Alessandro Greco - a.greco@sister.it   *Head of Business Intelligence Unit*
Francesco Pitto – f.pitto@sister.it   *SDMX team leader*

SISTEMI TERRITORIALI

The potential of SDMX for an IT company

n.3 new software solutions focused on SDMX to simplify the publication, research and reuse of statistical data

Sistemi Territoriali srl – Company profile

- IT company that was born in 1991 as a spin off of the CNR, the largest research center in Italy.

- It has more than 30 graduated employers, mostly in computer science (including PhD and university teachers), of which approximately 80% are senior profiles with very specialistic technical skills.

- Business partnership in Italy with ESRI (worldwide leader in GIS) and SAS (worldwide leader in Business Intelligence)

- One of the most active companies on the national scene in the field of big data and open data, and for such skills it is often engaged by some of the most important Italian IT companies (e.g. Engineering and Almaviva) and multinational corporations in joint projects for important customers.

SISTEMI TERRITORIALI

www.sister.it
info@sister.it

Sistemi Territoriali srl and SDMX

- Since 2013 Sistemi Territoriali has been developing solutions based on SDMX upon customer's request, specifically the statistical direction of the Lombardy Region.

- The potential of this standard has been immediately evident, so the company has continued in recent years to invest and to acquire growing know-how in this area.

- We are the only software company in Italy skilled in SDMX with a constantly growing team made up entirely of high-profile figures

- We are involved in n.3 projects that deal with complementary aspects such as the publication of data and their reuse:
  1. SAS for Open Hub
  2. Smart Data Cloud
  3. ISTAT Data & Meta Manager

SDMX Solution #1 : SAS for Open Hub

- Software platform, developed by Sistemi Territoriali and marketed in Italy by SAS Italy, which combines advanced linked open data functionality with SDMX data harvesting and publishing

- It integrates SAS and open source components to provide to public and private organizations all the necessary to set up a "data hub", i.e. a centralized catalog of data with search, data processing, display and open data publication

- The data hub can be automatically loaded with internal and external sources such as open data portals and SDMX endpoints

- The data hub has also a built-in SDMX wrapper: all the open statistical data loaded into the hub are automatically exposed also in SDMX, hiding SDMX complexity to the users

SAS for Open Hub – Case History

State General Accounting Office

Ministry of Labour and Social Policy
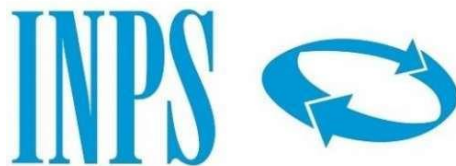
National Social Welfare Institute

Regione Lombardia

Regione Veneto

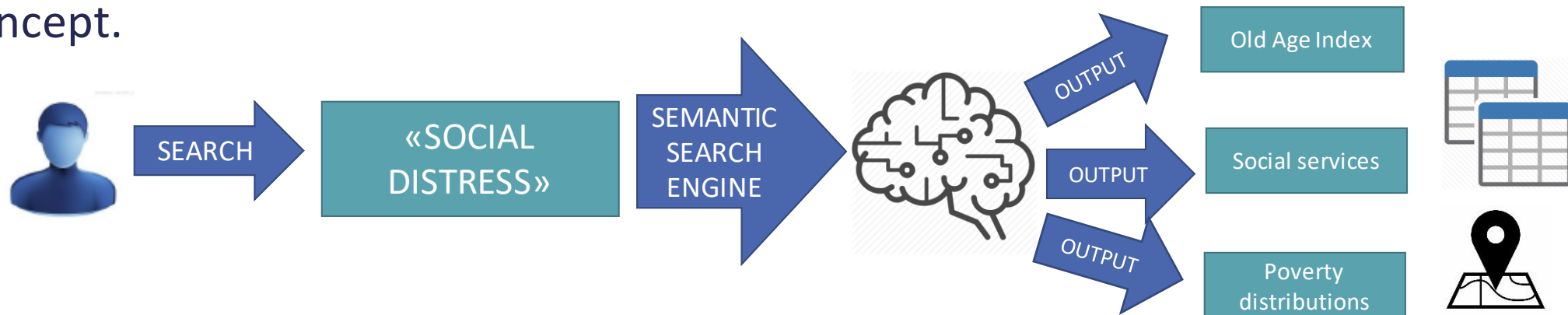
Istat Istituto Nazionale di Statistica

Work in Progress

SDMX solution #2 : Smart Data Cloud

- Smart Data Cloud is a Research and Development project co-funded by the Italian Ministry of Economic Development that provides for the creation of an innovative semantic data search engine based on ontologies.

- The main objective is to help final users in finding information of interest to them among those published as open data and in SDMX.

Smart Data Cloud - Semantic Search Engine

o It catalogues open data exposed by open data portal and SDMX endpoints

o It extracts from the data the most relevant semantic information

o It uses a Search Graph that links ontologies and data

o It uses Social Network Analysis (SNA) algorithms to perform semantic searches on the graph

o Returns very pertinent results to the user: not only data that contains the input text but also data semantically related to that concept.

SEARCH → «SOCIAL DISTRESS» → SEMANTIC SEARCH ENGINE → OUTPUT → Old Age Index

OUTPUT → Social services

OUTPUT → Poverty distributions

SDMX solution #3 : ISTAT Data & Meta Manager

- Designed and implemented on behalf of the Italian National Statistical Office (ISTAT). The main contractor is Almaviva, one of the Italian largest ITC companies

- It allows to streamline the dissemination and reporting business process, the metadata management and to facilitate open (statistical) data publishing following the requirements detailed by the European and Italian Digital Agenda

- With just few steps an organization can easily build a dissemination/reporting database driven by SDMX structural metadata, expose datasets through an SDMX Web Service and can disseminate the data catalog using the W3C Recommendation "DCAT" and CKAN v3 API

- A statistical organization can use the Toolkit for building "stand-alone" dissemination systems or a "distributed" data warehouse SDMX-based.

ISTAT Data & Meta Manager – Main features

- Full SDMX structural metadata management (Codelists, Concept Schemes, Category Scheme, Dataflows, DSDs, etc.)

- Artifacts management
  - Create, update, maintain, store, clone, delete and retrieve SDMX artifacts
  - Manage items with user friendly table/tree components

- Data publication in just 4 steps:
  1. Builder: Creation of a DDB scheme (data cubes) from DSDs and related artefacts
  2. Mapping: Creation of mappings between cubes (DSD) and CSV import files
  3. Loader: Loading data from CSV/SDMX-ML files
  4. Dataflow builder: Dataflow creation and publication

- Management of data embargo

- Full metadata management: production and exposure
  - Compilation of metadata reports
  - Machine-readable metadata in SDMX-JSON via REST API
  - Human-readable metadata in HTML via widget, sharable or embeddable in another website

- "Integration with the Open Data world"
  - Dedicated interface for the completion of DCAT-AP standard metadata
  - Federability with the EU Open Data Portal

SISTEMI TERRITORIALI
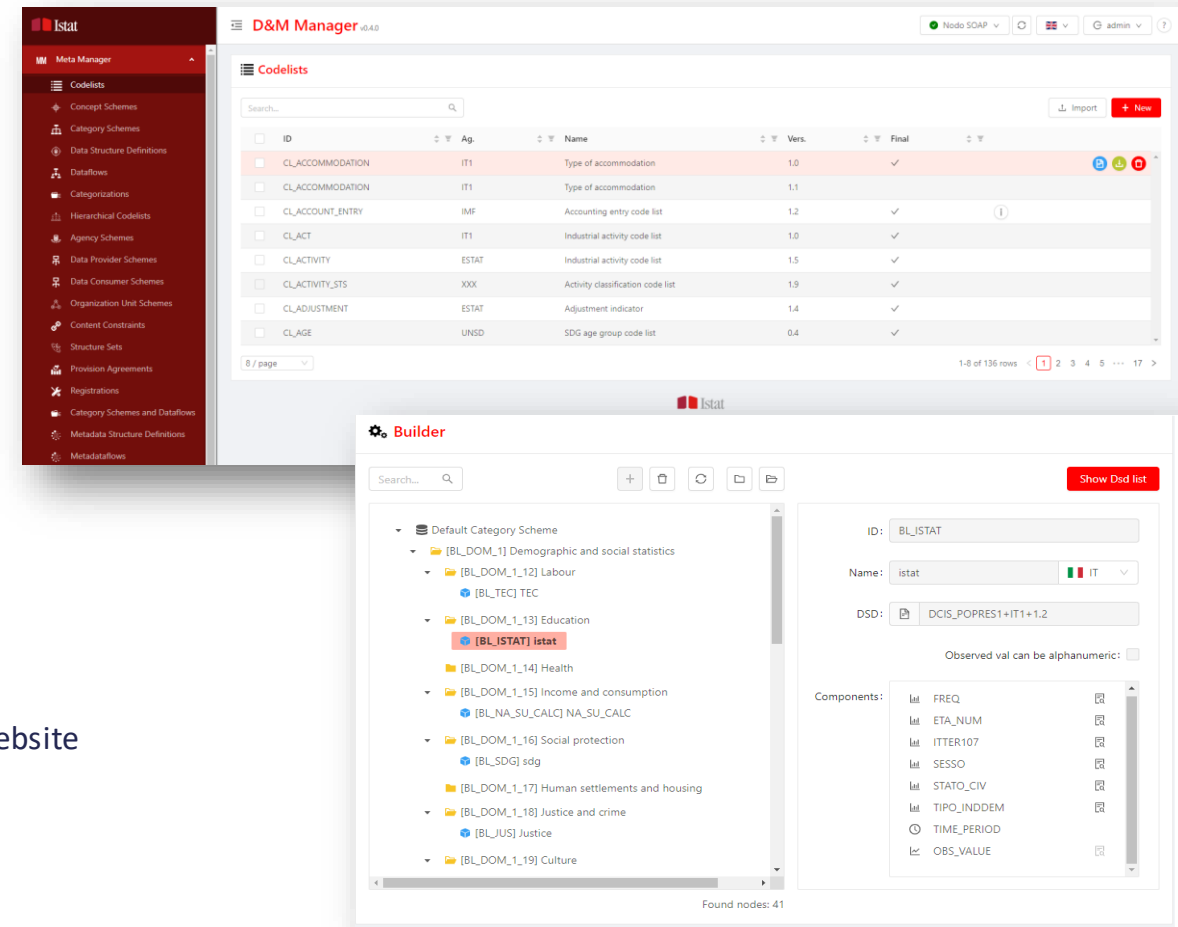
ISTAT SDMX Data & Meta Manager - Strengths

- Provides a wide range list of features for sdmx publishing
- Open source software (released with the EUPL), built using the SDMX Common API (SdmxSource.NET) and fully integrated with the SDMX-RI released by Eurostat. Will be released in Q4 2019.
- Very modern software fully developed with the latest IT tecnologies like React framework and .Net Core
- Based on long-term Istat's background and feedbacks received (previous version of the software is used by several statistical offices around the world)
- Parts of the development have been agreed with EuroStat and OECD on intensive exchanges of email and code
- User-friendly interface based on lessons learned
- Highly optimized data model (around 90sec for uploading 1 MLN of rows)
- Multilanguage native support
- Multiplatform and Multidatabase
- Easy to install

Alessandro Greco - a.greco@sister.it *Head of Business Intelligence Unit*
Francesco Pitto – f.pitto@sister.it *SDMX team leader*

SISTEMI TERRITORIALI

The potential of SDMX for an IT company

n.3 new software solutions focused on SDMX to simplify the publication, research and reuse of statistical data

Thank You!