

Issue 24: Different data sets

If I want to document that a data set was collected through a survey, how do I do it? Mode? Instrument? Instrument Implementation? Acquisition Activity? Data Channel? Channel Activity Specification?

The question arises because I am trying to distinguish between different types of microdata: event-history, timeseries, snapshot, survey, register in order to be able point the Data Consumer to more information.

GSIM v1.0 Annex C Definitions

Mode	A set of characteristics that describe the technique (the "how") used for the data acquisition through a given Data Channel based on a specific Instrument Implementation.	While the Data Channel describes the means used for data acquisition, the Instrument describes the "what" (i.e. the content, for example, in terms of questions in a questionnaire or a list of agreed time series codes in a data exchange template) and an Instrument Implementation describes the tool used to apply the Instrument; the Mode describes "how" the Data Channel is going to be used. The Mode is relevant for all types of Data Channels, Instrument Implementations and Instruments and can change over time. The list of Modes will potentially grow in the future and vary from organization to organization.
Instrument	A tool conceived to record the information that will be obtained from the Observation Units.	The Instrument describes the tool used to collect data. It could be a traditional survey, a set of requirements for a software collection program, a clinical procedure, etc. Instrument is described from the perspective of the statistical organization collecting the data. It includes the special type of Instrument used for the explicit purpose of gathering data through a questionnaire (Survey Instrument). The behavior and characteristics of a concrete Instrument is determined by an Instrument Implementation. Several implementations can be based in the same Instrument giving the possibility of using multiple channels and to apply different collection techniques (Modes) to gather data. An example of this is when a printed format to collect information for a survey is substituted by a software program; in both cases the Instrument will collect the data from the Unit but the behavior of the Instrument will be different accordingly with its implementation.
Instrument Implementation	A concrete and usable tool for gathering information based on the rendering of the description made by an Instrument.	This represents an implementation of an Instrument. It describes the way in which an Instrument has been translated from a design to a concrete tool. It could represent a printed form, a software program made following a specific technological paradigm (web service, web scraping robot, etc.), the software used by a specialized device to collect data, etc. When it describes a Survey Instrument, it can contain descriptions of how each construct (e.g. Questions, Value Domains, validation Rules contained in the Instrument) is implemented.

Acquisition Activity	The set of executed processes and the actual resources required as inputs and produced as outputs to acquire data about a given Population for a particular reference period. It includes the process and resource required to acquire data in a Statistical Program consisting of gathering data via one or more Data Channels in order to create or feed one or more Data Resources.	This object holds Statistical Activity information that relates specifically to data collection or acquisition. It inherits the relationships and attributes from the Statistical Activity type.
Data Channel	A means of exchanging data.	A Data Channel is an abstract object that describes the means for communicating with Data Resource(s). The Data Channel identifies the Instrument Implementation, Mode, and Data Resource that are to be used in a process. In some cases the Data Channel that is used by the Data Provider to send its responses could be different that the one used by the statistical office or organization to request information; the statistical office may put electronic formats that can be downloaded by the Data Provider and once answered returned by traditional mail. Two specialized objects are used to implement this abstract object: Channel Design Specification used at design time and Channel Activity Specification used at run time.
Channel Activity Specification	A structured, well-defined specification for a proposed change. The description of the Data Channel made at run time.	This object is a specialization of a Data Channel and is used to describe the behaviour of a Data Channel at execution time.