## **General Information**

Fund	MPTF_00249: Complex Risk Analyt	MPTF_00249: Complex Risk Analytics Fund (CRAF'd)								
FMP Record	MPTF_00249_00022: Risk Anticipa	MPTF_00249_00022: Risk Anticipation Data Hub								
MPTFO Project Id										
Start Date										
End Date										
Applicants	Status	Contact T	уре	Name	e-ma	e-mail			on Telephone	
	Active: 09-Jan-2024 2:08:00 PM	Project Manager		Corrado Scognamillo	corra	orrado.scognamillo@undp.org				
	Active: 17-Jan-2024 6:15:00 PM	Project Manager		Shouryadipta Sarkar	shour	houryadipta.sarkar@undp.org				
Signatories	Signature Process	Role	Name	of Organization		Name		User Email		
	Manual	Signatory	UNDP:	: UNDP (United Nations De P))	evelopm	ent Programme	Mr Turhan Saleh	tu	rhan.saleh@undp g	
Contacts	Contact Type	Name	lame e-mail			Position	Additional e-	mail	Telephone	
	Project Manager	Sun-joo Le	ee	sun-joo.lee@undp.org						
	Project Manager	Marcel Me	yer	marcel.meyer@undp.org						
Description	The project sime to transform an ex-	detine LINDO	intornal	data platform (Origin Diele	Daabba	ord) into a digit	م ما ما اما ام	by provi	ding a one oton	

### Description

The project aims to transform an existing UNDP internal data platform (Crisis Risk Dashboard) into a digital public good by providing a one-stop data warehouse for over 15,000 indicators combining a large set of public datasets to facilitate integrated multidimensional data analytics. This cost-effective solution enables efficient access to valuable data, promotes evidence-based decision-making, and strengthens response efforts in crisis situations, ultimately advancing sustainable development and humanitarian goals.

Key features of the project include:

- 1. Ready-to-Use Data Warehouse: UNDP's data warehouse provides cleansed and pre-processed data allowing analysts to directly use the data without additional Extract, Transform, Load (ETL) processes. The data are structured and linked for conducting integrated analyses, providing centralized access to one well-prepared data warehouse that links all key data sources with relevance for contextual risk analysis in the development sector for the wider community. This cost-effective approach saves time and effort, making data readily available for quick analysis and decision-making.
- 2. **Generative AI Front End:** Analysts will be able to interact with the data warehouse using natural language queries through an AI-driven front end. This innovative feature will empower analysts to extract insights without relying on traditional business intelligence tools, enhancing rapid risk analysis and reducing barriers to entry and costs.
- 3. On-the-Fly Data Analytics: Development and humanitarian partners will be able to perform data analytics on the fly, eliminating the need for expensive infrastructure and data storage. The user-friendly interface will facilitate real-time analysis, visualisation, and advanced modelling for informed decision-making.

Currently, in the data ecosystem, many data organisations and data providers exist with their individual respective datasets, data structures, and formats. However, this disparity in data structure/format leads to users having to pay a large data cleaning tax each time users want to use the data. This is because for practitioners, policymakers, and researchers to use these, they would be required to perform significant data cleaning, preparation and processing to be able to use these datasets for data analysis. The UNDP team has developed a large interoperable data warehouse that is currently internal and is ideal for being positioned as a digital public good for broader consumption to help remove/reduce this data cleaning tax on organisations.

This digital public good, the **risk anticipation datahub (RAD)** once developed will help with achieving two significant steps in significantly increasing the ease of use of data in the crisis risk data ecosystem:

- 1. Easy access to cleaned data: CRAF'd Funding would enable the team currently working on the UNDP internal platform to develop and deploy the infrastructure to democratize access to the platform. It would lead to building an API to open a gateway to a large repository of clean, ready-to-use datasets from myriad organisations such as the World Bank, ACLED, VDem, Fund For Peace, HDRO, and EU JRC, among others. This will in turn help policymakers, researchers, civil society organisations, and developers globally to easily explore and utilize these datasets for various data-driven applications without having to perform significant data processing.
- 2. Natural language to communicate with data: CRAF'd Funding will also enable the development and deployment of a natural language-based interface to interact with this data. This interface will simplify data interaction and increase data product adaptability and usage among non-data-savvv users and practitioners significantly.

The overall digital public good, once developed will enable several types of use cases including, strategic data-led horizon scanning, rapid risk analysis during crises, enhancing SDG monitoring, performing machine learning or other advanced analytics on indicators and fostering collaboration across the humanitarian, peace and development sectors. It will enable easier data-driven decision-making for organizations and governments and promote simpler data exchange for collaborative efforts in risk management and sustainable development.

Universal Markers	Gender Equality Marker	Risk
	GEM0 - The Key Activity is not expected to contribute to GEWE	• Low Risk

Optional Markers	Fragile Context			•	Yes				
Fund Specific Markers	Funding Window / Direct Cost	Funding Windows  • Window A: Critical datasets or and evaluation.	ı complex risks,	including dat	g datasets and initiatives that enable monitoring				
	Call for Proposals / Round	2023 • Third Call for Proposals (Esser	ntial Data Sets o	on Risk Events	rents and Affected Populations)				
Geographical Scope	Geographical Scope	Name of the Region		Region(s) Country			Country		
	Global/Interregional	<ul><li>Americ</li><li>Africa</li><li>Asia</li><li>Europe</li></ul>							
Participating	Participating Organizations	Government/ Multilateral/ NGO/ O	ther	New Enti	ities	Implemen	nting Partners		
Organizations and their Implementing Partners	UNDP - UNDP (United Nations Development Programme (UNDP))								
Programme and	Participating Organization Amount (in USD) Comments								
Project Cost	Budget Requested								
	UNDP	\$500,000.00	TRANCHE 1: Expected March 2024. Tranche 1 will cover API development and an MVP for the Large Language Models. TRANCH Expected January 2025. Tranche 2 will cover the web platform and iterative development / maintenance of the LLMs and API.				e Models. TRANCHE 2: web platform and		
	Total Budget Requested	\$500,000.00							
	Tranches								
	Tranche 1	Tranche 2							
	UNDP Total:	\$350,000.00 <b>\$350,000.00</b>				\$150,000.00 <b>\$150,000.00</b>			
	Other Sources (Parallel Funding)								
	UNDP	\$200,000.00	This involves funding currently deployed to maintain/support the infrastructure and staff for the current team. 200K over 2 years will come from UNDP Crisis Bureau core budget.						
	Total	\$700,000.00							
Thematic Keywords									
Programme Duration	Anticipated Start Date	31-Mar-2024							
	Duration (In months)	21							
	Anticipated End Date	31-Dec-2025							

## **Narratives**

Title	Text
Executive Summary	

Background and General Relevance

Whilst the importance of contextual data for evidence-based decision-making and anticipatory action has in recent years gained wide recognition in international development, a significant challenge looms in the fragmentation and high number of datasets and indicators from a variety of data providers that cover different areas of work.

Recognizing this need, UNDP (UN Development Programme), in collaboration with UNICC (UN International Computing Center), has developed a well-structed data warehouse for contextual risk analysis over the past eight years. This comprehensive repository facilitates internal risk analysis, informing decision-making and supporting policymaking at various levels within the organization. With the proliferation of diverse datasets in the wider community, there is an opportunity to scale and deploy the UNDP Risk Anticipation Hub's (RAH) internal Data Warehouse for contextual risk analysis on a broader scale.

Our project addresses critical data gaps across the humanitarian, development, and peacebuilding contexts by offering a cost-efficient digital public good: cleaned and well-structured data warehouse (Please find list in Documents Uploaded) linking over 15,000 indicators from diverse public datasets. These datasets include sources such as the World Bank, INFORM (Inter-Agency Standing Committee and European Commission), Climate Risk Indices, HDRO (Human Development Report Office), and others. The data are cleaned, pre-processed, and linked for comprehensive, integrated, multidimensional risk analyses. The various data indicators include key risk indices with a focus on development risks, political violence datasets, news datasets and disaster datasets among others organized at their respective geospatial and temporal resolutions. The user-friendly interface and API (Application Programming Interface) will allow easy access for streamlined analysis, visualization, and advanced modeling. The data will be provided in a structure which is cleaned, pre-processed, and linked to each other. This will enable easy to use data for comprehensive, integrated, multidimensional risk analyses, accessible via an advanced LLM (Large Language Model) interface and API.

Our project positions the Risk Anticipation Datahub (RAD) as a tool for broader crisis anticipation and response in crisis-affected and fragile environments, applicable across the <u>Humanitarian-Development-Peace</u> nexus by hosting datasets that serve all 3 dimensions of the nexus.

This approach minimizes the need for users to construct complex data ecosystems or possess specialized skills, promoting easier and streamlined integrated analysis, e.g. of correlations between different datasets. The platform offers clean and ready-to-use data, empowering users to access insights and conduct analyses without additional capacity building or technical barriers. To ensure complementarity and collaboration RAD will explore synergies with existing data platforms such as <a href="https://doi.org/10.100/JDMC.0016.0016">HDMC.0016.0016.0016</a>.

Theory of Change

The Risk Anticipation Data Hub addresses critical data gaps in the development context by offering a comprehensive new data warehouse for contextual risk analysis as a cost-efficient new digital public good. The Data Hub will provide a cleaned and well-structured data warehouse that links a multitude of indicators from a diverse set of public datasets to streamline analysis, visualization and advanced modeling. This will enable comprehensive, integrated, multidimensional risk analyses, accessible via an advanced LLM interface.

IF organizations and analysts in the development context require and utilize accessible, multi-dimensional information for contextual risk analysis

### THROUGH

- (1.1.) An API Accessible data warehouse, in which the API allows analysts to utilize pre-processed and cleansed data directly without the need for additional ETL operations;
- (1.2.) An Al-driven interface that enables users, including those without extensive data expertise, to access and
  use the data warehouse easily, enabling analysis through natural language queries facilitated by generative Al
  models with results produced as data visualizations and natural language;
- (1.3.) A web platform to host the Risk Anticipation Data Warehouse and associated analytics and visualization products

**THEN** the outcome of (1.) INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks will be achieved.

**AND IF** the broader multilateral community, also including humanitarian and peacebuilding actors with common information goals and requirements converges

## THROUGH

• (2.1.) A strengthened network and community of practice on harmonized data and standard practices, wherein partnerships and linkages among organizations in the HDP Nexus data ecosystem are fostered, providing a forum for collaboration, harmonization of data and standard practices, and reduction in duplicated efforts

**THEN** the outcome of (2.) LONG-TERM: A strong HDP Nexus data ecosystem on complex risks will be achieved.

### Methodology

#### 1. API Accessible Data Warehouse

The API will build on the UNDP Crisis Bureau's data warehouse and infrastructure, which already stores many cleaned datasets in well-structured tables using systematic processes. Therefore, only small adjustments will be made to the infrastructure and processes to accommodate API development, namely, 1) data source tagging and indexing, 2) minor schema design, and 3) full data ETL automation.

API development will take place in three stages: 1) API design, 2) implementation of user requirements, and 3) compilation of thorough documentation.

#### 2. Al-Driven Interface

The Al-Driven interface will enable users to input natural language and receive visuals and/or natural language in response. This will enable users with limited data or coding skills to easily access and query needed data and conduct analyses. Users will be able to input a question in text format and in response receive a visual representation of time series trends, comparisons between countries, or simple correlations across factors, etc. A text summary of the visual may also be provided.

We will explore partnerships and synergies with leading technology companies and use existing systems, products, and resources as much as possible to develop the most cost-effective and sustainable solution. The back-end infrastructure will include appropriate server and storage capacity and the bandwidth for a large model to operate. In addition, we will ensure that the infrastructure is conducive to scaling and periodic model tuning or re-training for future improvements. For the front end, we will develop an intuitive and user-friendly interface that accommodates natural language queries, and we will implement data visualization tools to dynamically display results in graphical formats

We will conduct rigorous performance testing through a pilot phase to mitigate some of the limitations of Al-driven interfaces. To address risks associated with Generative Al-based LLMs, we will use existing best practices that are trained on existing high-quality training data, implementing existing structured data templates, refining data sets, and prompting techniques with relevant safety agents. This will primarily include providing a language interface to translate natural language to query data from the Datawarehouse and thus reduce the risk of hallucinations as the Al will have a reduced function in generating narratives. Several other strategies can be utilized to mitigate some of the still-existing limitations of LLMs, such as hallucinations or incorrect responses. These include strategies like tagging and indexing the data, more extensive and targeted training of models, and developing extensive guidelines for users. We will choose the most appropriate combination of these strategies and conduct thorough testing through a pilot phase. We will then make adjustments, correct, and iterate to minimize the limitations of LLMs.

Currently, we only plan to offer the Al-driven interface in English, because all the meta-data is in English. We will explore translations and other language options in the future.

#### 3. Web Platform for Risk Anticipation Datahub

The content of the web platform will be comprised of key risk indicators and metrics, which will be identified in collaboration with major stakeholders in the international ecosystem such as UNDP's internal data teams, <a href="UNDP's Chief Digital Office">UNDP's Chief Digital Office</a>, CRAF'd's (Complex Risk Analytics Fund) stakeholers.

To build the web platform, we will use user-centered design and iterative development principles. We will focus on user needs throughout the design process, ensuring that each component and functionality of the dashboard meets user requirements and is user-friendly. Further, these functionalities will be developed and refined using feedback to iterate and enhance the dashboard continuously, adapting to evolving user needs.

Alignment with and Commitment to CRAF'd Principles

CRAF'd Data Ecosystem Impact & Use Cases

Proposed 3-part, roadmap for Launch/Familiarization

#### Pre Launch:

- 1. Host requirements gathering sessions with stakeholders and develop MVP (Minimum Viable Product)
- 2. Host feedback session once MVP is deployed:
- 3. Identify potential community of practice within the HDP (Humanitarian Development Peace) Nexus Ecosystem
- 4. Develop MVP into First Product

### Launch Event:

- 1. Virtual/In person launch event: This would be a public launch event
- 2. Monthly familiarization videos/case studies with users/community of practice
- 3. Host launch sessions at relevant conferences/events.

#### Post Launch:

- 1. Maintain version control of technology to upgrade/update product offering
- 2. Continuous community engagement/feedback sessions to improve tool

Together, the LLM and web platform of the UNDP Risk Anticipation Datahubub promises impact across these key areas:

1. Rapid risk analysis and expedited decision-making: Plug-and-use data warehouse enables fast risk analysis during crises, supporting timely responses and preparedness measures and contributing to savings in time and cost.

### Examples:

Peacebuilding Commission of Malawi	National Governments can plug into the API to easily build their own Early Warning/Risk Anticipation platforms
German Federal Foreign Office	Foreign office can use the cleaned data warehouse to easily deploy risk monitoring dashboards.

2. Efficient SDG (Sustainable Development Goals) monitoring and early warning signal enhancement: Seamless integration of risk indicators with SDG frameworks and the use of natural language will enable easy monitoring and tracking of SDG performance. Generative Al will also improve signal detection, easy to understand insights, and trend analysis, facilitating the identification of risks and vulnerabilities.

### Examples:

UN Secretariat/UNDP Regional Bureau/Desk officers	LLM based front end can be used to question datasets through natural language for non-data savvy users and derive NLG (Natural Language Generation) derived analysis to support Regional Monthly Reviews. (Regional Monthly Reviews)
INFORM Warning/EU Joint Research Committee	INFORM warning users can use LLM based front end to understand INFORM model outputs better through natural language

3. Unified hub for risk analysis and anticipation: Complementing existing mechanisms as a single access point, the RAD will strengthen risk management efforts and foster collaboration in development and humanitarian sectors, among others.

## Examples:

Humanitarian Peace Development	Users can use the RAD to access data in other data exchanges through an
Nexus Data Exchange	LLM for natural language-based analysis and vice versa.

Synergies with HDX/Similar Data Platforms: To increase complementarity and coverage of datasets, UNDP will work to strengthen the CRAF'd data ecosystem by working more closely with existing data exchange platforms/data portalss such as HDX and other platforms by:

Ensuring regular communication such as monthly meetings with HDX and other platforms on best practices and data standards

Co-organizing joint workshops with teams supporting existing data platforms, at least once a year.

Exploring the potential to build on the current sector-specific HDX data grids for humanitarian data for conceptualizing a broader HDP nexus-wide data schema that is better suited to incorporate both development and humanitarian data and could thereby serve as the basis for a common HDP nexus data platform.

Exploring the potential for data standardization using existing HLX language and other bilateral data standardization best practice implementation.

Exploring the potential for mirroring of public interface across similar data ecosystems such as HDX/IDMC/etc.

#### Sustainability

Leveraging over five years of expertise in data warehouse development and usage of the current internal version of the Risk Anticipation Data warehouse as part of the UNDP Crisis Risk Dashboard platform, this project is now poised for sustainable scaling and public deployment.

The Risk Anticipation Data Hub (RAD) currently employs customized Extract, Transform, Load (ETL) packages that operate seamlessly to guarantee the sustained and reliable delivery of data products. The ETL packages are designed to align with dataset refresh frequencies of the constituent datasets for automatic, up-to-date synchronization. This commitment to timely updates makes the Datahub a reliable, long-term risk related data solution.

We are committed to cultivating a community of practitioners dedicated to the sustainable utilization and growth of this newly established digital public good. Building on the central position of UNDP in the development sector and long-standing partnerships and experiences in our team, we will assemble a diverse community comprising think tanks, academics, organizations, and member states. Continuous engagement with users from the onset, following agile principles, and a consistent focus on user-friendliness are cornerstones of our strategy to build a substantial public user-base, which will ultimately ensure the sustainability of the new public digital good.

In addition to an active public user-community, we are focusing on strengthening partnerships with other leading data platforms in the humanitarian and peacebuilding sectors, most notably HDX and the Peace & Security Data Hub. These partnerships will serve as enablers for an integrated and sustainable data ecosystem serving the entire HDP nexus. Collaborating with HDP nexus partners and leveraging long-term partnerships, also around the newly developed INFORM Warning index, we aim at securing early follow-up funding to ensure long-term financial stability.

#### Scalability

The Risk Anticipation Datahub (RAD) relies on a robust, well-structured, and adaptable infrastructure hosted on the cloud, emphasizing practical scalability for effective deployment. It manages diverse datasets, varying from admin0 on an annual timescale to event-based datasets with lat/long admin granularity and a daily event update frequency. This framework forms a stable base for anticipated expansion in both dataset numbers and granularity, particularly in geospatial and temporal dimensions. Please see the data catalogue in the attachments as a resource. The aim is to increase both geospatial granularity as well as temporal coverage for relevant datasets especially along the following thematic areas:

- 1. Agricultural/Environmental Risk datasets
- 2. Conflict forecast-based datasets
- 3. Global News Monitoring datasets
- 4. National macroeconomic forecasts
- 5. National risk indicators

These datasets would be cleaned and deployed in a ready-to-use structure. Automated data cleaning and modeling routines, along with rigorous quality checks and appropriate citations, reinforce transparency and reproducibility. Future extensions aim to ensure user-friendly scalability for diverse stakeholder groups. Comprehensive documentation, including an extensive data catalog, will support an informed user community.

### Innovation

The project tackles a significant challenge faced by data users in the international development sector. Our solution seeks to solve the challenge of disparate datasets, by providing a new data warehouse that is specifically tailored to allow for integrated risk analysis and advanced analytics for development, taking into consideration the interconnectedness of different thematic areas and consistency over long timescales, by bringing together and connecting datasets from various sources in a well-structured data hub that is accessible via a user-friendly generative Al front end.

Utilizing the Large Data Warehouse for Advanced Machine Learning Models: The project harnesses the potential of the large data warehouse to build advanced machine learning models. By tapping into diverse yet pre-processed, well-structured and connected data sources, the platform can extract valuable insights and patterns, enabling predictive and prescriptive analytics.

Building Statistical Models to Reveal Interactions Between Datasets: The platform's ability to build statistical models goes beyond simple descriptive data analysis or visualization. It enables the overlay and identification of interactions between different datasets, highlighting nuanced correlations and interdependencies.

Rapid Deployment of Information Products through Templates and Clean Data Warehouse: The project employs innovative techniques for the rapid deployment of information products. By leveraging templates, the platform streamlines the creation of data-driven reports, visualizations, and insights.

The provision of a new digital public good for structured data in development fills a gap in the data ecosystem. The partnerships with humanitarian and peacebuilding data platforms that we aim at establishing as part of this project will ultimately form the basis for a new more integrated data ecosystem for the entire HDP nexus.

Cost Effectiveness

This project provides a digital public good that serves as a one-stop shop for development risk data that can be accessed by the public and the international development community. We are exploring partnerships with large technology companies, civil society, academia, and the international community to facilitate cost-effective collaborative approaches that eliminate duplication and redundancies.

As a one-stop shop for data on contextual analysis for development, the Risk Anticipation Data Hub (RAD) substantially reduces system-wide costs for data cleaning, pre-processing, management, and analysis. With a comprehensive and clean data warehouse, analysts and researchers can directly access reliable and ready-to-use data. RAD offers a centralized platform for data storage and hosting, reducing the need for multiple individual data hosting services. By consolidating data in a single platform, organizations can benefit from economies of scale. This eliminates the need for time-consuming and resource-intensive data cleaning processes, allowing organizations to allocate their budget and resources more efficiently towards higher-value activities and analysis.

RAD enables quicker turnaround times and reduction of processing costs. RAD's Generative AI front-end and cleaned data warehouse facilitate quicker turnaround times for data analysis and visualization. This efficiency translates into enhanced cost-effectiveness in the pursuit of sustainable development goals.

## **SDG Targets**

Target	Description					
Main Goals						
Goal 17. Strengthen the	Goal 17. Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development					
TARGET_17.16  17.16 Enhance the Global Partnership for Sustainable Development, complemented by multi-stakeholder partnerships that mobilize and share knowledge, expertise, technology and financial resources, to support the achievement of the Sustainable Development Goals in all countries particular developing countries						
Secondary Goals	Secondary Goals					
Goal 9. Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation						
TARGET_9.c 9.c Significantly increase access to information and communications technology and strive to provide universal and affordable access to the Internet in least developed countries by 2020						

## **SDG Indicators**

Indicator Code	Description
C171601	17.16.1 Number of countries reporting progress in multi-stakeholder development effectiveness monitoring frameworks that support the achievement of the sustainable development goals
C090c01	9.c.1 Proportion of population covered by a mobile network, by technology

## Contribution to SDGs

Participating Organization	% TARGET_17.16	% TARGET_9.c	% Total
UNDP	20	80	100
Total contribution by target	20	80	
Project contribution to SDG by target	20	80	100

## List of documents

Document	Document	Document	Document	Document	Classification	Featured	Status	Modified By	Modified	
	Туре	Source	Abstract	Date					On	

Risk Anticipation Hub Datawarehouse Indicator List 2024.02.13.xlsx	Other Docs	Project	List of indicators in the data warehouse of UNDP. These are datasets that have been collected from various existing public datasets, structured, cleaned and made ready for use.	13- Feb-2024	Confidential	No	Finalized	shouryadipta.sarkar@undp.org	11- Mar-2024 2:51:36 PM	
--	---------------	---------	--	-----------------	--------------	----	-----------	------------------------------	----------------------------------	--

# **Project Results**

Outcome	Output		Description							
INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis- affected settings.										
	1.1. An API endpoint for warehouse.	data		rs analysts to utilize pre-proces additional ETL operations.	sed and cleansed da	ta directly withou				
	Activities									
	Title	Description		Lead Participating Organization	Participating Organization	Other Organizations				
	1.1.1. Define API objectives and use cases in a scoping document.	Understand the cases and ident audience for AP establish the sp sets, functionali features that wi available throug	ify the target I access, and ecific data ties, and	UNDP - UNDP (United Nations Development Programme (UNDP))						
	1.1.2. Complete any necessary back end development or adjustments of the database systems and analysis processes to source and transform the data.			UNDP - UNDP (United Nations Development Programme (UNDP))						
	1.1.3. Develop the API in line with the use cases and functionalities identified in the scoping phase.			UNDP - UNDP (United Nations Development Programme (UNDP))						
	1.1.4. Maintain the API and supporting systems.	Resolve bugs, d version control, updates, and im system over tim consistent upda based on user f technical impro- recommended b partners.	implement aprove the e using ite sprints eedback and vements	UNDP - UNDP (United Nations Development Programme (UNDP))						
	1.1.5. Develop thorough documentation of the data in the Risk Anticipation Data Warehouse and guidance material on API usage.			UNDP - UNDP (United Nations Development Programme (UNDP))						
	1.2. An Al-driven interfactusers to access the data easily.		Large Language models will enable users, including those without extensive date expertise, to access and use the data warehouse easily, enabling analysis throug natural language queries facilitated by generative AI models with results product as data visualizations and natural language.							

ng Other Organizations

Outcome	Output		Description		
	1.2.9. Maintain the LLMs and deployment systems.	Resolve bugs, d version control, updates, and im system over tim consistent upda based on user fi technical improver recommended to partners.	implement approve the se using ate sprints eedback and evements	UNDP - UNDP (United Nations Development Programme (UNDP))	
	1.2.10. Conduct training and outreach activities to provide guidance on utilization of dataoutput and visual output LLMs and the API.			UNDP - UNDP (United Nations Development Programme (UNDP))	
	1.3. A web platform to he Anticipation Data Wareh associated analytics and products.	ouse and			

Outcome	Output		Description						
	Activities								
	Title	Description		Lead Participating Organization	Participating Organization	Other Organizations			
	1.3.1. Define user requirements by conducting user requirement assessments, and identify product roadmap ensuring connections to the API and LLM components.			UNDP - UNDP (United Nations Development Programme (UNDP))					
	1.3.2. Conduct initial and ongoing user centred design activities to ensure that the web elements meet user needs and remain in line with UN and UNDP design principles.			UNDP - UNDP (United Nations Development Programme (UNDP))					
	1.3.3. Define visuals and interactive data UX front end, and conduct information design activities to ensure data structure, processing and web rendering is useful and intuitive and produced in a visually beautiful format.			UNDP - UNDP (United Nations Development Programme (UNDP))					
	1.3.4. Deploy a frontend minimum viable product using kanban/agile management principles with repeated iterations.			UNDP - UNDP (United Nations Development Programme (UNDP))					
	1.3.5. Collect and ingest feedback for MVP and deliver a functioning platform for the data hub in line with user needs, user experience research, and testing.			UNDP - UNDP (United Nations Development Programme (UNDP))					
	1.3.6. Maintain the web platform front end.	Resolve bugs, d version control, updates, and im system over tim consistent updates based on user f technical improver recommended in partners.	implement approve the using ate sprints eedback and evements	UNDP - UNDP (United Nations Development Programme (UNDP))					
2. LONG-TERM: A strong HDP Nexus data ecosystem on complex risks.									
	2.1. A strengthened network community of practice of data and standard pract	n harmonized	are fostered,	and linkages among organizati providing a forum for collabora d reduction in duplicated effort	tion, harmonization o				

Outcome	Output		Description		
	Activities				
	Title	Description	Lead Participating Organization	Participating Organization	Other Organizations
	2.1.1. Establish a network of organizations aiming at the strengthening of a harmonized HDP Nexus data ecosystem.		UNDP - UNDP (United Nations Development Programme (UNDP))		
	2.1.2. Convene periodic workshops and meetings to foster knowledge-sharing and collaboration.		UNDP - UNDP (United Nations Development Programme (UNDP))		

## Signature Indicators

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
No signature inc	dicators available.											

# Imported Fund Outcome / Output Indicators

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
Datasets provided by the project.		This indicator aims to measure the provision and dissemination of datasets by the project to stakeholders.	Internal tracking.	Capacity	Yearly	Global	Number	70	2024	100	2026	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.1. An API endpoint for data warehouse.
	Datasets provided with granularity at the sub- national level or below (spatial resolution).		Internal tracking.	Capacity	Yearly	Global	Number	10	2024	15	2025	
	Datasets provided with at least monthly granularity, (temporal resolution).		Internal tracking.	Capacity	Yearly	Global	Number	5	2024	10	2025	
	Datasets provided with open access.		Internal tracking.	Capacity	Yearly	Global	Number	0	2024	100	2025	

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
	Datasets provided in non- proprietary formats,	E.g., csv, json, xml, txt, sql (not dta, spss or similar proprietary file formats).	Internal tracking.	Capacity	Yearly	Global	Number	5	2024	7	2025	
Publications produced as part of this project.		This indicator aims to measure the number and quality of publications produced by the project, which may include scientific reports, best practices, guidelines, and other types of knowledge products. The indicator reflects the extent to which the project has generated new knowledge, shared best practices, and disseminated findings related to the project's goals and objectives.	Internal tracking.	Capacity	Yearly	Global	Number	0	2024	1	2026	Outcome: 1. INTERMEDIAT Better data, analytics, knowledge, ar capacity on complex risks in fragile and crisis-affecter settings. Output: 1.1. / API endpoint for data warehouse.
Analytics products provided by the project.		This indicator aims to measure the provision and dissemination of analytics products by the project to stakeholders.	Internal tracking.	Capacity	Yearly	Global	Number	0	2024	1	2026	Outcome: 1. INTERMEDIA Better data, analytics, knowledge, a capacity on complex risk: in fragile and crisis-affecte settings. Output: 1.2. An Al-driven interface tha enables user to access the data warehou easily.
	Analytics products with open access.		Internal tracking.	Capacity	Yearly	Global	Number	0	2024	1	2026	

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
Project partners involved in the implementation of the project.		This indicator aims to measure the number project partners ('participating organizations' and 'implementing partners') involved in the implementation of the project.	Internal tracking.	Capacity	Yearly	Global	Number	0	2024	1	2026	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.2. An Al-driven interface that enables users to access the data warehouse easily.
Downloads and/or users of project outputs.		This indicator aims to measure the use and dissemination of project outputs by tracking the number of downloads and/or users of the project outputs.	Surveys, interviews, internal statistics.	Capacity	Yearly	Global	Number	500	2024	7500	2026	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.2. An Al-driven interface that enables users to access the data warehouse easily.
Stakeholders that use project outputs to support crisis action.		This indicator aims to measure the extent to which entities use project outputs for crisis action, including for programming, decision-making, and resource allocation.	Surveys, interviews, analysis of public policy documents/ emergency response plans/ reports, other documents.	Capacity	Yearly	Global	Number	0	2024	10	2026	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.3. A web platform to host the Risk Anticipation Data Warehouse and associated analytics and visualization products.

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
Knowledge and capacity building Initiatives conducted as part of the project.		This indicator aims to measure the provision of knowledge and capacity building initiatives by the project to stakeholders. The indicator reflects the extent to which the project has supported the development of skills, knowledge, and expertise related to the project's goals and objectives.	Internal tracking.	Beneficiaries	Yearly	Global	Number	2	2024	4	2026	Outcome: 2. LONG-TERM: A strong HDP Nexus data ecosystem on complex risks. Output: 2.1. A strengthened network and community of practice on harmonized data and standard practices.
Participants in knowledge and capacity initiatives as part of this project.		This indicator aims to measure the number of individuals who have participated in knowledge and capacity building initiatives provided by the project. The indicator reflects the extent to which the project has engaged stakeholders in the development of skills, knowledge, and expertise related to the project's goals and objectives.	Surveys, registration statistics.	Beneficiaries	Yearly	Global	Number	0	2024	40	2026	Outcome: 2. LONG-TERM: A strong HDP Nexus data ecosystem on complex risks. Output: 2.1. A strengthened network and community of practice on harmonized data and standard practices.

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
Funding allocated for crisis action with the support of project outputs.		This indicator aims to measure the extent to which the project outputs are used to facilitate funding decisions related to crisis action. The indicator focuses on the amount of funding allocated to crisis action that can be directly / indirectly attributed to the use of project outputs, such as data, evidence, and analysis, in decision—making processes.	Surveys, interviews, analysis of public policy documents/ emergency response plans/ reports, other documents.	Investment	Yearly	Global	Number	0	2024	100000	2025	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.3. A web platform to host the Risk Anticipation Data Warehouse and associated analytics and visualization products.
	Funding allocated for crisis action specifically in fragile settings.	This sub- indicator aims to measure the extent to which the project outputs are used to facilitate funding decisions related to crisis action specifically in fragile contexts. The indicator focuses on the amount of funding allocated to crisis action that can be directly / indirectly attributed to the use of project outputs, such as data, evidence, and analysis, in decision- making processes.	Surveys, interviews, analysis of public policy documents/ emergency response plans/ reports, other documents.	Investment	Yearly	Global	Number	0	2024	500000	2025	

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
People in fragile and crisis-affected settings benefitting from earlier, faster, more targeted and dignified assistance as a result of project outputs.		This indicator aims to measure the extent to which the project outputs have contributed to supporting people in fragile and crisis-affected settings earlier, faster, and in a more targeted and dignified way.	Surveys, reports, other documents, assessments, statistics etc.	Beneficiaries	Yearly	Global	Number	0	2024	1000	2025	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.1. An API endpoint for data warehouse.
Multilateral funding instruments and other entities that use project outputs to facilitate funding decisions.		This indicator aims to measure the extent to which the project results are used by multilateral funding instruments and other entities to inform funding decisions. The indicator focuses on the use of project outputs, such as data, evidence, and analysis, to support the decision-making processes of funding instruments and other entities involved in crisis action.	Surveys, interviews, analysis of public policy documents/ emergency response plans/ reports, other documents.	Investment	Yearly	Global	Number	0	2024	1	2025	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.3. A web platform to host the Risk Anticipation Data Warehouse and associated analytics and visualization products.
Understanding of the datasets / analytical tools by the key stakeholders.		This indicator aims to measure the level of comfortability and technical understanding of the datasets or analytical tool provided as part of the project.	Surveys, interviews, internal statistics.	Capacity	Yearly	Global	Percentage	0	2024	60	2025	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.3. A web platform to host the Risk Anticipation Data Warehouse and associated analytics and visualization products.

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
External reports and other tangible products that feature data or analytics from the project.		This indicator aims to measure external reports and other tangible products that feature data or analytics from the project.	Internal tracking.	Other	Yearly	Global	Number	0	2024	3	2025	Outcome: 1. INTERMEDIATE: Better data, analytics, knowledge, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.3. A web platform to host the Risk Anticipation Data Warehouse and associated analytics and visualization products.

# **Project Indicators**

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
Website Unique Visitors		Number of Unique Visitors	Online Website Numbers	Capacity	Yearly	Global	Number	0	2024	15000	2025	Outcome : 1. INTERME DIATE: Better data, analytics, knowledg e, and capacity on complex risks in fragile and crisis- affected settings. Output: 1.1. An API endpoint for data warehous e.

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
Training Sessions		Number of Training Sessions held to familiarize users with platform	Internal Tracking	Capacity	Yearly	Global	Number	0	2024	4	2025	Outcome: 1. INTERME DIATE: Better data, analytics, knowledg e, and capacity on complex risks in fragile and crisis- affected settings. Output: 1.3. A web platform to host the Risk Anticipati on Data Warehou se and associate d analytics and visualizat ion products.
	No components a	vailable.										
User Experience Surveys		Number of UI/UX Surveys/Feedback Sessions Conducted	Internal Tracking	Capacity	Yearly	Global		2	2024	2	2025	Outcome: 1. INTERME DIATE: Better data, analytics, knowledg e, and capacity on complex risks in fragile and crisis-affected settings. Output: 1.3. A web platform to host the Risk Anticipati on Data Warehou se and associate d analytics and visualizat ion products.

Indicator Title	Component Title	Description	Means of Verification	Category	Cycle	Scope	Value Type	Baseline Value	Baseline Year	Target Value	Target Year	Linked Outcome / Output
	No components a	vailable.										
API Calls		Number of API Calls Yearly	Online Platform Numbers	Capacity	Yearly	Global	Number	0	2024	6000	2025	
	No components a	components available.										

## **Risks**

Event	Category	Level	Likelihood	Impact	Mitigating Measures	Risk Owner
High rates of hallucinations and other inaccuracies/error rates from LLM based front end	Operational     Strategic	High	Possible	Major	Regularly validate and calibrate the LLM model to minimize hallucinations. Implement thorough testing and monitoring. Optimize for accuracy and increased error handling.	UNDP
LLM interface with data warehouse/API is not developed accurately	Operational     Organizational     Political	Low	Unlikely	Major	Invest in continuous improvement of the model's data communication capabilities. Regularly update and adapt to new data formats. Identify technical partners who have already implemented similar solutions and implement models that have high standard of performance.	UNDP
There is bias and lack of objectivity in the product development process. Inherent bias and systemic racism exist in LLM models already trained	<ul><li> Operational</li><li> Organizational</li><li> Political</li></ul>	Low	Possible	Moderate	UNDP and data implementing partners must ensure that the design process is inclusive of all voices and considers inherent bias in LLM and actively implement anti bias mechanisms to address various stages of LLM Development during user testing periods.	UNDP
Different data source owners can deploy their own LLMS to interact with data warehouses	Strategic	High	Likely	Major	Focus on differentiation by highlighting unique features and advantages. Collaborate with other providers early on to enhance offerings	UNDP
Reduced demand for cleaned/structured based API datasets	Operational     Strategic	Low	Unlikely	Minor	User surveys and feedback sessions to understand preferences. Providing options for both raw and cleaned data, accompanied by educational efforts on the benefits of cleaned data.	UNDP
Reduced accessibility due to availability of LLMs in Machine rich languages such as English	Operational	Medium	Very Likely	Minor	Investigate translation-based features and large language models capable of interacting in the official languages of the UN. Investigate a phased approach to developing increased language access.	UNDP
Improper engagement strategy in deploying platform	Operational	Low	Unlikely	Moderate	Discuss with other similar platforms, implementing partners and have regular feedback sessions, and develop project with user-centric design principles	UNDP
Lack of language support impacting access to relevant users	<ul><li>Political</li><li>Strategic</li></ul>	Low	Possible	Moderate	Introduce translation agents in natural language input mechanisms and data output mechanisms.	UNDP

# Budget by UNSDG Categories: Over all

Budget Lines	UNDP (7%) *	Total
1. Staff and other personnel	\$335,000.00	\$335,000.00
2. Supplies, Commodities, Materials		\$0.00
3. Equipment, Vehicles, and Furniture, incl. Depreciation		\$0.00
4. Contractual services	\$96,186.92	\$96,186.92
5. Travel	\$16,000.00	\$16,000.00
6. Transfers and Grants to Counterparts		\$0.00
7. General Operating and other Direct Costs	\$20,102.80	\$20,102.80
Project Costs Sub Total	\$467,289.72	\$467,289.72
8. Indirect Support Costs	\$32,710.28	\$32,710.28
Total	\$500,000.00	\$500,000.00

## **Budget by UNSDG Categories: 2024**

<b>Budget Lines</b>	Fiscal Year *	Description	UNDP (7%) *	Total
1. Staff and other personnel	2024	Project Management: \$120,000 LLM Data Engineering & Deployment Expert: \$120,000	\$240,000.00	\$240,000.00
2. Supplies, Commodities, Materials	2024			\$0.00
3. Equipment, Vehicles, and Furniture, incl. Depreciation	2024			\$0.00
4. Contractual services	2024	API Development: \$51,000	\$51,000.00	\$51,000.00
5. Travel	2024	Workshops: \$16,000	\$16,000.00	\$16,000.00
6. Transfers and Grants to Counterparts	2024			\$0.00
7. General Operating and other Direct Costs	2024	Admin and operating costs: \$20,103	\$20,102.80	\$20,102.80
Project Costs Sub Total			\$327,102.80	\$327,102.80
8. Indirect Support Costs			\$22,897.20	\$22,897.20
Total			\$350,000.00	\$350,000.00

# **Budget by UNSDG Categories: 2025**

Budget Lines	Fiscal Year *	Description	UNDP (7%) *	Total
1. Staff and other personnel	2025	Project Management: \$45,000 LLM Data Engineering & Deployment Expert: \$40,000	\$95,000.00	\$95,000.00
2. Supplies, Commodities, Materials	2025			\$0.00
3. Equipment, Vehicles, and Furniture, incl. Depreciation	2025			\$0.00
4. Contractual services	2025	Web Development: \$45,187	\$45,186.92	\$45,186.92
5. Travel	2025			\$0.00
6. Transfers and Grants to Counterparts	2025			\$0.00
7. General Operating and other Direct Costs	2025		\$0.00	\$0.00
Project Costs Sub Total			\$140,186.92	\$140,186.92
8. Indirect Support Costs			\$9,813.08	\$9,813.08
Total			\$150,000.00	\$150,000.00

## Performance-based Tranches Breakdown

Tranche			Total
Tranche 1	UNDP	\$350,000.00	\$350,000.00
Tranche 2	UNDP	\$150,000.00	\$150,000.00
			\$500,000.00

# **Programme Outcome Costs**

Outcome	Output	Activity	Implementing Agent				Time	Frame			
					20	24		2025			
				1	2	3	4	1	2	3	4
1. INTERME	DIATE: Bet	ter data, analytics, know	ledge, and capacity on complex risks	in fragile and crisis	-affected	settings.					
	1.1. An AF	PI endpoint for data ware	house.								
		1.1.1. Define API object	ives and use cases in a scoping docu	ment.							
			UNDP	✓	<b>V</b>						
		1.1.2. Complete any net the data.	cessary back end development or adj	ustments of the da	abase sy	stems and	analysis	processes	s to sourc	e and trar	ısform
			UNDP		<b>V</b>						
		1.1.3. Develop the API i	n line with the use cases and function	nalities identified in	the scopi	ng phase.					
			UNDP			<b>V</b>	<b>V</b>				
		1.1.4. Maintain the API	and supporting systems.								

Outcome	Output	Activity	Implementing Agent				Time	Frame				
				2024 2025 1 2 3 4 1 2 3								
				1	2	3	4	1	2	3	4	
			UNDP					V		<b>V</b>	<b>V</b>	
		1.1.5. Develop thorou	igh documentation of the data in the Risk Anticipa	tion Data	Warehous	se and gu	idance m	aterial on	API usag	e.		
			UNDP					<b>V</b>				
	1.2. An A	I-driven interface that	enables users to access the data warehouse easil	y.								
		1.2.1. Establish a sma	all working group of technical partners, implement	ing partn	ers, and a	user feed	dback gro	oup to guid	de develo	pment.		
			UNDP									
		1.2.2. Conduct a wor	kshop to develop, guide and establish the method	lology for	LLM deve	elopment.						
			UNDP									
		1.2.3. Define the object desired outcomes.	ectives of generating data-output to natural langu	age input	s in a scop	oing docu	ment, en	suring alig	ınment w	ith user ne	eeds and	
			UNDP		<b>V</b>							
		1.2.4. Develop a data	a-output minimum viable product.									
			UNDP		<b>V</b>	~						
		1.2.5. Conduct a wor	kshop to review the data-output LLM model qualit	ty and de	ployment,	and itera	te develo	pment wit	h collecte	ed feedba	ck.	
			UNDP				V	<b>V</b>				
		1.2.6. Define the object	ectives of generating visual-output to natural lang	uage inpu	uts in a sco	oping doc	ument, e	nsuring al	ignment v	with user r	needs	
			UNDP									
		1.2.7. Develop a visua	al-output minimum viable product.									
			UNDP									
		1.2.8. Conduct a wor	kshop to review the visual-output LLM model qua	lity and d	eploymen	t, and iter	ate devel	lopment w	ith collec	ted feedb	ack.	
			UNDP									
		1.2.9. Maintain the LL	_Ms and deployment systems.									
			UNDP							<b>V</b>	<b>V</b>	
		1.2.10. Conduct train	ing and outreach activities to provide guidance or	utilizatio	n of data-	output ar	nd visual	output LLI	Ms and th	ie API.		
			UNDP								<b>V</b>	
	1.3. A we	b platform to host the	Risk Anticipation Data Warehouse and associated	analytics	and visua	alization p	roducts.					
		1.3.1. Define user req LLM components.	uirements by conducting user requirement assess	sments, a	and identif	y product	roadmap	o ensuring	connecti	ions to the	API and	
			UNDP			<b>V</b>						
		1.3.2. Conduct initial UNDP design princip	and ongoing user centred design activities to ensoles.	ure that t	the web el	ements m	eet user	needs and	d remain i	n line with	UN and	
			UNDP				~					
			and interactive data UX front end, and conduct in and intuitive and produced in a visually beautiful fo		design ac	ctivities to	ensure o	data struc	ture, proc	essing an	d web	
		,	UNDP				~	<b>V</b>				
		1.3.4. Deploy a front-	end minimum viable product using kanban/agile r									
			UNDP					<b>V</b>	~	<b>V</b>		
		1.3.5. Collect and ing	est feedback for MVP and deliver a functioning pl									
		and testing.	UNDP						<b>V</b>	<b>V</b>	<b>V</b>	
		136 Maintain the w	eb platform front end.									
		1.5.5. Walltain the W	UNDP								<b>V</b>	
2 LONG-T	FRM: A stro	ing HDP Nexus data ec	cosystem on complex risks.									
			d community of practice on harmonized data and	standard	practices							
		_	work of organizations aiming at the strengthening		-		data ecc	svstem.				
			UNDP					Ø System.				
		2.1.2. Convene period	dic workshops and meetings to foster knowledge-					_				
		2 2 3 3 3 5 6 6 7 6	UNDP						<b>V</b>		<b>V</b>	
								_				

**Signatures** 

UNDP: UNDP (United Nations Development Programme (UNDP)) (Manual) Mr Turhan Saleh turhan.saleh@undp.org	SIGNATURE:
	28-Jun-2024 DATE:

