

# Contestina: A visibly understandable path toward more effective data dissemination

Mariagrazia Zottoli (zottoli@contesti.info), Sara Laurita (laurita@contesti.info), Francesco Monteleone (f.monteleone@sdlab.it)<sup>1</sup> and Andrea Pugliese (andrea.pugliese@unical.it)<sup>2</sup>

**Keywords:** digital storytelling, inferential engine, data dissemination, statistical literacy

## 1. INTRODUCTION

Today we live in a world literally awash with data. For years, the process of data dissemination has ended with placing various data into complex databases or ready-made tables to provide an accurate reading of the phenomenon of interest. As a result of such a process, it has rendered the most interesting types of stories completely underwhelming, or worse, difficult or impossible to understand [1] for the average user. Adelaida Trujillo echoes this sentiment, stating: "The important discussion around data is that there's a big gap in terms of how data is communicated to the ordinary citizen"[2].

At the present time, the vast majority of people are not skilled to fully understand the insights illustrated from data. United Nations Economic Commission for Europe (UNECE) states: "*What is ou[r] information worth, even if it is constantly mentioned in Tv news and newspapers, if many people... [don't]... even understand the meaning of the figures?*" (2012). Therefore, it is crucial to address this lack of statistical literacy –and more generally the limited ability to reason and apply simple numerical concepts (lack of numeracy)– in order to fulfill the democratic function of Official Statistics.

### **Contestina Approach**

One way in attempting to fill the gap would be to focus on applying more innovative storytelling techniques, such as that of digital storytelling, to guide the user in a step-by-step reasoning process. A new and innovative inferential engine that does just this, called **Contestina**, is a platform designed to tell adaptable stories based on the given data. We believe that this tool is an effective way to allow users "...to have a good understanding of the context in which indicators are presented to properly use and interpret the data" [3]. When people come across statistical information either on the internet, in libraries or through the media, they need to be able to consult, understand and check information without the need to navigate through the vastness of the web, and around parallel types of content, while being forced to visit several sites. Thus, **Contestina's** overall goal is to help users of all levels (also called "tourists" according to [4]) in understanding official statistics.

## 2. METHODS

As storytelling is the milestone of **Contestina's** system, it relies on a reasoning mechanism that automatically builds a scenario analysis according to the parameters chosen by the users. The system is based upon a set of semantic rules, which exploring the available data stored in a database, helps users to discover and model hidden patterns related to the phenomenon of interest. Moreover, this approach incites a more critical navigation by way

---

<sup>1</sup> Contesti srl, Via della Resistenza, 23, Rende (CS), Italy.

<sup>2</sup> Università della Calabria, Ponte Piero Bucci, Arcavacata di Rende (CS), Italy.

of a snapshot to display Official Statistics data, in an attempt to enrich the user experience through metadata and interactive visualization techniques.

Another important method *Contestina* uses is the concept of tutoring, which we believe to be an important step in the storytelling process, not normally found within traditional concepts of giving statistical information. The tutoring element is based on a software agent---that with an initial implementation of domain experts' command---guides users through a step-by-step rational: an initial objective reading of the data presented in the final report. This critical step helps users reason with the data, allowing them to be able to differentiate between various numerical values.

The platform architecture is developed into three main layers: *i)* Analytical Information Subsystem (containing all the metadata, the report layouts, and the tutoring rules); *ii)* Jdbc layer (a definition of protocols for the communication with the previous layer); *iii)* presentation layer (a modelling of the graphical representation of the objects). The definition of the report is carried out into three stages:

- 1) The construction of the *snapshot*: starting from a conceptual graph model, where the nodes are the report elements and the edges are the semantic relations among them.
- 2) The compilation of the *tutoring rules*: for the selection of the relevant data.
- 3) The automatic generation of the *storytelling*.

One of the biggest advantages of the technological solution of *Contestina* is the possibility to apply it to several domains, regardless of the Business Intelligence (BI) platform used for the reporting activity.

### 3. RESULTS (PROCESS)

The *Contestina* approach facilitates the comprehension of statistical information for all users, and in particular "tourist" users, thanks to the availability of metadata and figures joined together and formed into a comprehensible story. To publish official statistics through an assisted content creator (an important component of *Contestina's* digital storytelling), while emphasizing dynamic data visualization, represents a key advantage of this effective exploratory method. *Contestina* begins by guiding users through the reasoning process to allow them to easily gather insights from the data. Starting from a snapshot composed by two main areas: a chart with data and a graph (histograms, bar plots, pie charts, and so on), the tutoring agent, with its interactive storytelling, guides them through the report interpretation. The story gets built through the use of metadata (related to the source, as well), which helps the users navigate through report content, giving them the chance to find additional information by means of some interactive features (such as tooltip, brushing and highlight), as well as links to other web pages of interest. Additionally, in order to analyse the data in a comprehensive scenario, *Contestina* also suggests other useful reports to the users, improving the contextualization of the data, making it more understandable and attractive.

### 4. CONCLUSIONS

The *Contestina* platform introduced in this proposal launches a new paradigm in accessing information: while technology has increased the proficiency in tools to work with data,

there remains lacks in capabilities, and *Contestina* successfully addresses that. “*Statistics can tell people something about the world they live in. But not everyone is adept at understanding statistics by themselves*” [5]. Our approach has always been to address this challenging factor in the ever-growing *datasphere*, so that information may not only be accessible, but comprehensible and transparent. Through the employment of visual storytelling, statistical information becomes relatable to all types of users, thus, visibly understandable to all.

## REFERENCES

- [1] Knaflic, Cole Nussbaumer. *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Hoboken, New Jersey: John Wiley & Sons, 2015. Print.
- [2] Alan Stanley, *Data Brokers and the Demand-led Digital Revolution*. - Institute of Development Studies. Web. 20 Oct. 2016.
- [3] Peter Van De Ven (OECD), *Government Finance Indicators: The Truth about the Myths* (8/02/2013). The TSEconomist, Toulouse School of Economics (TSE) Students Magazine. Web. 20 Oct. 2016.
- [4] Vale, Steven (UNECE). *The role of Metadata. UNECE Training Workshop on Dissemination of MDG Indicators and Statistical Information*. UNECE- Statistical Division, Astana, Kazakhsan, 2009.
- [5] United Nation Economic Commission for Europe (UNECE). *Making data meaningful. Part 1: A guide to writing stories about numbers*. United Nations, Geneva, 2012.