NITI AAYOG

ADDENDUM NO. 1 TO THE

REQUEST FOR PROPOSALS (RFP)

FOR

DEVELOPMENT AND OPERATION OF

NATIONAL DATA AND ANALYTICS PLATFORM (NDAP)

The following is the modification to the RFP for Selection of Consultant for Development and Operation of National Data and Analytics Platform (NDAP). The deletions from the earlier text of the RFP are indicated as strikethroughs and additions are underlined.

S. No.	Clause No	Provision of the Request for Proposals
(i)	2.3.4 of the RFP	An Applicant eventually appointed to provide Consultancy for this Project, and its Associates, shall be disqualified from subsequently providing goods, works or services related to the development and operation of the same Project and any breach of this obligation shall be construed as Conflict of Interest; provided that the restriction herein shall not apply after a period of 5 (five) years from the completion of this assignment or to consulting assignments granted by banks/ lenders at any time; provided further that this restriction shall not apply to consultancy/ advisory services performed for the Authority in continuation of this Consultancy or to any subsequent consultancy/ advisory services performed for the Authority in accordance with the rules of the Authority. For the avoidance of doubt, an entity affiliated with the Consultant shall include a partner in the Consultant's firm or a person who holds more than 15% (~~five~~ fifteen per cent) of the subscribed and paid up
(ii)	Schedule-1: Term of Reference	Contents Attachment -I: ~~Indicative l~~List of Reports
(iii)	Schedule-1: Term of Reference: Clause 3.1.1	The NDAP would be developed in accordance with global best practices in line with similar initiatives internationally in order to achieve the above objectives. The following minimum features shall be complied with:

Table within row (iii):

S. No.	Category	Features
(1).	**Non Functional Requirements**	(a) Reliable 24/7 hosting with automated website backup. (b) Cross-browser compatibility. (c) Responsive web design across web interfaces, mobiles, and tablets (e.g. disabled images, various resolutions/font sizes). (d) Fast-loading pages (page load time involving less than 10 indicators should be less than 3 seconds for bandwidth of 512 kbps).

				(e)	Should be compliant with the Guidelines for Indian Government Websites (GIGW) and Interoperability Framework for E-Governance in India issued by Ministry of Electronics and Information Technology (MeitY).
		(2).	**User Sensitive Access Requirements**	(a)	Data catalogue for users to browse and access all available datasets.
				(b)	Based on the past history of search by user, display dynamic dashboards at any specified granularity of time and location by collating data in various pre-defined categories.
				(c)	User profiling and personalization based on user's previous visits for registered and logged in users. Some of the features of the portal will only be available to registered users. Display of generic content for anonymous user.
				(d)	Registered users can receive automated reports based on analysis performed directly at their inbox at scheduled intervals – snapshot of the report would be attached with the email and linked to the detailed report on the NDAP portal. For the same, NIC email server can be used.
				(e)	Displays trending/recently viewed/most popular segments and datasets of site.
				(f)	Provision of site map, listing all the contents of the website including static content, training materials, etc.
				(g)	Collects data on user searches, eventual datasets that users view, and user profiles to learn over time to produce more relevant searches.
		(3).	**Search Requirements**	(a)	The NDAP shall be designed and operated such that it functions as a search engine for retrieving the data hosted on it. The retrieval and display of data shall be in accordance with good industry practices followed by other eminent websites operating in India and abroad.
				(b)	Allows users to search content across a variety of criteria (sector, indicator, location) from search bar, through map-based search, and from data catalogue.
				(c)	Provides context and/or query-based suggested/related/recommended links and understanding of synonyms. Multiple options based on the query can be presented to the user if there is no single query interpretation.
				(d)	Approximate string matching while the user types in the query keywords.
				(e)	Employs natural language understanding of search query (not token keyword search) to provide:

<table>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>

(i)   Simple and intuitive Input Interpretations of the textual query; and

(ii)   Suggested Alternatives for each query subcomponent based on function of user's profile and previous similar searches;

(iii)   Relevant Representations of results (e.g. time series graphs, comparative graphs/tables) based on query of user.

</td>
</tr>
<tr>
<td></td>
<td></td>
<td><u>(4).</u></td>
<td></td>
<td>**Data Requirements**</td>
<td></td>
<td>

The data requirements are categorised into:

(1) Coverage of Data Sources and Formats

(2) Data Aggregation

(3) Data Analysis

(4) Data quality and standardisation

(5) Data refresh frequency

(6) Log management

(7) Data archival and retention features

(1)   **Coverage of Data Sources and Formats**

  (a)   Integrates Application Programming Interfaces (APIs) provided by the Authority to read data into the platform.

  (b)   Integrates initial corpus of data into website from the following sources:

     (i)   Website: data.gov.in;

     (ii)   Websites of 50+ Ministries and Departments of Central Government;

     (iii)   Websites of all departments and Ministries of State Governments not exceeding 250;

     (iv)   Autonomous and affiliated bodies of the following Ministries - Agriculture, Education, Finance, and Health; and

     (v)   Reports for key sectors as shown in the ~~indicative~~ list at Attachment-I.

  (c)   Have an automated system for alerts on addition of new datasets for any of the above sources, necessary for updating of NDAP.

  (d)   Fetch data in structured and semi-structured format from all the Government websites (both State and Central) and reports (indicative list is at Attachment-I)

</td>
</tr>
</table>

(e)   Regularly visit the websites of the Ministry(ies) and States to identify new data uploads. Fetch any new data that has been uploaded on the websites.

(f)   The Authority may include more data sources and reports over time.

(2)   **Data Aggregation**

(a)   Extraction of accurate data from the various Government websites.

(b)   Conversion of reports into machine readable format and extraction of the numeric information into machine readable format.

(c)   All data points ingested into NDAP should have a common schematic theme including location, time, and sector as specified in sub-paragraph (g).

(d)   Automatic Extract, Transform, Load (ETL) process to map incoming data to designed data model.

(e)   Metadata management tool to track ETLs, transformations and changes made to data.

(f)   Ability to process data from national to village/town level resolution (based on various codes like LGD codes, PIN codes, Census codes, etc.).

(g)   Ability to aggregate information from individual data sets to higher aggregation levels based on location (e.g. village/ district/ city/ state) or time (daily/ weekly/ monthly/ quarterly/ yearly).

(h)   Ability to input structured and semi-structured (e.g. XML, JSON) data formats.

(i)   Seamlessly integrate an initial data capacity of 1 TB, which should be scalable up to 10 TB upon payment of recurring costs by the Authority for provision and use of data capacity in excess of 1 TB.

(j)   Ensure quarterly incorporation of all non-API data within one week of publishing on the parent website of the data sources.

(k)   System should be able to handle 1,000 concurrent users and should be scalable to 2,500 concurrent users upon payment of recurring costs by the Authority for and in respect of concurrent users in excess of 1,000.

(l)   System should be able to give a response to any search query involving ten or less indicators (including their analysis and visualization) in 3 seconds or less.

(m)   Portal to be procured and hosted on a Government of India Cloud (MeghRaj) or on a Ministry of Electronics and Information Technology (MeitY) empanelled cloud

Consultant. Details about the cloud can be found on http://meity.gov.in/content/gi-cloud-meghraj

(n) For the purpose of traceability, all downloaded reports from other Ministries and State Governments websites would be saved in the database along with the date and time of the fetch activity.

(3)   Data Analysis

(a) Inbuilt visualization of data from datasets (e.g. bar chart, pie chart, bubble chart, donut chart, area chart, cartogram, info-graphics, 3D visualizations, radar chart, geo-mapping) with customizable data ranges and properties.

(b) Inbuilt analytics (e.g. correlations, multivariable plotting) with drag and drop customization functionality/the ability to choose suitable graphs intelligently.

(c) Animated visualizations on variable factors, including dynamic visualization with drill-down and drill-up features at all applicable granularity.

(d) Ability to compare, analyse and reason data across datasets using spatial (location), temporal (time period), domain (sector) and measurement indicators.

(e) Ability to generate customizable reports by cross-linking datasets.

(f) Downloadable/shareable datasets at a customizable resolution, range, time interval, etc. in standard formats (e.g., xml, json, csv, etc.).

(g) All visualizations exportable in formats as customized by user (e.g., jpeg, png, pdf, etc.).

(h) Ability to share visualizations/data on third party websites while keeping the metadata, including the source attribution, traceable.

(i) Provides metadata information (source, year, methodology, etc.) for datasets.

(4)   Data quality and standardisation

(a) SLAs and frameworks need to drive data quality as this will enable accuracy in the final reports.

				(b) Ability to correct mistakes in spellings, inconsistencies, capitalisations, and abbreviations. (c) Pre-built libraries for standardisation of data, capability to identify duplicates and cluster records. **(5) Data refresh frequency** (a) Most of the indicative reports are annual, but periodic refresh of the data for any ad-hoc reporting will need to be ensured. **(6) Log management** Needs to maintain a log feature accounting for all the data access and actions performed by various users- updates, changes, refreshes etc. **(7) Data archival and retention features etc** This will drive the temporary storage requirements for logs, cache data etc. to ensure faster loading of all the tables.
		(5).	**Interactivity and Accessibility**	(a) Built-in terminals and software for blind/visually impaired users. The navigation and usability aspects of the site should be friendly to blind/visually impaired people. (b) English/Hindi language translations for audio/video tutorials with speech. The tutorials should have content in both English and Hindi, both as speech and written sub-titles. (c) Multi-lingual website covering all 22 major Indian languages. Only the static navigation content needs to be made available in the major Indian languages. (d) Context sensitive 'Help' feature. (e) Integrated social media/share content function (email, Facebook, Twitter, etc.) as embedded graphics and embeddable code. (f) Platform tutorial videos and media. (g) Optional user registration for personalized features like "follow" datasets, sector, location and receive automated email/mobile alerts and notifications.
		(6).	**Site Analytics & Feedback**	**Site Analytics** (a) Tracks site traffic (page views, visits, pages per visit, bounce rate, new visits, and average time on site, most popular datasets, most downloaded datasets, etc.).

(b)   Tracks user demographics (access browser, operating systems used, languages, state/country of access).

(c)   Tracks user journey to understand how the website is being used.

(b)(d)   Generate a heat-map of clicks on various elements of the page, to understand feature usage.

(c)(e)   Tracks cookies and IP addresses.

(d)(f)   Tracks which parts of the backend database are accessed and at what frequency to understand dataset relevance.

(g)   Site statistics are displayed on the website with the user having options to filter the statistics on simple parameters like location, date, sector, data source, etc.

(h)   Providing fortnightly reports to the Authority on website usage, aggregating analysis from sources such as Google Analytics.

Feedback

(a)   Every page of the Portal would have an option for user Feedback as a floating icon on the page.

(b)   This floating icon for user feedback would be accessible while browsing various pages of the entire Portal.

(c)   Once clicked, pop-up window opens up.

(d)   There is a 5-Star rating of the site that the user can provide as feedback.

(e)   The pop-up window would have drop-down option on the reason for the feedback like: Incorrect search results, Data not available, In-appropriate visualization, In-appropriate analysis, data error, Other etc. Other opens up a blank box to enter the reason for the feedback.

(f)   Feedback details are compulsory to be entered by the user, including the Ministries who might have the missing data.

(g)   User name, email and mobile number can be collected as optional fields for unregistered users.

(h)   The captured feedback details along with the automatically generated screenshot (with details of the page, query and response) of the content currently displayed on the Portal is mailed to the admin of both the Vendor and NITI Aayog PMU along with being stored in the database as feedback logs.

(i)   The list of new data items missing from the site along with number of requests would be made available on the NDAP portal on a monthly basis and also reach out to the Ministries and State Governments requesting that data.

		(7).	**Security**	Produce an "Application Security Audit Clearance Certificate" for the solution to be deployed and should ensure adherence to approved security guidelines published by the Indian Computer Emergency Response Team (CERT-In), the Standardization, Testing and Quality Certification (STQC) by MeitY, or any other guideline issued by a Government of India agency. The Consultant should ensure the appropriate risk and contingency planning for project. Vendor is responsible for ensuring the timely completion of these activities in co-ordination with the certified vendors. Details of the same can be found on the following website: http://www.cert-in.org.in/.
		(8).	**Disaster Recovery and Business Continuity**	(a) The Consultant shall ensure that the NDAP system is protected against natural disasters, software failures, human error, and other contingencies that could interrupt services. (b) The Consultant shall establish and maintain a disaster recovery environment suitable to enable business continuity in the event of a disaster that affects the operational environment. (c) The Consultant shall forthwith create a Business Continuity Plan (BCP) that shall include a Disaster Recovery Plan as well as security policies and procedures. The BCP shall contain information regarding how the NDAP systems and data will be protected, what steps would be taken to replicate/restore/recover any lost data and the timeframes involved, frequency of maintenance procedures/hardware and software upgrades and their impact on system performance, capacity and/or uptime. (d) The Consultant shall conduct an exercise to test the BCP twice a year and meet recovery criteria of 2 Hours for NDAP applications. Test results shall be compiled by the Consultant in a report to be shared with the Authority. In addition, emergency drills shall be conducted once a year.
(iv)	Schedule-1: Term of Reference: Attachment-I	**Attachment-I** ~~Indicative~~ **List of Reports** * * * * * * * * * *		

(v)	Schedule-2: Form of Agreement: Clause 3.2.2	Consultant and Affiliates not to be otherwise interested in the Project
		The Consultant agrees that, during the term of this Agreement and after its termination, the Consultant or any Associate thereof and any entity affiliated with the Consultant, as well as any Sub-Consultant and any entity affiliated with such Sub-Consultant, shall be disqualified from providing goods, works, services, loans or equity for any project resulting from or closely related to the Services and any breach of this obligation shall amount to a Conflict of Interest; provided that the restriction herein shall not apply after a period of five years from the completion of this assignment or to consulting assignments granted by banks/ lenders at any time; provided further that this restriction shall not apply to consultancy/ advisory services provided to the Authority in continuation of this Consultancy or to any subsequent consultancy/ advisory services provided to the Authority in accordance with the rules of the Authority. For the avoidance of doubt, an entity affiliated with the Consultant shall include a partner in the Consultant's firm or a person who holds more than 15% (~~five~~ fifteen per cent) of the subscribed and paid up share capital of the Consultant, as the case may be, and any Associate thereof.