

# Exploiting crowd sourced platforms for statistical purposes

**Donatella Fazio and Marina Signore**  
**Italian National Statistical Institute (Istat)**

## **Abstract**

ICT tools have been widely used for supporting statistical activities from data collection to data dissemination. More recently, digital networks and platforms have been exploited for sharing information and data to foster the debate over “GDP and Beyond”, such as the e-FrameNET, launched in 2012 within the EU FP7 e-Frame project (*European Framework for Measuring Progress*) and hosted by Wikiprogress.org (OECD).

Recent developments concern the possibility of producing statistics using crowd sourced data. Indeed, the need to bridge the top down approach with a bottom up approach, at different stages, has been now recognized as a potential for statistical production. Against this awareness, the EU FP7 project Web-COSI (*Web COmmunities for Statistics for Social Innovation*), built up on the increasing trust in collectively - generated statistics, aims at fostering the engagement and access of society at large for progress statistics.

The paper will present the potentialities and expected results of this approach where the opportunities of Web2.0 are exploited for implementing tools to collect, produce and visualize information and data. Furthermore, it will discuss the main quality issues in view of a better integration of official and non- official data as well the governance needs for supporting such a process.

## **1. ICT tools for the collection, production and visualization of progress statistics**

Internet era has radically changed the approach in which information is produced and shared. The process is increasing its impact on statistical data collection, dissemination, visualization and retrieval. The Digital Agenda for Europe (DAE) [2], launched within the Europe 2020 strategy [3], has endorsed the process with the aim to deliver sustainable economic and social benefits from a digital single market based on fast and ultra-fast internet and interoperable applications. All the actors belonging to the European Statistical Systems are pushed to adopt the DAE actions to leverage the Internet improving ways from data collection to data dissemination.

Web 2.0 platforms are the “tools” in an era of mass-collaboration and participation. Blogs, wikis, Facebook, as well as, the innovative concepts of “collective intelligence” and “crowd

sourcing” have transformed the manner in which information is accessed. Web 2.0 allows the interactivity between producers and consumers of data, implementing the opportunities for producing ‘new data’ in a global perspective. Web 2.0 brings to the involvement of communities for data collection (Big Data, grass root generated data). The process redefines statistical construction by bridging top-down and bottom-up approaches for better statistics beyond GDP which allows for non-official sources to be standing along-side the official.

Over time the “beyond GDP” debate has been strongly empowered by Web 2.0 creating a global movement willing to voice its opinions. European and global initiatives have been set up to enlarge the discussion and to better interact with the communities.

Recent notable bottom-up initiatives include the OECD’s Better Life Index [7], which was launched in 2012 to answer one of the world’s most pressing questions: are our lives getting better? The index is designed to invite users to visualise and compare some of the key factors – like education, housing, environment – that contribute to well-being in OECD countries. The aim is to allow understanding what drives well-being of people and nations and what needs to be done to achieve greater progress for all. With the awareness that there is more to life than the cold numbers of GDP and economic statistics, the Index is an interactive tool that allows seeing how countries perform according to the importance the user gives to each of the 11 topics that make for a better life.

On the side of the EU NSIs, in order to develop official statistics beyond GDP, stakeholder’s consultations and discussions with society at large began to be carried out. In 2010, the Italian National Institute of Statistics (Istat) and the Italian National Council for Economics and Labour (CNEL) created the ‘BES’ initiative ([www.misuredelbenessere.it](http://www.misuredelbenessere.it)) designed to measure ‘Equitable and Sustainable Wellbeing in Italy’. A deliberative process for the definition of a theoretical framework for the measurement of well-being and the definition of the indicators which are composing it was carried out. The selection of the indicators has been done through the dialogue between a scientific commission of experts, a national steering committee with entrepreneurs, unions and the civil society at large, supported by public meetings, a national survey, a blog and an online questionnaire. The result of the consultation has provided a decision co-established by the Italian society at large.

## **2. Contributions from e-Frame project**

As seen, ICT developments are increasingly influencing the way by which data are generated, disseminated and shared both at national and supranational level, also in the context of measuring

well-being and sustainable development. In 2012, the European Commission funded the FP7 project e-Frame - *European Framework for Measuring Progress*<sup>1</sup> - with the overarching aim of streamlining current initiatives and fostering the on-going debate on progress statistics. By looking together at the social, economic, environmental components of “GDP and beyond” statistics, e-Frame contributes to empower the European dimension by stocktaking existing results, developing effective communication strategies and proposing the way forward.

Among the various project activities, e-Frame also exploits the usage of ICT tools for the production and dissemination of statistics on “GDP and beyond”.

On the one side, the project concentrates on National Statistical Institutes (NSIs) and the systems in place for the measurement of progress, well-being and sustainable development together with the identification of specific challenges NSIs face in accomplishing such a task. In particular, the e-Frame report “*ICT delivering tools. Catalogue of user tools and discussion relating to implementation*” ([www.eframeproject.eu/fileadmin/Deliverables/Deliverable8.1.pdf](http://www.eframeproject.eu/fileadmin/Deliverables/Deliverable8.1.pdf)) identifies and recommends best practices to NSIs in the presentation and visualisation of official statistical information on the Internet. Actually, ICT innovations and opportunities offer richer ways to disseminate statistical information to an increasing number of users and stakeholders. In addition, ICT tools may significantly help in the challenge to account for the multidimensional nature of well-being.

The report illustrates how good data visualisation can benefit NSIs. It catalogues the types of data visualisations found across European NSIs, grouping tools and technologies into four broad categories. The report discusses the benefits and drawbacks of each group. Case studies are used to demonstrate which tools and technologies are most suitable to different types of NSIs. The report ends with some recommendations on how NSIs can move forward and use data visualisations and interactive graphics to enhance their outputs and engage with a wider audience.

Among key points, the report stresses:

- The need for a well thought out data visualisation strategy, taking into consideration the organisations key products, user requirements, dissemination strategy, available funds and skill set, before considering the most appropriate tools and technologies.

---

<sup>1</sup> e-Frame is an FP7 project (2012-2014) funded by the European Commission, DG Research and Innovation (GA 290520). The project is based on a Consortium of 19 partners (NSIs, civil society organisations, academia, research centres, OECD) coordinated by ISTAT, the Italian National Statistical Institute and Statistics Netherlands ([www.eframeproject.eu](http://www.eframeproject.eu)).

- Data visualisation is an inter-disciplinary activity, the creators of visualisations using the raw programming approach need to be designers, programmers and statisticians working together.
- Rather than create entirely new networks, there is potential for existing pan-European dissemination groups (for example the EUROSTAT Dissemination Working Group) to provide a discrete forum for networking, sharing and best practice.

On the other side, Web 2.0 infrastructures are employed as networking tools, documentation repositories and dissemination platforms. The e-Frame network (e-FrameNET) on Wikiprogress platform represents the core networking tool at European level, as further described in next section.

In order to better coordinate and interlink European social monitoring and reporting initiatives, the European Social Monitoring & Reporting Web-Platform was developed ([www.gesis.org/en/social-indicators/products-of-the-zsi/european-social-monitoring-and-reporting](http://www.gesis.org/en/social-indicators/products-of-the-zsi/european-social-monitoring-and-reporting)). It allows to present continuously update social monitoring and reporting information as well as research results on individual and societal wellbeing in Europe and EU Members States, including news on conferences, publications, policies, etc.. The platform may also be used for communication and exchange among network members.

Finally, the project website [www.eframeproject.eu](http://www.eframeproject.eu) has been conceived as a virtual forum for all consortium partners and as a major tool for networking and spreading information and documentation on project activities and events.

### **3. The experience of the “European Network on Measuring Progress”**

In 2012, among the many digital initiatives carried out for sharing information and data over “GDP & Beyond”, the *European Network on Measuring Progress* (e-FrameNET) was set up by the Italian National Statistical Institute (Istat) as an offshoot of the e-Frame EU FP7 Project.

The e-FrameNET, launched at the OECD European Conference on Measuring Well-Being and Fostering the Progress of Societies, 26-28 June 2012- Paris [8], hosted by Wikiprogress platform<sup>2</sup>, is built up with the objective to contribute to the instituting of a European position on the issues related to methodological and theoretical aspects of new indicators for the measurement of societal

---

<sup>2</sup> The Wikiprogress platform was launched at the 3rd OECD World Forum on “Statistics, Knowledge and Policy” held in Busan, Korea in 2009. It’s an open source global platform crowd sourced allowing users to generate content by developing information and data. It counts: 30,000 subscribers and about 20,000 monthly visits ([www.wikiprogress.org](http://www.wikiprogress.org))

progress and well-being. At the same time, it aimed to drive the debate towards the highly global Wikiprogress platform hosted by OECD.

The Network's structure allows members to communicate through the e-Frame website and the Wikiprogress platform in order to connect stakeholders, researchers, organisations, citizens and policymakers in the on-going debate on what constitutes the most "accurate" measurement of well-being and societal progress. It involves stakeholders in the e-Frame project by facilitating the dissemination of relevant content. Members of the e-FrameNET are organised in a database, centrally managed by ISTAT to respect the privacy rules, which permits for the creation of mailing lists and for disseminating details about relevant activities. The Network offers interactivity in a global dimension through online discussions, blogs, eBrief, monthly news alerts, aiming to upraise the visibility of the findings of pertinent projects on progressing beyond GDP.

The position of e-FrameNET on the Wikiprogress platform represents its key strength driving the European local, national and international debate towards a highly global perspective. It constitutes the European Network, which functions alongside the Regional Networks of Africa and Latin America. It contributes to discussion at a regional and subject-specific level.

The Network forms a community of academics, analysts, opinion leaders, citizens and society at large around the subject of measuring well-being and societal progress in European countries. Key members include National Statistical Institutes, civil society organisations, academia and research centres belonging, till now, to eighteen European countries.

#### **4. Web-COSI project**

Against the background provided by the discussion on "GDP & beyond" and capitalizing on the Web 2.0 opportunities, a new FP7 project, coordinated by Istat, Web-COSI - *Web Communities for Statistics for Social Innovation*- kick offed in January 2014<sup>3</sup>. Web-COSI is funded by the EC-DG CONNECT within the FP7 ICT Work Programme 2013, responding to the Collective Awareness Platforms for Sustainability and Social Innovation (CAPS) call ICT10 [1, pag.63].

Following the request of the call, Web-COSI – based on the idea that collaboration through crowd-sourced platforms can produce solutions for a wide range of social needs – is designed with the general objective to foster the engagement of citizens and society at large in the area of the new measurements of wellbeing, societal progress and sustainability. Built up on the increasing trust in locally generated data, it has the specific objective to implement tools for collecting, producing and

---

<sup>3</sup> Web-COSI has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 610422 ([www.webcosi.eu](http://www.webcosi.eu)).

visualizing information and data towards a better integration of official and non-official statistics, empowering crowd sourced data.

The project is based on a four-partner consortium that sees collaboration between two relevant institutions (Istat and the OECD- Organisation for Economic Co-operation and Development), alongside Lunaria – the Italian Association for Social Promotion ([www.lunaria.org](http://www.lunaria.org)), and i-genius - the Social Entrepreneur Business and Enterprise Community in the UK ([www.i-genius.org](http://www.i-genius.org)), which represent society at large. The partners are linked by complementing expertise but, above all, each one represents a strategic actor in the progress of the statistics debate. The consortium as a whole is multidisciplinary and well balanced so as to work towards creating synergies and critical mass, thereby leveraging the network effect towards the implementation of new ways of collecting and producing information and data and to involve and motivate citizens to contribute to and access statistics (both official and non-official).

The coordination of the project in charge of the Italian NSI, sets the key question on how NSIs, producers of official statistical information, can make a good use of the new Web 2.0 sources of data in the era of “Statistics 2.0” [6]. These new processes arise the trade-off between having more and real-time information and the quality of the information produced. The matter is to integrate different information and disseminate data and metadata, granting a standard of quality for non-official data. A key element for the integrated usage of official and non-official statistics is to clearly distinguish the latter (label) to take into consideration their representativeness.

The two-year Web-COSI work plan foresees a number of activities carried out by the four partners: beginning with the mapping of beyond GDP existing digital initiatives, thus distilling the best practices of the involvement of the communities for the progress statistics (experiences carried out by the NSIs, by civil society organisations and social entrepreneurs). In the following months the work plan will carry on specific initiatives and targeted online campaigns with the aim of empowering the engagement of communities in the debate on progress statistics, reaching a high level of citizens’ involvement at the end of the project.

The development of a Wiki of progress statistics, at the mid-term point of the project, will constitute a tool to manage the collection of civil society, grass root locally-generated non-official data which can be integrated/complemented with official data.

Moreover, five stocktaking and reporting documents on the involvements of communities on statistics beyond GDP in Europe and abroad will be delivered. Finally communication and dissemination through face-to-face events will be organised: five workshops, four focus groups and a final conference.

The activities of the project will be carried out and reported using Internet platforms managed by the partners, using social networks, online discussions, blogs, e-brief and an extensive network and local grassroots partnerships creating synergies in order to move towards a growing integration of top-down and bottom-up approaches. A strong communication strategy on the on-going work will be carried out through videos and short films for social media distribution.

The consortium will be supported by an advisory board<sup>4</sup>, composed by outstanding known personalities, who will provide advice and guidance for the development of the project to ensure high quality and excellence.

Web-COSI will work in the context of the CAPS projects<sup>5</sup>, creating synergies with the diverse categories of the communities of users and stakeholders. A strong collaboration between the Consortium and the European Commission (DG CONNECT, Eurostat, DG Research and Innovation) will be ensured to foster the impact of the achievements of the project. The project will take advantage from the results of on-going or just completed related EU FP7 initiatives and Eurostat initiatives. Connections and interactions with projects will be fostered.

Achieving its results Web-COSI will pave the way for future research needs identifying the open issues for better statistics beyond GDP towards the integration and complementarity of official data with non-official data to meet the increasing users' data needs.

## **5. Quality challenges for statistics beyond GDP**

Among e-Frame policy outcomes, the *Roadmap for future research needs* ([www.eframeproject.eu/index.php?id=28&tx\\_wfqbe\\_pi1%5Bshowpage%5D%5B1%5D=2](http://www.eframeproject.eu/index.php?id=28&tx_wfqbe_pi1%5Bshowpage%5D%5B1%5D=2)) sets out next steps for moving forward on “GDP and beyond” agenda. The *Roadmap* identifies research needs and information gaps identified during the project activities. It suggests areas for further investments in order to empower the development, dissemination, understanding and use of statistics on well-being and sustainability. As requested by the EU Commission, the *Roadmap* also proposes topics for future research projects in the context of Horizon 2020 as well as future EU projects according to the needs of the European Statistical System (ESS).

In particular, the *Roadmap* identifies research needs and quality concerns for statistics beyond GDP in face of NSIs with regard to official statistics as well as challenges related to the frontier

---

<sup>4</sup> The Web-COSI advisory board is coordinated by Istat. It is chaired by Enrico Giovannini, professor in Economic Statistics, Department of Economics and Finance, University of Tor Vergata, Rome, former Italian Minister for Labor and Social Policies.

<sup>5</sup> CAPS projects are selected under the first call ICT 10 of objective 5.5. CAPS – Collective Awareness Platforms for Sustainability and Social Innovation and objective 1.7b – Internet Science (<http://ec.europa.eu/digital-agenda/en/caps-projects>)

represented by the availability of non-official data (e.g. big data, open data, crowd-sourced data) and their possible use in an official context.

The ESS relies on a sound framework represented by the EU Code of Practice [4], on the one side and by the recommendations of the Sponsorship Group on Measuring Progress, Well-being and Sustainable Development [5], endorsing the Stiglitz report [9], on the other one. The latter set out the goals and timetable for the development of EU harmonized system of measurements on well-being and sustainability. Thus, official statistics represents the main source of reliable information on well-being, progress and sustainable development both at national and European level. Such a background, including data availability, contributed to a well-established agenda on well-being at European level.

Nevertheless, as identified in the *Roadmap*, there is room for improvements alongside the different quality dimensions.

Generally speaking, quality issues related to official statistics on “GDP and beyond” mainly concern further developments to improve data availability, hence relevance, timeliness, coherence and comparability. Research investments are deemed for assessing accuracy when aggregate or composite indicators are used. A fully understanding of how the multi-dimensional nature of well-being might affect the quality framework and assessment requires an overall reflection, besides investments on specific subject matter areas or indicators.

From a users’ perspective and to foster policy use of well-being indicators, the key requirement is to improve data timeliness at the different geographic levels (local, national and global), as it clearly arose from e-Frame outcomes.

Relevance would benefits from enlargements of statistical production by covering additional topics (e.g. sustainability indicators), by expanding subjective measurements to complement objective information and by increasing the geographical detail and the disaggregation for target groups of interest.

Further enhancements for comparability are also necessary, despite the efforts done so far at EU level. Indeed, there are still great differences among the EU framework and the national systems adopted by Member States to measure well-being (e.g. dimensions, indicators). Comparability becomes even more important when measuring well-being at the local level, where locally collected data are extremely important but very often very specific to the territorial area. Whilst it is not feasible to reach an overall comparability at the different levels, it would be highly beneficial for research purposes to provide metadata allowing for a proper use of data.

Data accessibility does not represent a specific challenge for beyond GDP statistics but could take advantage by exploiting ICT tools and technologies as discussed in section 2.



Additionally, NSIs are challenged by the opportunity that non-official data sources provide to enhance statistical information on “GDP and beyond” at least with regard to timeliness and to data detail (e.g. geographical/target groups of interest). However, their use in an official context is not straightforward. Actually, many demanding issues are raised such as quality assessment, labelling, rules and governance, for which no recipes are available yet but there are animating an on-going debate at EU level involving national and international organisations.

The challenges ahead official statistics are quite demanding. as experienced in the EU projects (e.g. FP or ESSnet), it is vital to collaborate with the statistical community to join research efforts, to share good or best practices, to harmonise findings for a better understanding of quality of life through enhanced statistics and a consistent use of indicators.

## **6. Some final remarks**

The usage of non-official statistics and new sources of data is not “the solution” for the definition of better statistics beyond GDP but it is surely an opportunity to catch to foster the data related to the complex phenomena of the measurement of well-being, societal progress and sustainability proposing an integrated vision for addressing the unmet “user” needs.

The ESS has to set up strategies in order to plan and program at mid-term the usage of non-official data to integrate and complement progress statistics. Aside the crowd-sourced platforms, other liquid information and data available on line are there to be exploited. The challenge for the usage of Big Data and Open Data is becoming urgent for the EU NSIs. On this side many initiatives are currently growing up and many NSIs are carrying out pilot projects in partnership with other branches of government, with academia and with private data providers.

Besides the methodological, technological and quality issues, the usage on new sources of data brings crucial issues on various aspects. Among these: legislative aspects for the access and use of data; privacy aspects managing public trust and acceptance of data re-use; financial aspects to evaluate the costs of sourcing data vs. benefits; management aspects to set up policies and directives about the management and protection of the data; training aspects about the new skills needed.

At last but not at least the risk assessment of the use of non-official liquid data has to be strictly considered (human resources, stability of the sources,...).

The potentiality of this new sources of data is recognized at the high level of the research European world and within the ESS.

In particular, on the thematic of “Big Data for Official Statistics” since 2012, high level initiatives coordinated by the UNECE Secretariat are taking place with the collaboration of national

and international experts. In March 2013, it was clearly stated that "it is unlikely that National Statistical offices will lose the official statistics trademark, but they could slowly lose their reputation and relevance unless they get on board" of the Big Data era<sup>6</sup>.

To move concretely for the exploiting of the new sources of data it is necessary to have a large consensus at EU political level to make it clear the importance of the value added for Official Statistics. It is necessary to invest in research and human capital to move towards technological innovations getting fully involved in the Internet age.

## References:

- [1] European Commission (2013). FP7 Cooperation Work Programme 2013: Information and Communication Technology. Available at: <http://cordis.europa.eu/fp7/ict/docs/ict-wp2013-10-7-2013-with-cover-issn.pdf>
- [2] European Commission (2010a). A Digital Agenda for Europe. COM (2010) 245 final/2. Available at: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0245:FIN:EN:PDF>.
- [3] European Commission (2010b), "EUROPE 2020, A strategy for smart, sustainable and inclusive growth". Available at: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:EN:PDF>
- [4] European Statistical System (ESS) (2011a). Code of Practice (revised edition). Available at: [http://epp.eurostat.ec.europa.eu/cache/ITY\\_OFFPUB/KS-32-11-955/EN/KS-32-11-955-EN.PDF](http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-32-11-955/EN/KS-32-11-955-EN.PDF)
- [5] European Statistical System (ESS) (2011b). Final Report on the Sponsorship Group on Measuring Progress, Well-being and Sustainable Development. Available at : [http://epp.eurostat.ec.europa.eu/portal/page/portal/pgp\\_ess/0\\_DOCS/estat/SpG\\_Final\\_report\\_Progress\\_wellbeing\\_and\\_sustainable\\_deve.pdf](http://epp.eurostat.ec.europa.eu/portal/page/portal/pgp_ess/0_DOCS/estat/SpG_Final_report_Progress_wellbeing_and_sustainable_deve.pdf)
- [6] Giovannini E. (2010). "Statistica 2.0: the next level", Introductory speech at the 10th National Conference of Statistics, Rome. Available at: [http://www3.istat.it/dati/catalogo/20120621\\_00/atti\\_decima\\_conferenza\\_nazionale\\_statistica.pdf](http://www3.istat.it/dati/catalogo/20120621_00/atti_decima_conferenza_nazionale_statistica.pdf)
- [7] Organisation for Economic Co-operation and Development (OECD) (2012a). Better Life Initiative: Measuring Well-Being and Progress. Available at: <http://www.oecd.org/statistics/betterlifeinitiativemeasuringwell-beingandprogress.htm>
- [8] Organisation for Economic Co-operation and Development (OECD) (2012b). European Conference on Measuring Well-Being and Fostering the Progress of Societies, 26-28 June 2012- Paris. Available at: <http://www.oecd.org/site/progresseurope/>
- [9] Stiglitz et al. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. Available at: [http://www.stiglitz-sen-fitoussi.fr/documents/rapport\\_anglais.pdf](http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf).

---

<sup>6</sup> "What does Big Data mean for official statistics, 8 March 2013, Task Team, High-Level Group for the Modernisation of Statistical Production and Services (HLG)  
<http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=77170614>