

New trends in the dissemination of statistical services

Abstract

This presentation aims to illustrate the innovations introduced in the field of statistics dissemination to meet users' information needs on various economic and social phenomena.

People deal with statistics every day, even without realizing it, so it is increasingly important to make them statistically literate.

Official statistical information offers a fundamental support to public decisions, both for ex-ante choices and ex-post evaluations, and in general provides answers to all its users.

The usual tools used to this end (one among all: the websites of the public organizations officially charged with the production and dissemination of statistics) alone are no longer sufficient.

Times change and so do the users of statistical information, so the tools to reach them must be adapted.

(SLIDE 5 – SLIDE 6 – SLIDE 7) Introduction

Increasingly, organizations and individuals are recognizing the importance of using statistical results to make evidence-based decisions.

At the same time, social media and self-publishing have made great strides over the past decade, and the general shift to digital has transformed the playing field.

Content publishers are therefore facing a growing set of challenges: disseminating high-quality content using faster, more accurate, and more agile processes.

To do this, it is important to use the right tools.

Consumers of statistical information spend less time on a single website than in the past.

These days, users move from platform to platform and digital site to digital site to satisfy their thirst for information and easily share content with friends and family. That's why it is important to use new technologies so as not to decrease your chances of reaching your target audience.

But despite consumers spending less time on a given website, people still spend less than 7 hours a day interacting with Internet (January 2022 source we are social <https://wearesocial.com/it/blog/2022/01/digital-2022/>). This a lot of time that people spend listening, reading, watching and generally interacting with digital media.

Therefore, it is imperative to communicate effectively to achieve three important dissemination goals:

- **inform** the public about the latest releases of official statistics and reports on the social, economic and general conditions of the various Countries of Interest
- demonstrate the **relevance of statistical information** to both individuals, organizations and businesses in the public and private sectors to more effectively support decision-making in every domain of society
- **increase public awareness** and their support for statistical programs and services.

(SLIDE 8) What is driving these innovations?

- the richness of information provided by Big Data
- increasingly advanced software and technologies
- innovative methods of disseminating statistical information (Open Data, Linked Open data)
- web publishing products realized with technologies that simplify the

interaction with the reader, guaranteeing quality and innovative optics while offering information panels able to present phenomena according to different points of view and linking economic, social and environmental aspects, adding at the meantime value to data in terms of representation and analysis

(SLIDE 9) At European level there are six shared principles regarding the dissemination of statistical information. They are:

Relevance

Information should be relevant to the social, economic and general conditions of the Country and meet the needs of public and private decision makers. For the media, relevance translates into newsworthiness. However, care must be taken to present information in a way that does not trivialize data or results. The goal is to inform the public about the availability of the data or information. Media coverage is desirable because it broadens the audience of the message and therefore will increase knowledge and stimulate debate among the broader public.

Confidentiality

The confidentiality of individual respondents, whether individuals or businesses, must be protected at all times for all data collected. Any information that identifies an individual or group should not be released without their prior consent. Nor should information be released that compromises the confidentiality of respondents. This restriction applies to the media in the same way it applies to any other respondent.

Independence and objectivity

Information should be presented objectively and impartially and be independent of political control or influence. The Fundamental Principles of Official Statistics

establish the criteria by which independence and objectivity can be judged.

Timeliness

Information should be up-to-date and released as soon as possible after the reporting period. The timeliness of information will influence its relevance.

Accessibility and transparency

All users should have equal access to data and metadata. Information should be available to the public in appropriate formats, through suitable distribution channels and be written in simple, understandable language that is tailored to the level of understanding of key user groups.

It should also be ensured that the media, as well as all other users, are able to access and correctly interpret information about statistical methods, concepts, variables and classifications used to produce statistical results.

Coherence

The use of standard concepts, classifications and target populations promotes consistency and credibility of statistical information, as does the use of a common methodology across surveys (brand awareness).

Moreover, the adherence to the Fundamental Dissemination Principles will increase the credibility of suppliers of official statistics and build public confidence in the reliability of the information provided.

Among other things, the technological revolution of recent decades is also a potential cause of user-side confusion in the search for accurate statistical information recognizable as being from an official source. Relationships have been redefined within the "information society" and established paradigms of communication and dissemination have been irreversibly challenged. Indeed:

- the network has become an interactive environment: everyone can participate in communication processes.
- generation of "digital natives": born from 1990 onwards, growing up with digital technologies such as computers, the Internet, cell phones and MP3s.
- the relationship between producer and user of information becomes bidirectional: accelerated growth in demand for reliable information and data.

In order for quality information to emerge from the "noise" and deluge of data circulating on the web, a clear and well-defined communication and dissemination strategy is hence essential! As is the creation of reusable content.

(SLIDE 10 – SLIDE 11) Users' information needs

It is crucial to always improve the accessibility of existing official statistical data in order to meet a wider range of user needs and encourage the continued use of statistics in their choices.

What are the main categories of users of official statistics?

(SLIDE 12 – SLIDE 13 – SLIDE 14) The key audiences are:

- Researchers and data analyst
- Data journalists
- Politicians
- Public Administrations
- Companies
- Citizens

(SLIDE 15 – SLIDE 16 – SLIDE 17) Each of these categories has its own information needs that are different from the others. For this reason, they need to be provided with dissemination services tailored to their respective requirements to ensure they get the best use of the data they seek. The challenge today is to be able to serve this full range of different user needs, while also taking into account varying degrees of statistical literacy. What is needed are integrated web-based tools that are usable, responsive and accessible, suitable for mobile devices, that can support interactive content, conform to taxonomies and structured semantics of web-design, taking into account web analytics and user research on search engines.

This requires:

- Web services and APIs (application programming interfaces) based on shared standards and open formats
- Data visualization dissemination tools
- Digital editing consistent with innovation in content, formats and production processes, to be released on e-store
- Web services and APIs (application programming interfaces) based on shared standards and open formats
- Data visualization dissemination tools
- Digital editing consistent with innovation in content, formats and production processes, to be released on e-store

(SLIDE 18) Specifically:

USERS	TAILORED SERVICES & PRODUCTS
Heavy data users, data suppliers	Massive download tools, Linked Open Data, Microdata, Statistics in a machine-readable open data format
Researchers, analysts, financial community, data journalists	Microdata, Visualization, Digital Library
Institutions, policy/decision makers, central/local government	Statistical Reports with at regional level, Visualization, Census microdata
Media, old/new	Press releases, Video, Infographics, Training for data journalists, Social media, On line press room, podcast, App
Prosumers, influencers	Apps, Reusable contents to embed and comment, Social media
Large audience, general public, pupils, teachers, citizens	Web publishing, Social media, Infographic, On line press room, podcast, dedicated web section
Statistical producers / statistical community, Internal staff	E-learning system, Sharing platforms, Web seminars

(SLIDE 19) [Examples of dissemination services](#)

(FROM SLIDE 20 TO SLIDE 32 INCLUDED) Web sections, apps, videos, infographics, databases, data visualization, social media: all of these tools are also useful to fight misinformation and disinformation which today are viral on the network as much as the contents of proven official source.

(SLIDE 33) [Open data and Linked open data](#)

(SLIDE 34) Open data: Tim Berners-Lee, director of the World Wide Web Consortium (W3C), has ranked data formats on a five-star scale: one star corresponds to the minimum openness, five stars to the maximum:

*Unstructured data encoded in a proprietary format (examples: a pdf file; a jpeg image);

**Structured data but encoded in a proprietary format (thus easy enough to be processed by a computer application);

***Data structured in a non-proprietary format (for example, CSV format, which can be opened by any software);

****Data structured and encoded in a non-proprietary format and provided with a unique resource identifier (URI). An example is the RDF standard: it applies a shared meaning to data ("that data has the same meaning in any language, for any Country");

*****Open data linked to other open data sets (Linked data).

On the Berners-Lee scale, **DATA can be considered OPEN if they have at least three stars.**

Open data represent a discipline with a huge potential for (re)dissemination of official statistics.

(SLIDE 35) You can find European open data on the official European data portal <https://data.europa.eu/en>, here 170 catalogs from 36 different countries and 1.399.388 datasets are collected. The portal is in 24 languages.

(SLIDE 36 – SLIDE 37) LINKED DATA are structured data that are interconnected with other data in a way that becomes more useful through semantic queries. It builds on standard Web technologies such as HTTP, RDF, and URIs, but rather than using them to serve Web pages only to human readers, it extends them to share information in a way that can be read automatically by computers. Part of the vision of linked data is for the Internet to become a global database.

Linked data can also be open data, in which case they are usually described as **linked open data (LOD)**.

Once again Tim Berners-Lee outlined four principles to define linked data:

- Uniform Resource Identifiers (URIs) should be used to name and identify individual things.
- HTTP URIs should be used to allow these things to be searched, interpreted, and subsequently "dereferenced".
- Useful information about what a name identifies should be provided through open standards such as RDF, SPARQL, etc. When publishing data on the Web, other things should be referred to using their HTTP URI-based names.

(SLIDE 38) LOD have many benefits: improved interoperability, better access to data and metadata, the ability to handle multiple formats, and standardization.

(SLIDE 39) Eurostat disseminates the Nomenclature of territorial units for statistics (NUTS) classification in LOD format (<https://ec.europa.eu/eurostat/web/nuts/linked-open-data>). NUTS is a hierarchical system of subdivision of the economic territory of the EU for the purpose of:

- the collection, development and harmonization of European regional statistics;
- socio-economic analyses of the Regions;
- the framing of EU Regional policies (e.g. cohesion policy).

The system is based on population thresholds and encourages administrative division. The legal acts behind it began with Commission Regulation 1059/2003 specifying the stability of the classification for at least three years.

(SLIDE 40) Big data

We have just analyzed the characteristics of Open data, which are precisely public, open and interoperable data, but even if they represent an important volume that will grow continuously, they still represent a small part of data if we consider the more generic cauldron of **big data**. Big data also contains unstructured or semi-structured data.

Big data is a collection of data that is so extensive in terms of volume, velocity, and variety that it requires specific technologies and analysis methods to extract additional information value from it.

The progressive increase in dataset size is related to the need to simultaneously analyze a single dataset to extract additional information than could be obtained by analyzing partial and separate datasets, even if they cumulatively contain the same total amount of data.

An example can be the analysis conducted to probe the "mood" of markets and consumers: they require the identification of the main trends of society in a given historical moment and are obtained through the analysis and overcoming of the deluge of information that travels and transits through the Internet.

Big data represent, therefore, the interrelation of data potentially coming from heterogeneous sources: thus, not only structured data, such as those present in databases, but also unstructured data, such as images, emails, geo-localized or geo-referenced data, information taken from social networks or from downloads even if partial and not completed, etc.

(SLIDE 41) Even the experimental statistics use new data sources (Big data) and new methods in an effort to better meet the needs of our users in a timely manner.

Two examples of Big data-based Experimental Statistics are:

- Eurostat uses Wikipedia as a new source to produce statistics on visits to UNESCO World Heritage sites
- Italian statistical Institute uses Twitter (about 57,000 tweets per day) to elaborate the social mood on economy index

(SLIDE 42) The classification developed by the Unece Task Team on Big Data, in June 2013.

(SLIDE 43) **Open data** are certainly a subset of **Big data**, but the purposes and uses that characterize these two classifications of data are profoundly different:

- **Big data** are collected even without the knowledge of the person concerned to profile the tastes and trends of citizens and are used for private purposes and market analysis;
- **Open data** are data collected as part of the actions of Public Administrations, must be available, reusable and are made available to the community to encourage participation in the management of public affairs.

(SLIDE 44) **Web publishing**

And what about web publishing?

(SLIDE 45) Web publishing is the process of publishing original content on the Internet.

The process includes building and uploading websites, updating associated web pages, and publishing content to these web pages online. Web publishing includes personal, business, and community websites, as well as e-books and blogs.

Posting updates on social media such as Twitter, Facebook, etc. is not considered web publishing.

(SLIDE 46) Content intended for web publishing can include text, video, digital images and other forms of media.

(SLIDE 47) To create a web publishing product you must have:

- a website development software
- an Internet connection
- a web server to host the website

The website development software may be a professional web design application such as Dreamweaver or a simple web-based content management system such as WordPress or Drupal.

Larger sites may use a dedicated server to host them; however, many smaller sites usually reside on shared servers that host a variety of websites.

(SLIDE 48) Benefits

Since web publishing does not require physical materials such as ink and paper, **it costs virtually nothing to publish content.**

Thus, anyone who meets the above three requirements can become a web publisher. In addition, web publishing brings in countless visitors, as published content can be browsed by anyone who is authorized to access it. These benefits of web publishing have opened up a new era of personal publishing that was previously unimaginable.

E-book and blog publishers use almost the same web publishing tools used by website developers. People who do not have the necessary skills for web publishing seek the services of professional web publishing people or organizations to host, maintain, and

edit their websites, e-books, and blogs.

(FROM SLIDE 49 TO SLIDE 53 INCLUDED) Web publishing examples

We can find many examples of web publications on Eurostat
<https://ec.europa.eu/eurostat/web/main/publications/interactive-publications>

The European economy since the start of the millenium — a statistical portrait - 2021 edition - This interactive publication aims to show how main features of the economy of the European Union and its Member States have evolved since 2000 through a large range of statistical data giving both a micro- and a macro-economic perspective.

Throughout the publication, brief descriptions of the main findings are completed with interactive visualisations, where you can compare your country with other countries. This publication is aimed at those who would like to have an overview of the economic trends in the EU since 2000 in an easy and interactive way.

The lives of women and men in Europe - 2021 interactive edition - Another Eurostat publication is The Lives of Women and Men in Europe, which shows the differences and similarities between the lives of women and men in Europe. The latest edition of the interactive publication "The Lives of Women and Men in Europe" aims to compare women and men in their daily lives and also shows how similar or different the daily lives of women and men are in European countries. This interactive publication contains short texts, interactive visualization tools, infographics, etc. It was developed by Eurostat in collaboration with the national statistical institutes of the EU member states and EFTA countries.

Another type of Eurostat web publication is Statistics Explained.

Statistics Explained is an official Eurostat website that presents statistical topics in an easily understandable way.

Together, the articles form an encyclopedia of European statistics for everyone, supplemented by a statistical glossary that clarifies all terms used and numerous links to further information and the latest data and metadata, a portal for casual and regular users.

An example of Statistics Explained is Eurostat's Regional Yearbook; it is an online publication from Eurostat that can also be downloaded in PDF format and is available in print. In the web release version, however, the articles are updated or replaced once a year. Also, in the web version all maps can be explored interactively using Eurostat's statistical atlas.

(SLIDE 54) Tools

To date, there are many web publishing tools available.

Internet publishing is such a broad field that the tools developed for it meet a wide range of needs. Some web publishing tools are marketing and design oriented, while others are more interested in news publishing.

But how to select the best web publishing tool?

(SLIDE 55) Here are a few things that can make a web publishing tool the best:

1. SEO optimization options
2. Web analysis
3. Offline support
4. Subscription-based web publishing models

SEO-friendly

There is a lot of competition and you will need a technique to attract readers and stand out from the crowd. Your material should be optimized for search engines when it comes to digital publishing, so that it ranks high for a user-related query.

Web analysis

It is crucial to monitor the performance of your content and publishing platform if you want everything to work properly. Analytics can help to track user engagement, share status, downloads and other metrics so one can improve outreach strategies and get the most out of one's content.

Offline support

Many people prefer to save a piece of material and read it later. By taking advantage of your publishing platform, users should be able to download and access content offline.

Subscription-based web publishing models

Publishers often have a variety of material to publish. Some content is available for free, while other content may require readers to subscribe. For publishers who want to use subscription-based models, the platform they choose must obviously support them.

(FROM SLIDE 56 TO SLIDE 58 INCLUDED) Conclusions

We have seen and reiterated the importance of adapting one's data dissemination strategy to the main new IT tools that are available today

To re-cap, what are the main steps to follow?

- identify your users
- work to meet the needs and expectations of your users

- provide quality data and present your information in a clear and user-friendly way, in accordance with the statistical dissemination principles shared at European level
- finally, choose among the various dissemination tools available those that best suit our needs and those of our users

And remember:

‘Only statistics that are used are useful statistics’ (Petteri Bae)