# City Data Tools:
# Build or Buy?

STAE

# City Data Tools: Build or Buy?

## Preparing your city for the rise of real-time data

"Open data portals in cities are not a new thing, but many portals today have limited machine readability and therefore limited business value... the city becomes 'smart' when the data is collected and governed in a way that can produce valuable real-time streams, rather than just backward-looking statistics or reports."

**GARTNER, DECEMBER 2016[1]**

### MORE DEVICES, MORE DATA

Smart infrastructure is coming to cities --- and fast. **By 2020, urban IoT installations will grow to 9.7 billion devices, up from 1.7 billion in 2017** (an increase of over 470 percent).[2] These technologies hold the promise of changing cities for the better -- making them more productive, efficient and enjoyable places to live. These benefits can only be realized when a city is able to ingest, store, distribute, analyze and act upon the large amount of data generated by smart infrastructure.

**New smart infrastructure devices will reach their full potential only by communicating to each other in real-time.** For example, smart traffic lights can help alleviate traffic on their own, but when they can communicate in concert with public safety software and emergency response vehicles, they can do much more -- from dynamically rerouting traffic in response to an accident, to clearing the road for emergency response teams to arrive as quickly as possible.

Most cities are unprepared to handle multiple streams of real-time data to gain insights and take action. Existing data platforms and open city data portals are often difficult to leverage as data sets are irregularly updated, only available as flat files and cannot be easily combined. Furthermore, **if cities are unprepared for data to be centralized and real-time, they may find themselves with multiple vendor-specific tools that cannot talk to each other.**
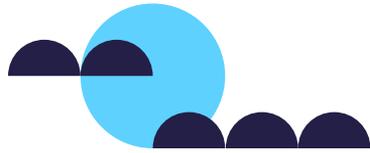
## CLOUD CONTROL

The growth of data in every connected city will rapidly outpace that of the general population or physical infrastructure. At the same time, the storage of this data is naturally moving away from a model of onsite hosting **(92% percent of all generated data will be in the cloud by 2020[3])** and cities will find themselves increasingly relying on third party services to host and distribute the real-time data generated by smart infrastructure.

Without leveraging an advanced data platform **cities are unable to use the same world-class tools that are deployed at scale in many of the most successful companies in the world**, and this problem will only become more pronounced as cities purchase smart devices that produce increasing amounts of real-time data. Nearly every Fortune 500 company uses real-time data and analytics to drive complex business decisions. Furthermore, for a typical Fortune 1000 company (average value is ~$5.1 billion dollars or the yearly budget of Houston, TX), just a 10% increase in data accessibility has resulted in more than $65 million additional net income.
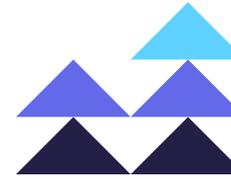
## ANTICIPATING EMERGENCE

It may prove that **the future livelihood of the city actually depends on the exchange of real-time data.** In the last ten years, services like Uber and AirBnB have extended and distributed city services by commodifying data exchange itself. **What new business models will emerge next that will require cities to be able to tap into and properly tax streaming services?**

On the other hand, the same technology that can help enhance cities can be harmful if not regulated. For example, in a city of autonomous vehicles where self-driving cars can roam unoccupied in the streets, who needs to pay for parking anymore? **And how will a city recoup that revenue?** Similar scenarios are likely to unfold in the years to come and cities in command of their data will be better poised to adapt to the emergence of new technologies and businesses that will disrupt the standard model of commerce and taxation.

**Cities are faced with a future where they must adopt an enterprise-grade data platform** in order to effectively manage their infrastructure and plan for the future. And the question cities find themselves asking is: "Do we build our own real-time data platform or do we buy one?" Below, we explore the process of choosing two paths.

[1] Gartner, "Gartner Predicts Fifty Percent of Citizens in Large Cities Will Share Personal Data With Smart City Programs by 2019," December 2016

[2] Gartner, "Gartner Says Smart Cities Will Use 1.1 Billion Connected Things in 2015," March 2015

[3] Cisco, "Cisco Global Cloud Index: Forecast and Methodology, 2015–2020," 2016

[4] Baseline, "Surprising Statistics About Big Data," February 2014

# Build Your Own Data Management Platform

Building a custom data management and analytics platform is a big project, even for established tech companies. Accurate estimates of the total cost are key. Asking these questions can help determine the short and long-term costs of ownership:
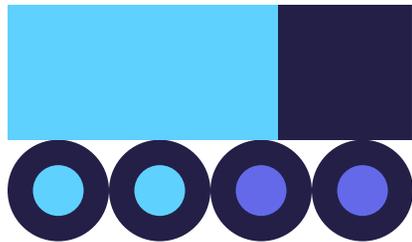
### DO YOU HAVE THE INTERNAL EXPERTISE?

Building a scalable data analytics stack requires specific expertise in ETL (Extract, Transmit, Load) and distributed computing architectures. Before considering building your own data management platform you should consider if you have that expertise on your team, and if not, **how long it will take to find and hire the right talent.** Data scientists and data engineers are among the most in-demand and well-compensated roles in tech.[5]

### DO YOU HAVE THE RESOURCES?

If you have the expertise, **do you have the time?** A full data management and analytics platform can take months or even years to ship. Decide if it's worth waiting or if you can achieve your goals faster with a purchase. You should also consider the importance of other projects—are you willing to take engineers away from other projects to build something bespoke?

### WILL YOUR DATA BE PUBLIC-FACING?

If you plan to expose real-time data to the public, are you well-positioned to support it? What is your SLA (Service Level Agreement) with public users and what security levels need to be enforced? Be certain you have factored in the **total cost and accountability** of supporting a public-facing portal for real-time data.

### WILL YOU BE ABLE TO SCALE?

**Data platforms are especially susceptible to scope creep and cost overruns.** When Gartner assessed the big data initiatives of 199 companies, they found that while nearly three quarters of organizations had invested or planned to invest in big data, the majority of the projects were indefinitely delayed in pilot phases. The diagnosis was that **too many initiatives were being built with "ad-hoc technologies and infrastructure that are not created with production-level reliability in mind."**[6]
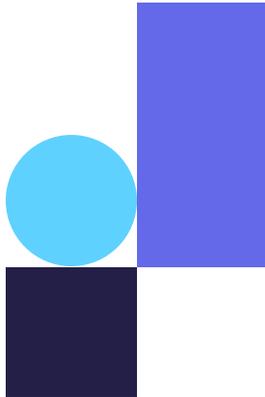
---

# Buy an Off-the-Shelf Data Management Tool

A city may decide that the total cost of ownership for building their own data management platform exceeds its budget or in-house abilities. When considering a third-party solution, this list of criteria is helpful to consider:

### WILL YOU HAVE REAL-TIME ACCESS?

Real-time access is required to **gain the true benefit of the emergent city data.** Does the solution you are considering support true real-time API access, and if it does, is there an additional cost?

### CAN THE TOOL DO WHAT YOU NEED?

Make sure the tool supports your use cases from the onset. Many support data storage, flat file export, API access, self-service API management, business intelligence and data visualization, but **few support all of those features.**

### CAN THE TOOL GROW WITH THE CITY?

The right tool should able to **adapt to a city's evolution.** Make sure that any features you may need in the near or long-term future are supported or on the product's roadmap.

### CAN YOU TAKE YOUR DATA WITH YOU WHEN YOU LEAVE?

You want your data to be available when the contract with the company is over. **Data should always belong to the city** and be easy to migrate or export at any time.

### IS THE PARTNER FOCUSED ON CITIES?

If it isn't, how important is city data management compared to the company's core business interest? Many large, stable enterprise companies may have a data management offering, but the **intricacies of city data, onboarding city workers, and ensuring public accountability** may not be a core competency.
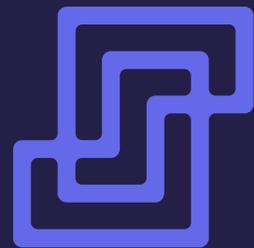
# STAE

Learn more about real-time city data tools at **stae.co**
or reach out to us at **we@stae.co**