The Generic Statistical Information Model (GSIM) and the Sistema Unitario dei Metadati (SUM): state of application of the standard

Cecilia Casagrande – casagran@istat.it

21 settembre 2016

SUMMARY

- Generic Statistical Information Model (GSIM) and Istat metadata system (SUM)

- The Istat metadata system SUM

  -- Population, variable and other concepts

  -- Classification

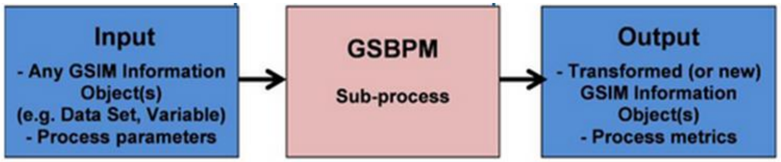
  -- Data content

  -- Data structures

- Functionalities, developments and benefits
  - – Search functionalities
  - – Implementation of GSIM
  - – Benefits and more

GSIM and Istat metadata system (SUM)

GSIM describes the “interfaces” in terms of subprocesses input and output, leaving to GSBPM the description of the subprocesses



Input
- Any GSIM Information Object(s)
(e.g. Data Set, Variable)
- Process parameters

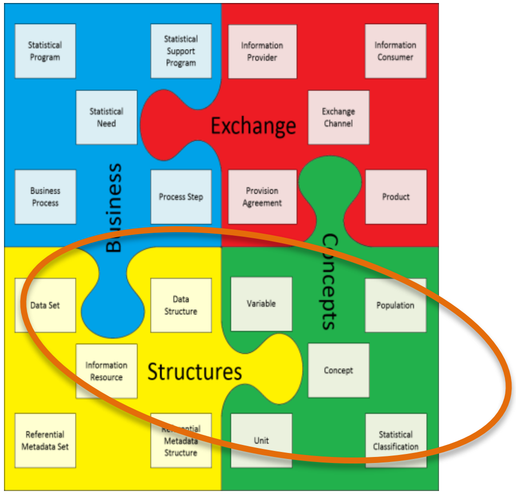
GSBPM
Sub-process

Output
- Transformed (or new) GSIM Information Object(s)
- Process metrics

1.In order to trace the data production process, Istat metadata system considered GSIM as a primary source of information.

2.Our current objective is to develop the part of Istat metadata system (SUM) related to data (structural metadata).

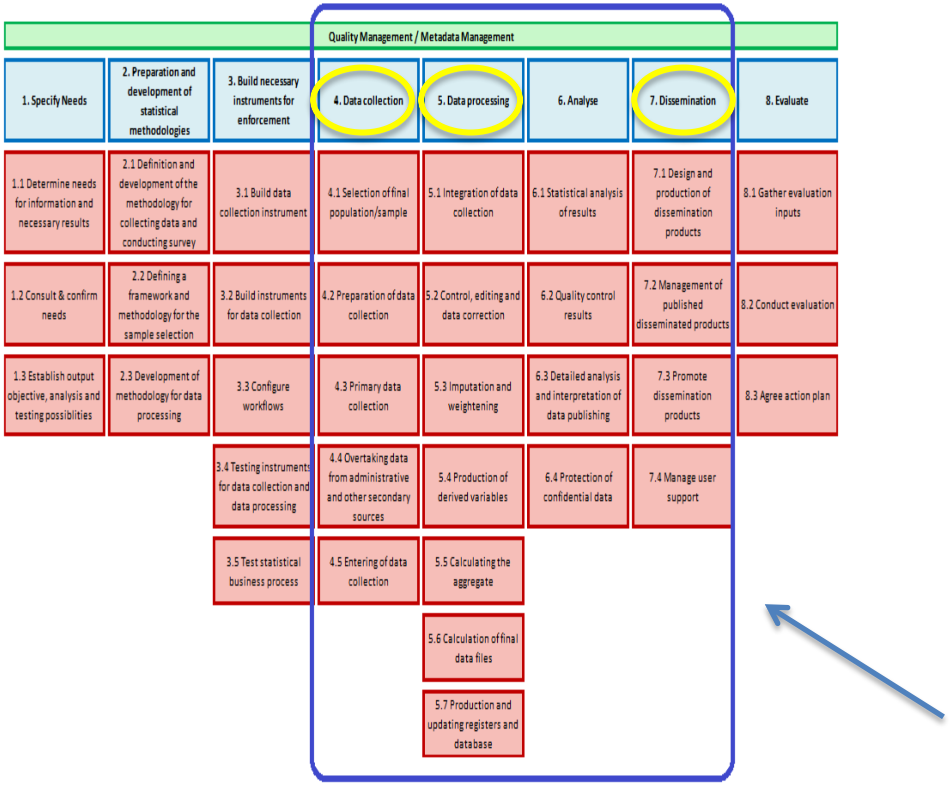
3.Hence, attention has been given mostly to a portion of GSIM, the part related to the “structure” and “concepts”



4.Istat metadata system SUM is coherent with GSIM for definitions and nomenclatures

The Istat metadata system SUM

- The Istat metadata system (SUM) is a centralized system which contains structural metadata. SUM aims to describe data and data connections able to trace the data production process from data collection (raw data) up to data dissemination.



Quality Management / Metadata Management

1. Specify Needs	2. Preparation and development of statistical methodologies	3. Build necessary instruments for enforcement	4. Data collection	5. Data processing	6. Analyse	7. Dissemination	8. Evaluate
1.1 Determine needs for information and necessary results	2.1 Definition and development of the methodology for collecting data and conducting survey	3.1 Build data collection instrument	4.1 Selection of final population/sample	5.1 Integration of data collection	6.1 Statistical analysis of results	7.1 Design and production of dissemination products	8.1 Gather evaluation inputs
1.2 Consult & confirm needs	2.2 Defining a framework and methodology for the sample selection	3.2 Build instruments for data collection	4.2 Preparation of data collection	5.2 Control, editing and data correction	6.2 Quality control results	7.2 Management of published disseminated products	8.2 Conduct evaluation
1.3 Establish output objective, analysis and testing possiblities	2.3 Development of methodology for data processing	3.3 Configure workflows	4.3 Primary data collection	5.3 Imputation and weightening	6.3 Detailed analysis and interpretation of data publishing	7.3 Promote dissemination products	8.3 Agree action plan
		3.4 Testing instruments for data collection and data processing	4.4 Overtaking data from administrative and other secondary sources	5.4 Production of derived variables	6.4 Protection of confidential data	7.4 Manage user support	
		3.5 Test statistical business process	4.5 Entering of data collection	5.5 Calculating the aggregate			
				5.6 Calculation of final data files			
				5.7 Production and updating registers and database			

The system should aim at easing:

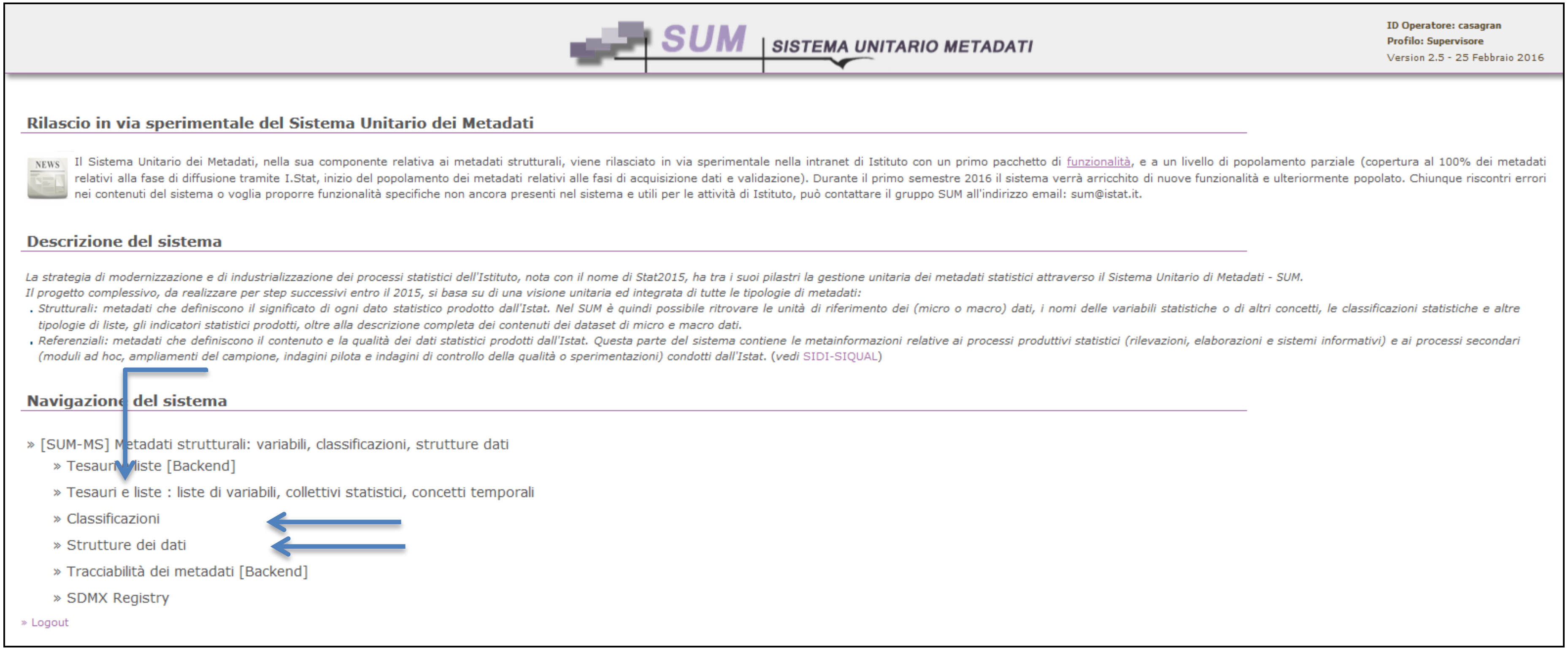
Harmonization: our Institute should speak using the same terms according to a common language structure, independently on the theme

Reuse: new processes should draw terms among those already existing

Search functionalities: in order to foster harmonization and reuse, there should be adequate search functionalities

Traceability: every step of the data production process is operationally defined so to allow reconstruction of the same data from another party (transparency)

The Istat metadata system SUM



ID Operatore: casagran
Profilo: Supervisore
Version 2.5 - 25 Febbraio 2016

Rilascio in via sperimentale del Sistema Unitario dei Metadati

Il Sistema Unitario dei Metadati, nella sua componente relativa ai metadati strutturali, viene rilasciato in via sperimentale nella intranet di Istituto con un primo pacchetto di funzionalità, e a un livello di popolamento parziale (copertura al 100% dei metadati relativi alla fase di diffusione tramite I.Stat, inizio del popolamento dei metadati relativi alle fasi di acquisizione dati e validazione). Durante il primo semestre 2016 il sistema verrà arricchito di nuove funzionalità e ulteriormente popolato. Chiunque riscontri errori nei contenuti del sistema o voglia proporre funzionalità specifiche non ancora presenti nel sistema e utili per le attività di Istituto, può contattare il gruppo SUM all'indirizzo email: sum@istat.it.

Descrizione del sistema

La strategia di modernizzazione e di industrializzazione dei processi statistici dell'Istituto, nota con il nome di Stat2015, ha tra i suoi pilastri la gestione unitaria dei metadati statistici attraverso il Sistema Unitario di Metadati - SUM.
Il progetto complessivo, da realizzare per step successivi entro il 2015, si basa su di una visione unitaria ed integrata di tutte le tipologie di metadati:

- *Strutturali: metadati che definiscono il significato di ogni dato statistico prodotto dall'Istat. Nel SUM è quindi possibile ritrovare le unità di riferimento dei (micro o macro) dati, i nomi delle variabili statistiche o di altri concetti, le classificazioni statistiche e altre tipologie di liste, gli indicatori statistici prodotti, oltre alla descrizione completa dei contenuti dei dataset di micro e macro dati.*
- *Referenziali: metadati che definiscono il contenuto e la qualità dei dati statistici prodotti dall'Istat. Questa parte del sistema contiene le metainformazioni relative ai processi produttivi statistici (rilevazioni, elaborazioni e sistemi informativi) e ai processi secondari (moduli ad hoc, ampliamenti del campione, indagini pilota e indagini di controllo della qualità o sperimentazioni) condotti dall'Istat. (vedi* SIDI-SIQUAL*)*

Navigazione del sistema

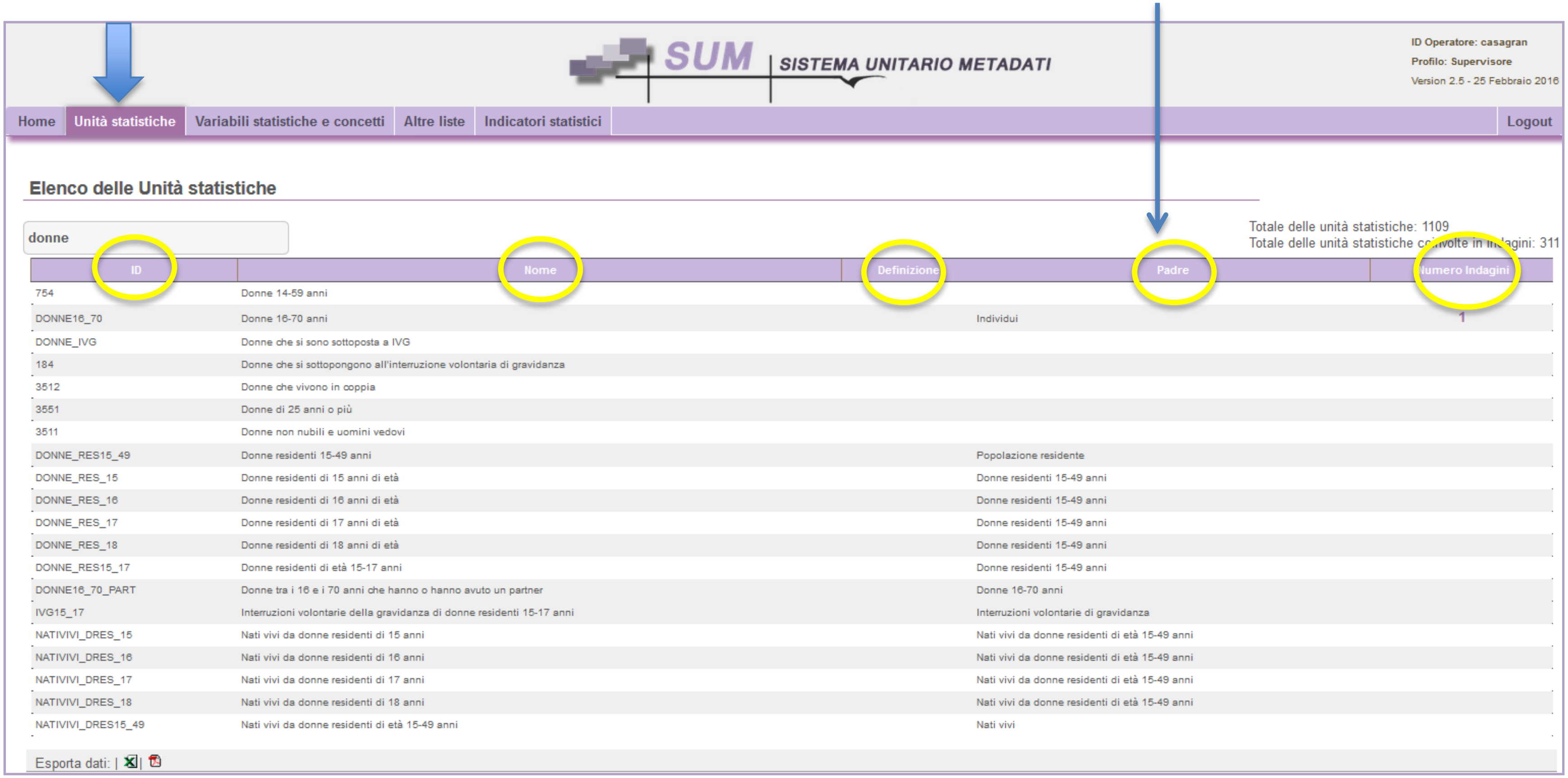
» [SUM-MS] Metadati strutturali: variabili, classificazioni, strutture dati
   » Tesauri e liste [Backend]
   » Tesauri e liste : liste di variabili, collettivi statistici, concetti temporali
   » Classificazioni
   » Strutture dei dati
   » Tracciabilità dei metadati [Backend]
   » SDMX Registry

» Logout

The Istat metadata system SUM: population, variable and other concepts

Population

Hierarchical relationship



ID Operatore: casagran
Profilo: Supervisore
Version 2.5 - 25 Febbraio 2016

Elenco delle Unità statistiche

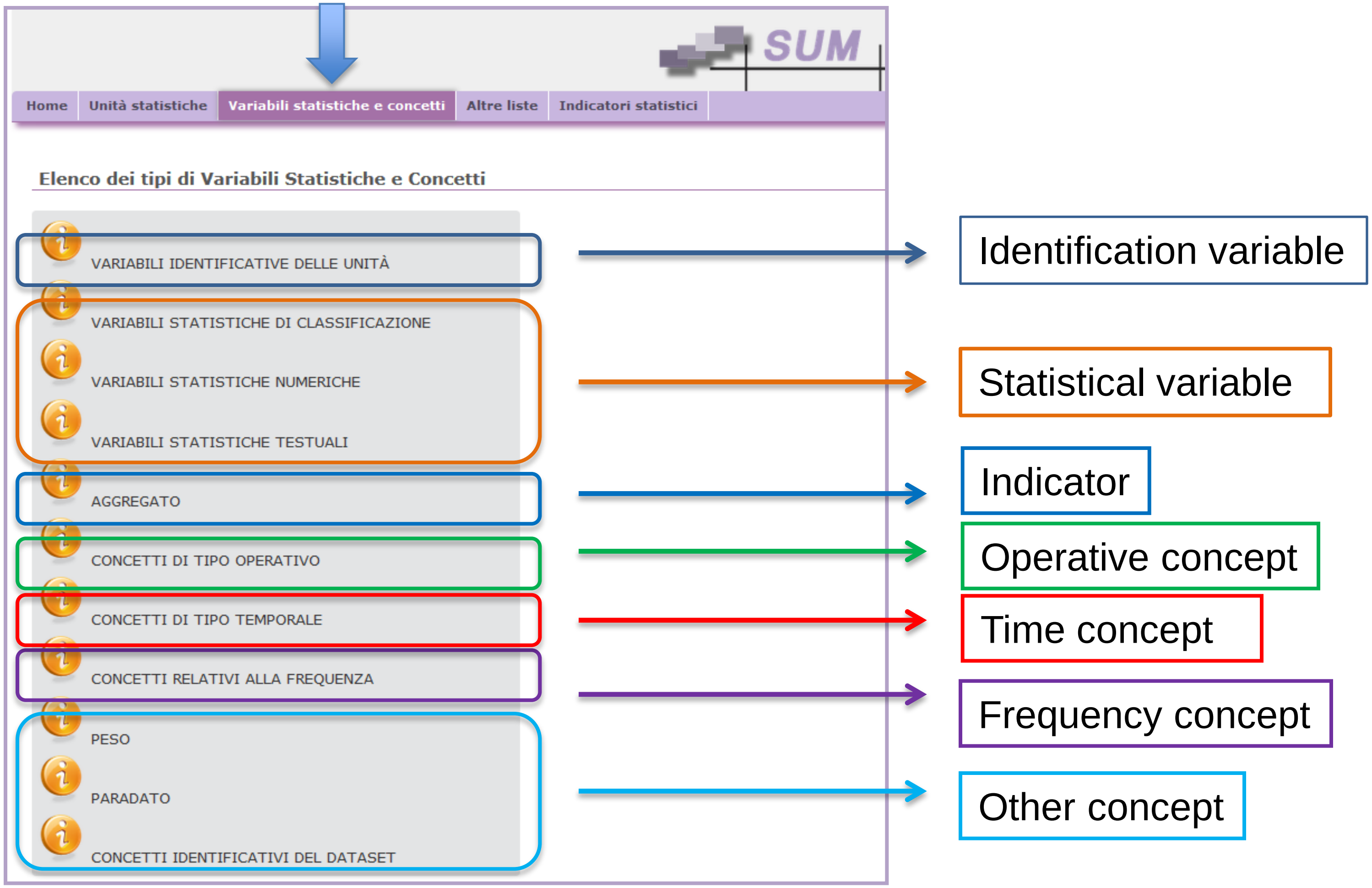
donne

Totale delle unità statistiche: 1109
Totale delle unità statistiche coinvolte in indagini: 311

ID	Nome	Definizione	Padre	Numero Indagini
754	Donne 14-59 anni			
DONNE16_70	Donne 16-70 anni		Individui	1
DONNE_IVG	Donne che si sono sottoposta a IVG			
184	Donne che si sottopongono all'interruzione volontaria di gravidanza			
3512	Donne che vivono in coppia			
3551	Donne di 25 anni o più			
3511	Donne non nubili e uomini vedovi			
DONNE_RES15_49	Donne residenti 15-49 anni		Popolazione residente	
DONNE_RES_15	Donne residenti di 15 anni di età		Donne residenti 15-49 anni	
DONNE_RES_16	Donne residenti di 16 anni di età		Donne residenti 15-49 anni	
DONNE_RES_17	Donne residenti di 17 anni di età		Donne residenti 15-49 anni	
DONNE_RES_18	Donne residenti di 18 anni di età		Donne residenti 15-49 anni	
DONNE_RES15_17	Donne residenti di età 15-17 anni		Donne residenti 15-49 anni	
DONNE16_70_PART	Donne tra i 16 e i 70 anni che hanno o hanno avuto un partner		Donne 16-70 anni	
IVG15_17	Interruzioni volontarie della gravidanza di donne residenti 15-17 anni		Interruzioni volontarie di gravidanza	
NATIVIVI_DRES_15	Nati vivi da donne residenti di 15 anni		Nati vivi da donne residenti di età 15-49 anni	
NATIVIVI_DRES_16	Nati vivi da donne residenti di 16 anni		Nati vivi da donne residenti di età 15-49 anni	
NATIVIVI_DRES_17	Nati vivi da donne residenti di 17 anni		Nati vivi da donne residenti di età 15-49 anni	
NATIVIVI_DRES_18	Nati vivi da donne residenti di 18 anni		Nati vivi da donne residenti di età 15-49 anni	
NATIVIVI_DRES15_49	Nati vivi da donne residenti di età 15-49 anni		Nati vivi	

The Istat metadata system SUM: population, variable and other concepts

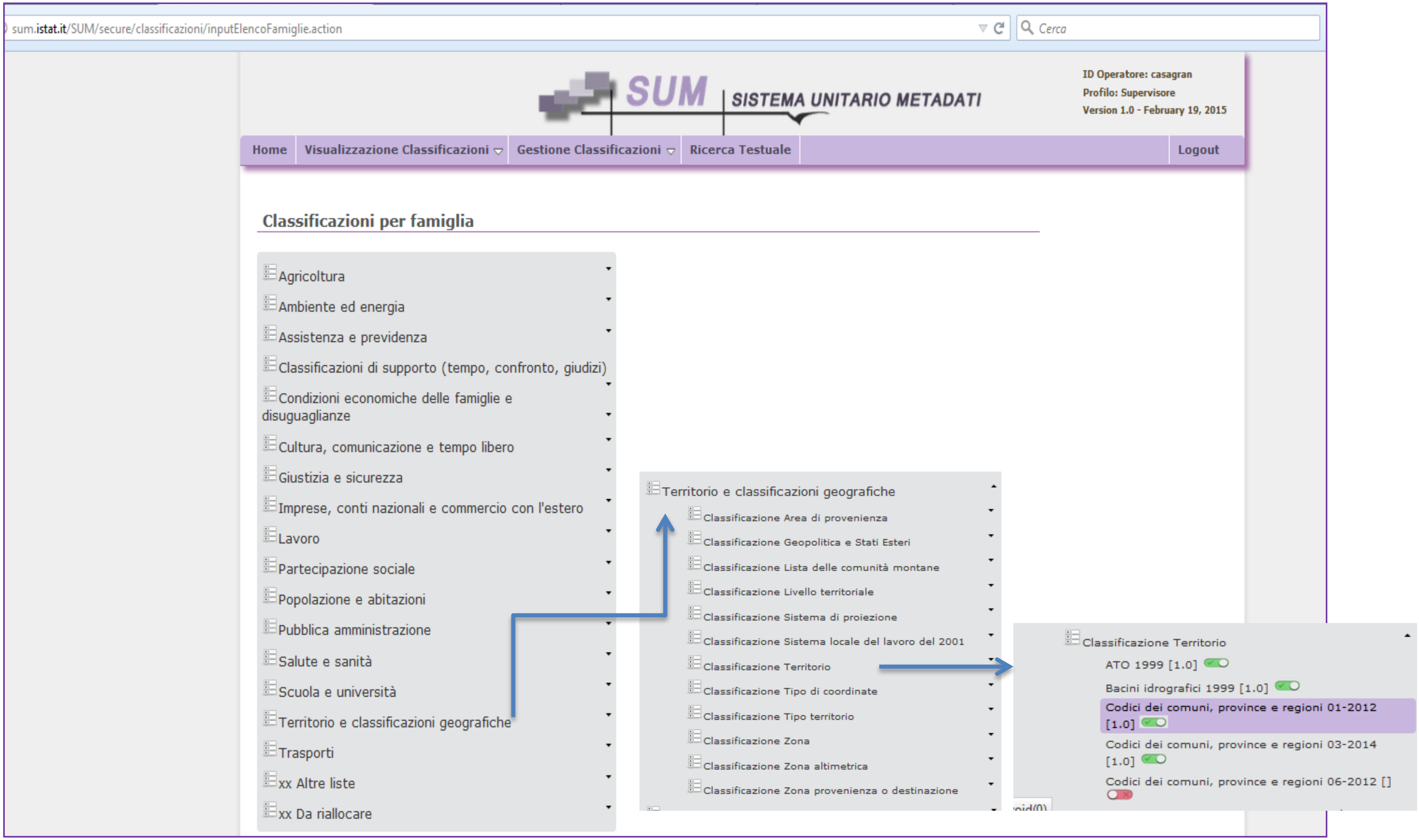
Variable



Elenco dei tipi di Variabili Statistiche e Concetti

- VARIABILI IDENTIFICATIVE DELLE UNITÀ → Identification variable
- VARIABILI STATISTICHE DI CLASSIFICAZIONE
- VARIABILI STATISTICHE NUMERICHE
- VARIABILI STATISTICHE TESTUALI → Statistical variable
- AGGREGATO → Indicator
- CONCETTI DI TIPO OPERATIVO → Operative concept
- CONCETTI DI TIPO TEMPORALE → Time concept
- CONCETTI RELATIVI ALLA FREQUENZA → Frequency concept
- PESO
- PARADATO
- CONCETTI IDENTIFICATIVI DEL DATASET → Other concept

Classification


sum.istat.it/SUM/secure/classificazioni/inputElencoFamiglie.action
Cerca
SUM
SISTEMA UNITARIO METADATI
Home
Visualizzazione Classificazioni
Gestione Classificazioni
Ricerca Testuale
Logout

ID Operatore: casagran
Profilo: Supervisore
Version 1.0 - February 19, 2015

Classificazioni per famiglia

- Agricoltura
- Ambiente ed energia
- Assistenza e previdenza
- Classificazioni di supporto (tempo, confronto, giudizi)
- Condizioni economiche delle famiglie e disuguaglianze
- Cultura, comunicazione e tempo libero
- Giustizia e sicurezza
- Imprese, conti nazionali e commercio con l'estero
- Lavoro
- Partecipazione sociale
- Popolazione e abitazioni
- Pubblica amministrazione
- Salute e sanità
- Scuola e università
- Territorio e classificazioni geografiche
- Trasporti
- xx Altre liste
- xx Da riallocare

Territorio e classificazioni geografiche

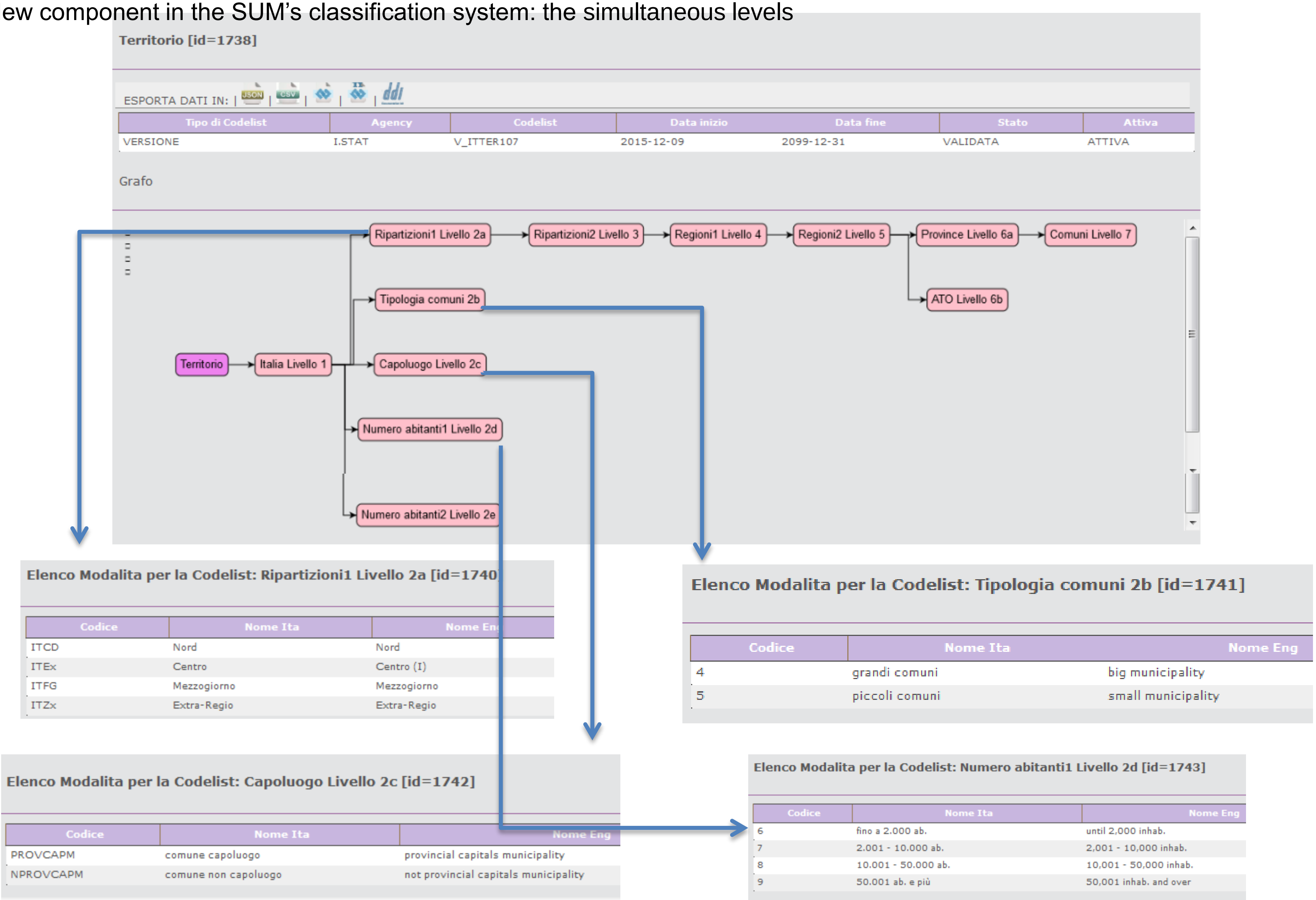
- Classificazione Area di provenienza
- Classificazione Geopolitica e Stati Esteri
- Classificazione Lista delle comunità montane
- Classificazione Livello territoriale
- Classificazione Sistema di proiezione
- Classificazione Sistema locale del lavoro del 2001
- Classificazione Territorio
- Classificazione Tipo di coordinate
- Classificazione Tipo territorio
- Classificazione Zona
- Classificazione Zona altimetrica
- Classificazione Zona provenienza o destinazione

Classificazione Territorio

- ATO 1999 [1.0] ✓
- Bacini idrografici 1999 [1.0] ✓
- Codici dei comuni, province e regioni 01-2012 [1.0] ✓
- Codici dei comuni, province e regioni 03-2014 [1.0] ✓
- Codici dei comuni, province e regioni 06-2012 [] ✗

The Istat metadata system SUM: classification

New component in the SUM's classification system: the simultaneous levels



Territorio [id=1738]

ESPORTA DATI IN: | JSON | CSV | | | ddi

Tipo di Codelist	Agency	Codelist	Data inizio	Data fine	Stato	Attiva
VERSIONE	I.STAT	V_ITTER107	2015-12-09	2099-12-31	VALIDATA	ATTIVA

Grafo

Elenco Modalita per la Codelist: Ripartizioni1 Livello 2a [id=1740]

Codice	Nome Ita	Nome En
ITCD	Nord	Nord
ITEx	Centro	Centro (I)
ITFG	Mezzogiorno	Mezzogiorno
ITZx	Extra-Regio	Extra-Regio

Elenco Modalita per la Codelist: Tipologia comuni 2b [id=1741]

Codice	Nome Ita	Nome Eng
4	grandi comuni	big municipality
5	piccoli comuni	small municipality

Elenco Modalita per la Codelist: Capoluogo Livello 2c [id=1742]

Codice	Nome Ita	Nome Eng
PROVCAPM	comune capoluogo	provincial capitals municipality
NPROVCAPM	comune non capoluogo	not provincial capitals municipality

Elenco Modalita per la Codelist: Numero abitanti1 Livello 2d [id=1743]

Codice	Nome Ita	Nome Eng
6	fino a 2.000 ab.	until 2,000 inhab.
7	2.001 - 10.000 ab.	2,001 - 10,000 inhab.
8	10.001 - 50.000 ab.	10,001 - 50,000 inhab.
9	50.001 ab. e più	50,001 inhab. and over

The Istat metadata system SUM: data content

Data content

In SUM, we introduced the “data content” concept. It is defined according to the GSIM (specification) lines 47-50: *Each **data** is a result of a Process step through the application of a Process method on the necessary Inputs.* Hence it is modelled by specifying:

Example:
Monthly average
household expenditures

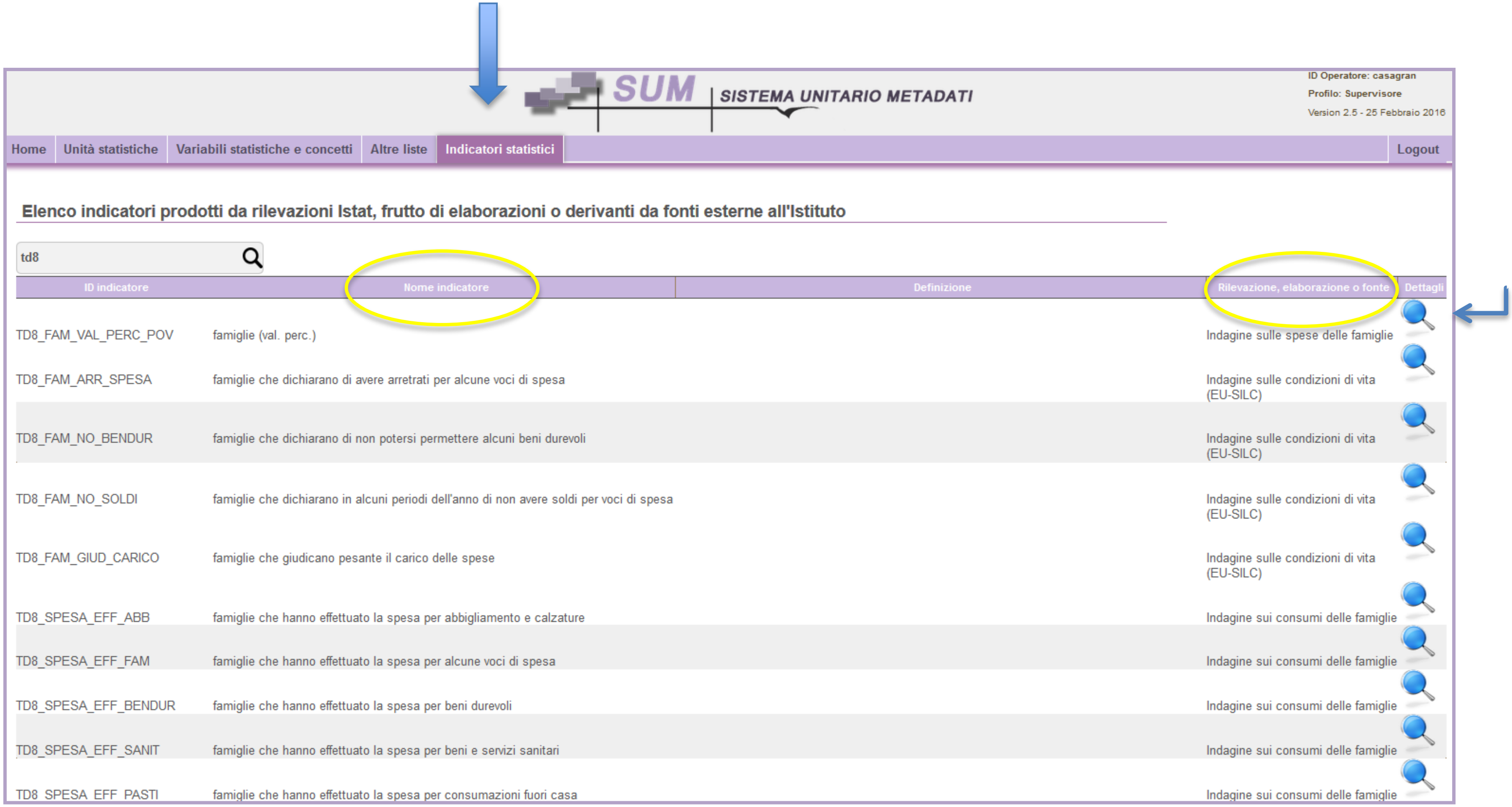
-Statistical Program and Statistical Program Cycle → Labour force survey + Time dimension and Freq dimension

-Process Step (phase) → Dissemination

-Process Method → Average

-Inputs → Population (Households) + Numeric variable (monthly expenditures)

The Istat metadata system SUM: data content



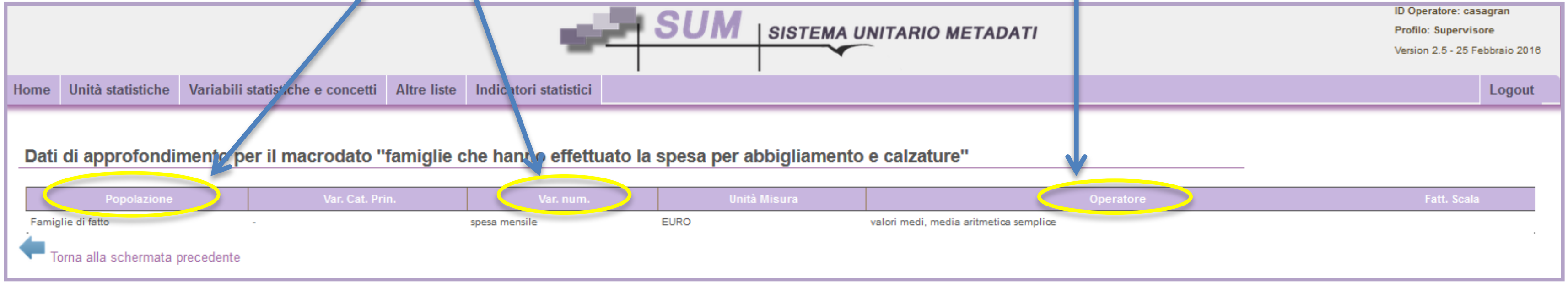
Elenco indicatori prodotti da rilevazioni Istat, frutto di elaborazioni o derivanti da fonti esterne all'Istituto

ID indicatore	Nome indicatore	Definizione	Rilevazione, elaborazione o fonte	Dettagli
TD8_FAM_VAL_PERC_POV	famiglie (val. perc.)		Indagine sulle spese delle famiglie	🔍
TD8_FAM_ARR_SPESA	famiglie che dichiarano di avere arretrati per alcune voci di spesa		Indagine sulle condizioni di vita (EU-SILC)	🔍
TD8_FAM_NO_BENDUR	famiglie che dichiarano di non potersi permettere alcuni beni durevoli		Indagine sulle condizioni di vita (EU-SILC)	🔍
TD8_FAM_NO_SOLDI	famiglie che dichiarano in alcuni periodi dell'anno di non avere soldi per voci di spesa		Indagine sulle condizioni di vita (EU-SILC)	🔍
TD8_FAM_GIUD_CARICO	famiglie che giudicano pesante il carico delle spese		Indagine sulle condizioni di vita (EU-SILC)	🔍
TD8_SPESA_EFF_ABB	famiglie che hanno effettuato la spesa per abbigliamento e calzature		Indagine sui consumi delle famiglie	🔍
TD8_SPESA_EFF_FAM	famiglie che hanno effettuato la spesa per alcune voci di spesa		Indagine sui consumi delle famiglie	🔍
TD8_SPESA_EFF_BENDUR	famiglie che hanno effettuato la spesa per beni durevoli		Indagine sui consumi delle famiglie	🔍
TD8_SPESA_EFF_SANIT	famiglie che hanno effettuato la spesa per beni e servizi sanitari		Indagine sui consumi delle famiglie	🔍
TD8_SPESA_EFF_PASTI	famiglie che hanno effettuato la spesa per consumazioni fuori casa		Indagine sui consumi delle famiglie	🔍

The Istat metadata system SUM: data content

Input

Process Method


SUM
SISTEMA UNITARIO METADATI
Home
Unità statistiche
Variabili statistiche e concetti
Altre liste
Indicatori statistici
Logout
Torna alla schermata precedente

ID Operatore: casagran

Profilo: Supervisore

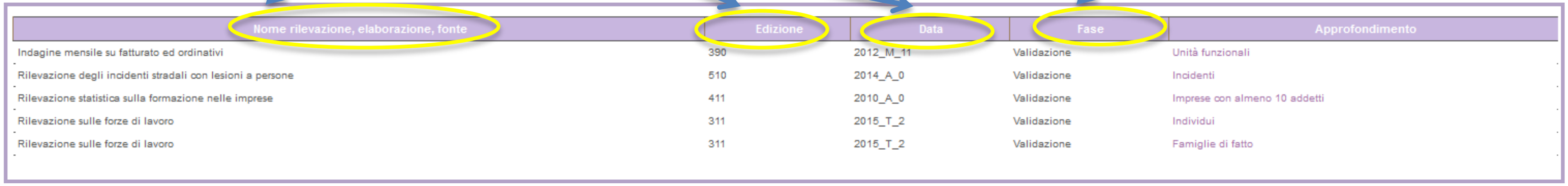
Version 2.5 - 25 Febbraio 2016

Dati di approfondimento per il macrodato "famiglie che hanno effettuato la spesa per abbigliamento e calzature"

Popolazione	Var. Cat. Prin.	Var. num.	Unità Misura	Operatore	Fatt. Scala
Famiglie di fatto	-	spesa mensile	EURO	valori medi, media aritmetica semplice	

Statistical program and
Statistical program cycle

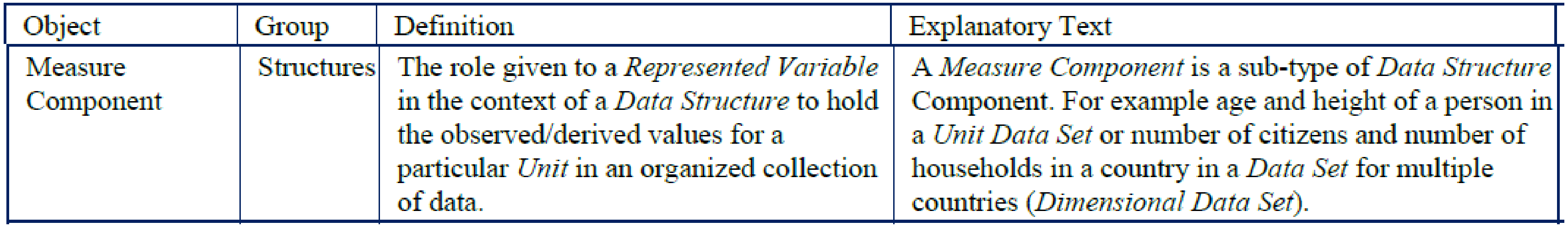
Process Step



Nome rilevazione, elaborazione, fonte	Edizione	Data	Fase	Approfondimento
Indagine mensile su fatturato ed ordinativi	390	2012_M_11	Validazione	Unità funzionali
Rilevazione degli incidenti stradali con lesioni a persone	510	2014_A_0	Validazione	Incidenti
Rilevazione statistica sulla formazione nelle imprese	411	2010_A_0	Validazione	Imprese con almeno 10 addetti
Rilevazione sulle forze di lavoro	311	2015_T_2	Validazione	Individui
Rilevazione sulle forze di lavoro	311	2015_T_2	Validazione	Famiglie di fatto

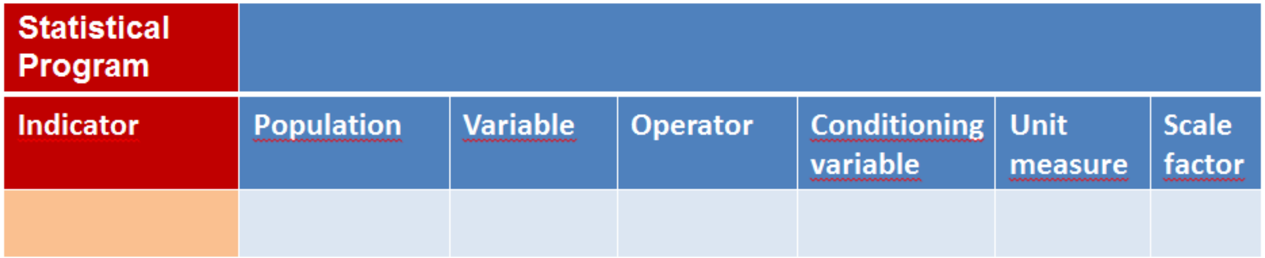
The Istat metadata system SUM: data content

Data Content feeds the GSIM concept **Measure** in a Data Structure.



Object	Group	Definition	Explanatory Text
Measure Component	Structures	The role given to a *Represented Variable* in the context of a *Data Structure* to hold the observed/derived values for a particular *Unit* in an organized collection of data.	A *Measure Component* is a sub-type of *Data Structure* Component. For example age and height of a person in a *Unit Data Set* or number of citizens and number of households in a country in a *Data Set* for multiple countries (*Dimensional Data Set*).

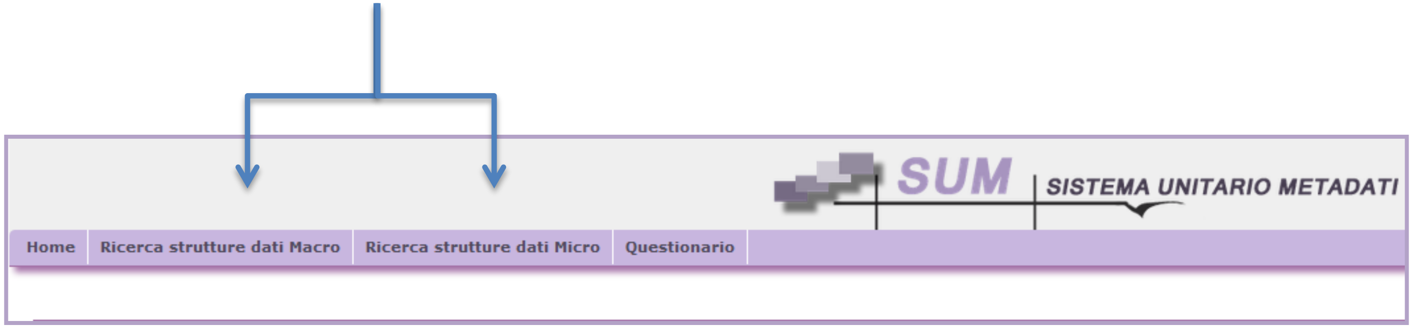
- ❑ SUM maintains a special code list “Data Content” where each item contains all the previous details.
- ❑ In this way a user can find the meaning of the data in a hypercube in a unique place.
- ❑ Furthermore a data producer has a form to complete for describing any new data content (*data content structure*).
- ❑ Any data producer should describe the “Data Content” according to the same “structure”.



Statistical Program						
Indicator	Population	Variable	Operator	Conditioning variable	Unit measure	Scale factor

- ❑ If the “Data Content” is not well defined, the meaning of data is not easy to understand and massive use of mappers should be foreseen for data and metadata exchange.

Data Structures


SUM
SISTEMA UNITARIO METADATI
Home
Ricerca strutture dati Macro
Ricerca strutture dati Micro
Questionario

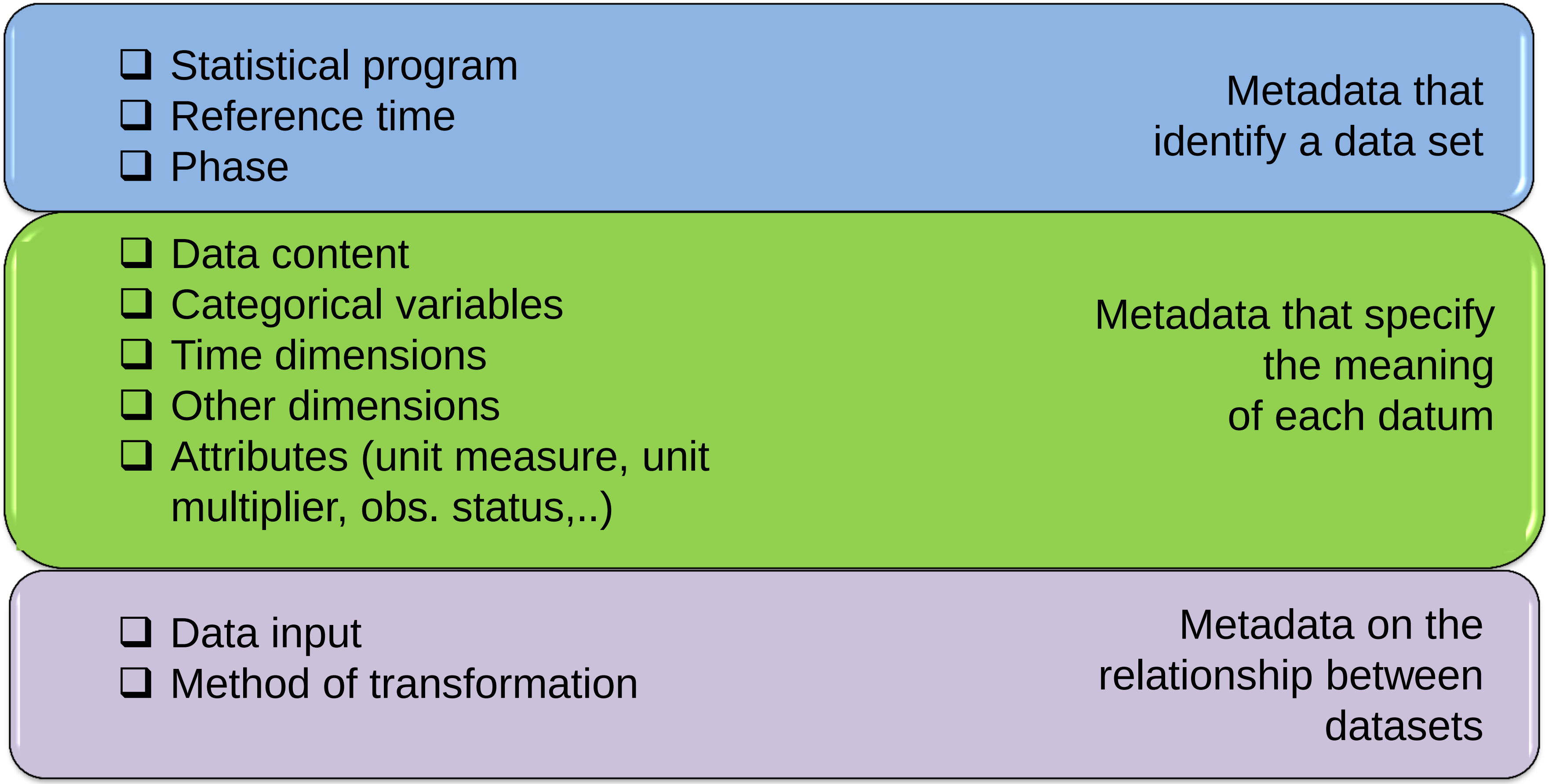
Dimensional Data Structures (macro)

Unit Data Structure (micro)

GSIM states that “A *Data Structure* describes the structure of a *Data Set* by means of *Data Structure Components* (*Identifier Components*, *Measure Components* and *Attribute Components*).”

Functionalities, developments and benefits: data structures

Dimensional Data Structures (macro)



- □ Statistical program
- □ Reference time
- □ Phase

Metadata that
identify a data set

- □ Data content
- □ Categorical variables
- □ Time dimensions
- □ Other dimensions
- □ Attributes (unit measure, unit multiplier, obs. status,..)

Metadata that specify
the meaning
of each datum

- □ Data input
- □ Method of transformation

Metadata on the
relationship between
datasets

Functionalities, developments and benefits: data structures

Unit Data Structure (micro)

□ Statistical program
□ Reference time
□ Phase
□ Reference population

Metadata that identify a data set

□ Unit identifier
□ Numerical/quantitative variables
□ Categorical variables
□ Textual variables
□ Macrodata requested at the microdata level
□ Attributes (maximum number of answers, rules, structural zeros,…)
□ Paradata

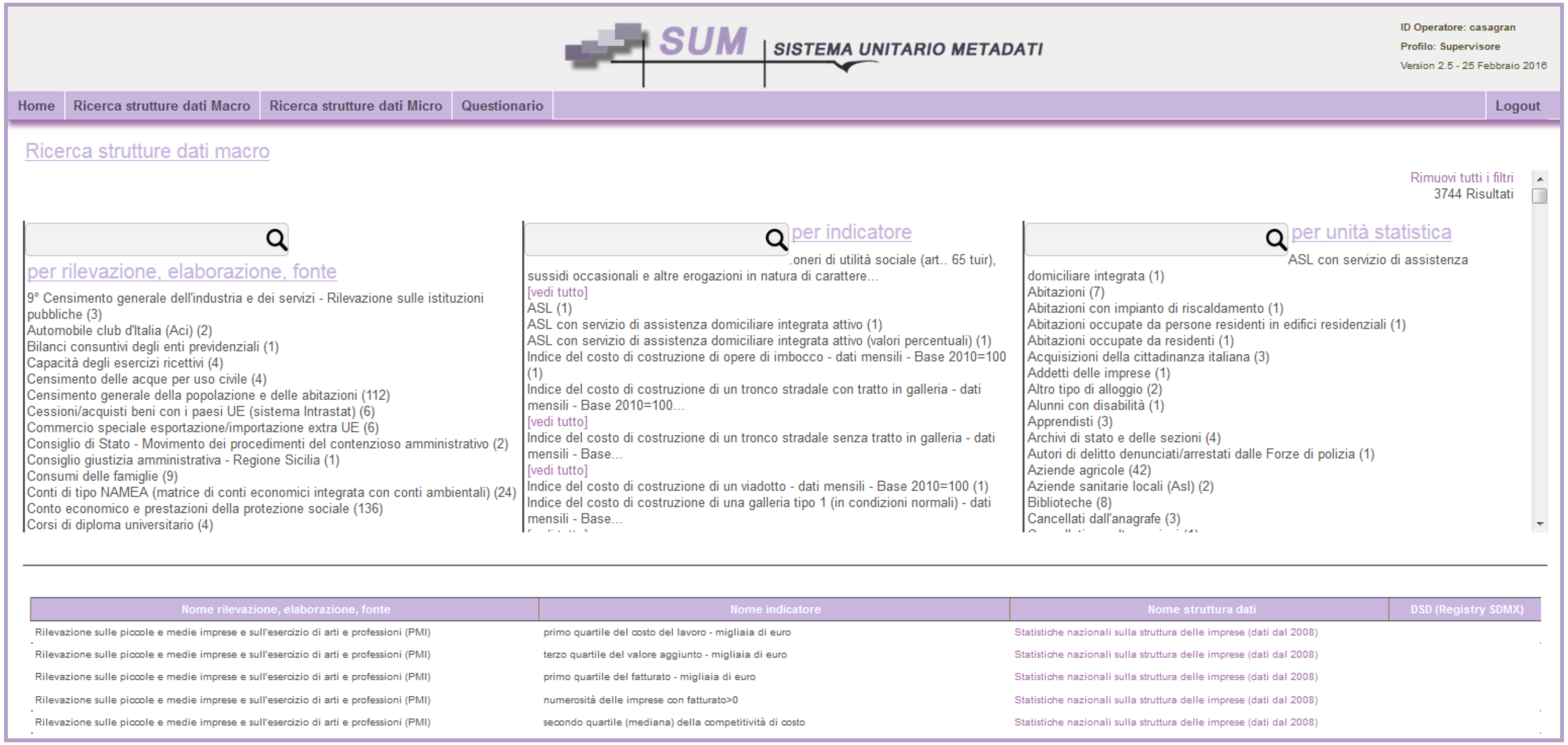
Metadata that specify the meaning of each datum

□ Relationship between the investigated populations
□ Data input
□ Method of transformation

Metadata on the relationship between datasets

Search functionalities

Dimensional Data Structures (macro)


SUM
SISTEMA UNITARIO METADATI
ID Operatore: casagran
Profilo: Supervisore
Version 2.5 - 25 Febbraio 2016
Home
Ricerca strutture dati Macro
Ricerca strutture dati Micro
Questionario
Logout

Ricerca strutture dati macro

Rimuovi tutti i filtri
3744 Risultati

per rilevazione, elaborazione, fonte

9° Censimento generale dell'industria e dei servizi - Rilevazione sulle istituzioni pubbliche (3)
Automobile club d'Italia (Aci) (2)
Bilanci consuntivi degli enti previdenziali (1)
Capacità degli esercizi ricettivi (4)
Censimento delle acque per uso civile (4)
Censimento generale della popolazione e delle abitazioni (112)
Cessioni/acquisti beni con i paesi UE (sistema Intrastat) (6)
Commercio speciale esportazione/importazione extra UE (6)
Consiglio di Stato - Movimento dei procedimenti del contenzioso amministrativo (2)
Consiglio giustizia amministrativa - Regione Sicilia (1)
Consumi delle famiglie (9)
Conti di tipo NAMEA (matrice di conti economici integrata con conti ambientali) (24)
Conto economico e prestazioni della protezione sociale (136)
Corsi di diploma universitario (4)

per indicatore

.oneri di utilità sociale (art.. 65 tuir), sussidi occasionali e altre erogazioni in natura di carattere...
[vedi tutto]
ASL (1)
ASL con servizio di assistenza domiciliare integrata attivo (1)
ASL con servizio di assistenza domiciliare integrata attivo (valori percentuali) (1)
Indice del costo di costruzione di opere di imbocco - dati mensili - Base 2010=100 (1)
Indice del costo di costruzione di un tronco stradale con tratto in galleria - dati mensili - Base 2010=100...
[vedi tutto]
Indice del costo di costruzione di un tronco stradale senza tratto in galleria - dati mensili - Base...
[vedi tutto]
Indice del costo di costruzione di un viadotto - dati mensili - Base 2010=100 (1)
Indice del costo di costruzione di una galleria tipo 1 (in condizioni normali) - dati mensili - Base...

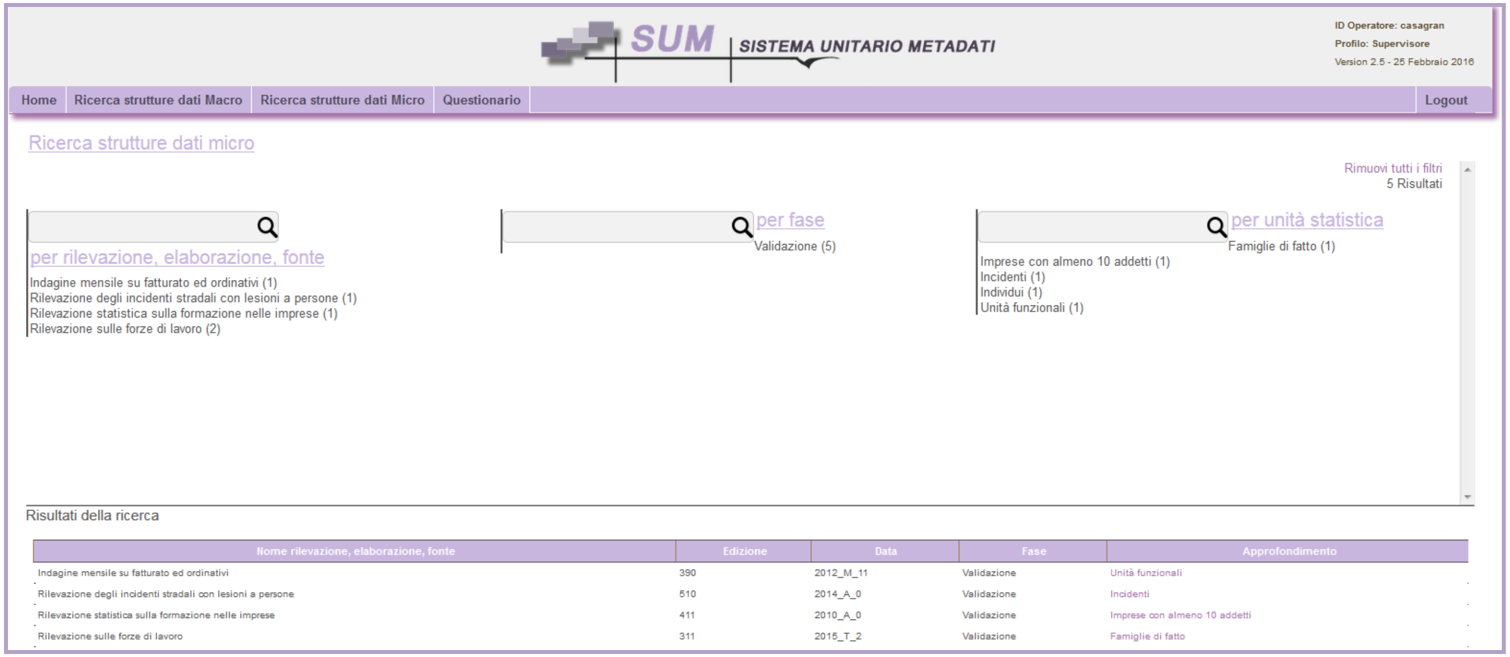
per unità statistica

ASL con servizio di assistenza domiciliare integrata (1)
Abitazioni (7)
Abitazioni con impianto di riscaldamento (1)
Abitazioni occupate da persone residenti in edifici residenziali (1)
Abitazioni occupate da residenti (1)
Acquisizioni della cittadinanza italiana (3)
Addetti delle imprese (1)
Altro tipo di alloggio (2)
Alunni con disabilità (1)
Apprendisti (3)
Archivi di stato e delle sezioni (4)
Autori di delitto denunciati/arrestati dalle Forze di polizia (1)
Aziende agricole (42)
Aziende sanitarie locali (Asl) (2)
Biblioteche (8)
Cancellati dall'anagrafe (3)

Nome rilevazione, elaborazione, fonte	Nome indicatore	Nome struttura dati	DSD (Registry SDMX)
Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI)	primo quartile del costo del lavoro - migliaia di euro	Statistiche nazionali sulla struttura delle imprese (dati dal 2008)	
Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI)	terzo quartile del valore aggiunto - migliaia di euro	Statistiche nazionali sulla struttura delle imprese (dati dal 2008)	
Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI)	primo quartile del fatturato - migliaia di euro	Statistiche nazionali sulla struttura delle imprese (dati dal 2008)	
Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI)	numerosità delle imprese con fatturato>0	Statistiche nazionali sulla struttura delle imprese (dati dal 2008)	
Rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI)	secondo quartile (mediana) della competitività di costo	Statistiche nazionali sulla struttura delle imprese (dati dal 2008)	

Functionalities, developments and benefits: search functionalities

Unit Data Structure (micro)


SUM
SISTEMA UNITARIO METADATI
Home
Ricerca strutture dati Macro
Ricerca strutture dati Micro
Questionario
Logout

ID Operatore: casagran
Profilo: Supervisore
Version 2.5 - 25 Febbraio 2016

Ricerca strutture dati micro

Rimuovi tutti i filtri
5 Risultati

per rilevazione, elaborazione, fonte

Indagine mensile su fatturato ed ordinativi (1)
Rilevazione degli incidenti stradali con lesioni a persone (1)
Rilevazione statistica sulla formazione nelle imprese (1)
Rilevazione sulle forze di lavoro (2)

per fase

Validazione (5)

per unità statistica

Famiglie di fatto (1)
Imprese con almeno 10 addetti (1)
Incidenti (1)
Individui (1)
Unità funzionali (1)

Risultati della ricerca

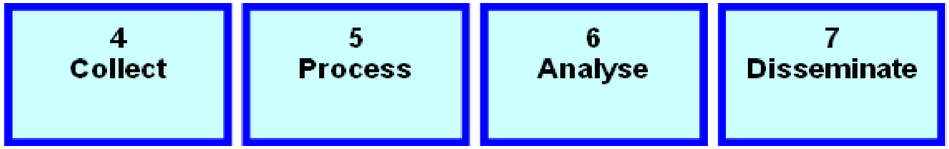
Nome rilevazione, elaborazione, fonte	Edizione	Data	Fase	Approfondimento
Indagine mensile su fatturato ed ordinativi	390	2012_M_11	Validazione	Unità funzionali
Rilevazione degli incidenti stradali con lesioni a persone	510	2014_A_0	Validazione	Incidenti
Rilevazione statistica sulla formazione nelle imprese	411	2010_A_0	Validazione	Imprese con almeno 10 addetti
Rilevazione sulle forze di lavoro	311	2015_T_2	Validazione	Famiglie di fatto

Comparisons and implementation of Gsim

GSIM concept	**SUM presence**
Population	Yes
Unit type	Not yet
Conceptual variable	Not yet
Represented variable	Yes
Instance variable	No
Classification	Yes (slight different definition)
Code list	Yes
Category set	Not yet
Data set	Yes (fixing roles for dimensions)
	Furthermore: Data content (for macro data structures)

Benefits and more

- ❑ The use of GSIM concepts helps in harmonizing the description of a dataset between each phases.



4 Collect

5 Process

6 Analyse

7 Disseminate

Output = transformation (input)

- ❑ Among the concepts already available in GSIM, an additional concept (the "**Data Content**") could be useful in order to feed in a standard and complete way a Measure of a Data Structure (of macrodata).
- ❑ This is what we have done in Istat. The corporate DWH (I.Stat) has more than 3300 "data contents". In SUM it is possible to search data through different facets:
  - ✓ Statistical program
  - ✓ Reference population of the data
  - ✓ Numerical variables used for the production of a data content
  - ✓ Categories of a categorical variable used in data structures

Thanks for your attention