Item No	Concept name	Definition
1	**Contact**	**Individual or organisati**
1.1	Contact Organisation	Organisation of the contact point(s) for the data or metadata.
1.2	Compiling agency	Organisation collecting and/or elaborating the data being reported
1.3	Custodian agency	Refers to an institution or agency which has responsibility of managing use, disclosure and protection of data/statistics.
1.4	Contact Details	The details of the contact points for the data or metadata.
2	**Data Description and Presentation**	**Descrip**

2.1	Data description	Metadata element describing the main characteristics of the Data Set in an easily understandable manner, referring to the main data and indicators disseminated.
2.2	System of Classification	Arrangement or division of objects into groups based on characteristics which the objects have in common
2.3	International/ National Standards Classification etc.	An international statistical standard is an internationally agreed statistical macro-socio-economic output framework or a cross-functional framework. International/ National standard classification, is the primary tool for collecting and presenting internationally/nationally comparable statistics by economic activity.
2.4	Sector coverage	Description of sectors covered by the statistics.
2.5	Concepts and definitions	Characteristics of statistical observations, or variables used

2.6	Unit of compilation	Entity for which information is sought and for which statistics are ultimately compiled.
2.7	Population coverage	Definition of the main types of population covered by the statistics or variables
2.8	Reference Period	Timespan or point in time to which the measured observation is intended to refer.
2.9	Duration and period of enumeration	Duration of enumeration refers to the time taken to conduct the enumeration process of all nature.
2.1	Sample size/Dataset Size	Sample size is the number of observations or individuals included in a study or experiment.
2.11	Data Confidentiality	Rules applied for treating the datasets to ensure statistical confidentiality and prevent unauthorised disclosure.
3	**Institutional Mandate**	**Law, set of rules or other formal set of instructio**
3.1	Legal acts and other agreements	Legal acts or other formal or informal agreements that assign responsibility as well as the authority to an Agency for the collection, processing, and dissemination of statistics.

3.2	Data sharing/Data Dissemination	Exchange of data and/ or metadata in a situation involving the use of open, freely available data formats and where process patterns are known and standard. Regular or ad hoc publications in which the data are made available to the public.
3.3	Release calendar	Schedule of release dates.
3.4	Frequency of dissemination	The time interval at which the statistics are disseminated over a given time period.
3.5	Data access	The conditions and modalities by which users can access, use and interpret data. Statistics should be easy to find and obtain, presented clearly and in a way that can be understood, and available and accessible to all users in line with open data standards.
4	**Quality Management**	**Systems and fram**
4.1	Documentation on methodology	Descriptive text and references to methodological documents available.
4.2	Quality documentation	Documentation on procedures applied for quality management and quality assessment.

4.3	Quality assurance	Guidelines focusing on quality in general and dealing with quality of statistical programmes, including measures for ensuring the efficient use of resources.
5	**Accuracy and Reliability**	**Statistics should accurately and reliably portray r measure. Reliability**
5.1	Sampling error	That part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a subset of the population is enumerated.
5.2	Measures of reliability	Reliability refers to the consistency of a measure.
6	**Timeliness**	**Statistics should be made available to users with the**
6.1	Timeliness	Length of time between data availability and the event or phenomenon they describe.
7	**Coherence and Comparability**	**Coherence is the degree to which data that are de**

7.1	Comparability – over time	Extent to which differences between statistics can be attributed to differences between the true values of the statistical characteristics.
7.2	Coherence	Extent to which statistics are reconcilable with those obtained through other Data Sources or statistical domains.
8	**Data Processing**	
8.1	Source data type	Characteristics and components of the raw statistical data used for compiling statistical aggregates.
8.2	Frequency of data collection	Time interval at which the source data are collected.
8.3	Mode and method of data collection method	The different combinations of data collection modes, pre-contact and follow-up modes. Method applied for gathering data for official statistics

8.4	Data validation	Process of monitoring the results of data compilation and ensuring the quality of the statistical results.
8.5	Data compilation	Operations performed on data to derive new information according to a given set of rules.
8.6	Data identifier(s)	The unique identifier for an administered item within a registration authority.
9	**Metadata Update**	**T**
9.1	Metadata last posted	Date of the latest dissemination of the metadata
9.2	Metadata last update	Date of last update of the content of the metadata.

Guidelines
ional contact points for the data or metadata, including informat
Provide the full name (not just acronym/code name) of the organisation responsible for the processes and outputs (data and metadata) that are the subject of the report
Provide the full name of the Department/Division under the organisation responsible for the processes and outputs (data and metadata) that are the subject of the report
Data/statistics custodians are agencies responsible for managing the use, disclosure and protection of source data used in a statistical ecosystem. Data custodians collect or authorise to collect and hold information on behalf of a data provider (defined as an individual, household, business or other organisation which supplies data either for statistical or administrative purposes). The role of data custodians may also extend to producing source data, in addition to their role as a holder of datasets.
Provide contact details of contact point(s) in following format: a. Name of Organisation owning the processes and outputs b. Author (if different from (a)) c. Disseminating Agency (if different from (a) and (b)) d. Name (first and last names) e. Designation f. Postal address g. email address (preferably designation based) h. Contact number i. Fax numberIf more than one name is provided, the details of main contact should be indicated. If the author of the report is different from the person(s) responsible for process and its outputs, provide this name also with his/her details
tion of the disseminated data which can be displayed to users as

This summary description should provide an immediate understanding of the data to users (also to those who do not have a broader technical knowledge of the Data Set in question).
Data can be displayed to users as tables, graphs or maps. According to the United Nations' Fundamental Principles of Official Statistics, the choice of appropriate presentation methods should be made in accordance with professional considerations. Data presentation includes the description of the Data Set disseminated with the main variables covered, the classifications and breakdowns used, the reference area, a summary information on the time period covered and, if applicable, the base period used.

List all classifications and breakdowns that are used in the data (with their detailed names) and provide links (if publicly available).
Type of dis-aggregation available in the data sets - for example rural-urban, male-female, etc. and whether data is available at the sub-national level or not, should be clearly specified.

An international statistical standard is an internationally agreed statistical macroeconomic output framework or a cross-functional framework. The framework consisting of concepts, definitions, classifications and inter-related tables for integrating broad set of statistics. The framework is internally coherent and consistent and externally consistent, to the extent possible, with other international statistical standards.

This metadata element describes all applicable sectors associated with the Data Set. An example is "Production units and households" in environmental accounts.

Define and describe briefly the main statistical observations or variables that have been observed or derived. Indicate their types.

Define the type of statistical unit about which data are collected, e.g. enterprise, household, etc.
The population coverage describes the types of population covered by the statistics or variables whenever applicable.
A reference period is the time period for which statistical results are collected or calculated and to which the data refer. The time period may be a calendar year (reference year), a fiscal year, a semester, a quarter, a month and even a day (reference date).
The enumeration period refers to the specific time of year when enumeration operations took place.
Sample size is a key feature potentially affecting the results of all experimental studies. A larger sample size can potentially enhance the precision of estimates, leading to a narrower margin of error. In other words, the results from a larger sample will likely be closer to the true population parameter. A larger sample size can also increase the power of a statistical test.
Describe the procedures that are used in protecting confidentiality, viz., anonymisation, legal provision, if any.

ns assigning responsibility as well as the authority to an organisa

The concept covers provision in law assigning responsibility to specific organisations for collection, processing, and dissemination of statistics in one or several statistical domains. In addition, non-legal measures such as formal or informal administrative arrangements employed to specific organisations for collection, processing, and dissemination of statistics in one or several statistical domains should also be described.

Data Sharing- In data sharing exchange, any organisation or individual can use any counterparty's data and metadata (assuming they are permitted access to it). This model requires no bilateral agreement, but only requires that data and metadata providers and consumers adhere to the standards. Dissemination-This metadata element provides references to the most important data dissemination done through paper or online publications, including a summary identification and information on availability of the publication means.
An advance release calendar is the schedule for release of data, which are publicly disseminated so as to provide prior notice of the precise release dates on which a national statistical agency, other national agency, or international organisation undertakes to release specified statistical information to the public. Such information may be provided for statistical releases in the coming week, month, quarter or year.
The frequencies with which data are released, which could be different from the frequency of data collection.
State the conditions and link on website from where the user can access the data. For easy access of users, following details should also be mentioned about the dataset: **Title:** Name by which the data is known **Dataset Edition:** Edition of data (ex: first, second, final etc) **Dataset Reference data type:** Type of data entered in the field (ex: .txt, .dbf, .xls) **Presentation Format:** Presentation format of the data (ex: document, map, table, etc.) **Dataset Language:** language of any text in the data. **Status/Version:** How updated is the data?

eworks in place within an organisation to manage the quality of s

This metadata element refers to the availability of documentation related to various aspects of the data, such as methodological documents, summary notes or papers covering concepts, scope, classifications and statistical techniques.
This metadata element is used to document the methods and standards for assessing data quality, based on standard quality criteria such as relevance, accuracy and reliability, timeliness and punctuality, accessibility and clarity, comparability, and coherence.

This metadata element refers to all the planned and systematic activities implemented that can be demonstrated to provide confidence that the data production processes will fulfil the requirements for the statistical output. This includes the design of programmes for quality management, the description of planning process, scheduling of work, frequency of plan updates, and other organisational arrangements to support and maintain planning function.

eality. Accuracy of data is the closeness of computations or estim
of the data, defined as the closeness of the initial estimated valu

Sampling errors are distinct from errors due to imperfect selection, bias in response or estimation, errors of observation and recording, etc. For probability sampling, the random variation due to sampling can be calculated. For non-probability sampling, random errors cannot be calculated without reference to some kind of model. The totality of sampling errors in all possible samples of the same size generates the sampling distribution of the statistic which is being used to estimate the parent value.

Reliability is how consistently a method measures some parameters. When one applies the same method to the same observation under the same conditions, one should get the same results.

shortest delay possible and be delivered on the promised, adverti
to be compiled.

Timeliness refers to the speed of data availability, whether for dissemination or for further processing, and it is measured with respect to the time lag between the end of the reference period and the release of data. Timeliness is a crucial element of data quality: adequate timeliness corresponds to a situation where policy-makers can take informed decisions in time for achieving the targeted results. In quality assessment, timeliness is often associated with punctuality, which refers to the time lag between the release date of data and the target date announced in some official release calendar.

rived from different sources or methods, but refer to the same to
compared over time and domain

Comparability aims at measuring the impact of differences in applied statistical concepts and definitions on the comparison of statistics between geographical areas, non-geographical dimensions, or over time. Comparability of statistics, i.e. their usefulness in drawing comparisons and contrast among different populations, is a complex concept, difficult to assess in precise or absolute terms. In general terms, it means that statistics for different populations can be legitimately aggregated, compared and interpreted in relation to each other or against some common standard. Metadata must convey such information that will help any interested party in evaluating comparability of the data, which is the result of a multitude of factors.

This metadata element is used to describe the differences in the statistical results calculated on the basis of different statistical domains, or surveys based on different methodologies (e.g. between annual and short-term statistics or between social statistics and national accounts).

Any processing undertaken to finalise the data

This metadata element is used to indicate whether the dataset is based on a survey, on administrative data sources, on a mix of multiple data sources or on data from other statistical activities. If sample surveys are used, some sample characteristics should also be given (e.g. population size, gross and net sample size, type of sampling design, reporting domain). If administrative registers are used, the description of registers should be given (e.g. source, primary purpose, etc.).

The frequencies with which the source data are collected and produced could be different: a time series could be collected from the respondents at quarterly frequency but the data production may have a monthly frequency. The frequency of data collection should therefore be described.

The consolidation of an integrated management system underlies the feasibility of several management choices, such as those concerning mode changes for a given household. Three modes of data collection (Computer Assisted Web Interviewing (CAWI), Computer Assisted Telephonic Interviewing (CATI), and Computer Assisted Personal Interviewing CAPI) based on a common questionnaire (unimode design), and with both sequential and concurrent mode organization features.
For each source of survey data:
· describe the method(s) used to gather data from respondents;
· annex or hyperlink the questionnaires(s).
For each source of administrative data:·
describe the acquisition process and how it was tested.
For all sources:·
describe the types of checks applied at the time of data entry.

Data validation describes methods and processes for assessing statistical data, and how the results of the assessments are monitored and made available to improve statistical processes. All the controls made in terms of quality of the data to be published or already published are included in the validation process. Validation also takes into account the results of studies and analysis of revisions and how they are used to improve statistical processes. In this process, two dimensions can be distinguished: (i) validation before publication of the figures and (ii) validation after publication.

In quality assurance frameworks, "Data compilation" refers to the description of statistical procedures used for producing intermediate data and final statistical outputs. Data compilation covers, among other things, the use of weighting schemes, methods for imputing missing values or source data, statistical adjustment, balancing/cross-checking techniques, and relevant characteristics of the specific methods applied.

Identifier is a sequence of characters, capable of uniquely identifying that with which it is associated, within a specified context.
In order to exchange statistical information, a central institution has to agree with its partners about which statistical concepts are necessary for identifying the series (and for use as dimensions) and which statistical concepts are to be used as attributes. These definitions form the data structure definition. Each data structure

'he date on which the metadata element was inserted or modified

The date when the complete set of metadata was last disseminated as a block should be provided (manually, or automatically by the metadata system).

The date when any metadata were last updated should be provided (manually, or automatically by the metadata system).

Response
ion on how to reach the contact points.
Ministry of Labour & Employment
SS-III
Ministry of Labour & Employment
(a). Ministry of Labour & Employment
(b). Ministry of Labour & Employment
(c). Ministry of Labour & Employment
(d). Shri. Minhaz Ahmad
(e). Deputy Secretary (eShram)
(f). Room No. 313, 3rd Floor, Shram Shakti Bhavan,Rafi Marg, New Delhi -110001
(g). Email Adress - minhaz.ahmad@gov.in
(h). 011-23710260
(i). Fax Number
tables, graphs or maps

Demographics: Information on age, gender, and regional distribution of registered workers. **Occupational Categories:** Types of occupation and industries where workers are employed. **Registration statistics:** Total number of workers registered, trends over time and occupation and state/ district - wise breakdown.
Various indicators are being captured on eShram at the time of registration on self-declaration by the unorganised workers. Key classifications being captured at eShram are as under: - Modes of registration (CSC, Self, SSK and Umang) - Gender (Male, Female & Other) - Age Group - State & district - Occupation (Primary as well as Secondary) The statistical data of eShram registration is publically available on website at https://eshram.gov.in/dashboard
Geographical Location - LGD code mapping Occupation - National Classification of Occupation (NCO) Pincode - Indian Post Office Gender - OECD
eShram targets to register all unorganised workers. To achieve this, around 400 occupations under 30 broad sectors including agriculture, construction, domestic workers, apparal, automobile etc. have been included.
Total Registrations (Type: Continuous), Age Distribution (Type: Categorical), Gender Distribution (Type: Categorical), Geographic Distribution (Type: Categorical), Occupational Categories (Type: Categorical), Mode of Registration (Type: Categorical)

The statistical unit about which data are collected on the eShram platform is individual workers in the unorganised sector. This includes various types of laborers, such as agricultural workers, construction workers and other unorganised sector workers. The data focus on personal, demographics, occupational details and social security access of these individual workers.
The target population of statistical units for the eShram platform includes individual workers in the unorganised sector across India. This encompasses labourers from diverse occupational backgrounds such as agricultural labourers, construction workers, street vendors, domestic workers, and other informal sector employees who typically lack formal employment benefits and social security coverage
The registration on eShram portal and unorganised workers data collection is continuous process through the registration module via CSC, Seva Kendra, Umang App and Self registration by the unorganised workers. Registration on the eShram portal commenced from 26th August 2021. eShram also allows workers to visit/ login on eShram and update their profile in case of change of address, occupation, contact details etc.
eShram is a registration portal, where registration commenced on 26 August 2021. Registration on eShram is a continous process
NA
Role Based Access Control, Least Privileged Principle, Transparent Data Encryption in database is enabled to ensure security of data. PI data is masked and not available publically.

tion for the collection, processing, and dissemination of statistics

The registation on eShram is done based on AADHAR authentication. Therefore, it is required that the eShram portal follows all the guidelines outlined by UIDAI. Accordingly all policies and procedures applied are based on UIDAI policies and framework.

MoLE in consultation with other key stakeholders including NIC, UIDAI etc. established Data Sharing Guidelines for sharing eShram registrants information in a secure manner with States/ UTs Governments and Central Ministries through Application Programming Interface (API) and Data Sharing Portal (DSP).
All entities with whom eShram data is being shared are expected to adhere to eShram Data Sharing Guidelines. Statistical information of eShram registration is available on eShram dashboard.
eShram anonymized data is also available through Open Government Data (OGD).

eShram dashboard with statistical information is made available at https://eshram.gov.in/dashboard for everyone.
Data on public dashboard is updated at a daily frequency on n-1 basis

Data on public dashboard as well as OGD platform is updated at a daily frequency on n-1 basis

Registered workers can access their data by logging into the eShram registration portal.
Public dashboard is available on eShram website, where statistics related to data is displayed. The data is being shared to State/ UT Governments as per data sharing guidelines using Bulk AP1 and Data Sharing Portal.
Data Reference Data Type - DB format

Open Government Data: https://www.data.gov.in/resource/district-wise-demographic-data-unorganised-workers-registered-eshram-previous-day

statistical products and processes.

ates to the exact or true values that the statistics were intended to e to the subsequent estimated value.

NA The concept described, sampling error arising from estimating population values using a subset, is not applicable to eShram, as the platform is designed to be a comprehensive, self-registration based National Database of Unorganised Workers. Unlike sample based surveys, eShram aims to capture actual data from individual workers across India, thereby minimising reliance on statistical estimation. Therefore, differences between population values and estimates due to sampling are not relevant in the context of eShram's data collection approach.

NA

ised or announced dates. The timeliness of the data collection release

Unorganized workers of eShram are registered online, the data is available on N-1 basis on the dashboard or to State/ UT Governments through API/ DSP. Workers registering on eShram and access and update their information on the real time.

pic, are similar. Comparability is the degree to which data can be

The concept, referring to the extent to which differences between statistics reflect actual differences in the true values of the underlying characteristics, is not directly applicable to eShram. Since eShram is a real-time, individual level registry of unorganised workers rather than a statistical survey. Its data is based on actual entries rather than inferred estimates. Therefore, variations in statistics derived from eShram are more likely to reflect real differences in worker profiles, rather than statistical variation due to sampling or estimation.
NA

The data is collected through registration/ updates through multiple sources like Self registration, CSC, Umang mobile Applicatuion and SSK.
Continuous
The data is collected through registration/ updates through multiple sources like Self registration, CSC, Umang mobile Applicatuion and SSK.

Unorganized workers data is kept in database after registration. This data is processed to be kept in the data warehouse for futher analysis purpose.
To restrict the duplicate records validation is done by applying constraints in the database.
The data is validated against the LGD codes for state, district and village information.
The statistical data is made available for public dashboard available at https://eshram.gov.in/dashboard
The aggregated data is validated by cross verifying the information with the output by querying the database.

Unorganized workers of eShram are registered online,which is stored in database. The transactional data is processed to
be kept in the data warehouse. To process data it is vallidated against the defined master data for eShram related to
parameters such as occupation, educational qualification, LGD codes, income slab, blood group, disability, skills, platform workers etc
The data wareshouse is used for analytical purpose and to share data with different government departments, states using APIs and Data Sharing Portal.
Data is aggregated to derive various statistics available in eShram public dashboard at https://eshram.gov.in/dashboard. Further, this data can be processed and analysed by government agencies and research bodies to generate insights such as sectoral distribution, migration patterns, skill mapping, and welfare eligibility.

After succesful registration on eShram portal worker get unique 12 digit Universal Account Number. Further, for identification with other source of data we are using aadhaar in encryped format.

l in the database.

Metadata is not exposed.

11-03-2025