

# OpenCoesione and Monithon - a transparency effort

**Lorenzo Canova<sup>1</sup>, Antonio Vetro<sup>1,2</sup>, Marco Torchiano<sup>1</sup>, Raimondo Iemma<sup>3</sup>, Federico Morando<sup>3</sup>**

*<sup>1</sup>Politecnico di Torino (Italy), <sup>2</sup>Technische Universität München (Germany), <sup>3</sup>Nexa Center for Internet and Society at Politecnico di Torino (DAUIN) (Italy)*

## Abstract

### Context

OpenCoesione is the first portal about the fulfilment of investments and projects planned by the Italian central government and by the Italian Regions using the 2007-2013 European Cohesion funds. Together with Monithon, it is a “transparency tool” whose aim is to foster participation of the citizens and efficiency of the public sector bodies in order to improve the implementation of development policies. By now it is one of the best Open Data portal in Italy quality-wise.

### Objective

Our goal is to show the utility of these portals, how this open information is supposed to help the civil society and how data quality might affect reuse.

### Method

We engage in the empirical observation on how data are exposed and used, discussing specific examples, and applying some data quality metrics.

### Results

We present some evidences on how open data can positively affect the public sector bodies and the spending of funds.

### Conclusions

Under-spending of EU Cohesion funds is a serious problem in Italy. OpenCoesione and Monithon can contribute solving this inefficiency, e.g., by presenting data in such a (standardised) way to enable their elaboration by third parties

## **OpenCoesione - a project of opening, transparency and reuse of data about the Cohesion policies' operations**

OpenCoesione<sup>1</sup> is the first portal about the fulfilment of the investments (both planned and actual) and planned projects by the Italian central government and by the Italian Regions using the 2007-2013 European Cohesion funds. The data about these projects are published on the portal in order to make citizens able to evaluate if the projects will satisfy their needs (and whether the achievements are consistent with the stated objectives) and if the related funds are employed in an efficient way. At the time of this writing, it contains data about 750.000 projects, worth 75 billion Euros of financings. Data and information about territorial Cohesion policies are available on the portal, specifically: projects fundamentals, amount of funding, locations, involved subjects and completion times.

OpenCoesione is designed to be used by normal citizens as well as by experts such as journalists, civil servants, policy makers, public managers, entrepreneurs: data can be viewed in an interactive way with graphs and maps (Fig.1) and they can be downloaded as raw data in the form of csv datasets.

---

<sup>1</sup> <http://www.opencoessione.gov.it/>

The aim of OpenCoesione is to foster transparency, participation of the citizens, tracking of expenditures (which is useful for the public sector itself, as well) and, more in general, economic development, in particular at the local level.

The portal also makes available a feedback procedure for suggestions, warnings for errors or missing projects, requests for specific elaborations of data, and for sharing examples of reuse of data.

The sources of the data are: in the matter of the EU structural funds<sup>2</sup>, the central monitoring systems fed by the funds' beneficiaries; in the matter of Italy's national Development and Cohesion Fund, also the resolutions of the Interministerial Committee for Economic Planning (CIPE), which assigns resources to single interventions<sup>3</sup>. In particular, the information about the 2007-2013 funds are gathered by a single monitoring system managed by the General State Accounting Department (RGS – Ragioneria Generale dello Stato) of Italy's Ministry of Economy and Finance. The regional and national governments that manage the funds therefore represent the main authoritative sources for the information published on the OpenCoesione portal.

The software used for OpenCoesione has not been released as open-source. Its data are released under the licence CC BY-SA 3.0<sup>4</sup>, that qualifies as an open license according to the Open Definition<sup>5</sup>. Therefore it is possible to reproduce, distribute and communicate data and analyses to the public, also for commercial purposes (as long as the business model of the relevant commercial activity is compatible with the requirement to distribute derivative works under the same terms of the original work, i.e., CC BY-SA, version 3.0 or later/compatible license).

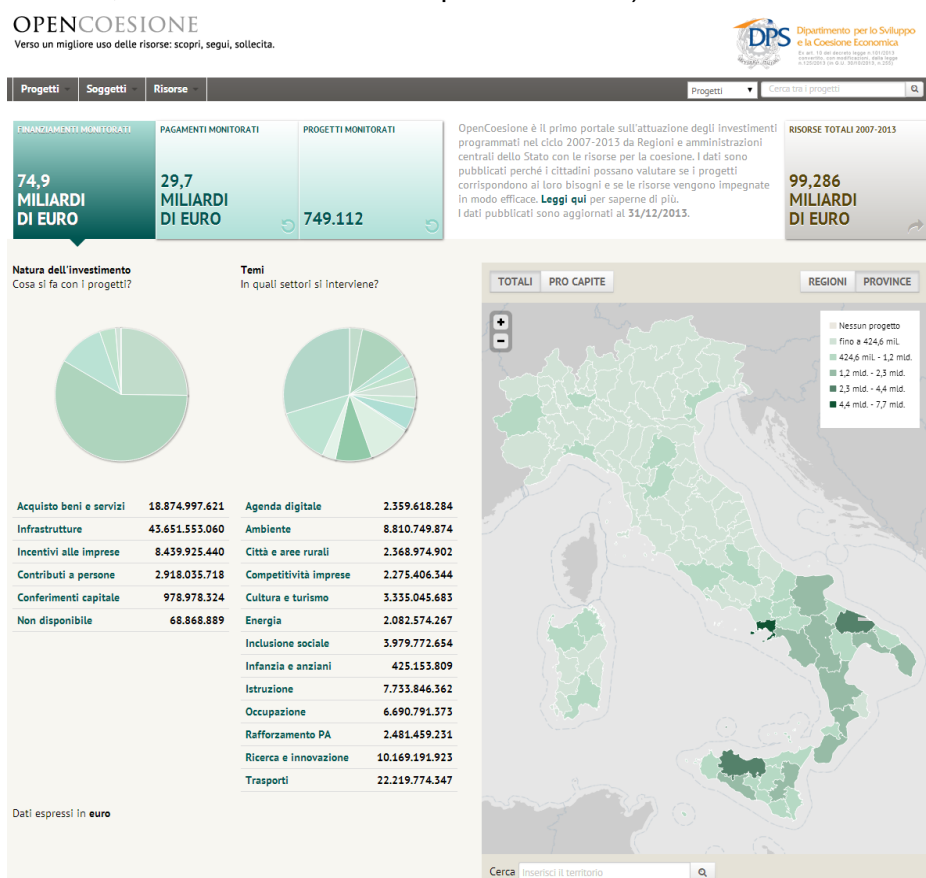


Fig.1 OpenCoesione's interactive interface and maps

<sup>2</sup> [http://en.wikipedia.org/wiki/Structural\\_Funds\\_and\\_Cohesion\\_Fund](http://en.wikipedia.org/wiki/Structural_Funds_and_Cohesion_Fund)

<sup>3</sup> source: <http://www.opencoessione.gov.it/progetto/en/>

<sup>4</sup> <https://creativecommons.org/licenses/by-sa/3.0/>

<sup>5</sup> <http://opendefinition.org/od/>

## Visits on OpenCoesione

Since its launch, in July 2012, OpenCoesione has had a good impact, at least in terms of awareness and media coverage, showing that there is interest in the open data made available by the Italian government. As a matter of fact, the visits to the Open Coesione portal, from July 2012 until March 2013, are:

- 700.000 pages have been visualized
- 140.000 unique visitors
- 180.000 visits
- 8% of visits from foreign countries
- 2 minutes average time on site
- 1000 downloads of the full “projects” file

## Monithon - citizens' collaboration to help transparency

Monithon<sup>6</sup> is an independently developed initiative to promote the citizen monitoring of projects funded by the Cohesion policy in Italy. Monitoring is possible thanks to the combination of citizens' collaboration and open government data. The goal is to check the thousands of projects funded by both the EU and the State through Cohesion Policy. The Monithon initiative takes this transparency challenge even further than OpenCoesione: it asks citizens to actively engage with open government data and to produce with them valuable information. The goal is to improve citizens' use of open data to make them feel closer to public policies and to let them know how the public money and EU funds have been spent. By contributing to Monithon, besides giving advantages to the whole community, citizens also contribute to the public mission themselves: they can signal errors, report on the actual development of projects, fill in missing information, and eventually assess the efficacy and efficiency of the funding. As explained in the website of Monithon<sup>7</sup>: ““Moni-thon” comes from “monitor” and “marathon”, and this is precisely what the platform seeks to help with an intensive activity of observing and reporting. Citizens, journalist, experts, researchers – or all combined – collect information on a specific project from the database and upload it on the website. This information – which can be in the form of interviews, quantitative data, pictures, videos, etc. – is then aggregated and geo-referenced on the Monithon platform, to create a form of civic, bottom-up, collective storytelling.” Moreover, thanks to the (third party's) Ushahidi<sup>8</sup> smartphone application, it is possible to participate in Monithon at any time. Anyone can do this individually, but if done collectively, during the Monithon Days, the whole activity increases its importance and therefore its effectiveness. The Monithon Days are basically events in which a group of people explores a specific area and gather information about projects of local interest.

## Benefits of Monithon

The lack of transparency on how public money is spent and the lack of control over the projects status are two of the main reasons for the slow pace in implementing development policies. This is a particularly hot topic in Italy at present given the low absorption rate of EU Structural Funds<sup>9</sup> (Katsarova, 2013). Monithon Days and singular reports have already occurred with success in several locations, even if Monithon is still in its beta version.

Over 43 major projects have been monitored, 22 of them in 2014. Problems have been detected in several projects, some of them have been stopped for many years due to project approval issues (e.g: the rail circuit in Palermo has been stopped for 9 years<sup>10</sup>), or they have suffered several delays

---

<sup>6</sup> <http://www.monithon.it/>

<sup>7</sup> <http://www.monithon.it/page/index/2>

<sup>8</sup> <http://en.wikipedia.org/wiki/Ushahidi>

<sup>9</sup> just 40% of the funds are absorbed, it's one of the worst in Europe, a useful picture is given here:

<http://www.insideurope.eu/node/403>

<sup>10</sup> <http://www.monithon.it/reports/view/95>

due to inefficiencies<sup>11</sup>. Once the problems and the defective projects have been discovered it is way easier to take the necessary steps for both the central government and the citizens.

## Quality of Open Data

In Italy most of the information is dispersed across a large number of managing authorities and lots of the available data are published in PDF format; they qualify for one star according to the Five Star Open Data evaluation (FSOD)<sup>12</sup>, in fact they lack machine readability. During the latest years there has been an increasing use of the Open Data by the public sector bodies<sup>13</sup>, even if the level of openness of other countries like the UK, is still far to reach.

In OpenCoesione, however, datasets are published in a machine readable and non proprietary csv format (deserving three stars FSOD). Moreover, together with the data, a good set of metadata is provided, making the understanding of the dataset easier to process and elaborate for all kind of users.

Different datasets in OpenCoesione are published, which regards distinct type of fundings<sup>14</sup> but share the same data structure, allowing comparative analyses among them. In each group there are different datasets: “projects” (representing the funded projects), “subjects” (that is either the Programming Authority, or the recipient of the funding), “locations” (representing projects’ location data) and “payments” (payments made for each project)<sup>15</sup>. For each dataset there is a standard set of attributes, for example for “projects” there are 74 attributes describing: financial values, description values, time values and flags (warning values for incoherences).

Before being published the data through pre-processing, to guarantee readability and coherence. This task is performed by the central monitoring system who manages the data coming from local governments.

During this process the data are verified taking into account different aspects, for each section (identification data, financial data, information over the subjects and procedures) completeness correctness and coherence<sup>16</sup> are controlled with several different tests. Analyzing 70.264 projects, several errors have been found doing this kind of controls: 530 errors in the ID of the project (either missing, or wrong), 4245 errors/incoherencies in financial sections, 1098 errors in the location of the projects, 185 incoherencies in procedures’ informations (Romano, 2013).

After these quality checks, some minor data quality issues have still been found by the authors of this paper on the published data:

- a small percentage of fundings (0.5%, but corresponding to approximately 50000 funding records) is below 1 Euro. They might indicate an error, if the numbers were expressed in thousands of Euro instead of Euro units (as declared in the metadata), the resulting error, corresponding to mis-tracked projects, would be of roughly 50 million Euro.
- there are some problems of completeness (missing values) in the sections regarding the start/end day of a project;
- concerning the description of the project, there are some problems in understanding the project’s categorisations (e.g., mismatches such as a project aimed at increasing women employment categorised as “youth and elderly”);
- the data domains are not specified in details, leading thus to confusion when trying to interpret empty cells and verifying the correctness of data.

These data quality issues do not invalidate the utility of the open information provided in OpenCoesione. In fact its level of quality is much higher in respect to the average quality in italian open data, which we are observing in an ongoing study (Vetrò, Torchiano, 2013).

---

<sup>11</sup> <http://www.monithon.it/reports/view/105>

<sup>12</sup> <http://5stardata.info/>

<sup>13</sup> <http://www.dati.gov.it/content/infografica>

<sup>14</sup>type of fundings: European Structural funds, funds for Cohesion and Development, CIPE funds

<sup>15</sup>further information here: <http://www.opencoesione.gov.it/progetto/en/>

<sup>16</sup>e.g.: between procedural data and physical data, or between start/end date

However, although it is still possible to keep track of the projects present in Open Coesione regardless of the observed quality issues, better controls during the input of the data in the system would be beneficial to increase its accuracy.

## Results and conclusion

The OpenCoesione portal together with the Monithon project provide analysis and monitoring on the use of Cohesion policy resources offering information, accessible to anyone, on what is funded, who is involved and where. The Monithon days have been particularly effective in finding inefficiencies and “zombie” projects<sup>8</sup>, this is one of the first evidences of this transparency effort’s usefulness. Ultimately, OpenCoesione represents an attempt to: firstly, improve policy effectiveness through better knowledge on which kind of investment projects are actually carried out in the territory; secondly, enhance coordination among the local governments responsible for implementing the projects; and, thirdly, provide a better public scrutiny on who benefits from the resources (also with the aim to avoid corruption and fraud).

Data quality is critical for the effectiveness of OpenCoesione; we have shown that despite the presence of pre-processing performed by central administration, a few other issues are still present, which may reduce reuse opportunities.

While data quality is still an underexplored theme in the open data landscape, these issues affect only a small portion of the OpenCoesione records without hampering the final goal of the project.

## Acknowledgements

We would like to thank Dr. Aline Pennisi for the useful information provided and Giuseppe Futia for proofreading.

## References

Antonio Vetrò, Marco Torchiano, Camilo Minotas Orozco, Giuseppe Procaccianti, Raimondo Iemma, Federico Morando (2013), *An Exploratory Empirical Assessment of Italian Open Government Data Quality*. Available at: [http://softeng.polito.it/reports/Softeng\\_TechReport\\_2014.pdf](http://softeng.polito.it/reports/Softeng_TechReport_2014.pdf)

Silvia Loddo (2006), *Structural funds and regional convergence in Italy*, CRENoS Univeristà di Cagliari. Available at: <http://veprints.unica.it/220/1/06-03.pdf>

Ivana Katsarova (2013), *The (low) absorption of EU Structural Funds*, Library of the European Parliament. Available at: <http://www.europarl.europa.eu/eplibrary/The-low-absorption-of-EU-Structural-Funds.pdf>

Saverio Romano (2013), *Quali condizioni per la trasparenza? Miglioriamo la qualità dei dati*, Workshop at Forum PA "Vieni ad aprire i dati sulle politiche di coesione" (Roma, 29 maggio 2013). Slideshow in Italian available at: <http://www.dps.tesoro.it/opencoesione/docs/documenti/20130529%20ForumPA%20Romano%20GRUE.pdf>