

Annexure I

DGQI Self-Assessment Questionnaire for Ministries/Departments of Gol (2021-22)

Detailed explanations for each question may be referred to in Appendix I provided after the questionnaire.
(Page 14-20)

Part -A (To be fed at Ministry/ Department Level)

A. Background Information

1.	Ministry / Department(M/D) Name:	
2.	Name of the Central Sector (CS) Schemes of the M/D:	a.
		b.
		c.
3.	Name of Centrally Sponsored Schemes (CSS) of the M/D:	a.
		b.
		c.
4.	Please enter any other non-schematic intervention (NSI) to be included for DGQI self-assessment:	a.
		b.
5.	Details of the nodal officer responsible for verifying authenticity of information provided in this form:	
	a. Name	
	b. Designation:	
	c. E-Mail ID:	

B. Data & Strategy Unit

1.	Has the M/D constituted a Data & Strategy Unit (DSU) as a central unit for developing data strategy? (as per the D.O. letter from Sh. Bhaskar Khulbe, Advisor to PM dated 02.02.2021)	<input type="checkbox"/> Yes <input type="checkbox"/> No
2.	(Respond if answer to 1 is 'yes', else skip to Q1 of next section) Who is the head of the DSU?	<input type="checkbox"/> AS <input type="checkbox"/> JS <input type="checkbox"/> Director <input type="checkbox"/> Below Director
3.	Please select the verticals established under the DSU of your Ministry/Department.	<input type="checkbox"/> Monitoring Unit <input type="checkbox"/> Statistics Unit <input type="checkbox"/> Technology Unit

	(as per the D.O. letter from Sh. Bhaskar Khulbe, Advisor to PM dated 02.02.2021)	<input type="checkbox"/> Analytics Unit
4.	Please provide the percentage of filled posts in DSU (number of posts filled up/ number of posts created by the Ministry/Department for the DSU) in the below provided table:	
		Enter % of posts filled up
		Monitoring Unit
		Statistics Unit
		Technology Unit
		Analytics Unit
		Total
5.	Is the terms of reference (ToR) for all units within DSU well defined and documented by the M/D to lay down their scope of work?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partial (to be selected if ToR development is in progress)
6.	Have any regular review meeting mechanisms at the level of the head of DSU and/or the Secretary been established for regular review of the work undertaken by the DSU?	<input type="checkbox"/> Yes <input type="checkbox"/> No
7.	<i>(Respond if answer to 6 is yes, else skip this question)</i> What is the frequency of regular review meetings/review reports?	<input type="checkbox"/> Daily <input type="checkbox"/> Weekly <input type="checkbox"/> Fortnightly <input type="checkbox"/> Monthly <input type="checkbox"/> Quarterly <input type="checkbox"/> Annually

C. Action Plan

1.	Has the M/D framed an action plan to improve its data preparedness levels? (as per the D.O. letter from Sh. Bhaskar Khulbe, Advisor to PM dated 02.02.2021)	<input type="checkbox"/> Yes <input type="checkbox"/> No
----	--	---

2.	<i>(Respond if answer to 1 is yes, else skip to Q1 of next section)</i> Does the action plan have all the sections as per the outline shared with all M/Ds? (As per D.O. letter from Sh. Bhaskar Khulbe, Advisor to PM on 02.02.2021)	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially. If partially, please specify how: <hr/>
3.	Does the action plan include data strategy for all CS/CSS schemes of the M/D?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially (Some schemes included) If partially, please specify which schemes are not included: <hr/>
4.	Are clear timelines for each action point identified under the strategy?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially (For some actions) If partially, please specify how and why: <hr/>
5.	Are the responsibilities for each action point clearly allocated to respective divisions for ensuring accountability?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially (For some actions). If partially, please specify how and why: <hr/>
6.	Please upload the action plan in PDF format.	
7.	Please enter the action points in the attached excel template. Scores based on exhaustiveness (number of action points per scheme) and timely completion/compliance on the action points against the timelines set by the M/D will get auto-calculated and displayed here.	

D. Data Management

1.	Are there data management guidelines/architecture, explaining how generated data is to be processed, stored, exchanged, archived and destroyed?	<input type="checkbox"/> Yes <input type="checkbox"/> No If yes, please briefly explain the scope implementation of these guidelines: <hr/>
2.	<i>(Respond if answer to in 1 is 'yes', else skip this question)</i> Is there a dedicated senior-level officer responsible to check the compliance of the data management processes?	<input type="checkbox"/> Yes <input type="checkbox"/> No

3.	Are data ownership norms clearly defined?	<input type="checkbox"/> Yes <input type="checkbox"/> No
4.	Is there a framework for assessing the risk and value of all the data collected by the M/D?	<input type="checkbox"/> Yes <input type="checkbox"/> No If yes, please explain how is this done: _____
5.	Is there a framework governing the ethical use of data, including the use of predictive algorithms, machine learning etc. by the M/D?	<input type="checkbox"/> Yes <input type="checkbox"/> No If yes, please explain how is this done: _____

Note: M/Ds may preferably fill up remaining sections of Part- A (given below from E-H) after completing Part – B of the questionnaire as these questions correspond to the third pillar of data driven outcomes.

E. Synergistic data use within the M/D

1.	Based on data analysis, has the M/D identified data gaps at M/D level that need to be plugged in from decision making/policy analysis perspectives?	<input type="checkbox"/> Yes <input type="checkbox"/> No If yes, please specify how: _____
2.	<i>(Respond if answer to 1 is yes, else skip this question)</i> Has the M/D made any implementation plan to overcome these data gaps to aid in decision making?	<input type="checkbox"/> Yes <input type="checkbox"/> No If yes, please specify how: _____
3.	Has the M/D created any systems for ensuring that data systems across scheme divisions are integrated so that data from different scheme divisions is shared with each other?	<input type="checkbox"/> Yes. <input type="checkbox"/> No <input type="checkbox"/> In progress <input type="checkbox"/> N/A If yes or in progress, please specify how: _____ If "N/A", please provide reasons why inter schematic division data integration is not applicable: _____

F. Inter-Agency Data Collaboration

1.	Has the M/D collaborated with other agencies (other M/Ds, private agencies, research organizations etc.) for improving their data systems wherever possible?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> In progress
2.	<i>(Respond if answer to 1 is yes, else skip this question)</i> Has the M/D undertaken any of the following steps to drive these inter-agency data collaboration initiatives?	
	<input type="checkbox"/> Sol, MoU, Partnerships with agencies	
	<input type="checkbox"/> API linking of MIS/Dashboards done to enable seamless data sharing between M/Ds	
	<input type="checkbox"/> Multiple data collection processes aimed at same target groups replaced by single synergistic process	
	<input type="checkbox"/> Integrated data storage/warehouses	
	<input type="checkbox"/> Collaboration with other M/Ds to use their data for developing own systems	
	<input type="checkbox"/> Collaboration with M/Ds to develop joint systems for data gathering/use of non-conventional data sources/emerging technologies	
	<input type="checkbox"/> Collaboration with private agencies for use of non-conventional data sources or emerging technologies	
	<input type="checkbox"/> Jointly conducting analysis using data from multiple M/Ds	
	<input type="checkbox"/> Partnerships/Collaborations for data security related measures	
	<input type="checkbox"/> Partnerships/Collaborations for capacity building of human resources	
	<input type="checkbox"/> Others - Please specify : _____	

G. Prescriptive Analytics

1.	Has the M/D gone beyond exploratory data analysis to cross-functional prescriptive analytics?	<input type="checkbox"/> Yes <input type="checkbox"/> In Progress <input type="checkbox"/> No If yes or in progress, please specify how: _____
2.	<i>(Respond if answer to 1 is yes, else skip this question)</i> How often is this being undertaken?	<input type="checkbox"/> Annually <input type="checkbox"/> Quarterly <input type="checkbox"/> Monthly
3.	<i>(Respond if answer to 1 is yes, else skip this question)</i> What is the mode in which this is being practiced?	
	<input type="checkbox"/> Mechanisms for regular prescriptive data analysis reports to be prepared and shared with decision makers at the highest level have been instated	
	<input type="checkbox"/> Committee formed to hold policy review meetings/review reports at regular frequencies	
	<input type="checkbox"/> Regular policy review meetings involving all scheme divisions/sections institutionalized	
	<input type="checkbox"/> Emerging actionables are undertaken, documented and disseminated via a separate newsletter/report/document/order etc. and tracked regularly	
	<input type="checkbox"/> Others - Please specify how: _____	

H. Good Practices - Please share any three good practices of how the M/D has taken measures to strengthen data-driven decision-making (non-schematic or scheme level) within the M/D along with its positive impact.

Good Practice 1
1a. Describe the problem statement faced by the M/D. (100 words)
1b. Describe how the M/D has used and implemented data systems and analytics to address the issue to drive smart, near real-time and granular decisions (100 words).
1c. Explain the positive impact generated with supporting evidence that indicated such impact due to the solution implemented (100 words).
Good Practice 2
1a. Describe the problem statement faced by the M/D. (100 words)
1b. Describe how the M/D has used and implemented data systems and analytics to address the issue to drive smart, near real-time and granular decisions (100 words).
1c. Explain the positive impact generated with supporting evidence that indicated such impact due to the solution implemented (100 words).
Good Practice 3
1a. Describe the problem statement faced by the M/D. (100 words)
1b. Describe how the M/D has used and implemented data systems and analytics to address the issue to drive smart, near real-time and granular decisions (100 words).
1c. Explain the positive impact generated with supporting evidence that indicated such impact due to the solution implemented (100 words).

Part -B (To be fed at CS/CSS/NSI Level)

To be fed by the Ministry/Department for each CS/CSS/NSI of the Department in Q.A of Part A of the Questionnaire

A. Data Generation

1.	Are the data requirements of the scheme well defined and documented?	<input type="checkbox"/> Yes <input type="checkbox"/> No
2.	Is data collected for all identified data requirements?	
	a. Input Data Points	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partial
	b. Output Data Points	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partial
	c. Outcome Data Points	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partial
3.	Is collected data reported digitally? (i.e. is there is a digital electronic database/MIS)?	<input type="checkbox"/> Yes <input type="checkbox"/> No i.e. On paper only If Yes, please provide the link: _____ If credentials are required for login, please provide some username and password: User - _____ Pw - _____
4.	<i>(Respond if answer to 3 is 'Yes', else skip to 1 of Q1 of data quality section)</i> At what granularity is data reported digitally for the scheme?	
	a. At the M/D (National)	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	b. State	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	c. District / City	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	d. Sub-District / Tehsil	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	e. Block	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	f. Village	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	g. Individual / Household	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	h. Facility	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	i. Project	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
5.	At what frequency is data reported digitally for the scheme?	
	a. Realtime or near realtime	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	b. Daily	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	c. Weekly/Fortnightly	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	d. Monthly	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	e. Quarterly	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	f. Half-yearly	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	g. Yearly	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
6.	How is this data collected at the ground level?	

	<input type="checkbox"/> Collected on paper by human resources and then fed on digital systems
	<input type="checkbox"/> Collected using digital modes (tablets/phones etc.) by human resources
	<input type="checkbox"/> Transactional data
7.	<i>(Respond if answer to 6 is 'second/third option', else skip this question)</i> Are any of the following technologies used?
	a. CAPI Surveys <input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	b. Geotagged information <input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	c. Geofenced information <input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	d. Others - Please specify which technology _____

B. Data Quality

1.	Are there pre-defined documented mechanisms to assess quality of incoming data?	<input type="checkbox"/> Yes <input type="checkbox"/> No
2.	How is data quality assessment done? <i>(If answer to 2 of data generation section is 'No', please select 'Manually')</i>	<input type="checkbox"/> Automatically <input type="checkbox"/> Manually <input type="checkbox"/> Hybrid <input type="checkbox"/> Not done If Hybrid, please specify how: _____
3.	<i>(Respond if answer to 2 is not "not done", else skip to Q1 of next section)</i> Are following protocols followed during data quality assessment?	
	a. Incoming data is filtered/cleaned after checking for missing values, logical flaws in data, incorrect values etc.	<input type="checkbox"/> Yes <input type="checkbox"/> No
	b. Summary statistics of incoming data are generated and checked for errors/abnormalities	<input type="checkbox"/> Yes <input type="checkbox"/> No
	c. Existence and accuracy of metadata for all the scheme's data is periodically checked (Schema is well defined)	<input type="checkbox"/> Yes <input type="checkbox"/> No
	d. There is a system for identifying duplicate data and removing redundancies	<input type="checkbox"/> Yes <input type="checkbox"/> No
	e. There is a system to ensure data is accurate, consistent and traceable to origin/source, whenever it is reproduced by any agency (data integrity)	<input type="checkbox"/> Yes <input type="checkbox"/> No
4.	Are following feedback mechanisms/backchecks also leveraged for data quality control?	
	a. Social audits	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	b. Telephonic backchecks/verification with beneficiaries	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	c. Multimedia data – citizen voice, video, images as evidence	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	d. Sample inspections based on data	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	e. Third party data verification/ data audits	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA

C. Data Analysis, Use and Dissemination

1.	What types of data analysis is undertaken on collected data?	
	a. Descriptive data analysis (e.g. basic cross tabulation, frequency distribution, mean, median etc.)	<input type="checkbox"/> Yes <input type="checkbox"/> No
	b. Exploratory data analysis (e.g. correlation etc.)	<input type="checkbox"/> Yes <input type="checkbox"/> No
	c. Inferential data analysis (Using a small sample of data to infer about a larger population)	<input type="checkbox"/> Yes <input type="checkbox"/> No
	d. Predictive analysis (Using historical or current data to find patterns to make predictions about the future)	<input type="checkbox"/> Yes <input type="checkbox"/> No
	e. Causal analysis (Looks at the cause and effect of relationships between variables, focused on finding the cause of a correlation)	<input type="checkbox"/> Yes <input type="checkbox"/> No
	f. Mechanistic Analysis (Understand exact changes in variables that lead to other changes in other variables)	<input type="checkbox"/> Yes <input type="checkbox"/> No
	g. Others - Please specify the name and the type of data analysis - _____	
2.	<i>(Respond if answer to any of the options in 1 is "yes", else skip to Q5)</i> Is cross-schematic/sectoral data also analysed, wherever needed?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA. If NA, please explain why: _____ _____
3.	How often is this data analysis well-documented (in reports/notes/publications)?	<input type="checkbox"/> Real-Time on a dashboard <input type="checkbox"/> Quarterly <input type="checkbox"/> Half-yearly <input type="checkbox"/> Annually <input type="checkbox"/> Never
4.	How often is this data analysis being used by the M/D officials for:	
	a. To re-design the schemes or activities undertaken under the scheme at the end of the tenure?	<input type="checkbox"/> Real-Time on a dashboard <input type="checkbox"/> Quarterly <input type="checkbox"/> Half-yearly <input type="checkbox"/> Annually <input type="checkbox"/> Never
	b. To do mid-course corrections through design or implementation changes ?	<input type="checkbox"/> Real-Time on a dashboard <input type="checkbox"/> Quarterly <input type="checkbox"/> Half-yearly <input type="checkbox"/> Annually <input type="checkbox"/> Never
	c. To guide intra-scheme funding decisions like inter-state allocations, inter-component allocations, etc.?	<input type="checkbox"/> Real-Time on a dashboard <input type="checkbox"/> Quarterly <input type="checkbox"/> Half-yearly <input type="checkbox"/> Annually <input type="checkbox"/> Never
	d. To guide inter-scheme budgetary allocations?	<input type="checkbox"/> Real-Time on a dashboard <input type="checkbox"/> Quarterly

		<input type="checkbox"/> Half-yearly <input type="checkbox"/> Annually <input type="checkbox"/> Never
	e. To decide quarterly releases to implementing agencies?	<input type="checkbox"/> Real-Time on a dashboard <input type="checkbox"/> Quarterly <input type="checkbox"/> Half-yearly <input type="checkbox"/> Annually <input type="checkbox"/> Never
	f. For fraud management and analysis	<input type="checkbox"/> Real-Time on a dashboard <input type="checkbox"/> Quarterly <input type="checkbox"/> Half-yearly <input type="checkbox"/> Annually <input type="checkbox"/> Never
	g. Day to day delivery and monitoring of implementation/ performance of the scheme	<input type="checkbox"/> Real-Time on a dashboard <input type="checkbox"/> Quarterly <input type="checkbox"/> Half-yearly <input type="checkbox"/> Annually <input type="checkbox"/> Never
5.	What other modes are used to disseminate the MIS/ paper-based data and related data analysis?	
	a. Dashboard	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	b. Mobile App	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	c. Social Media	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	d. SMS	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	e. Newspapers/ Magazines	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	f. Outdoor media (signages/ billboards)	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	g. Events	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	h. TV/ Radio	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	i. Others - Please mention the mode - _____	
6.	<i>(Respond if 'Yes' in 5a, else skip to 9)</i> What purposes are dashboards used for by the M/D officials?	
	a. Visual presentation of KPI/KRAs with drill-down capability to lowest level to gain total visibility	<input type="checkbox"/> Yes <input type="checkbox"/> No
	b. Capturing trends over time and identifying preempt trends	<input type="checkbox"/> Yes <input type="checkbox"/> No
	c. Measure efficiencies/inefficiencies in processes	<input type="checkbox"/> Yes <input type="checkbox"/> No
	d. User friendly one stop access to multiple automated reports	<input type="checkbox"/> Yes <input type="checkbox"/> No
7.	What types of Data Visualizations are used?	
	a. Bar chart/Histogram	<input type="checkbox"/> Yes <input type="checkbox"/> No
	b. Pie charts	<input type="checkbox"/> Yes <input type="checkbox"/> No
	c. Scatter plot	<input type="checkbox"/> Yes <input type="checkbox"/> No
	d. Heat maps	<input type="checkbox"/> Yes <input type="checkbox"/> No
	e. Treemaps	<input type="checkbox"/> Yes <input type="checkbox"/> No
	f. Gantt chart	<input type="checkbox"/> Yes <input type="checkbox"/> No

	g. Specialized visualizations- Stripe graphics, streamgraph, etc.	<input type="checkbox"/> Yes <input type="checkbox"/> No
	h. Others - please mention data visualizations used - _____	
8.	Does the Dashboard visualize information on maps?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
9.	<i>(Respond if 'Yes' in 3 of Data Generation section, else skip to Q6 of next section)</i> Does the MIS support multilingual features as per GIGW norms?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially (some norms followed but not all)
10.	Does the MIS support features for differently abled as per GIGW norms?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially (some norms followed but not all)
11.	How is the MIS data accessible for general population?	<input type="checkbox"/> Openly accessible without credentials <input type="checkbox"/> Accessible through credentials <input type="checkbox"/> Not accessible
12.	Is there an option on the MIS to download bulk data in excel, csv, dta files (machine readable formats)?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partial data download allowed
13.	Is the MIS data available on 'data.gov.in'?	<input type="checkbox"/> Yes <input type="checkbox"/> No

D. Use of Technology

1.	<i>(Respond if 'Yes' in 3 of Data Generation section, else skip to Q4)</i> Does the MIS of the scheme have linkages with PFMS?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
2.	<i>(Respond if answer to 1 is yes)</i> Is PFMS integration completed till the field-level implementation agency?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially
3.	Does the MIS of the scheme have linkages:	
	a. Aadhaar	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	b. Mobile numbers	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	c. Bank accounts	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	d. GSTN	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	e. Udyog Aadhaar	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	f. Others – please specify	
4.	Does the scheme use any of the following:	
	a. Remote sensing data	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	b. Night light data	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	c. Social media data	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	d. Private sector generated data	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	e. Others – please specify	

5.	<i>(Respond if 'Yes' in 3 of Data Generation section, else skip to Q6)</i> Is the MIS compliant with Local Govt Directory (LGD)?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
6.	Does the scheme apply/use any of the following:	
	a. Machine Learning	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	b. Artificial Intelligence	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	c. Blockchain	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	d. Internet of Things (IoT)	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	e. Big Data analytics	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA
	f. Drones	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> NA

E. Data Security and HR Capacity

1.	<i>(Respond if 'Yes' in 3 of Data Generation section, else skip to Q10)</i> Does the MIS follow regular antivirus updates?	<input type="checkbox"/> Yes <input type="checkbox"/> No
2.	Is the MIS regularly assessed by third party auditors for the online security?	<input type="checkbox"/> Yes <input type="checkbox"/> No
3.	Does the MIS/ website uses SSL certificate?	<input type="checkbox"/> Yes <input type="checkbox"/> No
4.	If "Yes" in previous question, is the SSL certificate at least 2048 bit SHA 256 encryption or higher?	<input type="checkbox"/> Yes <input type="checkbox"/> No
5.	Does the MIS use firewalls to secure access to data?	<input type="checkbox"/> Yes <input type="checkbox"/> No
6.	All external communication/ 3rd party integration/ API integration for the MIS is done through encrypted channel?	<input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> No external communication established
7.	What measures are undertaken to secure sensitive/personally identifiable information? (Multiselect)	
	<input type="checkbox"/> Single-factor/ Multi-factor authentication	
	<input type="checkbox"/> Access control list is maintained	
	<input type="checkbox"/> Data is encrypted	
	<input type="checkbox"/> Data is anonymized	
	<input type="checkbox"/> No such data	
8.	(If anonymization is selected in previous question) How do you protect de-identified data from re-identification risks?	
	<input type="checkbox"/> No efforts made	
	<input type="checkbox"/> Tighter security for databases that store anonymized information	
	<input type="checkbox"/> Implementation of Differential Privacy	
	<input type="checkbox"/> Generation of Synthetic Data that exhibits the statistical properties of the raw data, without allowing real individuals to be identified	
	<input type="checkbox"/> Others - provide details – _____	
9.	<i>(Respond if answer to 7 is any option other than "no such data", else skip to Q10)</i> Is permission taken from user to collect, store and use their personal data?	<input type="checkbox"/> Yes <input type="checkbox"/> No
10.	Is there a dedicated data quality assessment and management team for the scheme?	<input type="checkbox"/> Yes <input type="checkbox"/> No

11.	Is there a dedicated data analysis team for the scheme?	<input type="checkbox"/> Yes <input type="checkbox"/> No
-----	---	--

F. Data Management

<i>Respond if you have answered 'Yes' in 3 of Data generation section, else skip this section.</i>		
1.	Where is MIS data stored?	<input type="checkbox"/> On separate servers for different schemes (distributed storage) <input type="checkbox"/> On central server which is used for all schemes
2.	<i>(Respond if first option is selected in 1, else skip this question)</i> Are there mechanisms in place which can enable data sharing with other scheme divisions?	<input type="checkbox"/> Yes <input type="checkbox"/> No If yes, please explain how: <hr/>
3.	How is MIS data stored?	<input type="checkbox"/> Physical servers <input type="checkbox"/> Cloud Storage <input type="checkbox"/> Others
4.	<i>(Respond if "Cloud Storage is selected in 3, else skip this question)</i> Which cloud service is being used?	<input type="checkbox"/> NIC/ Gov cloud- Meghraj <input type="checkbox"/> Cloud Services directly from CSP (Cloud Service Provider) <input type="checkbox"/> Cloud Services through System Integrator <input type="checkbox"/> Cloud Services through Managed Service Provider (MSP)
5.	How is historical MIS data managed?	<input type="checkbox"/> Data is not backed up (i.e. it is destroyed) <input type="checkbox"/> Data is backed up and data is archived <input type="checkbox"/> Data history is well maintained including retention, destruction, and audit trail details

Appendix I

Detailed explanations to questions of DGQI Self-Assessment Questionnaire for Ministries/Departments of GoI (2021-22)

Part	Section	Question	Explanation
A	A	1	M/D name would be automatically filled up when the M/D logins using their credentials.
A	A	2	A pre-populated list of CS schemes of the M/D would be visible here.
A	A	3	A pre-populated list of CS schemes of the M/D would be visible here.
A	A	4	M/Ds to enter any other non-schematic intervention such as sector dashboards, sector level MIS, any other MIS/dashboards etc. that they would like to include for DGQI assessment using the self-assessment questionnaire.
A	A	5	M/Ds to enter details of DGQI nodal officer. He/she would be assumed to have verified the correctness and authenticity of the information filled in this self-assessment form.
A	B	1	Constitution refers to establishing the unit, hence, even if its staffing is ongoing, M/Ds can select 'yes' if they have established the admin structure of the unit and some members have been assigned to it.
A	B	4	M/Ds to undertake calculations at their end based on how many posts they have proposed to create for the DSU based on their requirement and how many of these posts have been filled up. Total will be auto-calculated using values entered in the table.
A	B	5	Documentation of terms of reference here refers to the documentation of detailed objectives, roles and responsibilities of the DSU specific to the M/D. Indicative ToR for guiding M/Ds was shared by DMEO earlier.
A	B	6	M/Ds to select yes if guidelines for a standard system for regularly scheduling review meetings (via OM etc.) has been issued.
A	C	1	M/Ds to select yes if they have completed preparation of exhaustive action plan to improve data preparedness levels of the M/D.
A	C	2	Action plan is to have 3 sections with all sub-sections: 1. Background, 2. Vision, Mission & Objectives, 3. Strategy – Scope, Overall approach, scheme wise strategy, non-schematic strategy, operational execution plan.
A	C	3	M/Ds to select yes if the action plan has separate action points for all CS/CSS schemes of the M/D (as per the list on this portal).
A	C	4	M/Ds to select yes if every action point has a corresponding mm/yy timeline by which it is aimed to be completed, clearly documented in the action plan.

A	C	5	M/Ds to select yes if every action point is mapped to unit/personnel within the M/D by whom it is expected to be completed, clearly documented in the action plan.
A	D	1	Data management guidelines/architecture explains how data is to be managed across its lifecycle, i.e., how is it to be collected, stored, processed? How will it be exchanged? What will be done with historical data?
A	D	3	Data ownership norms would define who would be the owner of data when data is shared with other divisions or M/Ds or in public.
A	D	4	Understanding the value of the data collected by the M/D from utility perspectives and comparing it to the associated data security and privacy risks to ensure there is a balance between the two.
A	D	5	Data ethics refers to systemizing, defending, and recommending concepts of right and wrong conduct in relation to data, particularly personal data. With use of machine learning and predictive algorithms, it becomes even more important to protect sensitive data.
A	E	1	Data gaps refer to data that is required by the M/D from decision making point of view, however, for some reasons, such data is not available with the M/D. M/Ds to select yes if they have identified such data gaps based on analysis of their current data.
A	E	2	After identification of data gaps, M/Ds must take reform actions to develop data capture mechanisms/exchange mechanisms to fill up data gaps. M/Ds to select yes if they have started planning these actions.
A	E	3	For schemes with similar target groups, data collection can be done together rather than separately. This is an example of integrated data systems for collection. Similarly, if one scheme is collecting data on some indicator which is required by another division on its portal, it should be able to get this data from the scheme division via suitable exchange systems. M/Ds to select yes if this is possible currently.
A	F	1	If M/Ds are collecting similar data or running similar interventions, data collaborations can be undertaken. If private sector has some useful data (let's say e-commerce or traffic data), data collaborations can be undertaken by M/Ds.
A	F	2	If some data collaboration has been undertaken, M/Ds to select how it has been done from the given options.
A	G	1	Prescriptive analytics is the final stage of analytical capabilities. While predictive analytics answers what, when and why something will happen, prescriptive analytics builds on this further by specifying what present actions need to be undertaken to achieve the predictions and how will these decisions affect /impact other outcomes. Therefore, it helps in taking advantage of a future opportunity or mitigating future risks. It can also improve the

			accuracy of predictions by continuously taking in new data to re-predict and re-prescribe.
A	G	2	M/Ds to select the frequency of prescriptive analytics.
A	G	3	M/Ds to select the modes/mechanisms by which they have institutionalized prescriptive analytics, to ensure it is continuously undertaken to inform policymaking, and not just undertaken on random basis.
A	H	1,2,3	M/Ds to enter good practices of how they have used data for policymaking and/or set up systems for institutionalizing data driven policymaking.
B	A	1	Data requirements refer to various input, output, and outcome data points/indicators that need to be monitored. They must be clearly documented for each scheme. M/Ds to select yes if this is done.
B	A	2	After gathering of data requirements, scheme division to select the indicators for which it is collecting data also. For e.g.: If scheme has multiple outcome indicators documented but the division is collecting data on only some of them due to various reasons, it must select Partial.
B	A	3	After data is collected, it must be collated and reported via paper or digitally through a MIS. Scheme divisions to accordingly choose Yes/No. Regarding credentials, this is optional. However, M/Ds are encouraged to create dummy login credentials for DMEQ with view-only rights. This shall stay confidential and not be used outside the government for unintended purposes.
B	A	4	Scheme division to select all the granularities at which data is reported on the MIS. For e.g.: if a scheme MIS has district, state as well as national level data, scheme division to select all three options.
B	A	5	Scheme division to select the frequency at which data is updated on the MIS.
B	A	6	This question is to essentially understand if the data reported by the M/D on the MIS is “collected” by humans or is it transaction-based collection. If it is collected by human resources, is it directly collected using digital tablets/mobiles etc. or is it the case that it is first collected on paper and then fed on computers by someone else.
B	A	7	If data is collected using digital modes or it is transactional in nature, use of survey tools and/or geotagging can improve the data reliability. Scheme division to select yes if the same is done.
B	A	7a	Computer-assisted personal interviewing (CAPI) refers to survey data collection by an in-person interviewer (i.e., face-to-face interviewing) who uses a computer to administer the questionnaire to the respondent and captures the answers onto the computer.

B	A	7b	A geotagged photograph is a photograph which is associated with a geographic position by geotagging. Usually this is done by assigning at least a latitude and longitude to the image, and optionally altitude, compass bearing and other fields may also be included.
B	A	7c	Geofencing is a location-based service which triggers some pre-programmed action like a survey when a mobile device or RFID tag enters or exits a virtual geographical boundary.
B	B	1	Data quality protocols and mechanisms should be clearly documented by the scheme division. Scheme division to select yes if the same is done.
B	B	2	Data quality assessment of collected data against data quality protocols can be undertaken automatically by advanced digital systems, manually or using a hybrid of both manual and automated systems.
B	B	3	Question to assess which protocols are included and followed by the scheme division in its data quality assessment.
B	B	3a	This is the first step of data quality where collected data is cleaned by checking missing values, incorrect responses etc.
B	B	3b	Next step is to generate summary statistics of data (like mean, median, trends etc.) to check for outliers
B	B	3c	Another important step is to ensure metadata is properly defined. Metadata is data about data – containing details on variables covered in the data, their number of observations, summary statistics, units etc. This must also be regularly updated if new data is collected.
B	B	3d	Next important protocol is to check collected data for duplicate values (this duplication may be in old data or new data) and remove any such redundancies
B	B	3e	Finally, ensuring data integrity. This means that if collected data is being reflected anywhere (on the MIS, on any other portal etc.), it must be ensured that the accurate and recent most value is reflected everywhere. It should not be the case that at one place, data is updated as of last month, but at other portal, it is updated as of last year or showing inaccurate value due to some error.
B	B	4	Apart from data qual assessment, backchecks may be deployed to further improve data quality and increase its reliability.
B	B	4a	Social audit is a form of citizen participation that focuses on government performance and accountability. If social audits are being used to improve scheme data, select yes.
B	B	4b	If telephonic backchecks are undertaken based on collected data to verify that data is correctly collected, select yes. E.g.: Based on PDS beneficiary data available on MIS, random sample of ppl are contacted on phone to validate data entries made on MIS.
B	B	4c	If there are provisions for citizens to submit multimedia evidence which is then used to improve the quality of data, select yes. For eg: People submitting photos of quality of roads built near their

			locations and this feedback data being used to reflect the quality of roads on MIS.
B	B	4d	Based on reported data on MIS, random inspections are made by MD officials to verify data on ground.
B	B	4e	Getting data on MIS verified/audited by third parties.
B	C	1	Scheme divisions to select all types of data analysis undertaken by them.
B	C	2	Apart from scheme data, if data from other schemes or sector level data is also used to complement scheme data for analysis purpose, select yes.
B	C	3	Data analysis must be documented in some manner. Select the frequency at which this is done.
B	C	4	M/Ds to select the uses for which data analysis is done.
B	C	5	Select different modes used for disseminating data and its analysis.
B	C	5a	DB is essentially a tool to display key KPIs from data and important analytics through interesting visualizations.
B	C	5b	Mobile apps can be used to share data with citizens and interact with them.
B	C	5c	Social media outlets can be used to share data with citizens and interact with them.
B	C	5d	SMS are often used to send details to users/beneficiaries with respect to the scheme activities.
B	C	5e	Such mass communication methods may also be used to share data with citizens.
B	C	6	M/Ds to select the purposes for which dashboard are being used by them.
B	C	9	This is important to ensure data is accessible to all.
B	C	10	This is important to ensure data is accessible to all.
B	C	11	All MIS may not be in public domain. Hence, scheme divisions to enter details on how can public in general can access MIS data.
B	C	12	Scheme divisions to check if there is an option to download all MIS data in machine readable formats by users on the MIS and accordingly select.
B	C	13	As per NDSAP, all non-personal data should be available on data.gov.in to facilitate easy access to all govt. data at one place. Scheme division to select yes if non-personal data of their MIS is available on this platform.
B	D	1	MIS linkage with PFMS means that latest status of funds being routed through PFMS should be linked with MIS of the scheme.
B	D	2	The field-level implementation agency is the last agency to which funds are to flow. For eg: if PFMS integration is done till state implementing agency level but fund flow below states is not PFMS integrated for a scheme where projects are implemented by city level agencies, integration is not completed till last mile.
B	D	3a	Applicable for beneficiary-oriented schemes
B	D	3b	Applicable for beneficiary-oriented schemes
B	D	3c	Applicable for beneficiary-oriented schemes

B	D	3d	Applicable for industry/firm oriented schemes
B	D	3e	Applicable for industry/firm oriented schemes
B	D	4a	Remote sensing is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted day-time radiation at a distance (typically from satellite or aircraft). Special cameras collect remotely sensed images, which help researchers "sense" things about the Earth. For e.g.: large forest fires can be mapped from space, Tracking clouds to help predict the weather or watching erupting volcanoes, and help watching for dust storms, tracking the growth of a city etc.
B	D	4b	Night-light data is basically the data of night-time lights emanating from the earth captured by satellites from outer space. These sources include moonlight, light directly emitted by a source (e.g., buildings and transport), and light reflected by the ground. It has several use cases - aid in disaster mitigation, estimating economic activity etc.
B	D	4c	This data is collected from social media networks to see how people are engaging on specific topics of interest. Scheme divisions may use the same to check for behavior change etc.
B	D	4d	Scheme may use data generated by private sector also as per requirement. For eg: mobility data from private cab aggregators, economic activity data from e-commerce websites etc.
B	D	5	Unique LGD codes have been created for each state, distt, sub-distt, block, village and local body by GoI. All MIS must use the same codes so that data on different platforms is easily integrable.
B	D	6a	Machine learning gives computers the ability to learn and predict from data without being explicitly programmed. E.g.: predicting the probability that individuals commit crimes, targeting hygiene inspections by data-mining online restaurant reviews or estimating poverty levels based on satellite imagery.
B	D	6b	AI refers to intelligence demonstrated by machines and can have several use cases in governance and delivery of schemes. e.g.: Monitoring social media for public feedback on policies, Monitoring social media to identify emergency situations, Anticipating road maintenance requirements, Providing personalized education to students etc.
B	D	6c	Blockchain refers to having distributed ledgers or blocks of transactional data that are linked together. Using this structure, govt. can offer services with improved data security. For e.g.: electronic health records, e-registries etc.
B	D	6d	IoT refers to network of objects embedded with sensors and technologies for collecting and exchanging data over Internet. e.g.: IoT to measure air quality, IoT to monitor power consumption i.e., smart metering etc.
B	D	6e	The use of advanced analytic techniques against very large, diverse data sets that include structured, semi-structured and unstructured data, from different sources.

B	D	6f	Drones can be used for monitoring of various sectors like agri, infra projects, commerce, logistics etc.
B	E	3	An SSL certificate is a digital certificate that authenticates a website's identity and enables an encrypted connection. SSL stands for Secure Sockets Layer, a security protocol that creates an encrypted link between a web server and a web browser.
B	E	5	a firewall refers to a network device which blocks certain kinds of network traffic, forming a barrier between a trusted and an untrusted network.
B	E	7	Sensitive/PII contains personal information of individuals, firms etc. which are not freely accessible to all.
B	E	7a	Single-Factor Authentication (SFA) is a method of logging users by having them present only one way of verifying their identity (usually, username and password). Multi-factor authentication uses more than one way – such as OTP, Captcha etc.
B	E	7b	List of users of MIS along with details of which user has access to which type of data is regularly maintained.
B	E	7c	Encryption refers to conversion of data from readable format to encoded format. Encrypted data can only be read and processed after its decrypted by recipient if they have the codes.
B	E	7d	Data anonymization refers to the process by which personal data is altered in a way that the data subject can no longer be identified directly by data user.
B	E	8	With advancements in machine learning and big data analytics, it is becoming increasingly easier to de-identify anonymized data using indirect means. Hence, it is important to protect personal data from re-identification risks.
B	E	8b	Includes provisions for mandatory audit trails, controlled access, only central server logins allowed etc.
B	E	8c	Sharing information about a dataset by describing the patterns of groups within the dataset while withholding information about individuals.
B	E	9	Before using and putting personal data in public domain such as photographs, names, other details of individuals or firms, their consent must be asked for and documented.
B	F	2	Before using and putting personal data in public domain such as photographs, names, other details of individuals or firms, their consent must be asked for and documented.
B	F	3	Data may be stored on physical servers or cloud servers. Cloud servers offer better disaster recovery.
B	F	4	Select the cloud server used by the scheme MIS.
B	F	5	Historical data refers to data corresponding to previous time periods which may not be actively used at present.