerszych informacji dotyczących planowanego komponentu dot. AI  Nie określono, jaki zakres funkcjonalny ma być realizowany z wykorzystaniem AI, jakie modele lub klasy modeli zostaną zastosowane oraz czy celem użycia AI będzie generowanie informacji, wspieranie decyzji, automatyzacja procesów, analiza treści, czy inny sposób przetwarzania danych. Czy będzie stosowany OCR, HTR, NLP czy RAG, jak będzie realizowane wyszukiwanie semantyczne, jakie będą zasady trenowania modeli, jakie dane będą używane do treningu modeli, czy w ogóle zastosowane modele będą trenowane.	Proszę o analizę i korektę opisu założeń.		

			<p>Brakuje również omówienia kluczowych aspektów związanych z wykorzystaniem AI, w szczególności: compliance, bezpieczeństwo, ochrona danych, jakość wyników oraz odpowiedzialność za rezultaty generowane przez system.</p> <p>Należy uzupełnić informacje dot. AI, m.in. w tabeli pkt. 1.1 oraz w liście systemów w pkt 7.1 w budowanym opisie systemu, który będzie zawierał komponent AI (w przypadku systemów modyfikowanych - w opisie zmiany), a także w widoku komponentów w pkt 7.2, jeśli rozwiązanie dot. AI będzie stanowić wyodrębniony komponent systemu. W przypadku braku miejsca w dokumencie dodatkowe pisemne wyjaśnienie.</p>		
5.	RA IT	Uwaga ogólna	Planowane jest wdrożenie AI jako komponentu/modułu, czy rozważa się jego reużycie lub udostępnienie innym podmiotom?	Proszę o analizę i korektę opisu założeń.	
6.	RA IT	1.1. Identyfikacja problemu i potrzeb	Wymieniono system „dLibra Uniwersytetu Jagiellońskiego” – prawdopodobnie chodzi o „dLab UJ” (dLibra UJ to inna nazwa Biblioteki Cyfrowej UJ).	Proszę o analizę i korektę opisu założeń.	
7.	RA IT	1.1. Identyfikacja problemu i potrzeb 2.4. Produkty końcowe projektu 7.1. Widok kooperacji aplikacji	<p>W projekcie jest mowa o dLab UJ, dLibraUJ, Biblioteka UJ.</p> <p>Należy zweryfikować nazwy systemów i uspoźnić ich stosowanie w całym dokumencie. Nie należy używać synonimów.</p>	Proszę o analizę i korektę opisu założeń	
8.	RA IT	1.1. Identyfikacja problemu i potrzeb	Choć problemy interesariuszy zostały opisane szeroko i wyczerpująco, zwłaszcza w odniesieniu do Beneficjenta i partnerów projektu, wątpliwości budzi stopień, w jakim opisane przedsięwzięcie rzeczywiście odpowiada na ich eliminację. zgodnie z dobrymi praktykami OZPI problemy interesariuszy powinny wynikać z braku lub niewystarczającego rozwoju systemów teleinformatycznych, infrastruktury, oprogramowania lub innych produktów projektu wskazanych w pkt 2.2–2.4. W obecnym kształcie część problemów wydaje się mieć charakter ogólny i organizacyjny, przy ograniczonym powiązaniu z konkretnymi produktami projektu oraz zakresem planowanych modyfikacji systemów.	Proszę o analizę i korektę opisu założeń.	

9.	RA IT	2.1. Cele i korzyści wynikające z projektu KPI	KPI 13-17 – brak wyodrębnienia wartości docelowych. Dwukrotne wskazanie wartości aktualnych.	Proszę o analizę i korektę opisu założeń.	
10.	RA IT	2.1. Cele i korzyści wynikające z projektu	Wśród podstaw strategicznych projektu dla celu 1 i 2 wskazano projekt dokumentu strategicznego „Sprawne i Nowoczesne Państwo”, który nie stanowi obowiązującego dokumentu polityki państwa, gdyż nie został przyjęty przez RM. Wątpliwości budzi również podana poz. 4 Należy zrewidować niniejsze tj. usunąć lub zastąpić obowiązującymi dokumentami strategicznymi i programowymi, w które wpisuje się cel projektu.	Proszę o analizę i korektę opisu założeń.	
11.	RA IT	2.1. Cele i korzyści wynikające z projektu	Część celów projektu została sformułowana wokół technologii, funkcjonalności lub procesów, a nie produktów projektu. Tymczasem zgodnie z dobrymi praktykami OZPI cele projektu powinny odnosić się do docelowych produktów i rezultatów projektu oraz pozostawać spójne z przypisanymi KPI. Proszę o rozważenie przeformułowania celów, np. w kierunku: 1) „Zwiększenie dostępności oraz możliwości inteligentnego wyszukiwania i ponownego wykorzystania zasobów nauki agregowanych w Federacji Bibliotek Cyfrowych poprzez rozwój platformy FBC.AI”; 2) „Przygotowanie interoperacyjnych i przetwarzalnych maszynowo zasobów nauki w postaci AI-ready poprzez digitalizację, poprawę jakości danych oraz opracowanie skalowalnej metodyki przygotowania danych dla instytucji naukowych”	Proszę o analizę i korektę opisu założeń.	
12.	RA IT	2.1. Cele i korzyści wynikające z projektu	Nie wykazano wskaźników jakościowych pozwalających ocenić rzeczywiste efekty projektu, w szczególności jakość kopii cyfrowych, jakość OCR/HTR, jakość danych AI-ready, skuteczność wyszukiwania semantycznego, wykorzystanie warstwy konwersacyjnej, liczbę użytkowników korzystających z FBC.AI, liczbę zapytań realizowanych przez chatbot/warstwę konwersacyjną oraz stopień wykorzystania metodyki AI-ready. Wskaźniki mają głównie charakter obowiązkowy i ilościowy. Konieczność uzupełnienia.	Proszę o analizę i korektę opisu założeń.	
13.	RA IT	2.1. Cele i korzyści wynikające z projektu	Dla wszystkich wskaźników przyjęto częstotliwość pomiaru wyłącznie na zakończenie projektu. W przypadku części KPI, w szczególności dotyczących np. digitalizacji zasobów taki sposób monitorowania wydaje się niewystarczający z punktu widzenia zarządzania postępem rzeczowym projektu. Część wskaźników zwłaszcza dotyczących procesów realizowanych etapowo powinna podlegać bieżącemu monitorowaniu.	Proszę o analizę i korektę opisu założeń	
14.	RA IT	2.1. Cele i korzyści wynikające z projektu	Dla celu 2 wartości KPI dotyczące liczby zdigitalizowanych i udostępnionych dokumentów nie są w pełni spójne z liczbą zasobów wskazanych w pkt 2.3. OZPI (po zsumowaniu 23 950)	Proszę o analizę i korektę opisu założeń	
15.	RA IT	2.3 Udostępnione informacje sektora publicznego i	W odniesieniu do czasopism oraz kart katalogowych UJ nie wynika jednoznacznie, czy zasób jest już zdigitalizowany, czy będzie digitalizowany w projekcie, ani na czym konkretnie będzie polegać „podniesienie jakości” zasobów. Proszę o doprecyzowanie, które zasoby będą	Proszę o analizę i korektę opisu założeń	

		zdigitalizowane zasoby	digitalizowane, które będą jedynie udostępniane, a które będą podlegały poprawie jakości np. poprzez dodanie warstwy OCR, wraz ze wskazaniem zakresu tej poprawy.		
16.	RA IT	2.4 Produkty końcowe projektu	Produkty końcowe wymagają uporządkowania chronologicznego.	Proszę o analizę i korektę opisu założeń	
17.	RA IT	3. Kamienie milowe	W harmonogramie kamieni milowych w obszarze podnoszenia jakości i publikacji danych wskazano głównie kamień dotyczący digitalizacji 100% zasobu. Brakuje wyraźnych kamieni milowych dotyczących poprawy jakości danych, przygotowania danych AI-ready, wdrożenia produkcyjnych usług FBC.AI dla użytkowników końcowych oraz zakupu infrastruktury do digitalizacji i opracowania zasobów.	Proszę o analizę i korektę opisu założeń	
18.	RA IT	5.1 Ryzyka wpływające na realizację projektu	Brakuje ryzyk związanych z wykorzystaniem modeli AI, w szczególności dotyczących ochrony danych, transparentności modeli, ryzyka halucynacji, jakości generowanych odpowiedzi. Z uwagi na charakter projektu niniejsze jest konieczne do uzupełnienia.	Proszę o analizę i korektę opisu założeń	
19.	RA IT	6. Otoczenie prawne	W przedstawionym otoczeniu prawnym projektu nie uwzględniono części kluczowych regulacji związanych z wykorzystaniem technologii AI oraz przetwarzaniem danych, w szczególności Rozporządzenia Parlamentu Europejskiego i Rady (UE) 2024/1689 ustanawiającego zharmonizowane przepisy dotyczące sztucznej inteligencji (AI Act) oraz Rozporządzenia Parlamentu Europejskiego i Rady (UE) 2016/679 (RODO).	Proszę o analizę i korektę opisu założeń	
20.	RA IT	7.1. Widok kooperacji aplikacji	<p>Lista systemów z diagramu kooperacji nie jest odzwierciedlenia w Liście systemów (powinno być 1:1 zgodne)</p> <p>W Liście systemów wykazuje się wszystkie systemy teleinformatyczne z diagramu kooperacji aplikacji odrębnie, z:</p> <ul style="list-style-type: none"> <li>• uwzględnieniem systemów będących produktami projektu,</li> <li>• pominięciem grup systemów</li> </ul> <p>W Liście systemów w kolumnie:</p> <ul style="list-style-type: none"> <li>• „Nazwa systemu” należy zaprezentować skrótowiec lub ew. nazwę systemu teleinformatycznego zgodnie z wykazaną na diagramie kooperacji</li> <li>• „Gestor systemu” należy wprowadzić nawę podmiotu będącego właścicielem systemu</li> <li>• „Opis systemu”</li> </ul> <p><b>Opisy systemów powinny zawierać:</b></p> <ul style="list-style-type: none"> <li>○ Pełną nazwę systemu (na początku pełna nazwa potem sformułowanie „to system wspierający...”)</li> <li>○ Cel utworzenia systemu</li> </ul>	Proszę o analizę i korektę opisu założeń	

			<ul style="list-style-type: none"> <li>○ Informacja o prowadzonych w systemie rejestrach publicznych, jeśli nie jest to zawarte w celu</li> <li>○ Opis głównych grup funkcjonalności (może być lista modułów z krótkimi opisami)</li> <li>○ Informacja o integracjach z innymi systemami krajowymi lub zagranicznymi</li> <li>○ w przypadku grup systemów wstawiamy na początku "System wielokrotny"</li> <li>○ Opis: <ul style="list-style-type: none"> <li>▪ powinien być sformułowany w czasie teraźniejszym, niezależnie od tego czy jest planowany, istniejący czy modyfikowany,</li> <li>▪ nie powinien zawierać odniesień do jakichkolwiek projektów czy przyszłych lub zrealizowanych działań,</li> <li>▪ nie powinien zawierać odniesień do wyjątkowości systemu.</li> </ul> </li> <li>• „Status systemu” należy prezentować wartości zgodne z prezentowanymi na diagramie: <ul style="list-style-type: none"> <li>○ planowane (na diagramie: planowane w projekcie, planowane w innym projekcie)</li> <li>○ modyfikowane (na diagramie: modyfikowane w projekcie, modyfikowane w innym projekcie)</li> <li>○ istniejące (na diagramie: istniejące)</li> </ul> </li> <li>• „Krótki opis ewentualnej zmiany” należy: <ul style="list-style-type: none"> <li>○ wypełnić pole tylko dla systemów „modyfikowanych” w projekcie</li> <li>○ wskazać nazwy modułów budowanych i włączanych do modyfikowanego systemu budowanego/modyfikowanego modułu, o których mowa w pkt 1.1 i w pkt 2.4, jako uzupełnienie nazwy głównego produktu</li> <li>○ opisać budowane/modyfikowane w projekcie funkcjonalności/podmoduły tych modułów</li> </ul> </li> </ul>		
21.	RA IT	7.1. Widok kooperacji aplikacji	<p>Występują niespójności pomiędzy diagramem kooperacji aplikacji, listą systemów oraz listą przepływów danych. Należy stosować te same nazwy systemów na diagramie i tabeli. Muszą być opisane wszystkie przepływy, które są przedstawione na diagramie. Część systemów wskazanych w tabeli nie została przedstawiona na diagramie (m.in. Węzeł Krajowy), natomiast część systemów obecnych na diagramie nie została jednoznacznie odzwierciedlona w tabelach lub występuje pod inną nazwą, np. na diagramie wskazano GRAPHIA, podczas gdy w tabeli systemów figuruje SSH Knowledge Graph.</p> <p>Wątpliwości budzi również przedstawienie na diagramie Biblioteki Cyfrowej UJ i Biblioteki Cyfrowej UWr jako odrębnych systemów, podczas gdy z opisu tabelarycznego wynika, że są to</p>	Proszę o analizę i korektę opisu założeń	

			instancje/komponenty systemu dLibra. Status systemu Kronika jest ponadto niespójny pomiędzy diagramem a tabelą systemów. Na liście systemów brakuje również opisanych systemów udostępniających dane zdigitalizowane oraz Europeany.		
22.	RA IT	7.1. Widok kooperacji aplikacji	W przepływach danych wskazano system FBC.AI, który został częściowo rozbity na komponenty, mimo że na diagramie kooperacji występuje wyłącznie FBC. Zgodnie z informacją z p. 1.1. OZPI „FBC.AI”. FBC.AI nie jest odrębnym systemem, lecz zestawem komponentów wykorzystujących sztuczną inteligencją (SI).	Proszę o analizę i korektę opisu założeń	
23.	RA IT	7.2 Diagram komponentów	Diagram częściowo miesza funkcjonalności biznesowe systemów z warstwą integracyjną, co utrudnia jednoznaczną interpretację architektury rozwiązania. W diagramie komponentów dla większości systemów partnerskich wskazano nowe komponenty w postaci „Wymiany danych”. Jednocześnie z opisów systemów zawartych w tabelach 7.1 oraz 4.2 OZPI wynika, że zakres planowanych modyfikacji polega głównie na integracji z komponentami FBC.AI oraz rozbudowie systemów o API udostępniające dane lub treści. Wątpliwości budzi zatem, czy „Wymiana danych” stanowi rzeczywiście nowy komponent biznesowy systemów, czy raczej mechanizm integracyjny i/lub interfejs API służący komunikacji pomiędzy systemami wrażany w ramach modyfikacji już istniejących komponentów biznesowych wskazanych systemów.  Ponadto wątpliwości budzi, czy systemy Omeka, IZZ, dLab oraz biblioteki cyfrowe (które wydają się stanowić komponenty systemu dLibra) rzeczywiście posiadają wyłącznie pojedynczy komponent biznesowy związany z udostępnianiem lub wymianą danych. Czy może systemy te mają więcej niż jeden komponent.  Konieczna jest ponowna analiza i korekta diagramu komponentów oraz doprecyzowanie charakteru modyfikacji poszczególnych systemów zgodnie z dobrymi praktykami OZPI.	Proszę o analizę i korektę opisu założeń	
24.	RA IT	7.3 Przyjęte założenia technologiczne	Tabela nie zawiera żadnych założeń, co jest niezbędne w przypadku produktów systemowych, w szczególności modyfikowanych. Dotyczy to również standardów wymiany danych, o których m.in mowa w rozporządzeniu KRI i ustawie o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego	Proszę o analizę i korektę opisu założeń	