urbanNext Lexicon
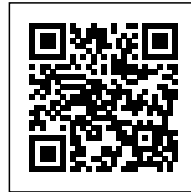
Sense and the City. Towards a New Digital Common.

Carlo Ratti + Daniele Belleri

SENSE AND THE CITY. TOWARDS A NEW DIGITAL COMMON

Posted on July 31, 2019 by content

Categories: [Actar](), [Carlo Ratti](), [Daniele Belleri](), [Essay](), [expanding design practices](), [Imminent Urban Commons](), [No Density](), [SENSEable City Lab](), [Senseable Technologies](), [Technology and fabrication]()

Tags: [Architect's role](), [Bigdata](), [Data](), [Data as mediation](), [Dataviz & simulation](), [Digital technologies](), [Emergent technologies](), [Environmental technologies](), [Essay](), [Future Commoning Cities](), [Global Awareness](), [Imminent Commons](), [New technologies](), [Planetary urbanizations](), [Responsive Environment](), [Responsive technologies](), [Sensing](), [Sensors](), [Seoul Biennale 2017](), [Technology](), [Urban Paradigms]()

Authorship: Research by Carlo Ratti + Daniele Belleri.

urbanNext Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

The world today is awash with data. In 2016 alone, people produced as much information as was created in all of human history. Every time we send a message, make a call, or complete a transaction, we leave digital traces behind. We are quickly approaching the creation of what Italian writer Italo Calvino omnisciently called the "memory of the world": a complete digital copy of our physical universe.

Such a scenario raises fundamental questions related to both who has access to data, and what data can be used for. As increasing distrust towards political institutions is apparent all over the world, our society finds itself at a turning point: data can become either an instrument exploited for private, adversarial interests or a tool to constitute a new positive "commons." In other terms, to borrow Richard Buckminster Fuller's words, we are at a "utopia or oblivion" crossroads.

To foster a debate on the issues at stake, we should first take a step back from the heated debate on the relationship between democracy and the emerging "dataville." What we want to suggest here is a reflection on the different types of data available today, their taxonomy, and their possible uses. The fundamental premise is that Big Data can also provide us—as planners, engineers, designers, and, above all, citizens—with new tools to understand and transform the spaces we live in. If we take the right steps today, the city of tomorrow could evolve into an open platform to foster civic engagement—a kind of new commons based on the shared knowledge of the city.

As a starting point towards that goal, what we would like to propose here is a classification of data—focusing on its acquisition method and its urban usage. After having described the proposed classification system, we will illustrate them using case studies taken from the present and past work of the MIT Senseable City Lab.
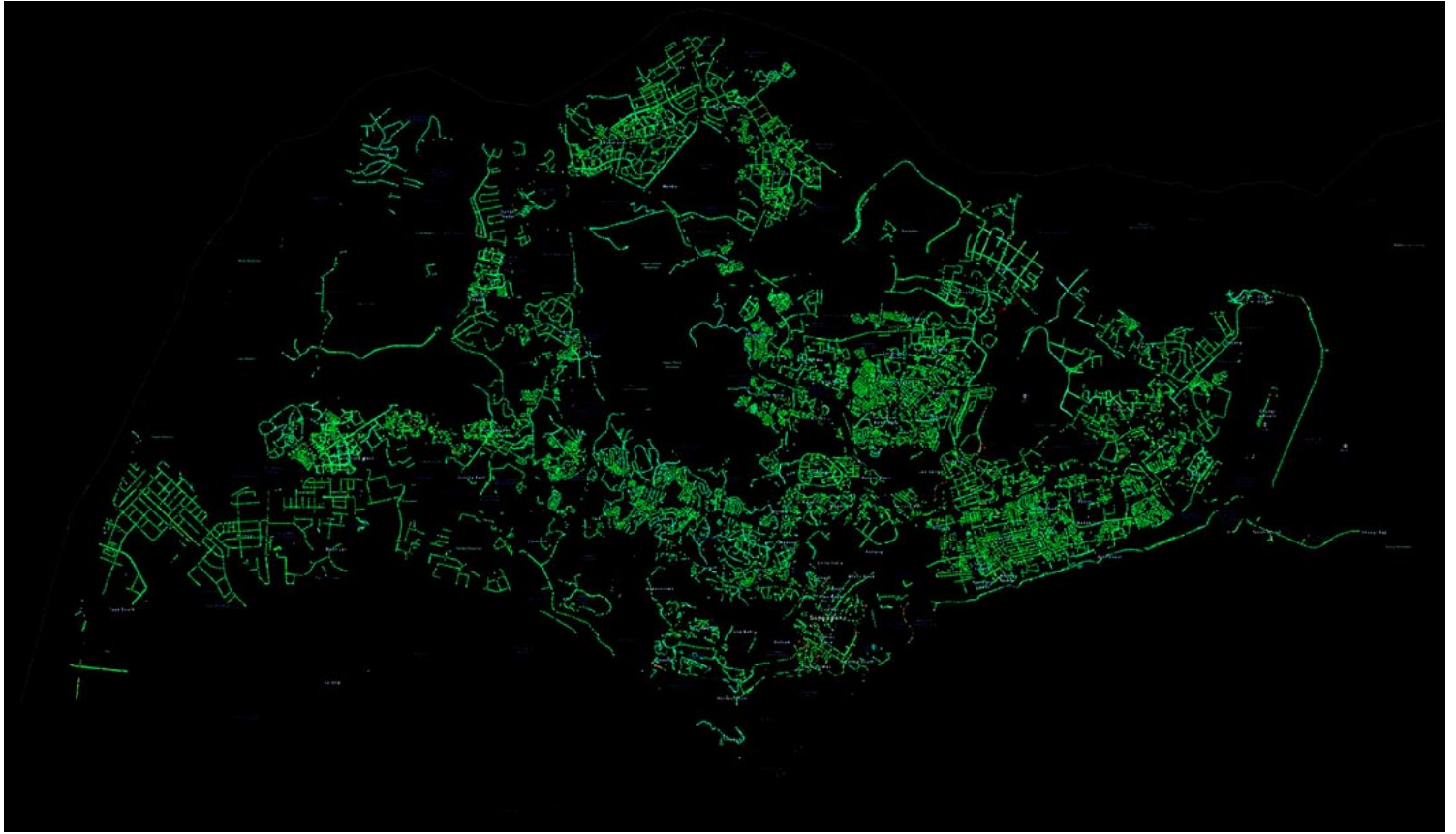
Treepedia sources Google Street View (GSV) panoramas, which are analyzed by artificial intelligence to measure "canopy obstruction". This result is the formulation of a "Green View Index (GVI)" to compare human perception of the urban green from the street level in different cities. Image: Singapore

DIFFERENT TYPES OF BIG DATA

There are many different ways we could classify urban Big Data. One could start with its applications—say in fields as diverse as transport, energy, production, etc. Alternatively, one could take a more in-depth approach and look at the structure of the data itself, such as the taxonomy of all the possible fields that it contains. In most urban data, for instance, two commonly recurring fields are time and location, latitude, and longitude coordinates.

Below we propose a classification that starts with the very nature of the data by considering its source. Much of the current research on Big Data is largely agnostic regarding the origin of urban data sets. But data sets do not just appear out of nowhere; the conditions of their generation need to be examined in detail. We think that starting with the "modes of production" is imperative if we want to better understand the politics inherent to Big Data—and work towards a future condition where Big Data can evolve into an open, urban commons.

We could distinguish between three different types of data acquisition. First, we have what can be referred to as "opportunistic data." This is data that is collected by running some kind of system, but that can be "opportunistically" used for something else. Think about data collected by cellphone companies to run their operation. In recent years, thanks also to our research, this very fine-grain recording of human life has become a powerful tool for understanding the city and its dynamics. In general, we can say that "opportunistic data" is a byproduct of some large information infrastructures. To analyze them means taking the generating system as a proxy for another phenomenon of interest. The data sets are uniform, follow a consistent logic, and reflect the properties of the system that generated them. Elaborated data-sharing agreements are often required with the owner of the collection infrastructure. In case of cellphone data, credit card data, and other similar types of data, this process can be rather tedious, governed by detailed data-sharing agreements.

The second type of information acquisition deals with "user-generated data"—such as data produced on social media platforms. Every tweet, Facebook post, or Flickr upload can provide valuable information to better understand cities and society. Access conditions vary: for instance, everyone can access a percentage of all tweets that are produced online—while the possibility to access "all" tweets on a certain subject or geographic area requires ad hoc permissions or payment. However, "user-generated data sets" are generally very large, and even if only partially accessible, can become a valuable input into different types of analytics.

The third and final category of data is "purposely sensed data." Its acquisition is achieved by deploying sensors ad-hoc, in order to better understand a specific phenomenon. If the two previous categories dealt primarily with the "hunter gathering" of data, we could say that the third one refers to the new space of "data farming." With sensors becoming increasingly inexpensive and self-powered—as we enter the new era of "smart dust"—more and more sensors in our cities and buildings will provide an increasing amount of data in real time.

urbanNext Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

CASE STUDIES

I. OPPORTUNISTIC DATA

TREEPEDIA

Increasing a city's tree canopy contributes to lowering urban temperatures by blocking shortwave radiation and increasing water evaporation. In addition to creating a more pleasant microclimate, trees also help mitigate air pollution caused by everyday urban activities. However, how can we measure the tree canopy? Treepedia—a project developed in collaboration with the World Economic Forum's Global Agenda Council on the Future of Cities and the World Economic Forum's Global Shapers community—uses Google Street View (GSV) panoramas. Thanks to artificial intelligence, images are analyzed and canopy obstruction measured. As a result, the Green View Index (GVI) is calculated to evaluate and compare urban areas. The GVI presents human perception of the urban green from the street level (as opposed to other methods based on satellite images) and allows the comparison of canopies among most cities all over the world—virtually all of those scanned by GSV. In 2015, the World Economic Forum's Global Agenda Council on the Future of Cities included increasing green canopy cover on their list of top ten urban initiatives: "Cities will always need large-infrastructure projects, but sometimes small-scale infrastructure—from cycle lanes and bike sharing to the planting of trees for climate change adaptation—can also have a big impact on an urban area." Treepedia shows how to use opportunistically collected data by Google to better understand the green canopy—and to use this information to allow citizens to take action.
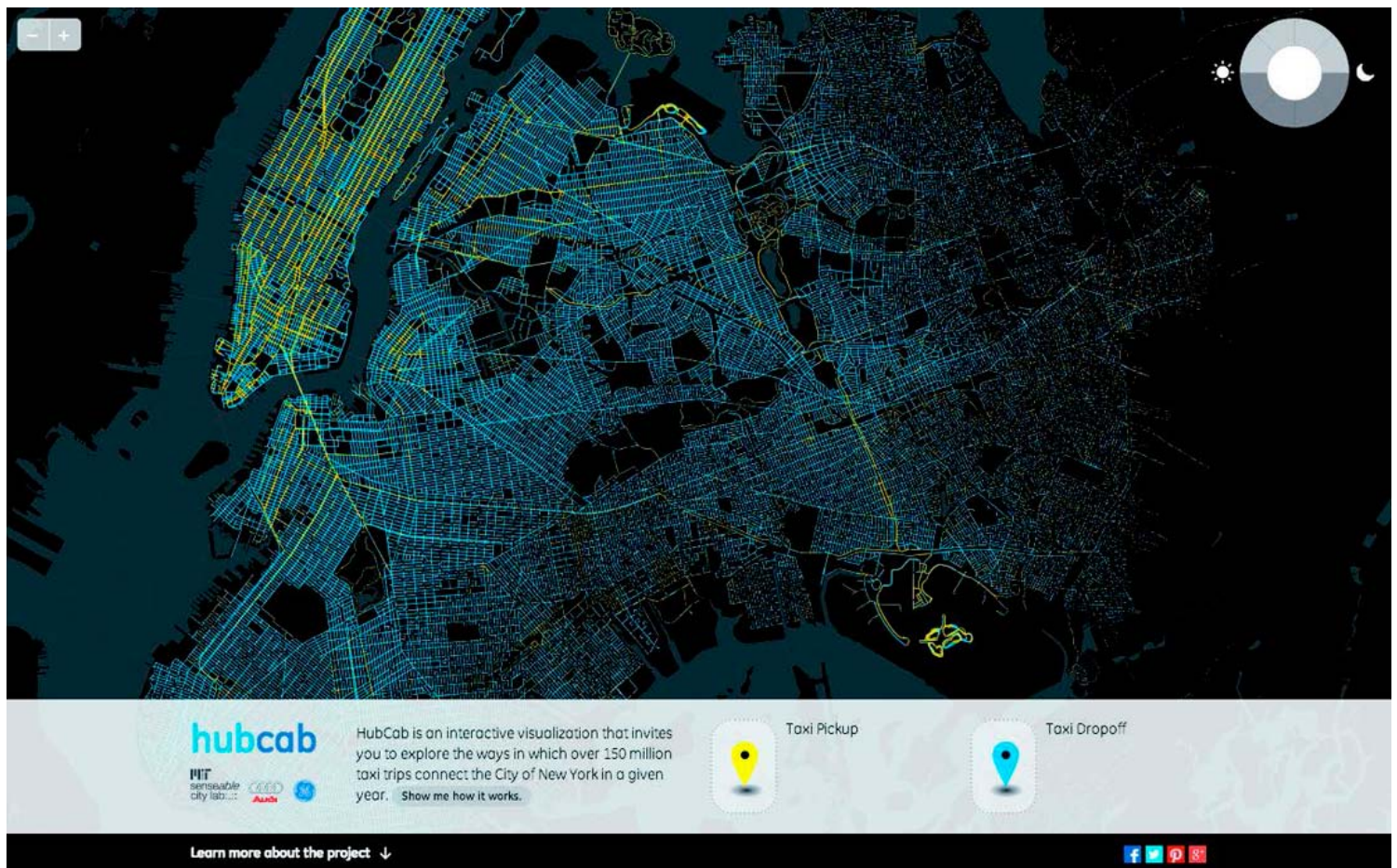
HUBCAB

HubCab is an interactive visualization that explores the ways in which over 170 million taxi trips connect the City of New York in a given year. The basis of the HubCab tool is a data set of over 170 million taxi trips by over 13,000 Medallion taxis in New York City: GPS coordinates of all pick-up and drop-off points and corresponding times. The HubCab interface provides a unique insight into the inner workings of the city from the previously invisible perspective of the taxi system. HubCab investigates exactly how and when taxis pick up or drop off individuals and identifies zones of condensed pick-up and drop-off activities. The HubCab tool expands and changes the perception of

urban space by using a large-scale data set. Furthermore, the analysis of the data shows the vast potential of taxi sharing. Our mathematical method introduces the novel concept of "shareability networks" that allows for efficient modeling and optimization of trip-sharing opportunities. Such an approach could lead to less traffic congestion, reduced operating costs and split fares, and to a less polluted environment. An interactive map shows the total fare reduction to passengers, the distance saved in miles travelled, and emission savings in kg of CO2 that come from potentially shared trips. Quantitative results demonstrate how taxi sharing could reduce the number of trips by 40% with only minimal delays for passengers.

Screenshot of HubCab, showing pickups and drop offs of all 170 million taxi trips over one year in New York City

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

URBAN EXPOSURE

Vast digital data sets are also changing how we predict the impacts of the urban environment on human health. In this project we have been looking at this space using a premier example of "opportunistic" data: cellphone information, which is collected for the sake of running a telecommunication infrastructure but at the same time provides invaluable content to better quantify human mobility patterns. Until recently, much of our understanding of the impact of air pollution on population health has been based on the relationship between air quality and mortality and/or morbidity rates in a population which is assumed to be at their home location all the time. Accounting for the movements of people can improve our understanding of this relationship. In this project we quantified human exposure to air pollution at an unprecedented scale thanks to data aggregated from cellphones. We examined 121 days of data from April through July 2013, using many types of wireless devices from a variety of providers, and blending the phone data with pollution information from the New York City Community Air Survey. We mapped the movements of several million people using ubiquitous cellphone data, and intersected this information with neighborhood air pollution measures. Covering the expanse of New York City, the study reveals where and when New Yorkers are most at risk of exposure to air pollution—with major implications for environmental and public health policy. The study broke New York City into 71 districts and found that exposure levels to particulate matter (PM) in 68 of the districts were significantly different when the daily movement of 8.5 million people was accounted for.

II. USER-GENERATED DATATWEET BURSTS

Social media has fully pervaded our lives. Thanks to its widespread uptake, it has become possible to study massive data streams in which people express their sentiments, often towards a specific topic. MIT Senseable City Lab, in partnership with Ericsson, has undertaken a visual and scientific exploration of how people express emotions online, and how this information could improve our understanding of human behavior. The study raises a number of important questions: Are people doing this independently, or in response to seeing other short messages? Are people following the herd? Could we use these insights to learn more about financial bubbles by measuring more impulsive, less rational responses? And can we design better communication services?

In this study, researchers used several large data sets of online messages collected from different

urbanNext Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

media sources. Data set 1 contained around 410,000 messages from Twitter during the 2012 Masters Tournament, a major championship in golf, held between April 5 and 8, 2012, in Augusta, Georgia. Data set 2 includes almost 20,000 messages posted in one thread of a popular online forum, the Something Awful (SA) forums (forums.somethingawful.com), during the U.S. presidential election night of November 6, 2012, and a smaller number of messages posted the week before the election night. A third data set includes the well-known Enron email corpus, containing roughly 250,000 emails exchanged between the employees of the Enron Corporation over four years, between October 30, 1998, and July 19, 2002. Data set 4 includes over 200,000 tweets and 40,000 posts on the online social networking service Facebook, with the common topic of the snow storm called "Nemo" which struck the northeastern coasts of the United States and Canada on February 8 and 9, 2013. The last data set contains the entire corpus of almost 3 million posts from the Twitter-like microblogging service app.net over a six-month period.

The results were often unexpected: researchers discovered, for example, that emotional tweets are very short. During the most exciting moments, when Twitter is bursting with short and emotional tweets, the average length drops substantially from 90 characters to 60 characters.

The more excited we are, and the more intense the flurry of messages in the collective, the shorter our messages become.

LOS OJOS DEL MUNDO

Los ojos del mundo (*The world's eyes*) illustrates the photos that people visiting Spain leave behind as evidence of contemporary tourism in the country. Tourism in Spain is hardly quantifiable because tourists leave few tangible traces of their stay. As a consequence, citizens and local authorities struggle to identify what tourists see, what tourists enjoy, and where tourists travel to and from. *Los ojos del mundo* provides insights to these question from the digital photos publically shared on the web by people visiting Spain. Through data mining and visualization techniques, the study uncovers the presence and flows of tourists. As photos pop up, they reflect the intensity of tourist activity, thus uncovering where tourists are, where they come from, and what they are interested in capturing and sharing from their visit. The analysis and mapping of this data sheds light on the attractiveness of leisure cities and their hotspots. In contrast, it also reveals the unphotographed regions of Spain, still free from the tourist buzz.

When posting photos online, users of the photo-sharing platform Flickr transmit to the world their perspective of a place or event through the lens of a digital camera. Each digital photo file codes

urbanNext
Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

both the time that photo was taken and the location it captures. Analyzing this information allows us to follow each Flickr photographer as they travel through Spain. Also, about 60% of Flickr users disclose information about their home country. Analysis of the time and location data embedded in their digital photo files allows us to examine the Flickr photographers' geographic presence and trails over time, and to differentiate locals from visitors. Researchers, for example, could easily understand that Britons who visited Barcelona in Fall 2007 stayed on the beaten paths delimited by the city's main elements such as Parc Guell and Sagrada Familia, with Passeig de Gracia and the Rambla acting as main arteries. Another possible conclusion that can be deduced from Flickr data is linked to spaces of activity. Photographers often attach descriptions and tags when posting their photos on Flickr. The data mining of these tags allows us to infer what kinds of activities these photos capture. Spaces of activity reveal the regions and cities that host memorable parties in Spain over the course of a year.

urbanNext Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/
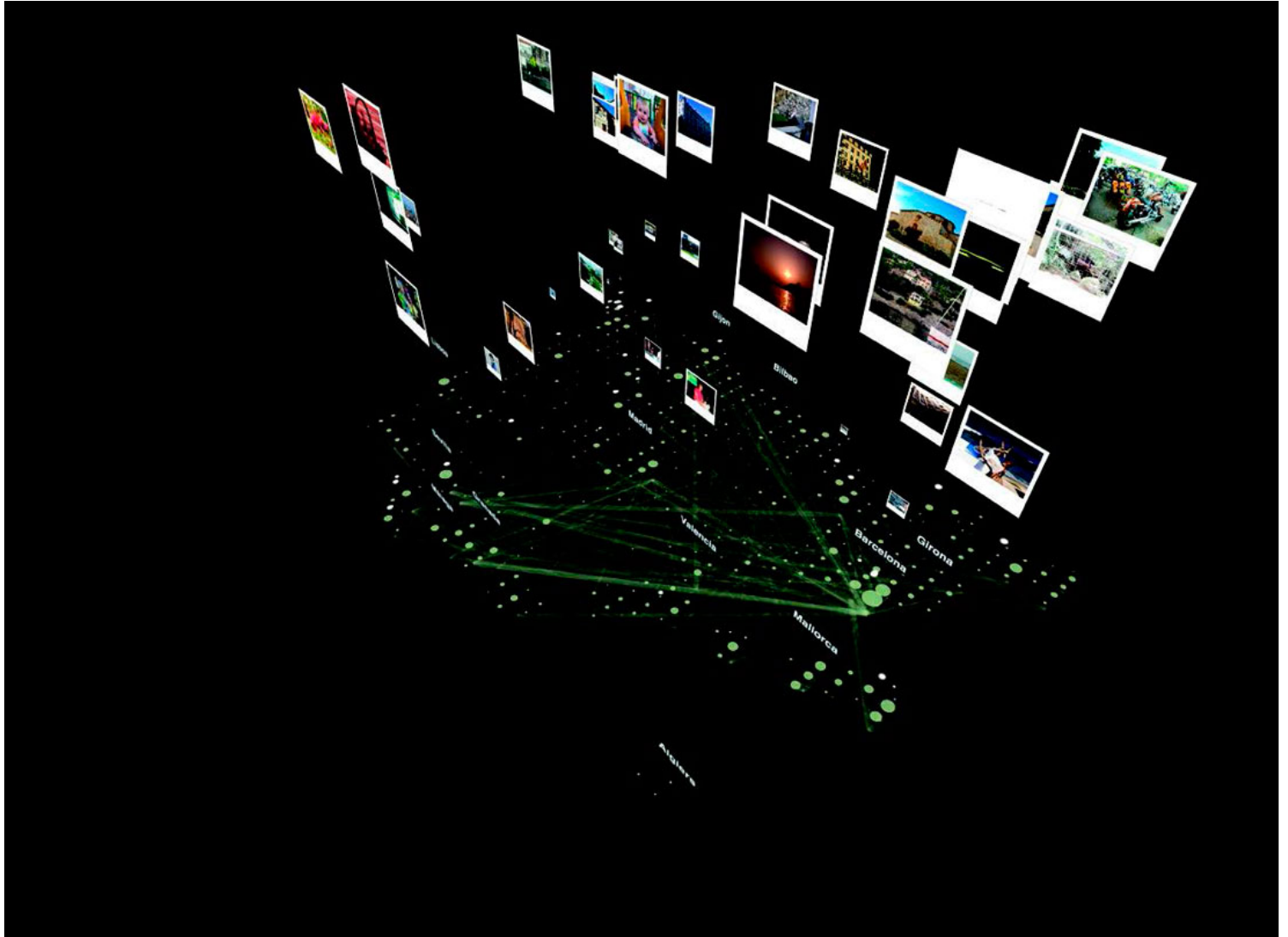
urbanNext
Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

Density and flows of photographers in Spain in 2007; Partying in Barcelona

Los ojos del mundo (the world's eyes) was one of the first projects to employ big data sourced from the web to quantify tourism—and particularly, tourists' paths and choices. Los ojos del mundo provided insights to these issues by mapping digital photos publically shared on the web by people visiting Spain

urbanNext
Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

III. PURPOSELY SENSED DATA

TRASH TRACK

The Trash Track project investigated the geographic dimension of urban waste systems by following the movement of individual trash items, thus tracing the flows of an urban infrastructure that is often hidden. Over the course of the project, we used ad-hoc sensors to record the trajectories of 3,000 trash items discarded in households, most of them in the metropolitan area around Seattle. We did not want to limit the experiment to assumptions about possible waste destinations; therefore, the project required an active sensing technology that is capable of autonomously reporting back from any location. Active location sensing means that an electronic location sensor is attached to the object, with the sensing device being slightly smaller than a cellphone. Location is acquired and reported by using the cellphone network infrastructure. The deployment of the sensors relied heavily on the involvement of volunteers. Initially, Trash Track was not designed as a participatory project, yet this aspect soon became the most important part. Volunteers eager to learn about the structure of the waste system contacted us, and contributed their ideas, time, and materials to the project. Among the tracked objects were packaging made from metal, glass, paper, or plastic; cellphones, TVs, and computers; books, clothing, furniture, toys, and many other items. After the tagged objects had entered the waste stream, the sensors started reporting their movement at regular intervals via the cellular network. We traced the movement of the discarded objects over a period of six months, until the batteries of most sensors had expired. The aggregated traces conveyed a rich picture of the waste removal chain; facilities including transfer stations, recycling centers, and landfills could clearly be made out as frequented nodes in the network. The project was an initial investigation into better understanding the "removal-chain" in urban areas—a first step towards making it more efficient and promoting behavioral change in society at large.

urbanNext
Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

urbanNext
Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

urbanNext
Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

Composite Map of the Recorded Traces; Plastic Container of Liquid Soap in New York

The Trash | Track project investigated the geographic dimension of urban waste systems by following the movement of individual, sensor-laden trash items, thus tracing the flows of an urban infrastructure that is often hidden

ART TRAFFIC

From sensors to track waste, to sensors to track people. Museums often suffer from "hyper-congestion," wherein the number of visitors exceeds their capacity. This can potentially be detrimental to the quality of visitors' experiences. Although this situation can be mitigated by managing visitors' flow between spaces, a detailed analysis of visitor movement is required before being able to take action. In this pioneering study, we attempted to analyze visitors' behavior in one of the world's largest museums—the Louvre—from anonymized longitudinal data sets generated by noninvasive Bluetooth sensors. This data enabled us to unveil some features of visitor behavior and spatial impact that shed some light on the mechanisms of museum overcrowding.

In particular, the research team deployed seven Bluetooth sensors, with sufficient coverage to measure visiting sequences and duration at key representative locations. The sensors recorded a unique encrypted identifier that distinguishes each Bluetooth-enabled mobile device within its range, as well as time stamps for entry and exit times. Assuming that a mobile device belongs to a person, we can relate the movement of the device to that of the visitor. The study was conducted over a twenty-four-day period with a high volume of visitor traffic. During this period, the array of sensors recorded the presence of 24,452 unique devices. The findings increased the understanding of the unpredictable behavior of visitors, which is key to improving the museum environment and experience.

UNDERWORLDS

Sewage contains important health data. Such is the idea behind Underworlds, a cross-disciplinary, open-data platform for monitoring urban health patterns, shaping more inclusive public health strategies, and pushing the boundaries of urban epidemiology. Underworlds consists of a physical sensing infrastructure and biochemical measurement technologies to analyze sewage. The Underworlds project is the first of its kind, and a proof of concept that cities can use their waste water system to do near real-time urban epidemiology and to understand human health and

urbanNext Lexicon

Sense and the City. Towards a New Digital Common

https://urbannext.net/sense-and-the-city/

behavior with a fine spatio-temporal resolution. Early warnings of the presence of new flu strains in urban centers could significantly reduce a community's medical costs and even help mitigate outbreaks. In addition, smart sewage could impact the way noncommunicable diseases are studied; for instance, biomarkers for diseases such as obesity and diabetes can be measured at an unprecedented scale and temporal resolution. The implications of this platform extend beyond just disease surveillance to the development of a new type of human population census. Analyzed in tandem with demographic data, this platform can study the aggregate health of a city to the health of a particular neighborhood.


CONCLUSIONS AND NEXT STEPS

The three different categories of data illustrated above each comes with its own set of challenges and particularities. However, looking at these examples, it's easy to understand the importance of combining and aggregating data sets from multiple sources. The aim is an organic and complete perspective on our cities, in order to identifying their patterns.

In doing so, we need to go well beyond our inquiry on accessing data, and must strategize on how to make it accessible to the public. We actually need to intervene between the accumulation of data and releasing it for public usage. Access to information allows the urban public to see hidden patterns that are not otherwise observable. In keeping citizens in the loop, we can improve data volume and quality by expanding its audience. This offers citizens a way to critique the data and its embedded assumptions, leading to better methods for generation, acquisition, and aggregation of data sets. Also, we can instigate a sense of responsibility for the city and its shared public goods—meaning the urban data that allows the decoding of its daily dynamics. Data can promote behavioral change and bottom-up initiative. If managed positively, it can constitute a new and relevant collective voice—a new common among society's more established ones.
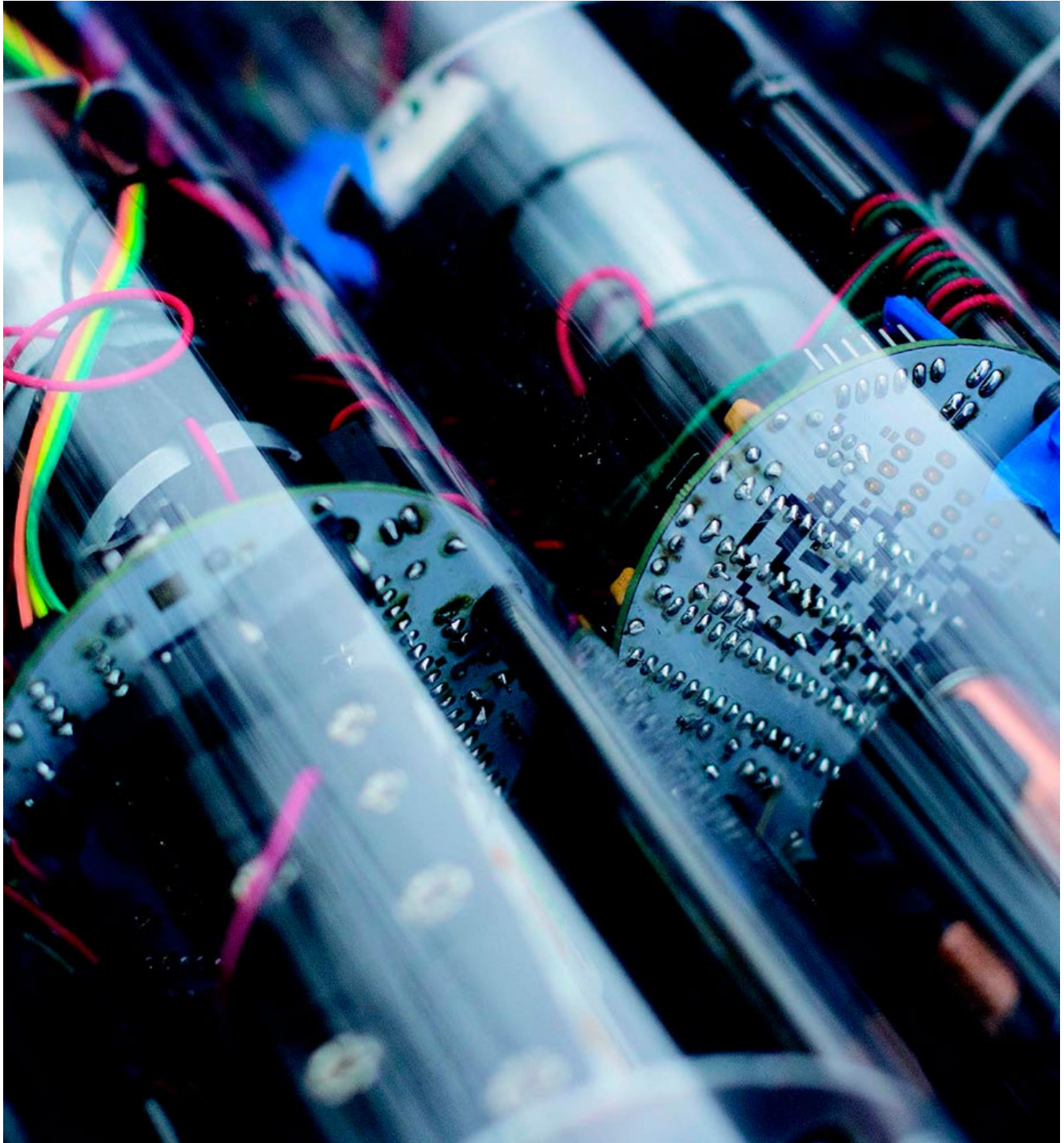
urbanNext
Lexicon

A detail of the portable robots used by MIT Senseable City Lab to sample sewage in Boston

Sewage contains important health data. Such is the idea behind Underworlds: a cross-disciplinary, open-data platform for monitoring urban health patterns, shaping more inclusive public health strategies, and pushing the boundaries of urban epidemiology

urbanNext
Lexicon