

# Domain based data editing: the INSPECTOR project

Photis Stavropoulos

Liaison Systems SA, Akadimias 77, 10678, Athens, Greece,  
Photis.Stavropoulos@liaison.gr

George Petrakos

Liaison Systems SA, George.Petrakos@liaison.gr

Gregory Farmakis

Liaison Systems SA, Gregory.Farmakis@liaison.gr

**Keywords:** data editing, domain of variable, dataset template, enhanced dataset, abstract data model, segmentation

## Abstract

INSPECTOR is a project in the area of statistical data editing, funded under the 5<sup>th</sup> Framework Program and undertaken by a consortium of two research and development companies, two universities and two national statistical offices ([www.liaison.gr/projects/inspector](http://www.liaison.gr/projects/inspector)). While editing is usually implemented by coding edits in ad-hoc computer programs or in statistical and survey packages (e.g. SAS, Blaise, etc.), INSPECTOR investigates domain-based data editing, an alternative approach to handle edits computationally. The project aims to design and develop a generic, distributed and flexible system which will implement this approach and will be integrated in the IT infrastructure of the European Statistical System.

The project presently spans two years, with another six months remaining for its completion. It has formulated its approach to applying data editing and is currently developing a prototype of the system. The approach in essence expresses each edit as a definition of the domain of a variable. During these two years the project defined a number of new concepts and formulated methodological 'tools' which are needed for the adoption and more efficient use of the approach; such concepts and tools include the dataset template, which is a description of the structure of an enhanced dataset, the abstract data model which is a description of the relationships between survey variables and a method for the semi-automatic derivation of edits based on segmentation.

The aim of this talk is to present in its totality the approach of INSPECTOR to data editing and the functionality that will be offered by the software system. It will also present the new concepts and methodological tools, and will show their role in the approach and therefore in the software too.