Making the switch to real-time analytics

If your company is like most, you’re likely collecting customer data and feedback every second of every day – and this data must be analyzed and acted upon quickly. After all, as technology has made communication easier and more rapid, expectations for how quickly complaints and issues should be resolved have also increased. Companies can no longer get away with analyzing their data once every few weeks or even less often.

Most companies realize this and are taking steps towards implementing real-time, stream-based analytics. Indeed, in 2016, 65% of companies were already using real-time data pipelines, and an additional 24% were planning to implement them before the end of the year.¹

Of course, going from sporadic batch data processing to streaming, real-time data analytics is much easier said than done. Acquiring and setting up streaming tools is the easy part. What is much more difficult, and where many companies stumble, is in adjusting their business processes to be able to handle all that data as it comes in.

This white paper will offer advice on the ways in which businesses can overcome the barriers (or perceived barriers) to real-time and near-real time analytics. It will also offer tips on how to ensure you’re getting the best from your provider.

Table of Contents

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>An introduction to real-time analytics</td>
<td>2</td>
</tr>
<tr>
<td>Barriers to implementing real-time analytics</td>
<td>3</td>
</tr>
<tr>
<td>Things to consider before making the switch</td>
<td>3</td>
</tr>
<tr>
<td>Success story: Using real-time analytics to resolve customer issues before they wake up</td>
<td>5</td>
</tr>
<tr>
<td>What to look for when selecting a provider</td>
<td>6</td>
</tr>
<tr>
<td>About Luminoso</td>
<td>7</td>
</tr>
</tbody>
</table>
An introduction to real-time analytics

Until very recently, most data was analyzed no more frequently than once every few weeks – or more likely, only every quarter or year. This was largely due to the difficulty in analyzing unstructured, text-based data, including call center transcripts, customer emails, and product reviews. Such data had to be hand-coded or analyzed using Boolean queries or keyword searches, a resource-consuming endeavor.

Recent advancements in artificial intelligence (AI) and natural language processing (NLP) have significantly changed how data, especially unstructured data, can be processed and analyzed. Keyword lists and ontologies are no longer necessary; using AI and NLP, a machine can automatically sort through hundreds of thousands (or more) units of unstructured data and identify key concepts and new trends. This shift has made real-time analytics (processing data within seconds) and near-real-time analytics (processing data within minutes or hours) possible. More companies have begun looking at their data as a (near-)real-time stream rather than sporadically in batches. In addition to enabling them to make use of their data before it becomes stale, real-time analytics makes it easier for companies to:

**Uncover and resolve issues**

Customers have moved towards airing their grievances on public forums and social media. In 2015, 56% of customers said they complained via an online forum or community, and 43% said they tweeted at a company to complain – numbers that have steadily increased in the past few years.

The rapid pace at which such complaints come in, and their public nature, make it essential for companies to quickly process, triage, and handle issues that crop up. Real-time analytics enables companies uncover and respond to issues as they happen, nipping problems in the bud, before they become an even bigger problem.

**Identify and monitor trends**

Just as near- and real-time analytics makes it easier to find and resolve issues, it also makes it easier to quickly unearth trends and track them over time. Many companies have moved to a real-time, streaming data model to identify references to their company in the news and on social media and/or monitor public brand perception. As a result, they can quickly act should a trend suddenly begin to spike or drop off.

Despite these benefits and clear use cases, less than a quarter of data and analytics decision-makers can currently access data and generate new reports within days (as opposed to weeks or months). Though real-time analytics solutions are readily available, companies often balk or stumble when it comes to implementation.
Barriers to implementing real-time analytics

Roadblocks to implementing streaming, real-time analytics fall into two categories: actually building streaming data pipelines, and adjusting existing business processes to handle all that streaming data.

In the first category, companies most struggle with the complexity of such systems, especially higher frequency of code changes and pushes, and their relative lack of experience with handling streaming data. These challenges, while not immaterial, tend to be overcome fairly quickly and with minimal disruption.

The second category, adjusting existing business practices to handle streaming data, causes companies far more difficulty and frustration. This is largely because many companies don’t consider their data collection and business practices at all when diving into real-time analytics. They set up the proper software, but aren’t prepared to handle all that data because the correct processes aren’t in place.

Things to consider before making the switch

To successfully implement real-time analytics, companies must:

1. **Automate data collection, consolidation, and analysis**

   Processing and analyzing data in real-time is impossible if any stage of the process relies on manual pulls or other intervention. This seems obvious, of course, but many companies try to make the leap to real-time without ensuring that the needed data can be pulled automatically and programmatically. The result is a fully functioning analytics system, but without a reliable data stream.

   Companies must confirm that:

   - Data sources can be accessed programmatically and can handle frequent pull requests. If the goal is to analyze data in the aggregate from multiple sources, data formatting should be reviewed to ensure that data can be combined seamlessly.
   - Their real-time analytics software can plug into existing systems so that data transfer is automatic and instantaneous.
   - The data processing time and any required human analyst intervention is minimal.
Ensure that your analytics solution can handle unstructured text – flexibly and in real-time.

The word “data” conjures up images of columns upon columns of numbers and figures, but the reality is that an estimated 90% of all digital data is unstructured (i.e. in text-based formats like contact center transcripts, customer emails, product reviews, and social media posts). Companies often overlook this when choosing an analytics provider. This is to their detriment, as unstructured data often provides the “why” and “how” behind the raw numbers in structured data.

But choosing a provider who can “handle” unstructured data isn’t enough on its own – they must also be able to analyze unstructured data in real time. This effectively means that they must use artificial intelligence (AI)- and natural language processing (NLP)-based technology. AI and NLP has replaced the legacy approaches of keyword lists and ontologies, which require significant training time, take longer to process data, and require human intervention to adapt to changes in the marketplace. But not all providers have made the switch – so be sure to ask them in detail about their approach before making a decision.

Legacy approach:
Queries or keyword lists manually written and updated

Modern approach:
AI and NLP technology automatically maps out and understands concept relationships

Proactively build out processes for collecting data and acting upon insights.

You might be able to analyze your data in time, but if you don’t have a system or process in place to take immediate action on the insights you find, you might as well not have analyzed the data at all. It’s important to have a plan in place for how to ingest and respond to those insights before you pull the trigger on real-time analytics.
**Proactively build out processes for collecting data and acting upon insights.**

Some questions you should answer before you get started include:

- Where will the data come from? What existing barriers do you need to remove to get that data and be able to pull it automatically?

- How often should reports be generated to be useful and effective? Who will review the reports and take action? Do you have the manpower available to review the reports and act on them at that frequency?

- Who is responsible for acting on the findings of those reports? How will any actions taken be communicated?

No, these questions aren't always fun to think through – and it can be easy to put off. However, it's critical to have an answer and a plan in place for every question so that you'll be fully prepared once you make the leap to real-time analytics.

Success story: Using real-time analytics to resolve customer issues before they wake up

The mobile game industry is a fiercely competitive one. Revenues and consumer loyalty can be phenomenal for a company that executes; but if a mobile app game becomes buggy or doesn’t perform well, gamers will flock to other alternatives. As a result, resolving any issues that come up as quickly as possible is incredibly important.

A $10B mobile game development company with only 180 employees turned to real-time analytics in order to categorize, triage, and resolve their issue tickets. During a normal day, this company would receive 2,500 issue tickets per hour, or roughly 60,000 tickets a day – often rising to 5,000+ tickets per hour (more than 120,000 tickets each day) after software updates. Their old processes of prioritizing issue resolution based on feedback from support agents simply wasn’t enough.

The company’s first step was to automate as many processes as possible. They selected a real-time analytics vendor and utilized the vendor’s API to link their data sources to the vendor’s systems, enabling streaming data. The vendor’s system used AI and NLP to identify the key concepts and ideas in each ticket, categorize the tickets, and track which issues were becoming more or less prevalent over time.

This enabled the company to isolate new and emerging issues, measure the number of players experiencing that issue, and potentially fix problems within minutes or hours – often before the players woke up in the morning. This also lowered the overall number of issue tickets because many players would never realize there had been a problem with the update in the first place.
What should your real-time analytics vendor provide?

Real-time, streaming analytics will help your company become more agile and insights-driven. But choosing the right real-time analytics vendor is critical. Here are some essential questions you should ask before you choose a vendor:

1. **How complex is the process for integrating the unstructured data analytics solution with your existing systems?**
   
   The more complicated or longer the process is, the more problems there will be in truly automating your system – which is critical to analyzing data in real-time.

2. **How will your data be uploaded into their system?**
   
   Can you hook up your files or connect directly via an API, or will your data require cleanup and processing before you can even upload it (or, worse case, requires manual upload)? If the process isn’t seamless, programmatic, and automatic, then it won’t be able to handle streaming data in real-time.

3. **Can the vendor analyze unstructured data in real time?**
   
   As discussed above, unstructured data holds more value and information on intent and the “why” behind data than structured data does – but it’s often overlooked. Your provider should be able to handle this data in real time, without needing to employ a team of consultants to create and train ontologies and conduct ongoing maintenance.

4. **How long will it take to get set up and running with their system?**
   
   Apart from slowing down your transition to real-time analytics, a long setup time raises other red flags. For example, if it’s longer than a week or two, this is a sure sign that the company is using ontologies and training data sets to analyze unstructured data.

5. **How long is the data processing time?**
   
   For real-time analytics, you need a data processing time that’s as close to instantaneous as possible – and certainly shorter than a couple of hours. True AI and NLP solutions can process data in a manner of minutes. If it’s much longer than that, it’s a good sign that your vendor is using a team of data coders or consultants – increasing both the cost to you and the time it takes to find insights.

6. **Do they ask for sample or training data sets (or say they provide their own)?**
   
   If they need sample data sets before they can get you set up, or can’t answer the question of how they handle data that may change over time, you’re probably working with a team of consultants – not a system based on AI and NLP.
It is now estimated that by 2020, spend on streaming analytics software alone will approach $2.6 billion. Near- and real-time analytics can help drive better customer service, HR and sales decisions. Are you getting the most from yours?

References

5. OpsClarity.
7. Forrester. (7/2016). Streaming Data from the Internet of Things will be the Big Data World’s Bigger Second Act.

About Luminoso

Luminoso Technologies is a leading natural language understanding company that allows clients to rapidly discover value in their unstructured text data. With roots at the MIT Media Lab, Luminoso’s artificial intelligence-based software uniquely produces an accurate, unbiased, real-time understanding of what people are saying, including insights that were not anticipated. These insights are used to increase marketing performance and build better customer experiences.

The company provides multilingual, flexible software that can be deployed to meet client needs in either a standalone Cloud or On Premise solution or integrated into an end-to-end client platform via an API. Luminoso serves clients such as Staples, Sprint, and Scotts Miracle-Gro, as well as a growing set of channel partners including Publicis.Sapient and Basis Technologies.

Luminoso is privately held with headquarters in Cambridge, MA. For more information, please visit www.luminoso.com.