

# An Open Data platform for decision making in local public administration

Giuseppe Sindoni, Matteo Massenzio

## 1. Introduction

This paper presents the Milan Open Data (OD) platform as a means of providing statistics and data in the framework of “Data-Driven Milan”, a city where policy decisions are taken in an “informed and aware” way using data. Such a strategic approach is enabled by the enormous amount of data available to public administrations. This includes not just the well-known big data, but also all the data automatically produced by digitalized systems, such as citizen relations systems or systems issuing permits for occupation of public areas. The former can be analysed in real time to understand citizens’ needs and adjust service development policies accordingly, while the latter, integrated with maps of the city, enable every event to be kept under control and any clashes between events of a different nature to be managed more efficiently.

Data-Driven Milan has been implemented since 2016 through a data exploitation strategy aimed at developing a digital platform system to collect and safely integrate data for use in analysis reports, dashboards and geographic intelligence applications and to publish easily accessible open data to share knowledge with citizens.

OD are ever more important in providing citizen communities with useful information. Over the last 10 years, the municipality of Milan has developed its OD platform from an experimental portal to a fully-fledged portal with more than 1,600 datasets, implemented a Linked Open Data (LOD) system and 8 advanced data visualization projects, and produced OD policies and operating guidelines.

## 2. The open data portal for Open Government Data

According to the Organisation for Economic Co-operation and Development, “Open Government Data (OGD) is a philosophy - and increasingly a set of policies - that promotes transparency, accountability, and value creation by making government data available to all.” [OECD, 2020]. OGD is about using public data to enforce the transparency of public administration, which generates trust and in turn improves citizen participation and collaboration between public and private organizations.

Citizen participation is about getting feedback, suggestions, ideas and help through public debates on the development of public policies. Collaboration must be implemented by tearing down watertight compartments and hierarchical structures inside and between organizations, by working “horizontally” and locally between organizations with service design tools and flexible methods, and through the involvement of citizens and promotion of cooperation.

In this context, data can help to enforce transparency through the monitoring of public policies, for example through data-based communication strategies and impact indicators, and through citizen education, using advanced data visualization and explaining the governance process with data and infographics.

Figure 2.1 shows how Milan’s OD strategy has developed from the launch of the portal 10 years ago to the publication of the first Report on the Council’s results, as well as its constantly

increasing number of datasets.

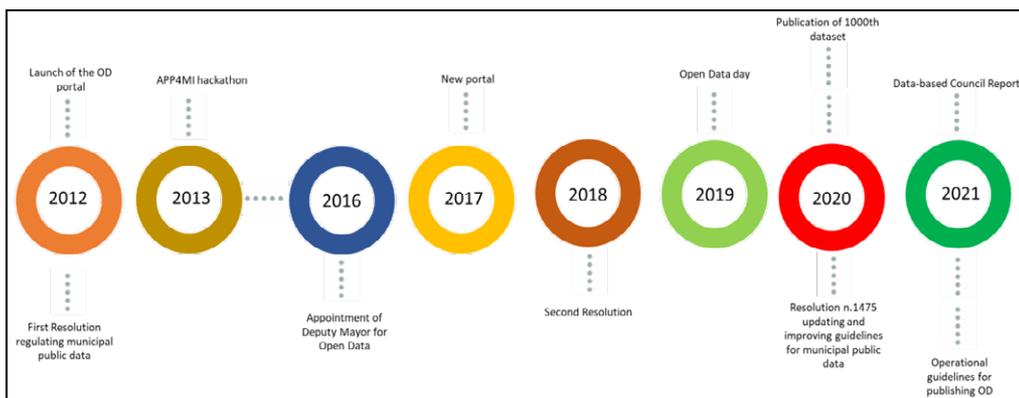


Fig. 2.1 – The history of the Milan Open Data strategy.

The constant rise in the number of published datasets<sup>1</sup> is depicted in Figure 2.2. This growth is due to the publishing strategy, by which all datasets representing evolving phenomena are updated whenever a new version is available, and new datasets are created from internal sources or collected from external sources.

The push for the qualitative and quantitative improvement of the municipality of Milan’s public information assets in open format arises from the provisions of European legislation and the digital administration code (so-called CAD [CAD, 2005]) as well as, since 2012, from a series of municipal council resolutions regulating the open data sector.

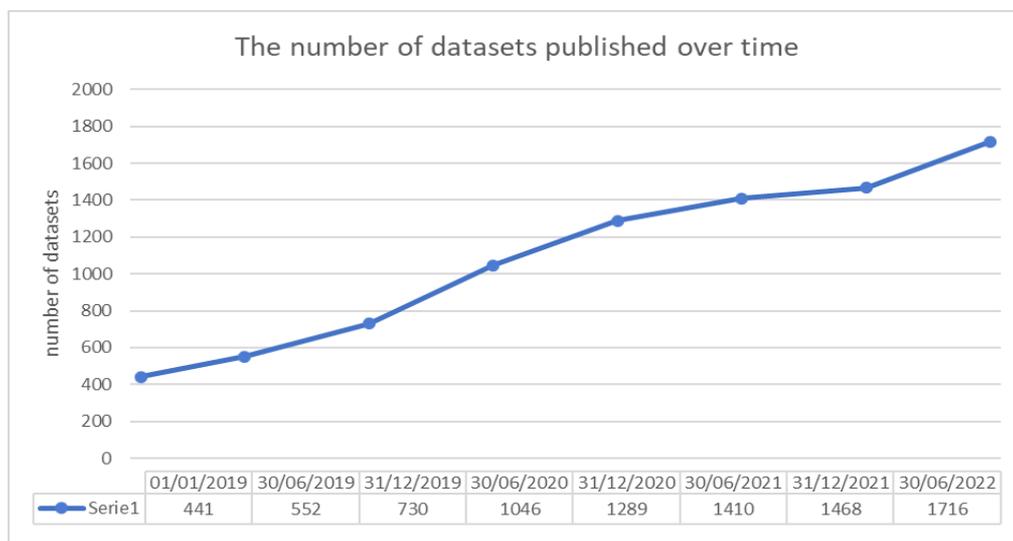


Fig. 2.2 – Increase over time in published datasets.

The increasing number of datasets produced and maintained by Milan has seen it become a national leader in Open Data (winner of the ICity rank editions 2020 and 2021 [ICity rank 2020, 2021]) and place it on a par with major international cities such as London (data.london.gov.uk): 1817 published datasets; Paris (opendata.paris.fr): 335; and New York (data.cityofnewyork.us): 3589.

<sup>1</sup> The Open Data portal is available at: [dati.comune.milano.it](http://dati.comune.milano.it)

The datasets cover the themes of the DCAT-AP-IT (DCAT, 2022) metadata profile. Figure 2.3 shows how they are distributed across the themes.

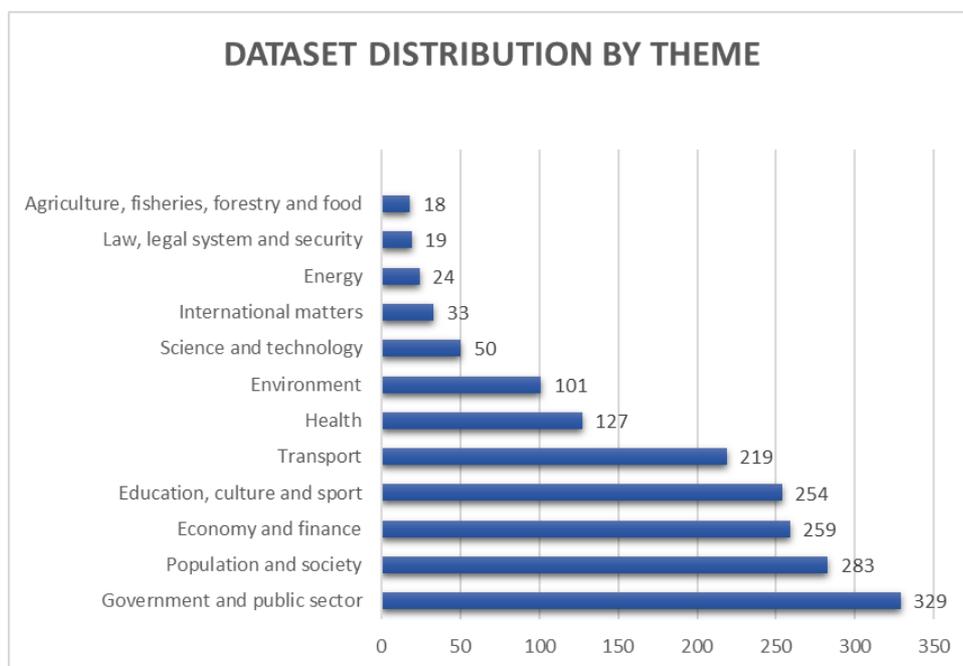


Fig. 2.3 – Distribution of the datasets by theme.

The highly biased distribution is due to the fact that most datasets come from internal sources - for the main part, the digital systems supporting the administration's processes and services. The distribution hence reflects the core themes of the various services.

The portal is based on CKAN technology, which makes datasets available via both download and Application Programming Interfaces. It currently has about 9,500 visits per month. Data are also published as tables on the statistical portal and maps on the geo-portal.

### 3. Linked Open Data

Linked Open Data are semantically enriched machine-readable data that help data interoperability between distributed systems. The international community classifies OD on a 5-star scale based on 3 characteristics: information, access and services. The stars represent an increasing level of usability and accessibility, with 5 stars awarded to the most valuable data: Linked Open Data, which enables both human and automated access to data.

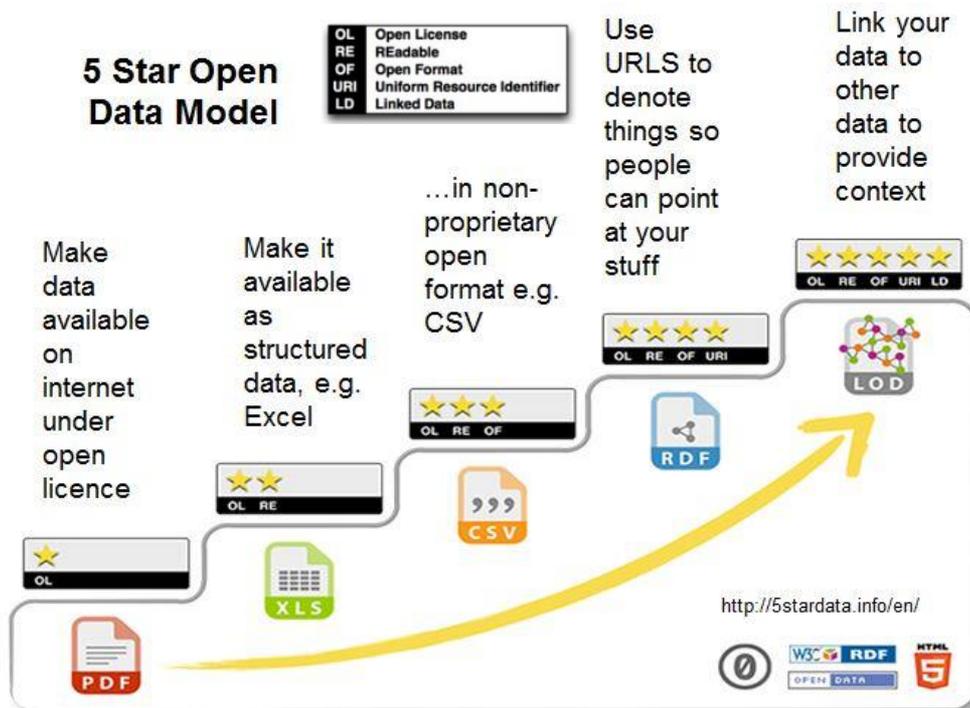


Fig.3.1 – The five stars of Open Data.

LOD are semantically enriched and interlinked, so they enable the development of very efficient data services based on data mashups, where datasets can be used machine-to-machine with automatic integration made possible by their semantic representation through ontologies. The Milan LOD platform<sup>2</sup> is based on 6 ontologies allowing semantic access to the datasets available for the topics covered by each ontology: libraries, public acts, schools, consumer prices, limited traffic zones and sports facilities.

The working model for the design of the ontologies and the implementation of LOD is based on three phases: ontology design, data census and preparation, data loading and graph generation. An ongoing LOD automation project aims to improve the current system by minimizing manual operations in the dataset lifecycle.

#### 4. Data visualization projects

Data visualization projects are part of Milan’s strategy for “data democratization”, i.e. making data accessible and usable to the greatest possible audience. This includes people without specific data manipulation skills who just need objective, easy-to-understand information about the council’s activities and performance. In addition to LOD, seven more special projects have been carried out in the last 4 years to better exploit the open data assets for the benefit of citizens.

<sup>2</sup> The LOD platform is accessible from the OD portal or directly at: [dati.comune.milano.it/sparql/home.html](http://dati.comune.milano.it/sparql/home.html)

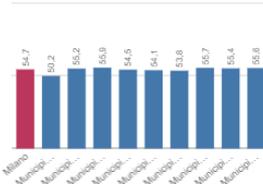


Fig. 4.1 –Openbilancio - current expenses

## Politica e Istituzioni

Partecipazione elettorale (primo turno elezioni comunali 2016)

per 100 aventi diritto



Percentuale di persone che hanno votato al primo turno alle elezioni comunali sul totale degli aventi diritto.

% Donne negli organi decisionali

41,7<sup>2016</sup>  
Anno

Percentuale di donne assessori comunali sul totale degli assessori.

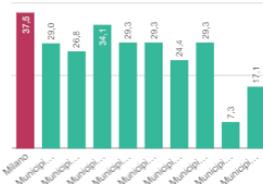
% Donne e rappresentanza pol...

37,5<sup>2016</sup>  
Anno

Percentuale di donne elette nei Consigli Comunali sul totale degli eletti.

Donne e rappresentanza politica a livello locale (consiglio comunale 2016)

per 100 eletti



Percentuale di donne elette nei Consigli Comunali sul totale degli eletti.

Età media negli organi decisionali

45,2<sup>2016</sup>  
Anno

Età media degli assessori comunali alla data delle elezioni

Età media consiglieri comunali

46,9<sup>2016</sup>  
Anno

Età media dei consiglieri comunali alla data delle elezioni

Elaborazione dell'Unità Statistica su fonte dati: Comune di Milano

Fig.4.2 Local analysis of the BES index (sustainable equitable well-being)

The most important data visualization projects are Open Budget and the Council's Mandate Report. Open Budget is an advanced project for the publication of both the final balance and anticipatory municipal budget data. It provides a very advanced user experience and, from this perspective, can be seen as a true data democratization effort. The site was made public in 2018 and is based on data from the Management Executive Plan, published as open datasets since 2013. It has been continuously improved ever since, to provide better usability and more data

transparency.

The Council's Mandate Report is another data democratization project aimed at reporting, through data, the results achieved by the Council during its 2016-2020 term of office (a 2021 update is ongoing). The web site complements the traditional document-based report and offers readers a quantitative view of the Council's performance.

Milan's open data is also widely used by various socioeconomic operators to create their own applications. From this point of view, the Municipality tries to anticipate the needs of stakeholders right from the "Demand" stage, by carrying out various thematic meetings. These meetings have shown that the major users are universities, companies and citizens, who use the data to better direct their choices.

Since 2018, the municipality of Milan has constantly monitored and published information on individual accesses to each dataset:

[https://dati.comune.milano.it/dataset/ds916\\_accessi\\_unici\\_ai\\_dataset](https://dati.comune.milano.it/dataset/ds916_accessi_unici_ai_dataset)

Political decision-makers also make extensive use of open data.

## 5. Next Steps

As part of the broader project to create a data-driven administration, the Municipality of Milan intends to continue strengthening the Opendata Portal. The cornerstone of this approach will be the creation of datasets based on ontologies and glossaries in order to develop an increasing number of high-quality datasets that can be easily made available as Linked Open Data.

## References

OECD (2020). Open Government Data <https://www.oecd.org/gov/digital-government/open-government-data.htm> Web page

AGID (2022). *DCAT-AP\_IT v1.1 – Profilo italiano di DCAT-AP*. <https://dati.gov.it/content/dcat-ap-it-v10-profilo-italiano-dcat-ap-0>. Web page

ICity rank (2021). <https://www.forumpa.it/citta-territori/icity-rank-2021-firenze-milano-e-bologna-le-citta-piu-digitali-ditalia/> Web page

ICity rank (2022). <https://www.forumpa.it/citta-territori/icity-rank-2020-firenze-bologna-e-milano-sono-le-citta-piu-digitali-ditalia/> Web page

CAD (2005) Law decree 7 March 2005, n. 82