RUNNING UP THAT HILL
Italian civic addresses and the quest for accuracy

By Valerio Zunino

The quality of toponymy data has come increasingly under the microscope in recent years in direct correlation to the surge in the number and functional versatility of new products available to the geographical information market. These augmentations infiltrating the sector require extremely precise and comprehensive databases, most crucially in respect of, for example: the automatic driving project, 118 emergency services, applications inherent in civil protection, and other civil service sectors responsible for urban and environmental intervention.

STUDIO SIT Srl was conceived of in 1991, the dim and distant past for many, when the few precious and elusive resources available were employed to ask the most talented hackers to find us the formidable MapInfo, the former, and in all likelihood best desktop GIS software ever. At the time a version had just been released for MS Windows.
Ever since then the company has co-operated with pertinent networks and participated in the development of flagship products released by the different market players; initially concentrating on the world of Autodesk and then expanding its focus to include the production of applications for the benefit of local authorities and others dedicated to the territorial survey activities on civic addresses (from 1998).

In 2007/8 two aircraft were acquired in the service of aerial mapping, both endowed with autonymous equipment and restitution software. In 2010 one of the first small-to-medium-sized drones available to civilians was purchased and subsequently developed - this having been produced in Germany and adapted to the photogrammetric instrumentation already available.
The same year saw the resumption of the surveying and mapping of civic numbers, so as not to be aban-doned. At that time the company's software development activity concentrated first on the Environment Map Server, then QGIS, and finally the world of ESRI - a culmination of nearly thirty years of development and experience.
For suppliers such as HERE Global Content B.V. and

TOMTOM Global content B.V. involved with the surveying and mapping of Italian civic numbering, the Ligurian company has been able meet current requirements for very high-level accuracy, for the positioning of points and for the entirety of national built territory. These needs derive fundamentally from two causal factors, manifested gradually over the last two decades – the catalyst being predominantly the evolution of the toponymy data market and mobile communications sector, which has seen its scope of application expand from the initial world of car navigation, fleet management and market research, into the Multiutilities (the planning and management of utilities) i.e. emergency services, politics and the electorate. These latter areas of interest demand the availability of a complete and precise geo-referenced civic numbering database, as well as an enhanced sense of responsibility towards its users, who have over the past few years become much more sophisticated, finally reaching a necessary qualitative standard, inexistent until a few years ago and therefore not applied to either supply or demand.

It's fair to say that the causal factors that have mobilized the market in question initially were the evolution of a product which is today able to reach more and more people comprising a wide range of end-users, along with the ensuing awareness of the need to finally have geographic data that is much more reliable and complete than in the past.

In order to maintain a high level of satisfaction among its customers, STUDIO SIT has remained true to the initial specialities of its own surveying activities – fieldwork that covers the entirety of our national territory - through a network of professionals that now numbers almost 50, assembled through programs of education and experience pertaining to rigid selection criteria. A network consistently looking to cooperate on integrative interventions with reference to the dynamic regulatory framework, in-depth cartographic data and software applications used in the context of digitalizing, editing and autovalidating operations and data pretesting.

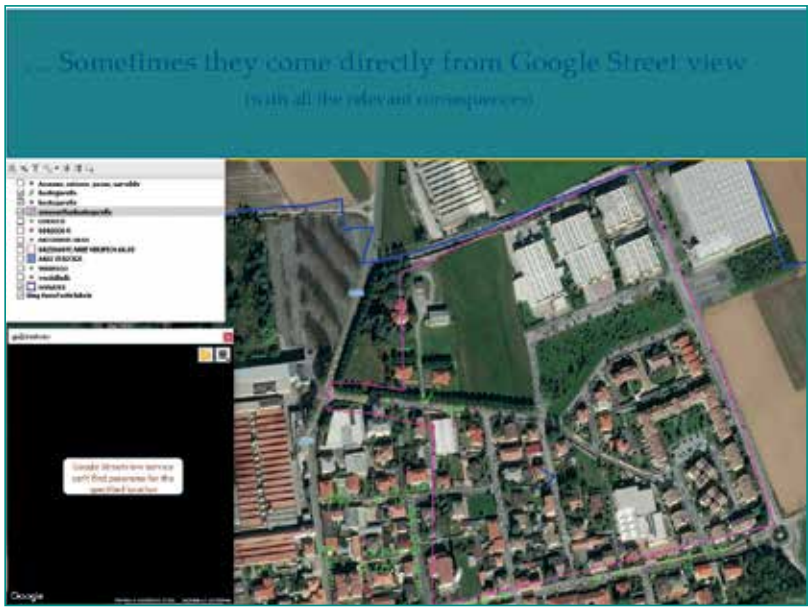
The resulting qualitative system exceeds a 95% degree of accuracy, completeness and high-level updating, a unique case in our country.

These three elements already mentioned combine to apply the same values to the sphere of geographical toponymy. The issue of quality, until a few decades ago passively suffered by the end-users that characterized the demand, cannot be disregarded and has today assumed crucial importance, even for the very survival of some areas of operation.

Stemming from the three qualitative factors above, a comprehensive survey of the whole municipal territory, accuracy in the positioning of the numbers and the degree of updating, where located and verified, generates a greater interest on the part of a wide nucleus of users, who in the past did not deem certain geographical information sufficiently detailed for their own purposes. Among those looking to exploit the mapping of addresses are the management of multiutilities (in the context of the maintenance operations of networks and utilities, as well as those involved in administrative management), professional services firms, car sharing companies,

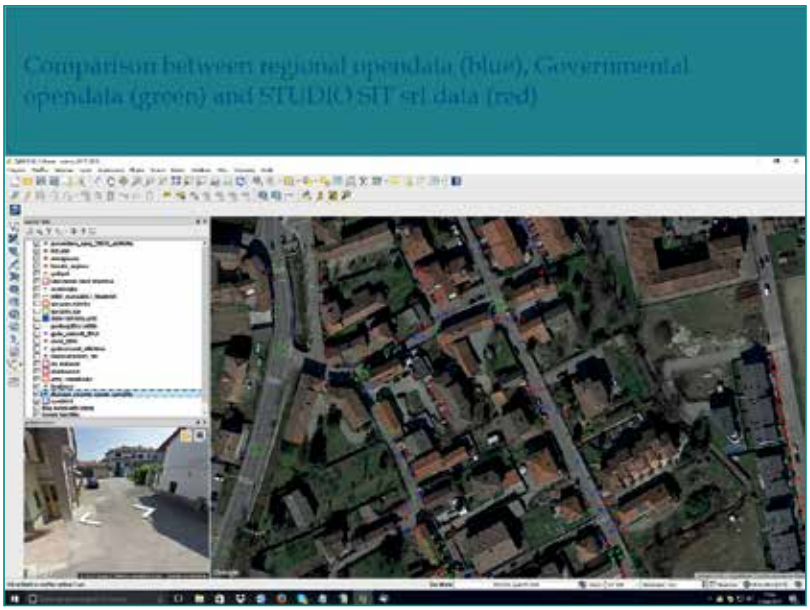
large commercial distribution activities and, furthermore, some departments responsible to the Ministries of the Republic.

Also relevant is the fact that in Italy there is an increasingly sharp contrast between the perceived quality of geographical toponymy data at central government level, compared to that experienced by most sectors of industry, including multinationals and the SME world. This first example, central government, does not anticipate a need for geographical data pertaining to addresses that is comprehensive and geometrically precise. While the second category, industrial players (including their end-users) of car navigators and other mobile mapping application, up to the most cutting-edge apps employed by those in creative fields, insist on yet higher quality in the form of accuracy, completeness and maximum updating of information.



Public Administration	territorial coverage	quality	average updating
Ministry of Environment	95-98%	very low	1999
Aosta Valley Region	NO	//	//
Piedmont Region	20-30%	very low	2010-2012
Liguria Region	NO	//	//
Lombardy Region	95-98%	low	2010-2014
Veneto Region	20-25%	n.a.	n.a.
Friuli Venezia Giulia Region	94-96%	quite good	2013-2014
South Tyrol Province	95-98%	good – very good	2010-2011
Trento Province	3-5%	n.d.	n.d.
Emilia-Romagna Region	80-85%	quite good	2011-2016
Tuscany Region	88-92%	low	2012-2016
Le Marche Region	NO	//	//
Umbria Region	90-92%	quite good	n.d.
Lazio Region	NO	//	//
The Abruzzi Region	NO	//	//
Molise Region	NO	//	//
Campania Region	NO	//	//
Basilicata Region	NO	//	//
Apulia Region	NO	//	//
Calabria Region	NO	//	//
Sardinia Region	20-25%	low	n.d.
Sicily Region	80-85%	low	2005-2007
Municipalities	< 1%	variable	variable

At the forefront are the major players involved in car navigation who, along with Google, currently have a stranglehold on the market due to a competitive advantage accumulated in now well over two decades of data collecting and acquisition; these majors being involved in perpetual competition to adapt the quality of their data, augmenting its reliability and completeness, often to the point of rebuilding their geographic databases from scratch. All this is driven by the demand from the multiutilities, which have identified the need to maintain management services using GIS, which are based on the civic addresses in their territories of interest. At the same time, some smaller companies have specialized in the field of geographical toponymy data, making their services widely available through distribution to a clientele that ranges from transport and logistics, modern geomarketing, civil protection and emergency services. In all these areas of demand, there is now a need to have all civic numbers geo-referenced in the correct position, given that this information plays a fundamental and decisive role, for example, even the saving of human lives.



The market's drive towards improving the quality of the data is fuelled by two imperatives: one coming from a product where each frequent new version renders the previous version obsolete, and one from the demand by increasingly more discerning users - an awareness that requires a higher and higher quality level, composed of accurate, up-to-date and comprehensive data to ensure the development and maintenance of modern, toponymy-based, mobility applications.

As regards the availability of open data, some Italian regions and large cities offer a good quality open data service, which also includes civic numbering. However, these are rare exceptions since the process of surveying and mapping civic numbering is hampered by the fact that it has often been based within some regional procure-

ment procedures awarding these contracts to construct topographic databases, which were containing a preponderance of cartographic details and a paucity of requirements for qualitative mapping components, such as civic addresses.
These same contracting companies are traditionally tied to pure aerial photogrammetry, with no other corresponding competencies and therefore consider/have considered the logical level of civic numbering as an accessory, an appendix of the photogrammetric activity of restitution and cartographic dressing.
This situation allows companies to interpret this meagre specification to their advantage, detailing just the position of the existing:
- without the insertion of the corresponding toponym;
- without the inclusion of buildings which were displaying no number.
Unfortunately it is also necessary to take into account the fact that often the contracting companies produce georeferenced house numbers according to mere mathematical algorithms (or Google Street View...), in which numbers are distributed in an unreal, uniform way, adhering to the vagaries of the road network which is in itself already inaccurate in terms of length and route.
All this inevitably leads to an average quality reduction of around 65% of the value of the geographical civic toponymy data. As a consequence, italian toponymy open data, is largely unusable on many of today's most widespread applications, whose procedures require data capable of withstanding a level of use almost unimaginable until a few years ago.
As of today (30/06/2018), considering a total of 19.5 million civic numbers, STUDIO SIT srl has comprehensively mapped territory equal to 12 million of them (whose place names are taken from official municipal routes). This coverage comprises 5000 municipalities with their approximately 42 million inhabitants. The average update of data refers to the year 2016. Among provincial capitals the coverage is 98%, that of the 130 most populous cities is close to 96%, and that of municipalities with more than 40,000 inhabitants is now 89%. By 2020 we predict that we will reach 16 million civic numbers contained within 6900 municipalities. We're still running up that hill !

KEYWORDS
GIS; MAP; LOCALIZATION; CIVIC ADDRESS; ITALIAN TOPONYMY DATA; GEOREFERENCED; MULTIUTILITIES

ABSTRACT
The quality of toponymy data has come increasingly under the microscope in recent years. These augmentations infiltrating the sector now require extremely accurate, up-to-date and comprehensive data, most crucially in respect of, for example, the automatic driving project, emergency services, and applications inherent in civil protection.
STUDIO SIT Srl was conceived of in 1991, the dim and distant past for many. For suppliers such as HERE and TOMTOM involved with the surveying and mapping of Italian civic numbering, the Ligurian company has been able to exceed a 95% degree of accuracy, completeness and updating, unique in this country, which traditionally lacks the availability of good quality toponymy opendata.

AUTHOR
VALERIO ZUNINO
VALERIO.ZUNINO@STUDIOSIT.IT
STUDIOSIT SRL