A Data Framework for the Real Estate Industry

Dr. Carter Casady[[1]](#footnote-1) and Dr. Herman Donner[[2]](#footnote-2)

Stanford School of Engineering

*Disruptive Technology and Digital Cities Program*

About the Disruptive Technology and Digital Cities Program

The **Digital Cities research initiative** creates a focused research effort bringing together cross-discipline expertise including data analytics, institutional investment on urban sustainability, and exploration of how emerging technologies will change the way we think about business model development for government and enterprises engaged in digital cities. Collaboration with corporate affiliates will ensure market connection, validation, and new ways of transforming disruption into competitive advantage.

The **Stanford Affiliate Program** in Digital Cities is one of the first programs of its kind in the United States. Its goal is to fundamentally change the way governments and companies think and operate as the world moves toward truly digitized urban centers - by harnessing focused research, commercial opportunity, and new ways of thinking about governance and sustainability.

[1. An Overview of Data Driven Disruption 4](#_Toc34818616)

[2. Impact of Data on Urban Systems 9](#_Toc34818617)

[*2.1* *Big Data* 10](#_Toc34818618)

[*2.2* *Alt-Data* 12](#_Toc34818619)

[*2.3* *Origination of Big Data and Alt-Data* 14](#_Toc34818620)

[*2.4* *Strategies for Third-Party Data Acquisition and Within-Organization Data Origination* 16](#_Toc34818621)

[*2.5* *Organizational Considerations and Implementation Challenges* 17](#_Toc34818622)

[3. Understanding the Role of Cities 19](#_Toc34818623)

[*3.1* *Existing Examples of Digitalized Environments* 19](#_Toc34818624)

[3.1.1 Toronto (Eastern Waterfront), Canada 21](#_Toc34818625)

[3.1.2 Belmont, Arizona (USA) 22](#_Toc34818626)

[3.1.3 Songdo, South Korea 23](#_Toc34818627)

[3.1.4 Dubai, UAE 24](#_Toc34818628)

[3.1.5 Neom, Saudi Arabia 25](#_Toc34818629)

[4. Data Applications and Use-Cases in Urban Settings 26](#_Toc34818630)

[*4.1* *The Potential Power of Predictive Analytics in Real Estate* 28](#_Toc34818631)

[*4.2* *Use-Cases of Predictive Analytics in Real Estate* 30](#_Toc34818632)

[4.2.1 Predicting Store-Front Demand and Closures Using Street View Imagery 31](#_Toc34818633)

[4.2.2 Tracking Pricing of Consumer Goods to Identify Urban Change 32](#_Toc34818634)

[5. A Strategic Framework for Characterizing Data in the Real Estate Sector 33](#_Toc34818635)

[*5.1* *Aggregated Market Observables* 34](#_Toc34818636)

[*5.2* *IoT and Smart Building Data* 35](#_Toc34818637)

[*5.3* *Third-Party Alt-Data Streams* 36](#_Toc34818638)

[6. Conclusions 38](#_Toc34818639)

## 1. An Overview of Data Driven Disruption

Across industries, internet and communications technologies (ICT) have transformed the business landscape. Increased speeds of communication and data gathering have made markets more efficient, increased competitiveness, and led to entirely new business models for product and service delivery. Such improvements have had an enormous impact on the economy. Between 2005 and 2010 alone, improved communications technologies accounted for 21% of GDP growth in mature economies.[[3]](#footnote-3) Today, scholars estimate “the digital economy is worth $11.5 trillion globally, equivalent to 15.5 percent of global GDP and has grown two and a half times faster than global GDP over the past 15 years.”[[4]](#footnote-4) However, as the exponential pace of technological innovation advances and begins to augment everyday life, a new era of big data analytics is taking shape, one that portends to offer even broader and more far-reaching market disruptions.

As the pervasiveness of Moore’s Law has grown across technology sectors, advancements in cloud storage technology, computing power, and sensor technology have led to exponential advancements in technology. While storage of big data was very costly less than a decade ago, today it is fairly inexpensive to collect information on millions of data points (e.g. car rides, transactions, viewers, etc.), store them, and run algorithms on the data to draw conclusions. Low cost sensors, networks and cloud storage now make it possible for virtually all devices to gather data. With Internet-of-Things (IoT) connections expected to grow by 30% a year until 2023,[[5]](#footnote-5) the IoT revolution is driving massive increases in data generation—i.e. the global amount of data generated in 2020 is projected to be 40 times larger than the data produced in 2017.[[6]](#footnote-6)

Today, a variety of industries are now using this big data to predict market trends and gain customer understanding.[[7]](#footnote-7) This type of analysis is also beginning to reach unprecedented levels of granularity. As services become increasingly digitalized, they also provide data points for individual consumer profiles. Companies are monetizing this data through tailored advertising, content, and even pricing.[[8]](#footnote-8) For example, insurance and finance firms are analyzing behaviors to individually assess risk characteristics of individuals,[[9]](#footnote-9) retailers are analyzing responses to offers to enable behavioral based pricing,[[10]](#footnote-10) and manufacturers can now estimate component reliability so accurately that what was once sold as a good can now be sold as multi-year service contact.[[11]](#footnote-11)

In the world of urban investment and development, these potential impacts of data are now being discussed through “Smart Cities,” “Intelligent Cities,” and “Digital Cities” frameworks. Although there remains relatively little consensus on the meaning of these terms across academia and the private sector,[[12]](#footnote-12) the discourse around digitalized urban environments often focuses on the provision of services and information to the public.[[13]](#footnote-13) To date, most of the extant research on IoT technology pertains to public sector management rather than private sector adoption and implementation.[[14]](#footnote-14) However, truly “digital cities” serve as the heart of analytics across industries. In cities, companies collect data on where people move using mobile, location-based-services, where money is spent via credit cards, digital wallets and online orders, and how much value is created through digitalized public and private records (e.g. taxes, building permits, retail reward memberships, review platforms, and street view imagery). Similarly, the usage of buildings, roads, transit, and other utilities is monitored by sensors. This data can be used to improve the efficiency and reliability of infrastructure delivery, such as energy distribution, traffic congestion, public transport ridership, and wastewater management. Most of this data is collected by public or private sector organizations for specific purposes. However, this data often finds secondary uses which can create novel insights. When data is used in this unconventional way, it becomes alternative data (alt-data).[[15]](#footnote-15)

Although alt-data and big data tend to not be mutually exclusive and classifications around these data types remain fluid, the pervasiveness of this information is growing across industries. The real estate industry, however, lags behind others in its use of these types of data.[[16]](#footnote-16) This is because the real estate market is generally characterized by low levels of innovation, limited transparency (relative to the financial markets), and a strong emphasis on relationships and experience, among other factors.[[17]](#footnote-17) In this setting, both big data and alt-data will likely offer significant and specific value for real estate market actors. For instance, in urban environments, the health and safety of various neighborhoods can be monitored by measuring traces of drugs in local sewage pipes, tracking flu outbreaks using Google searches, or relying on Yelp reviews to direct health inspectors to restaurants with complaints. The city of New York even analyses the addresses of asthma patients to identify buildings where mold and dust mites issues might be present.[[18]](#footnote-18) Additionally, the sentiment of financial markets can now be measured through Twitter activity,[[19]](#footnote-19) and real estate markets can use any and all forms of data associated with time and location to help identifying the next “hot” neighborhood where people want to live, work, and play.

In this light, secondary uses of data can be exponentially more valuable when converged with other information—such as real estate market data covering rents, prices, yields and occupancy rates—and applied in novel ways to discover economic relationships. Similarly, as more variables are tracked, the likelihood of such discovery increases.[[20]](#footnote-20) Cities at the heart of this digitalization should therefore be conceptualized as layers of convergent data that hold tremendous amounts of latent value.[[21]](#footnote-21)

However, today many companies and government entities fail to recognize and understand the enormous value of this data and its disruptive potential of existing business models.[[22]](#footnote-22) This is because, in reality, the convergence of data streams makes analyses of business impacts on urban settings particularly complex. Moreover, although terminology such as Business Intelligence, Big Data, and Predictive Analytics are used fairly interchangeably,[[23]](#footnote-23) most companies routinely fail to monetize or create value from this existing data,[[24]](#footnote-24) and a majority of big data projects tend to lack tangible benefits.[[25]](#footnote-25)

A recent study among institutional investors found that although a majority believe that alt-data can improve financial returns, they either lack alt-data strategies, or have “not well developed” strategies.[[26]](#footnote-26) Similarly, the consultancy Deloitte found that only 13% of real estate companies are prepared for new technologies such as robotics, AI, and advanced manufacturing—yet, a significant proportion of firms mistakenly believe that they are making competitive investments.[[27]](#footnote-27) This provides opportunities for market actors that are early adopters.

In the real estate industry, failure to monetize and create value from market data is particularly pronounced. Although aggregated market data, such as local cap-rates, rents, and property values, is readily available through services such as CoStar and Zillow, applications of advanced data analytics in property management and development remain extremely sparse because the industry has yet to master the development of data-driven business models. This report thus aims to address a portion of this gap between the enormous potential of data-driven decision-making and the under-utilization of data-driven business models by exploring how data and analytics may impact the real estate sector and market actors within real estate investment, development, and management.

In the sections to follow, we begin by describing the impact of data on urban systems and specifically focus on the generation, convergence, and value of big data and alt-data in the real estate industry. We then discuss ongoing digitalization efforts in cities and the role they play generating large quantities of untapped yet valuable data. Subsequently, we explore data applications and use cases in the real estate sector. Finally, we conclude by offering a framework for data characterization and conceptualization in the industry and discuss how real estate companies can work to better implement data driven strategies.

## 2. Impact of Data on Urban Systems

As indicated above, digital transformation is already disrupting a variety of industries, including real estate. In this market, the impacts associated with this transformation can be categorized into two broad themes. The first broad-based impact stems from ICT. ICT has increased the efficiency and variety of ways real estate information is collected, aggregated, and transmitted. As result, ICT advancements have created large knowledge spillovers which now facilitate diverse and frequent exchanges of information amongst individuals at relatively little cost.[[28]](#footnote-28) In the context of digital cities, these impacts are particularly relevant because increasing amounts of ICT data are being collected and utilized to enhance and predict economic activity. Although the individual causal relationships of these impacts remain difficult to disentangle to date, this connection between data and economic activity has created another thematic sea change—i.e. disruption of existing business models.

Today, new technologies, data sources, and web-enabled applications are already enabling companies to map, analyze, and micro-target individual demand profiles and user preferences.[[29]](#footnote-29) In real estate, the opportunity to use data on movement, demand, and preferences to create novel market insights is particularly promising. However, real estate currently remains insulated from these advanced analytical practices. This is because real estate is a fairly heterogenous asset which is traded with incomplete information. At its core, real estate can be characterized as a search-market where it is costly, and time consuming for buyers to find information and assess property values. Sellers similarly put in effort to set reservation prices and decide on a marketing strategy. Naturally, this dynamic creates market conditions where value is derived from relationships and informational asymmetries. There is in fact a long stain of academic literature that reviews this link between sales price and sales mechanisms[[30]](#footnote-30) and studies have shown that sellers who research the market tend to lock in higher sales prices because they find better matches with prospective buyers.

However, if all commercial lease activity was public knowledge, would tenants require consultants to aid them in the search process? And if so, would the value added be as significant? The short answer is no, because increased transparency and information dissemination will erode this value proposition. In this light, the real estate industry stands to benefit tremendously from digital transformation and the proliferation of data because increased transparency, improvements in market efficiency, and reductions in transaction costs amongst buyers and sellers in the marketplace will disrupt long-standing business models based on relationships and information arbitrage.

While listing sites such as Zillow, Redfin, RightMove and market aggregation services such as CoStar already serve as direct evidence of an emerging industry transformation—breaking down informational advantages held by real estate professionals—recent trends in the real estate are starting to pivot toward improved analytics and decision support systems. For example, start-ups such as Realeflow and Opcity offer software applications that improve the matching of buyers and sellers. Other companies like Cadre provide individuals and institutions with direct access to large commercial real-estate properties via a technology-enabled, institutional investing platform. Taken together, these new business models tend to share one common feature which helps them generate value: lots and lots of data. Whether its improved property valuations, investment decisions, or market predictions, each of these businesses relies on insights from big data and alt-data, both of which we define in more detail below.

### *Big Data*

Although big data is often cited a major disrupting force across industries, there remains little consensus on the definition of big data. Some tend to focus on the fact that data is *complex,* and that it can reveal patterns and create insights which would not be possible without the data. Others like Jim Sempere of Real Capital Analytics focus on the convergence of data by emphasizing “[b]ig data as a variety of data, put together from various sources to come together in a meaningful way”[[31]](#footnote-31) while

scholars such as Fabio Duarte of MIT’s Senseable City Lab stress that big data is “any human activity that leaves digital traces, which does not have an immediate meaning, but when combined together gives some meaningful pattern.”[[32]](#footnote-32)

In this light, a few studies have offered frameworks for characterizing big data in relation to real estate. Glaeser et al. (2018),[[33]](#footnote-33) for example, define a set of data categories that can be used for analyses of urban environments: These include:

1. *Digital Exhaust* (e.g. under-utilized trails of data left by internet-users);
2. *Open Government Records* (e.g. crime statistics, transit ridership, etc.); and
3. *Corporate Records* (e.g. sales, revenues, etc.).

Donner et al. (2018)[[34]](#footnote-34) offer a similar framework for big data applications in real estate. In their work, they categorize this data into four main categories, namely:

1. *Digitalized Records,* both public and private (e.g. credit-card data tracking of local changes in spending over time);
2. *Application Provided Information on User Preferences (*e.g.identifying local demand through Zillow searches by cost and geography, Facebook groups membership; Yelp comments; etc.);
3. *Sensor Information in the Urban Environment* (e.g. measuring foot traffic, noise/air pollution; etc.); and
4. *Sensor Information on the Urban Environment (e.g.* created by network-connected cars, public transit and devices such as smartphones and tablets).

Taken together, these data categorizations generally encapsulate the three fundamental V’s of big data[[35]](#footnote-35)—i.e. Volume (the data is so large that it cannot be store on a standard hard drive), Velocity (the data is updated in real-time, or near real-time) and Variety (the data is especially complex for analysis because it takes on a mix of structured and/or unstructured forms and encompasses a diverse set of formats (e.g. numbers, text, photos, audio, etc.)). These three V’s tend to play a key role in crossing big datasets so that preferences and consumption can be put in geographical and temporal context. They also collectively align with our proposed view of digital cities as layers of private and public data from which convergence can create insight.

### *Alt-Data*

Like big data, alt-data also has similar properties for creating insight. For example, Kolanovic and Krishnamachari (2017) focus on the origins of the alt-data and divide it into the following three non-exhaustive categories:

1. *Individual Processes,* encompassing data such as social media, news and web-searches;
2. *Business Processes,* encompassing transactions, public-, and private records; and
3. *Sensors,* that cover geolocation data from satellites and devices.

However, other scholars note that alt-data tends to take on a variety of additional dimensions. For instance, Kitchin (2015) offers a framework which builds on the 3-Vs for categorizing big data (i.e. not specifically alt-data) by adding *comprehensiveness, granularity[[36]](#footnote-36), relationality,[[37]](#footnote-37)* and *flexibility.*[[38]](#footnote-38)Monk, Prins, and Rook (2019) also offer a six-dimensional model of alt-data based on *reliability, granularity, freshness, comprehensiveness, actionability,* and *scarcity* while Dannemiller and Kataria (2017) avoid discrete categories altogether and define alt-data on a continuous scale from structured to unstructured.

Naturally, across these frameworks, the process of determining whether a dataset is indeed alternative remains challenging. This is because the thresholds that divide conventional and unconventional data (e.g., scarcity, novelty, or another relevant quantitative or qualitative dimension) are not always so clear. As a result, characterization of alt-data is often a more achievable and actionable set. At the very least, characterizing enables organizations to determine whether certain alt-datasets align with specific firm capabilities and strategic objectives. Although the literature is fairly limited on this subject (see above), real estate firms looking to glean insights from alt-data may find the framework developed by Monk, Prins, and Rook (2019) to be the most informative.

In their characterization schema, Monk, Prins, and Rook offer a six-dimensional model of alt-data based on reliability, granularity, freshness, comprehensiveness, actionability, and scarcity (see Table 1).

**Table 1:** A Six-Dimensional Characterization of Alt-data

|  |  |
| --- | --- |
| **Dimension** | **Explanation** |
| *Reliability* | How accurate, precise, and verifiable the data are (e.g., error-free, unbiased, checkable) |
| *Granularity* | The scale covered by specific data points or entries (e.g., continental, industry-wide) |
| *Freshness* | Age of the data (i.e., when collected/generated) relative to the phenomena they reflect |
| *Comprehensiveness* | What portion of a given domain the data cover (e.g., 25% of households in Canada) |
| *Actionability* | Degree to which significant actions or decisions can be made based on the data |
| *Scarcity* | How widely or readily available the data are to other (especially competing) organizations |

**Source:** Monk, Prins, and Rook (2019)

These dimensions offer a succinct process for identifying alt-data streams that align with a firm’s operating resources, existing internal capacities, and ongoing investment strategies. Characterization of data in this way will also enable organizations in the real estate market to more efficiently isolate data feeds that yield value in their investment, development, and management operations.[[39]](#footnote-39)

Regardless of the framework, each of these conceptualizations makes one thing clear: the use case, structure, and relevance of data is key for generating value. This means the origination and value proposition of alt-data and big data needs further examination.

### *Origination of Big Data and Alt-Data*

Much of what constitutes big data and alt-data is the application of information in secondary use cases—i.e. the data was collected for a different purpose. Exponential increases in available data are closely related to the proliferation of IoT devices and the ubiquity of Location-Based-Services (LBS) that enable corporations to understand movement patterns and directly micro-target markets, demand and risk preferences, and create unique consumer profiles.[[40]](#footnote-40)

This type of information is collected, largely by technology companies, and used to achieve “Context Based Intelligence.” For instance, Starbucks might send an offer to your smartphone for a cup of coffee on your way to work if your Fitbit noticed you had a bad night’s sleep.[[41]](#footnote-41) Likewise, a company could send you an offer for flowers if they know that your wedding anniversary is coming up in your calendar.[[42]](#footnote-42) Data convergence is essential in this context, and from an implementation perspective, this highlights the need for companies to analyze how data impacts their business strategy and objectives. This kind of decision simplicity is also becoming increasingly important for corporations because it makes consumers “sticky”—i.e. they follow through on an intended purchase, repeat the purchase, and recommend the product or service to others.[[43]](#footnote-43) Moving forward, additional data-points will only increase the accuracy with which consumers can be targeted. As more and more everyday items become data-gathering devices (e.g. toothbrushes tracking brushing habits, smart-scales monitoring BMI and weight, smart TVs noting what ads you have been exposed to, smartphones documenting transport usage, etc.), companies will increasingly offer more tailored services to what customers want and will want in the future.

In real estate, the abundance of big data and alt-data generated may be particularly powerful because understanding patterns of movement and consumption as well as preferences for entertainment, food and drink, and social-networks will directly affect strategic development and investment decisions. Because the built environment generally encapsulates both the delivery of public services and the growth of private enterprises, opportunities abound for converging public and corporate layers of data to create truly digitalized environments. Notably, Glaeser et al. (2018) indicate that data can provide valuable insight when converged with exogenous events—e.g. identifying the commercial impacts of making a road “pedestrian only” using credit-card transactions or analyzing how walking patterns change if a park is built.[[44]](#footnote-44) By crossing public data layers (e.g. public transport ridership, city permits records, demographics and employment information, etc.) with non-traditional and unstructured data streams (e.g. internet searches, Yelp reviews, credit-card transactions, satellite imagery, etc.), city planners, real estate developers, and investors have tremendous opportunities to generate value from previous untapped insights.

### *Strategies for Third-Party Data Acquisition and Within-Organization Data Origination*

Within this convergence of data streams, there is, however, an important distinction to be made between data originating from third-parties—such as IoT and application exhaust, data originating from within the industry, and data originating from within an organization.

From a strategic perspective, real estate companies need to focus not just on strategies for acquisition of big data and alt-data, but also on data generation from within the organization and how it relates to third-party data. Technologies such as sensors, digital twins (BIM), and increasingly structured databases with contracts such as leases, appraisals and financing agreements make it possible to measure and converge real estate market data with other sources of big data and alt-data.

Within-organization data can be conceptualized at two levels, the building-level, and the organizational level, with the former referring to measurement and analysis of building usage and performance, and the latter the ability to collect information that creates value through a better understanding of organizational processes.

The value of data convergence requires data strategies to address both third-party data and within-organization data origination. Given the cost of acquisition and associated competencies, these strategies need to address how these data sources will create joint value. Given the rapid digitalization of the urban environment, the real estate industry faces specific challenges in converging data from digitalized physical assets—e.g. sensors, BIM models, digitalized records and visualization tools—with third-party data-streams such as exhaust from applications, imagery, and digitalized records.

A key consideration, in terms of data strategy, is that internal data is associated with different challenges than third-party data. The former requires new technology to be deployed within an organization, while the latter requires acquisition and partnering and is likely to be less familiar. Initially, real estate organizations should focus on internal data gathering, and real estate industry data as organizational capabilities are likely to be far stronger in relation to processing this type of real estate market data and building usage data. As the industry is lagging in terms of digitalization, more efficient means of exhausting building-level data can provide significant competitive advantage. Organizations should also consider how internal data relates to goals and convergence with third-party data as both identifying relevant alt-data to converge with real estate data, and implementation is likely to be challenging.

### *Organizational Considerations and Implementation Challenges*

From an organizational perspective, real estate also faces industry-specific challenges, as the decentralized and relationship-based model of real estate makes data collection more difficult. Adding specific challenges is the heterogenous nature of real estate—all real estate is local, and analysis requires local context.

Decentralized organizations for real estate management, leasing and transaction and dependence on personal networks results in losses of valuable information. Digitalized records of the sales, leasing, and funding process in real estate would enable knowledge retention and give firms the ability to identify larger organizational patterns when analyzed at scale.

Centralized data would enable for organizations to apply algorithms that detect patterns in activities such as leases and property management.[[45]](#footnote-45) New capabilities that make it possible to structure unstructured data and create insights from activities within an organization could shift some of the relational capital held by employees to the organization.[[46]](#footnote-46) This could have a truly transformative impact on the real estate industry. Identification of patterns of organizational activity can also create significant efficiency improvements and better risk management—with one example being identifying dependence of certain employees and networks.

Developing these kinds of capabilities, however, ultimately requires an organization to first identify its goals and develop data competencies in tailored ways that support these objectives. However, such developments cannot be completed without recognition and consideration of the increasing levels of digitization occurring in urban environments. In many ways, these aforementioned strategies require firms to explore novel ways of extracting value from urban digitalization, both in the public and private sector. In this light, it is important to understand the role of cities as reservoirs of both traditional, real estate market data and other big/alt-data sources. In the next section, we highlight this increasing importance of cities and examine a handful of cases which push the boundaries of digital integration.

## Understanding the Role of Cities

The importance of data is closely related to that of cities and their increasing dominance in terms of population, economic output, and technological development. For example, in the United States, 50% of employment growth between 2010 to 2017 went to cities with only 30% of the population.[[47]](#footnote-47) Across the globe, cities are experiencing similar growth and contributing disproportionately to GDP—i.e. by 2025, 600 cities with 25% of the global population are predicted to account for 60% of global GDP.[[48]](#footnote-48)

Because social and economic activities that create data are a predominantly urban phenomenon, governments are increasingly paying attention to the value that data can provide for improving transportation system efficiency, supporting better urban planning, and increasing the efficacy of other city services.[[49]](#footnote-49) Today, a variety of cities are already exploring ways to extract value from their urban environments using digitalization. For example, in the U.S. alone, the number of “smart city” projects is projected to grow at 30% annually until 2030.[[50]](#footnote-50) However, these numbers are currently dwarfed by private R&D investment for data collection in cities.[[51]](#footnote-51)

Given these trends, we highlight in the next section a handful of cases which are pushing the boundaries of digital integration and embracing radically different digital cities strategies. To unpack this further, we next review a variety of digitalization efforts going on in cities around the world.

### *Existing Examples of Digitalized Environments*

Across the global, many urban centers are undergoing major digital transformations. Although there remains relatively little consensus around the meaning of a “smart,” “digital,” or “intelligent” city, efforts have emerged in recent years to track and ranking the relative “smartness” or “digitalization” of cities. For example, the 2016 European Digital City Index (EDCi), developed by the European Digital Forum, aimed to describe “how well different cities across Europe support digital entrepreneurs.”[[52]](#footnote-52) Likewise, the IMD Smart City Index 2019, which ranks 102 cities worldwide, “focuses on how citizens perceive the scope and impact of efforts to make their cities ‘smart’”[[53]](#footnote-53) while Bloom Consulting’s Digital Cities Index views digital cities as locations which have “triggered proactive interest from Global Citizens (tourists, investors and general citizens) towards that City in the Digital World.”[[54]](#footnote-54)

Naturally, none of these indices measure the same thing. However, they do signal a growing trend among cities around the world to begin digitizing and modernizing their urban environments in ways that improve the livability and quality of life in these settings. And yet, today only a handful of cities currently (or plan to) offer radically new approaches to technological integration in the built environment. Because time and space requirements preclude us from documenting ever illustrative case of radical urban digitalization, we chose to highlight a select number of greenfield and brownfield developments across various geographic regions, including Asia, the Middle East, and North America, which offer different levels of urban digitalization. The cases discussed include:

1. Toronto (Eastern Waterfront), Canada;
2. Belmont, Arizona (USA);
3. Songdo, South Korea;
4. Dubai, United Arab Emirates (UAE); and
5. Neom, Saudi Arabia.

We next discuss the developments of each of these cities in turn.

#### 3.1.1 Toronto (Eastern Waterfront), Canada

In partnership with the tri-government agency Waterfront Toronto and the local community, Sidewalk Labs (a subsidiary of Alphabet) is currently developing a 12-acre ‘smart city’ project in Toronto’s Quayside district. Although the project launched in 2017 and significant public outcry seemed to put the development in jeopardy, Sidewalk Toronto continued progression is widely regarded as the “proving ground for the latest thinking in sustainable design and technology integration into urban planning.”[[55]](#footnote-55) Meant to serve as a global hub for urban innovation, Sidewalk Labs’ Master Innovation and Development Plan (MIDP) for Toronto’s Eastern Waterfront boasts a wide variety of potential benefits. For example, the development expects to cut greenhouse gas emissions by 89%, enable three quarters of all mobility trips by transit, walking, or cycling, and create 44,000 direct jobs as well as $14.2 billion in annual economic impact by 2040.[[56]](#footnote-56)

For housing and real estate, Sidewalk Toronto is also exploring new construction methods and flexible building designs that will enable radical mixed-use, walkable neighborhoods and drastic cost reductions for both housing and retail space. For example, Sidewalk Toronto’s innovation plan[[57]](#footnote-57) includes:

Building design innovations that could accommodate the full range of live-work needs and respond nimbly as those needs change. These include adaptable “Loft” spaces — supported by flexible interior panels and a real-time code-monitoring system — designed to cut renovation times and help communities retain a lively mix of businesses and residents. For homes in particular, efficient units and co-living spaces could improve affordability while expanding options for all types of households.

Combined with other economic and sustainability-focused initiatives, the Quayside development expects to ultimately offer 40% of housing units at below-market rates. Moving forward, Sidewalk Labs and its partners plan to contribute up to $1.3 billion in funding and financing for the ‘smart city’ in hopes of catalyzing $38 billion in investment by other third parties. By undertaking this venture, Sidewalk Labs ultimately intends to earn its profits from real estate development, fees, and interest on infrastructure finance.

#### 3.1.2 Belmont, Arizona (USA)

Out in the Arizona desert just 45 miles west of Phoenix, technology entrepreneur and Microsoft founder Bill Gates is also developing a new ‘smart city.’ With Gates’ investment firm, Cascade Investment LLC, already committing $80 million to the project, Belmont Arizona is expected to cover 24,800 acres—3,800 devoted to office, commercial, and retail space and 470 acres aside for public schools, house 80,000 residences, and accommodate a potential population size of about 182,000, about the same size as Tempe.[[58]](#footnote-58) Designed with technology at its core, Belmont will feature “accommodation for autonomous vehicles, high-speed digital networks, high tech data centers, new manufacturing technologies and autonomous logistics hubs.”[[59]](#footnote-59) When completed, Belmont Partners claim the city "will [ultimately] transform a raw, blank slate into a purpose-built edge city built around a flexible infrastructure model."[[60]](#footnote-60)

#### 3.1.3 Songdo, South Korea

From its conception in 2001, the International Business District (IBD) of Songdo City in South Korea was slated to be “a sustainable, low-carbon, and high-tech utopia.”[[61]](#footnote-61) Built from scratch on land reclaimed from the Yellow Sea, the 1,500-acre development’s proximity to Seoul, the international airport, and the sea makes it both a transportation hub and the “gateway to Korea.” Designed with an eye toward sustainability, the city’s buildings and streets teems with sensors—over 20 million square feet of LEED-certified space—which monitor everything from traffic flow to energy usage.[[62]](#footnote-62) The colocation of most offices, schools and non-residential buildings to apartment buildings makes accessibility another key feature of the city. Because IBD’s design maximizes public transport and walkability, approximately 40% of Songdo remains dedicated to green space. Additionally, trash collection is conducted by pneumatic tubes that send waste from homes to an underground waste facility, where it’s sorted, recycled, or converted into energy. Everything from the lights to temperature in homes and apartments can also be adjusted via a central control panel or mobile device. (Poon 2018). Originally planned for completion in 2015, the $40 billion development project, led by Gale International, still remains a work in progress and the city its continuing to attract residents in pursuit of its original population goal of 300,000 goal.

#### 3.1.4 Dubai, UAE

In 2013, Shaikh Mohammad bin Rashid Al Maktoum, the vice president of UAE, initiated the Smart Dubai project. This initiative consists of over 100 projects aimed at making Dubai a ‘smart city’ by 2030. With IoT-enabled integration of the public and private sectors at its core, Smart Dubai boasts a host of digitalization initiatives. For example, the UAE’s Dubai Autonomous Transportation Strategy aims to institute driverless transit, fully digitalized government, business, customer transactions, and 5000 mobile hotspots by 2021.[[63]](#footnote-63) The city already hosts two mobile applications, mPay and DubaiNow, which offer integrated payment platforms for services ranging from traffic fines and utility bills to educational, healthcare, transportation, and business-related items. Likewise, the Smart Nol Card offers citizens a unified payment platform for all transportation services, including but not limited to the metro, buses, taxis, etc. Moreover, the city has started using Radio Frequency Identification (RFID) to levy road tolls and assigns each building a unique, Quick Response (QR) code that contains details about the building’s plot, location, eServices offered by the municipality.[[64]](#footnote-64)

In real estate, Dubai also now hosts a suite of technological services and innovations. For example, the Dubai Land Department (DLD) now offers an Investment Map Portal which provides investors with information on ongoing real estate projects. By streamlining transactions, this tool increases transparency for relevant stakeholders and allows investors to make more informed, data-driven decisions. Likewise, DLD also offers a trust account system app (Dubai Brokers) and Smart Pay Service to handle property/project registrations and producer payment processes more efficiently. Finally, the Dubai Financial Services Authority (DFSA) regulates a crowdfunding platform called Smart Cloud which uses a screening tool and multiple other data points to assess the relative attractiveness of real estate investment opportunities. Once fund have been raised for a dedicated project, a Special Purpose Vehicle (SPV) is formed and shares are allocated proportionally to investors.

#### 3.1.5 Neom, Saudi Arabia

Finally, Saudi Arabia is building a new, futuristic mega-city on the Red Sea coast in the northwestern Saudi province of Tabuk. This 16-borough city is expected to cover 10,230 square miles, be 33 times larger than New York City, and cost Saudi Arabia's Public Investment Fund at least $500 billion, plus millions more in foreign investment (if it can raise it). Dubbed by some as "the world's most ambitious project," Neom is integral to the country’s Vision 2030—an ambitious agenda to revolutionize Saudi society, diversity its economy, and develop the kingdom into a technological hub.[[65]](#footnote-65) Accorinding to Nadhmi Al-Nasr, CEO of Saudi Arabia’s landmark project, Neom is “set to become the first completely digitalized city in the world,” connecting different destinations with big data and AI. Work on the city’s overall strategy is also near completion and will be unveiled in March 2020.[[66]](#footnote-66) However, the full build-out of the city is still years away. Although construction work has already begun on Neom Bay (phase one of project) and the new airport is nearly finished, Crown prince Mohammed bin Salman expects Neom to be completed by 2025 at the earliest.[[67]](#footnote-67)

## Data Applications and Use-Cases in Urban Settings

As more and more urban centers like the examples provided above undergo digital transformation, city planners, real estate developers, and investors stand to benefit significantly from untapped value and insights embedded in new, non-traditional/alternative data streams. As stated above, both big data and alt-data now make it possible to identify both social and commercial activities in a city so that various urban phenomena—notably real estate markets—can be predicted. It is now possible to predict GDP based on satellite imagery,[[68]](#footnote-68) local voting patterns through street views,[[69]](#footnote-69) and the likelihood of installing solar panels.[[70]](#footnote-70) Previous studies have also shown that a Machine Learning algorithm trained on Google Street View imagery could do out-of-sample prediction of income in both New York and Boston with explanatory power exceeding of 0.71 and 0.77 respectively.[[71]](#footnote-71) Similarly, when data on perceived safety is linked to imagery, it can be predicted with a high degree of accuracy.[[72]](#footnote-72) At the Stanford Global Projects Center, one project is currently developing these capabilities using an “Urban-to-Vector” method which links any urban characteristic to a vector of values. Currently, socio-economic characteristics of a neighborhood can be predicted with a high degree of accuracy when this framework is applied to Street View Imagery and Yelp Reviews. Similarly, using the same vectors, every neighborhood in New York can be assigned its most similar “twin” neighborhood in Chicago, matched only using Street View. Figure 1 below shows preliminary results of this analysis.

A picture containing newspaper, text

Description automatically generated

**Figure 1:** Similarity Analysis Using Urban-to-Vector on Street View Imagery and Yelp Data

As models predicting current real estate values begin to offer higher degrees of accuracy,[[73]](#footnote-73) we believe that there are a few main applications of non-traditional urban data that are high-value propositions. They include:

1. prediction of change using advanced analytics,
2. understanding the determinants of change, and;
3. enabling market transparency

Prediction of change in urban environments is dependent on innumerable intertwined developments that can be measured with IoT data—so that the when, where and who of local demand can be determined. Local changes in foot-traffic, spending and online-ordering are examples that can tell us the trajectory of a neighborhood. In digitalized cities, patterns and correlations can be identified and used for “what-if?” analysis so that the impact of policies, real estate developments can be modeled using predictive analytics.

### *The Potential Power of Predictive Analytics in Real Estate*

Today, prediction using both traditional and big/alt-data is used by both professionals and academics across a variety of industries. For instance, start-ups such as Descartes Labs and Orbital Insight utilize aerial images and other alternative data sources to help investors pick stocks. Likewise, some investors are now using real-time satellite images to count cars in mall parking lots and predict retailers’ sales.[[74]](#footnote-74) Across the real estate sector, a significant part of market price and performance predictions can now come from alternative data sources. McKinsey estimates this share of predictive power to be ~60%.[[75]](#footnote-75)

Because thousands of non-traditional variables can now be linked to a variety of diverging yet location-specific outcomes, it is now possible to apply and scale advanced analytics using alternative data sources to a variety of scenarios. For example, when forecasting rents, real estate investors and developers can also now:

1. stress-test expectations for individual properties within a variety of market conditions in order to improve their investment and divestment decisions;
2. identify individual assets that will maintain their value in volatile market conditions;
3. optimize capital expenditure decisions for specific properties; and
4. compare predictive-model outputs to traditional market forecasts from brokers.[[76]](#footnote-76)

For many in the real estate industry, the growing proliferation of data sources is making the accessibility and processing of these plentiful data streams cheaper and more user friendly. For example, in the last half decade alone, the tally of large-scale alt-data vendors has increased from only a few dozen to several hundred.[[77]](#footnote-77) As these types of data steadily becomes more mainstream, many organizations are beginning to view them as “untapped alpha.”[[78]](#footnote-78) Today, real estate developers, investors, and managers need to probe causality more deeply because:

correlations in conventional datasets often break down over longer horizons and typically do not reflect the entire spectrum of events that could occur over long periods of time. Alt-data can (partly) mitigate these shortcomings by supplying more context about how events in the wider world drive downside moves in markets.[[79]](#footnote-79)

Additionally, they can rely on novel forms of data, both traditional and alternative, to strategically inform their investment, development, and management decisions. For example, a firm that directly owns real-estate development projects in an emerging market might use aircraft or satellite images of shadow lengths from construction sites to measure the pace of their projects. This capacity, developed by Orbital Insights, may give planners, developers, and investors much greater clarity about the development progress and management of these assets. Likewise, other real estate firms may be able to leverage online price and ratings histories from sources like Airbnb, TripAdvisor, or Yelp to conduct due diligence on their competitors and candidate investments in leisure-related properties (e.g., hotels or casinos).[[80]](#footnote-80) Companies may also be able to predict crime and perceived safety in various neighborhoods using Google Street View,[[81]](#footnote-81) understand nuances in market rents using Starbucks locations,[[82]](#footnote-82) or predict future store closings.[[83]](#footnote-83) However, discriminating and “selecting which [data] deserve resources (e.g., money to acquire; time to store, prepare, and analyze; and capacity to be governed) . . . requires characterizing [datasets] to establish which will be most valuable for organizational needs.”[[84]](#footnote-84) Naturally, discriminating various forms of data for valuation, acquisition, storage, and analysis is a difficult task. For this reason, characterization and interrogation of uses cases becomes a crucial yet achievable step that makes it possible to determine whether certain datasets aligns with organizational capabilities and priorities. In the next section, we describe a handful of use cases which demonstrate the power of predictive analytics in real estate.

### *Use-Cases of Predictive Analytics in Real Estate*

At the Disruptive Technologies and Digital Cities Program at Stanford, we have identified several use cases were data can provide tangible value for real estate companies. These alt-data use-cases have emerged from research projects in the program and face varying needs of development in terms of data, analysis, and implementation. In the examples to follow, we describe the values and challenges to implementation of each of these use cases.

#### Predicting Store-Front Demand and Closures Using Street View Imagery

One recent project Stanford has been exploring follows the approach of studies that match Google street view imagery with geo-coded data and apply machine learning algorithms to make predictions.

In terms of implementation, this type of analysis could have significant value when real estate organizations converge internal data and third-party data—allowing for the prediction of rents, gentrification, shifts in demand and localization. Examples of these types of studies include predictions of local level crime using digitalized criminal records[[85]](#footnote-85) or the presence of a McDonalds.[[86]](#footnote-86) Predicting likelihoods for certain retail services and then comparing them with what is currently present can provide valuable insight into what to develop. The ability to quickly assess risk-characteristics of retail locations would be highly valuable for both real estate management, appraisal, and development.

The basic premise of this type of analysis involves the convergence of imagery and algorithmic development. This is also a good illustration of the value of improved internal, within-industry real estate data because building-level data (such as leases) is a key component in the analysis. Alt-data—in this case street-view imagery—then builds on these measures of market activity.

However, one data challenge associated with this type of analysis is availability. For instance, data on store closures is not readily available in urban settings. Companies that do track store closures across the U.S. do so manually and only cover big-box chains. That means there is no easily available geocoded data covering the closure of an independent store in San Francisco or restaurant in New York.

#### Tracking Pricing of Consumer Goods to Identify Urban Change

Identical goods oftentimes sell for different prices—for reasons such as local cost-levels and demand. This can tell us something about a neighborhood, and if tracked over time, provide insights about urban change.

On such example is gas-prices, which can exhibit large differences within a small geography. Understanding patterns of gas-prices across neighborhoods and changes of these patterns could provide insight into the trajectory of an area, price sensitivity, and the overall economy. For instance, decreasing gas-margins may potentially indicate a slower local-economy while areas with smaller spreads in gas-prices may have a more homogenous day-time population versus neighborhoods with larger spreads.

When converged with other local data, this type of alt-data can provide valuable insights into local market demand, as illustrated in the “Urban-to-Vector” framework above. Access to this type of data would, however, require partnering with platform-providers or retail chains, thereby illustrating the need for partnering when building alt-data capabilities.

## A Strategic Framework for Characterizing Data in the Real Estate Sector

The use cases outlined above offer useful input for real estate organizations to consider when evaluating their data strategies. However, as noted by earlier research, developing processes that relate data strategies to organizational objectives and financial returns is no easy task. We therefore propose a framework which disaggregates real estate value into tiers of data streams that are connected to the components of real estate value—e.g. impact on net operating income, risk, cap-rates or prediction (see Figure 2). This supports alignment of data strategies with organizational goals.

**Figure 2:** Disaggregating the Drivers of Real Estate Value

We separate data streams based on their origination, as real estate market data and smart-building data provide valuable insights, on top of which convergence with third-party, alt-data, enables novel analysis. We also emphasize that within-organization and real estate market data provide a foundation that organizations are likely to have greater capabilities to analyze. Initial focus should therefore be on market data collection and smart-building technologies that make current analysis more accurate and granular.

A key aspect when adding alt-data to the analysis is breaking down these aggregated measures of market observables. For instance, in may be possible to break down retail cap-rates into measures of volatility in credit-card spending across a neighborhood or derive office rents from geo-coded tax data that offers information about the state of various industries and the impact on an office building.

### *Aggregated Market Observables*

Analysis of real estate—for purposes spanning development, investment, or management—builds on expected revenue streams. Market analysis in terms of rents, vacancy-rates, cap-rates and operating costs represent aggregated market data that is typically used by real estate professionals. This data stems from aggregation of observed market events such as lease contracts and sales transactions and therefore needs to be collected and categorized to various degrees using processes that often are labor intensive.[[87]](#footnote-87) On a market-level, this data offers neighborhood and property-type information which can save real estate professionals considerable time and effort when gathering market information. The initial stages of prop-tech primarily focused on this form of collection and aggregation of data, with Co-Star a notable example of a widely-used data provider.[[88]](#footnote-88) In terms of innovation, it is likely that the process of collecting and aggregating this form of data will shift from labor intensive, manual market research towards increased automation when processes are digitalized and increasingly advanced algorithms can interpret written content. This will, in turn, make data more accessible, lower barriers of entry, and improve market transparency. Although aggregated market data can stem from both internal and external sources, it is where real estate organizations are likely to have most existing competencies and is therefore a good starting-point to build analytical capabilities.

### *IoT and Smart Building Data*

Besides the greater accuracy, volume, and granularity of real estate market data, real estate assets are also being digitalized with IoT technologies (e.g. sensors) that track when and where tenants use space as well as monitor building component energy usage and potential breakdowns. This type of smart-building data can add granularity to the real estate market data as described above. For instance, this information enables companies to benchmark building profitability more accurately and predict lease renewals/rents based on usage.

Data collection capabilities are typically driven by operational efficiency. As smart-building technologies become pervasive across the real estate industry, competitive advantages will diminish, and they will become a pre-requisite to be competitive in terms of building management and analysis of data for investment and development purposes. Competitive advantages will be gained through early adoption, better data collection or analytics, rather than novel analysis.

What separates this form of data from other novel and alt-data sources is that it is generated at the real estate asset itself. It is also data that real estate organizations are likely to be better prepared to analyze and interpret. Resources will therefore primarily have to be built in terms of technical implementation, including but not limited to the digitalization of real estate assets themselves, the identification of what data to collect, and development of processes to manage its convergence.

### *Third-Party Alt-Data Streams*

Finally, data not typically applied when analyzing real estate can provide novel insights when converged with building-level and real estate market data. Digitalized environments and the Internet-of-Things (IoT) can in theory make it possible to measure everything, everywhere, so that more accurate measurement allows for both prediction and understanding of real estate demand. This suggests that market observables such as rents can be further derived across the value-chain. For instance, deriving retail rents from foot traffic as is already being done by companies such as Safegraph.[[89]](#footnote-89)

Convergence of data sources can make this type of analysis far more detailed, especially when it becomes possible to identify temporal and geographical flows of demand. This may involve converging credit-scores and spending habits with mobility data or tracking individual spending over the course of a day using credit cards transactions. Over-time, time-series analysis would allow real estate firms to identify causal-economic relationships, early signs of urban change, and “what-if” analysis for real estate development.

More advanced analytics would also provide measures of volatile movement and spending patterns over-time and demonstrate the interconnectedness between certain streams of demand and real estate values. When connecting to cap-rates, this would provide better risk measures and scenario-analyses in situations where the impact of various predicted scenarios could be determined more accurately.

When this alt-data is related to drivers of real estate value, organizations are better prepared to understand both how data relates to organizational goals and use data to measure and create insight of causal economic relationships. This is in contrast to traditional data correlations that are far more unstable over time and therefore less useful for longer-term prediction (which is essential considering the long-term nature of real estate). Moreover, these traditional market data streams do not allow for “what if?” scenario-analysis, especially in situations when understanding of *why* a change occurs is lacking.

Naturally, application exhaust (Yelp reviews, Uber ridesharing data), congestion information, noise and pollution monitors, and location-based movement data are all examples of data sources that can provide measures of real estate market activity. Converging this data with building-level and real estate market data is, however, the foremost challenge associated with this form of analysis. That does not mean though this type of analytics is not within reach. Credit-card companies now merge their data with credit-scores. A Stanford research team is also currently creating a measure of wealth through its analysis of cell-phone movement data for 30 million+ users and estimated home values based on the location of where the person spends his or her night. This information is also being mapped with business locations and home values. Matching movement data with geo-corded credit-card transactions could also be done by overlaying several transactions and phone locations based on time and place.

## Conclusions

The nature of real estate—heterogenous assets, high transaction costs, limited transparency and a resulting high value on relationships and networks—is a likely explanation for why the sector has lagged other sectors in innovation and productivity.[[90]](#footnote-90) As noted by Wessel and Christensen, “[t]he more difficult the barrier, or the more barriers a disrupter faces, the more likely it is that customers will remain with incumbents.”[[91]](#footnote-91)Consequently, market participants are likely to preserve with status quo.

Moving forward, it is clear that players in the real estate market are keen and eager to reap the potential benefits of alt-data but remain ill-equipped to independently source, process, and maintain real estate market data, building-level data, and alt-data sources from urban settings. We believe that data-driven data accessibility will have a transformative impact on the real estate industry. Notably, the value of networks and relationships will decrease as information becomes more readily available, thereby making markets more efficient, decreasing barriers of entry and lowering transaction costs. Organizations with business models that rely on providing services which value is dependent on relationships and informational advantages will see the value of their offerings decrease. Increased transparency is driven by the development of tools for collection and categorization of data so that current forms of market data can be collected in a faster, cheaper and more accurate way.

Another driver of increased availability of real estate data is the development of smart-building technology and IoT, enabling data on building usage, maintenance, social- and commercial activities to be measured. In light of these shifts, real estate organizations need to build capacity to improve within-organization data origination and analysis of smart-building data and have a strategy for acquisition of third-party alt-data. Organizations are more likely to have better capacity to process and analyze the former, while novel alt-data is far more complex to analyze. Similarly, current streams of real estate data can be improved by IoT technologies that offer increased accuracy, volume, and granularity.

Improved real estate market data is the foundation on which third-party alt-data builds upon. Converging real estate data with alt-data such as imagery, movement data, application exhaust, and digitalized records can result in novel analysis that provides real estate organizations with a competitive advantage for investment, development, and risk-management. A distinction between real estate market data and alt-data is that the former is likely to be a necessity for all real estate organizations as the entire industry is disrupted. Competitive advantages on this front will thus be provided by first movers and likely more temporary in nature. Novel analysis using alt-data, however, is far more complex and can offer competitive advantages through insights that are accessible to fewer market participants.

Thus, when building organizational capacity, we emphasize that real estate organizations should focus on both internal organizational data and third-party data. Instead of surrendering these functions to external managers and consultants, real estate developers, investors, and managers, should partner for access to third-party alternatives data sources and use these partnerships to capitalize on the efficiencies of organizing and (pre-) processing alt-data. Naturally, using alt-data in the real estate industry comes with its own reservations about the uncertainty and murkiness of third-party data. However, such concerns should not discourage organizations from prioritize partnerships with alt-date platform providers/advisors, such as Safegraph, Neudata, and Quandl. Although point vendors may offer more limited, scarce, and/or proprietary data streams, platform vendors usually offer a much wider variety of alt-datasets that can be integrated or aided by complementary analytical tools. Overall, building internal capacity around alt-data through partnering mitigates or eliminates problems with relying on external managers, improves transparency, and limits an organization’s susceptibility to opportunistic tendencies. Moreover, by experimenting with different alt-data forms via platform products, real estate investors, developers, and managers have the ability to experiment with more configurations of alt-data when trying to incorporate these feeds into their strategies. In the end, this increases the odds of finding data that will generate meaningful insights produce both financial and operational value.[[92]](#footnote-92)

1. cbcasady@stanford.edu [↑](#footnote-ref-1)
2. hdonner@stanford.edu [↑](#footnote-ref-2)
3. Manyika and Roxburgh, 2011, McKinsey [↑](#footnote-ref-3)
4. Huawei & Oxford Economics. (2017). [Digital Spillover. Measuring the true impact of the Digital Economy](https://www.huawei.com/minisite/gci/en/digital-spillover/files/gci_digital_spillover.pdf). [↑](#footnote-ref-4)
5. Paul, F. 2019. “Six IoT predictions for 2019”. 2019. Networkworld [↑](#footnote-ref-5)
6. Paul, F. 2019. “Six IoT predictions for 2019”. 2019. Networkworld [↑](#footnote-ref-6)
7. Hashem et al. 2016 [↑](#footnote-ref-7)
8. Donner, Steep and Peterson 2019 [↑](#footnote-ref-8)
9. 6 Keys to Credit Risk Modeling in the Digital Age. SAS [↑](#footnote-ref-9)
10. Donner, H. Steep, M. and Peterson, T. 2019. “Crossing the Urban Datalayer: Mobility as a Data Generating Activity”. Global Projects Center, Working Paper Series. [↑](#footnote-ref-10)
11. Bigdeli, A. Z., Baines, T., Bustinza, O. F., & Shi, V. G. (2015). Holistic approach to evaluating servitization: A content, context, process framework. *EUROMA. Anais*. [↑](#footnote-ref-11)
12. Dameri and Cocchia (2013), Cocchia (2014) and Yin et al (2015). [↑](#footnote-ref-12)
13. Yin et al. 2015 [↑](#footnote-ref-13)
14. Yin et al. 2015 [↑](#footnote-ref-14)
15. Monk, A., Prins, M. and Rook, D. 2019. “Rethinking Alternative Data in Institutional Investment. *The Journal of Financial Data Science.* [↑](#footnote-ref-15)
16. [↑](#footnote-ref-16)
17. [↑](#footnote-ref-17)
18. Barkham, R., Bokhari, S. and Saiz, A. 2018. “Urban Big Data: City Management and Real Estate Markets”, Massacussets Institute of Technology / CBRE. [↑](#footnote-ref-18)
19. Zamani, M., and Schwartz, H. A. 2017. "Using twitter language to predict the real estate market." *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. [↑](#footnote-ref-19)
20. Monk, A., Prins, M. and Rook, D. 2019 “Rethinking Alternative Data in Institutional Investment. *The Journal of Financial Data Science.* [↑](#footnote-ref-20)
21. Donner, H. Eriksson, K. and Steep, M. 2018. “Digital Cities: Real Estate Development Driven by Big Data”. Global Projects Center, Working Paper Series. [↑](#footnote-ref-21)
22. As of February 3, 2019. The world’s most valuable companies are Apple, Alphabet (Google), Microsoft, and Amazon. Microsoft was founded in 1975, Apple in 1976, Amazon in 1994, Google in 1998, and Facebook in 2004. As of 2019, the average age of these companies is 30 years. [↑](#footnote-ref-22)
23. Trieu, V-H. 2017. ”Getting value from Business Intelligence systems: A review and research agenda”.

    *Decision Support Systems*. Vol. 93. pp 111-124. [↑](#footnote-ref-23)
24. Hui, G. 2014. ”How the Interent of Things Changes Business Models”. *Harvard Business Review*. [↑](#footnote-ref-24)
25. Gartner, 2015 [↑](#footnote-ref-25)
26. Monk, A., Prins, M. and Rook, D. 2019. “Rethinking Alternative Data in Insititonal Investment. *The Journal of Financial Data Science.* [↑](#footnote-ref-26)
27. Digital Disruption in Commercial Real Estate – Catalyst for Growth? 2016. Deloitte Canada. [↑](#footnote-ref-27)
28. Donner, Steep, and Eriksson (2018) covers this topic and relates the economics literature on knowledge spillovers to ICT technologies. [↑](#footnote-ref-28)
29. Donner, Steep and Eriksson, 2018 [↑](#footnote-ref-29)
30. Mayer, 1995, 1998 [↑](#footnote-ref-30)
31. Barkham, R., Bokhari, Sh. And Saiz, A. 2018. Urban Big Data: City Management and Real Estate Management. *MIT Center for Real Estate.* [↑](#footnote-ref-31)
32. Ibid. [↑](#footnote-ref-32)
33. Glaeser, E.L., Kominers, S.D., Luca, M. and Naik, N., 2018. Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry*, *56*(1), pp.114-137 [↑](#footnote-ref-33)
34. Donner, H., Eriksson, K., and Steep, M. 2018. Digital Cities: Real Estate Development Driven by Big Data. Global Projects Center, Working Paper Series. [↑](#footnote-ref-34)
35. Kitchen, R. 2014. The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. *Sage Publications: London.* [↑](#footnote-ref-35)
36. This refers to how detailed and granular the data is. [↑](#footnote-ref-36)
37. How much overlap the data has with other datasets in the organization. [↑](#footnote-ref-37)
38. How easy it is to add data-points to the data. [↑](#footnote-ref-38)
39. Ibid. [↑](#footnote-ref-39)
40. [↑](#footnote-ref-40)
41. Friedman, A. 2015. The Future Of Search Engines Is Context. *Search Engine Land*. [↑](#footnote-ref-41)
42. Reilly, M. 2017. Is Facebook Targeting Ads at Sad Teens?. *MIT Technology Review*. [↑](#footnote-ref-42)
43. Spenner, P. and Freeman, K. 2012. To Keep Your Customers, Keep It Simple. *Harvard Business Review.* [↑](#footnote-ref-43)
44. Glaeser, E.L., Kominers, S.D., Luca, M. and Naik, N., 2018. Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry*, *56*(1), pp.114-137 [↑](#footnote-ref-44)
45. A potential leasing application would be to create a similarity-index between tenants based on their industry, location and size which in turn could predict likelihoods of lease renewals—for if law-firms in building X have been downsizing it is likely that law-firms in building Y will downsize once the lease expires. [↑](#footnote-ref-45)
46. One example of technologies that enable this kind of shift is Natural Language Processing (NLP). Development of algorithms that analyze e-mail correspondence can provide significant value—i.e. letting a real estate company know patterns of key people in a certain metropolitan area, aggregating which people their employees reach out to regularly, and the outcomes of these interactions. Srivastava, Goldberg, Manian and Potts (2018) demonstrated the power of NLP analysis in their analysis of over 10 million emails within an organization. This analysis, among other things, determined if an employee was a cultural fit for the organization, predicted if an employee would stay in their employment, and estimated if a departure was going to be a voluntary or involuntary Sameer B. Srivastava, [Amir Goldberg](https://www.gsb.stanford.edu/faculty-research/faculty/amir-goldberg), V. Govind Manian, Christopher Potts. *Management Science*. 2018, Vol. 64, Issue 3, Pages 1348-1364). [↑](#footnote-ref-46)
47. Muro, M. and Whiton, J. 2018. Geographic gaps are widening while U.S. economic growth increas- es.

    The Brookings Institute. [↑](#footnote-ref-47)
48. Dobbs, R. Smit, S. Remes, J. Manyika, J. Roxburgh, C. and Restrepo, A. 2011. “Urban world:

    Mapping the economic power of cities “ McKinsey Global Institute. [↑](#footnote-ref-48)
49. The Future of Transportation: Mobility in the Age of the Megacity. 2019. Visa Inc.; Glaeser, E.L., Kominers, S.D., Luca, M. and Naik, N., 2018. Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry*, *56*(1), pp.114-137. [↑](#footnote-ref-49)
50. O’Rourke, B. 2019. “US Smart Cities Projects to Push Annual Revenue Over 700 Million in Five-Year Outlook. S&P Global. [↑](#footnote-ref-50)
51. Steep, M. 2019. Lecture. Stanford Disruptive Technologies and Digital Cities. [↑](#footnote-ref-51)
52. See <https://digitalcityindex.eu/>. [↑](#footnote-ref-52)
53. See <https://www.imd.org/smart-city-observatory/smart-city-index/>. [↑](#footnote-ref-53)
54. See <https://www.digitalcityindex.com/city-index-results-World>. [↑](#footnote-ref-54)
55. Shieber, J. 2019. [Sidewalk Labs (Alphabet’s grand experiment in smart cities) will move forward with Toronto project.](https://techcrunch.com/2019/10/31/sidewalk-labs-alphabets-grand-experiment-in-smart-cities-will-move-forward-in-toronto/) *TechCrunch.* [↑](#footnote-ref-55)
56. Sidewalk Toronto. 2019b. [Toronto Tomorrow: A new approach for inclusive growth](https://www.sidewalktoronto.ca/). Sidewalk Labs. [↑](#footnote-ref-56)
57. Sidewalk Toronto. 2019a. [The Innovation Plan](https://www.sidewalktoronto.ca/innovations/buildings-housing). Sidewalk Labs. [↑](#footnote-ref-57)
58. Weller, Chis. 2017. [Bill Gates is building a smart city in Arizona.”](https://www.weforum.org/agenda/2017/11/bill-gates-is-building-a-smart-city-in-arizona) *World Economic Forum.*  [↑](#footnote-ref-58)
59. SmartCityAZ 2019. [The Smartest City is Coming Soon](http://smartcityaz.com/). SmartCity, AZ. [↑](#footnote-ref-59)
60. Weller, Chis. 2017. [Bill Gates is building a smart city in Arizona.”](https://www.weforum.org/agenda/2017/11/bill-gates-is-building-a-smart-city-in-arizona) [↑](#footnote-ref-60)
61. Poon, L. 2018. [Sleepy in Songdo, Korea’s Smartest City](https://www.citylab.com/life/2018/06/sleepy-in-songdo-koreas-smartest-city/561374/). *CityLab.*  [↑](#footnote-ref-61)
62. This is the highest concentration of LEED-certified projects in the world. [↑](#footnote-ref-62)
63. KPMG. 2015. [Dubai - a new paradigm for smart cities.”](https://assets.kpmg/content/dam/kpmg/pdf/2016/04/Dubai-a-new-paradigm-for-smart-cities-uae.pdf) KPMG; Smart Dubai. 2019. [Smart Dubai 2021](https://2021.smartdubai.ae/). [↑](#footnote-ref-63)
64. United Arab Emirates (UAE). 2019. [Smart Dubai.”](https://government.ae/about-the-uae/digital-uae/smart-dubai) The Official Portal of the UAE Government. [↑](#footnote-ref-64)
65. Bostock, B. 2019. [Everything we know about Neom, a 'mega-city' project in Saudi Arabia with plans for flying cars and robot dinosaurs.](https://www.businessinsider.com/neom-what-we-know-saudi-arabia-500bn-mega-city-2019-9) *Business Insider.*  [↑](#footnote-ref-65)
66. [Saudi Arabia's NEOM Will Be The First Totally Digitized City In World](https://www.abouther.com/node/26506/people/features/saudi-arabias-neom-will-be-first-totally-digitized-city-world). 2020. *About Her.* [↑](#footnote-ref-66)
67. Bostock, B. 2019. [Everything we know about Neom, a 'mega-city' project in Saudi Arabia with plans for flying cars and robot dinosaurs.](https://www.businessinsider.com/neom-what-we-know-saudi-arabia-500bn-mega-city-2019-9) *Business Insider.* [↑](#footnote-ref-67)
68. Sutton, P.C., Elvidge, C.D. and Ghosh, T., 2007. Estimation of gross domestic product at sub-national scales using nighttime satellite imagery. *International Journal of Ecological Economics & Statistics*, *8*(S07), pp.5-21. [↑](#footnote-ref-68)
69. https://techcrunch.com/2018/01/01/researchers-can-now-make-neighborhood-voting-predictions-from-google-street-view-images/ [↑](#footnote-ref-69)
70. https://news.stanford.edu/2018/12/19/inventory-indicates-goes-solar/ [↑](#footnote-ref-70)
71. Glaeser, E.L., Kominers, S.D., Luca, M. and Naik, N., 2018. Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry*, *56*(1), pp.114-137 [↑](#footnote-ref-71)
72. Naik, N., Philipoom, J., Raskar, R. and Hidalgo, C., 2014. Streetscore-predicting the perceived safety of one million streetscapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 779-785). [↑](#footnote-ref-72)
73. Such as Donner, Song and Wilhelmsson (2016), having a R2 of .91 and .88 when analyzing residential apartment and single-family house prices, respectively. [↑](#footnote-ref-73)
74. <https://www.theatlantic.com/magazine/archive/2019/05/stock-value-satellite-images-investing/586009/> [↑](#footnote-ref-74)
75. Asaftei, G., Doshi, S., Means, J., and Sanghvi, A. 2018. Getting ahead of the market: How big data is transforming real estate. Mckinsey. [↑](#footnote-ref-75)
76. Ibid. [↑](#footnote-ref-76)
77. Monk, A., Prins, M., and Rook, D. 2019. [Rethinking Alternative Data in Institutional Investment](https://caia.org/sites/default/files/014-031_monk_jfds.pdf). *The Journal of Financial Data Science*, 1(1), pp.14-31. [↑](#footnote-ref-77)
78. See <https://www.quandl.com/>, a leading alt-data platform operator. [↑](#footnote-ref-78)
79. Monk, A., Prins, M., and Rook, D. 2019. [Rethinking Alternative Data in Institutional Investment](https://caia.org/sites/default/files/014-031_monk_jfds.pdf), pp. 4-5. [↑](#footnote-ref-79)
80. Ibid. [↑](#footnote-ref-80)
81. Dormehl, L. 2014. [This Algorithm Predicts A Neighborhood’s Crime Rate Using Google Street View](https://www.fastcompany.com/3036677/this-algorithm-knows-your-neighborhood-better-than-you-do). *Fast Company.* [↑](#footnote-ref-81)
82. Donner, H. and Loh, T.H., 2019. Does the Starbucks effect exist? Searching for a relationship between Starbucks and adjacent rents. *Property Management*. [↑](#footnote-ref-82)
83. See, e.g. <https://blog.thinkgeo.com/2019/03/21/thinkgeos-2019-u-s-retail-closings-map-how-we-built-it/> and <https://www.aggdata.com/future-store-closings>. [↑](#footnote-ref-83)
84. Monk, A., Prins, M., and Rook, D. 2019. [Rethinking Alternative Data in Institutional Investment](https://caia.org/sites/default/files/014-031_monk_jfds.pdf), pp. 7. [↑](#footnote-ref-84)
85. https://www.fastcompany.com/3036677/this-algorithm-knows-your-neighborhood-better-than-you-do [↑](#footnote-ref-85)
86. https://www.fastcompany.com/3036677/this-algorithm-knows-your-neighborhood-better-than-you-do [↑](#footnote-ref-86)
87. e.g. by property type and geography. [↑](#footnote-ref-87)
88. [↑](#footnote-ref-88)
89. https://www.safegraph.com [↑](#footnote-ref-89)
90. Digital Disruption in Commercial Real Estate – Catalyst for Growth? 2016. Deloitte Canada. [↑](#footnote-ref-90)
91. Wessel, M. and Christensen, C. 2012. Surviving Disruption. Harvard Business Review. [↑](#footnote-ref-91)
92. Ibid. [↑](#footnote-ref-92)