URBAN DATASCAPES[1]

Carlo Ratti,
Sensible Lab MIT and CRA, United States
ratti@mit.edu

In a short story written in the late 1960s, the Italian writer Italo Calvino imagined a dystopian society in which every detail and every moment is recorded for posterity. All information would be compiled into the greatest document ever conceived, blending the details of every individual life. The concept is problematized by intrigue and paradox surrounding information control and deletion, and a final twist puts an uncanny spotlight on the condition of absolute archiving. How will humanity remember itself? And how will it act when it knows that it is being recorded? These are prescient questions for a society that, today, is confronted with a similar situation of total recall.

Fifty years after the first publication of "Memory of the World", data is turning Calvino's fiction into a reality. As digital technologies become increasingly pervasive, permeating our physical spaces with the Internet of Things, every individual is generating a staggering amount of data. It is not always easy to precisely quantify how much aggregated data our digital society is producing. In 2016, it was famously estimated that, in that year alone, mankind produced as much information as had been created since the beginning of human civilization. One can also consider this sense of scale through our billions of everyday digital devices, most of which contain more computational power than NASA wielded in the time of Apollo 11. Every time we send a message, make a call, or complete a transaction, we leave digital traces. As digital information is captured and stored, our "Memory of the World" – the virtual copy of it – becomes more detailed.

At the forefront of this data-driven revolution there are cities. With their population density and sensor-laden streets, cities are the first places where we are experimenting with new ways of using data for improving our daily life. How can designers and urban analysts leverage this huge amount of information to improve the way we understand and design our cities? To foster a discussion, we are putting forward a reflection on the different types of data available today, and some examples of their possible uses, as drawn from the work of the MIT Senseable City Lab.

Any dataset collected for a specific purpose has an array of potential data by-products. Working with this deluge of urban information is what researchers often call opportunistic sensing: using data that has been generated for a specific reason and analyzing it in a different context to arrive at new conclusions. Datasets are often enriched with many dimensions, and whether or not every one of those dimensions was intended to have an explicit use when it was created, every aspect can be instrumentalized in creative ways. Credit card transaction data, for example, includes unique IDs for the vendor and the consumer. These tags allow researchers to filter the data by location and type of purchase (food, gas, clothing) to understand patterns of economic behavior in cities. Similarly, analysis of telecommunications data and social media have proven them both to be powerful tools for understanding human networks and social dynamics.

When sensors are distributed in a dynamic network, datasets can be considered individually, but far greater insights lie at their intersection and superposition. Geographic space is the common denominator that allows them to be linked. As early as 2006, our researchers at the MIT Senseable City Lab began compounding telecommunications with transportation data. The aggregate urban portrait that emerged – specifically during such extraordinary events as the final match, in Rome, of the 2006 soccer World Cup – revealed collective behavior tied directly to the event. Before the game, movement and communication were frenzied. Activity slowed almost to a stop during the match, spiked sharply at halftime, fell to almost zero during the tense final minutes, and exploded when the match was called. Communication traces during the following hours revealed mass movement into the downtown area to celebrate the national team's victory.

Subsequent projects in cities with more readily accessible data, such as Singapore, have compounded even more datasets. Data from weather, shipping, social media, public transit, cell networks, and more flow together to create a multidimensional portrait of cities and their patterns. Beyond opportunistic sensing, data can also be generated by deploying an array of sensors with specific intent. Embedding technology into the urban environment can yield robust and fine-grained data, whether to map an existing system, to reveal dynamics that have never been brought to light, or to gain a new understanding of humanity's fingerprint. On a macro scale, the Google Street View car, for example, has driven across the world photographing 360-degree panoramas. After its first five years of operation, the Street View team announced that their fleet had captured five million miles of road in thirty-nine countries, generating a staggering twenty petabytes of data – quadrillions of images.

As more and more of these digital elements are embedded in physical space, many other aspects of the urban environment can be revealed with a participatory approach to gain more information about our collective behavior and impact on the environment.The Senseable City Lab began a project, Trash Track, that brought the power of ubiquitous tracking to our waste disposal system. Researchers created geolocating tags and worked with residents of Seattle to attach them to thousands of ordinary pieces of garbage – effectively creating an "Internet of Trash" – to map the waste removal chain across the United States. Over the following months, the devices revealed a surprising network that had been completely unknown before. In the future, an accelerating diffusion of technology into urban space may offer an unprecedented understanding of systems like waste management dynamics and may create data that

ISSN online: 2239-0243 | © 2021 Firenze University Press | http://www.fupress.com/techne
DOI: 10.13128/techne-10679

can be used to optimize the entire system, even in real-time. As digital devices become increasingly lighter, smaller and able to process greater amounts of data, we are moving towards the full realization of a phenomenon that has been termed "smart dust" by Kristofer Pister. Physical space could be ubiquitously laced with nanosensors – scattered devices that are smaller than grains of rice. Large-scale networks of wireless sensors are becoming an active topic of research, promising a rich array of data in a future of ubiquitous. In fact, a pervasive network already exists in our cities today: citizens themselves. Internet-connected humans are producing huge amounts of data. The computer scientist Gordon Bell was one of the first to explore the idea of individual data in a practical way – in 1998 he began a project called Your Life, Uploaded, making himself the subject of the first full-resolution experiment in so-called life-logging. Bell created the hardware and software to capture every moment and every action of his life through photos, computer activity, biometrics, and more. The technology was primitive, and in many ways disruptive, but the project was successful in cataloging his existence for more than a decade. "The result?" he wrote. "An amazing enhancement of human experience from health and education to productivity and just reminiscing about good times. And then, when you are gone, your memories, your life will still be accessible for your grandchildren." What Bell initially set out to do as a full-scale scientific and sociological study is now the unconscious norm – the default condition – of an Internet generation. Our spatial and social activity is tracked and logged; in many cases, it requires more effort and determination to opt out of documentation than to opt in. Tweets, Uber calls, text messaging, restaurant reviews, and check-ins have become the natural activities of daily living.

The resulting trove of data can provide a deep understanding of how people interact with physical space as digital traces are mapped and overlaid-revealing, for example, the movement and activity of tourists. Using Flickr data, Senseable City Lab researchers started the project "Los Ojos del mundo" to map a crowd-generated cartography of Spain, showing how visitors and residents see and use their environment and identifying hotspots, or "visual magnets." Researchers could effectively borrow the eyes of the population in a continuing analysis, applying computer vision image-processing and color-matching to landscape photos. The user-generated photographic data began to reveal natural ecological conditions like drought and urban green spaces.

The categories of data collection in cities mentioned above – opportunistic sensing, ad hoc sensor deployment, and crowdsensing – can be hybridized. On the backbone of telecommunications networks, a new universe of urban apps has appeared, allowing people to broadcast and exchange geolocated information and reveal the city from their personal perspective. Air quality, for example, is poorly understood because data is collected in static and sparse ground-based stations. In a possible future, citizens themselves could carry a distributed network of sensors that create a real-time atmospheric map. Using smartphone-integrated sensing devices, pedestrian commuters could generate data at the human scale, as though a tracer were running through the veins of the cities, showing the urban environment that the commuters live in and move through. This concept may inspire manufacturers of consumer electronics to include environmental sensors and to publicly release the resulting data for analysis.

NOTES

[1] This text is an adapted elaboration of: Matthew Claudel and Carlo Ratti, The City of Tomorrow: Sensors, Networks, Hackers, and the Future of Urban Life (New Haven: Yale University Press, 2016).