

Unstructured Data Management: 3 Points to Consider

By Janet Wagner for AlchemyAPI, An IBM Company

In this recent [e-book](#), we cover why organizations cannot afford to ignore unstructured data. The amount of unstructured data generated by internet-connected devices is only getting bigger, and the speed at which it is growing is only getting faster. It is imperative that organizations start looking now for solutions to solve the challenges caused by the big data generation.



- *Which unstructured data problems are the most important and need to be solved first?*
- *What types of unstructured data hold the most value and will provide important business insights?*
- *What information is critical to the organization's success?*

Challenges

There are many obstacles associated with extracting insight from unstructured data. For most organizations, the answer is not found in a single solution, but rather a combination of internal tools and external services. Best-of-breed data management programs result in increased agility and ROI. In order to maximize impact, leaders should consider all of their options. Build a strategy by looking at data monitoring, capture, extraction, deduplication, normalization and storage. Read on to learn important aspects of any data-minded leader's plan.

Solutions

There are quite a few data monitoring and analysis platforms available today, most of which focus on social media networks. However, there are many platforms that also include the ability to extract and process data from other sources such as blogs, review sites, emails, video sites, forums, etc. These platforms tend to be designed for enterprises and often include a variety of features to analyze data sources such as real-time social media monitoring, search and discovery, trend detection, sentiment analysis, data visualization, data analytics and more. A few examples of leading data monitoring and analysis platforms include [Bottlenose](#), [Brandwatch](#), [DataSift](#), [Engagor](#), [Gnip](#), [Nexalogy](#) and [Viralheat](#).

Data Management Planning

While most of the aforementioned platforms offer a variety of features that help organizations gain insights from unstructured data, they focus primarily on external data sources and can be considerably expensive. Much of the data generated by enterprises and SMBs comes from emails, text documents, digital images, spreadsheets and other internal files. While it is important to gain insights from social media unstructured data and other external data sources (i.e. Twitter). Businesses should also implement a data management plan for internal unstructured data. They should determine where data will be stored, how long it should be stored, how the data should be analyzed, etc.



Budget, timeline and use case vary from organization to organization. Using APIs to enrich unstructured data as well as gain insight from that data can be far less expensive than feature-packed data monitoring and analysis platforms. APIs also allow for a customized, streamlined approach to handling unstructured data as opposed to a “one size fits all” solution. APIs are used to integrate data monitoring and analysis features with existing applications as well as to build custom unstructured data solutions that address the specific needs of an organization at a lower cost.

Storage

Storage is an important consideration when it comes to unstructured and structured data. There are a lot of cloud storage providers, too many to list here, that are fairly inexpensive and provide APIs. Cloud storage services are offered by [Amazon](#), [Bitcasa](#), [Box](#), [Dropbox](#), [EMC](#), [Google](#), [Microsoft](#), [Rackspace](#), [Red Hat](#), [Smartfile](#) and the list goes on.

Data Monitoring and Analysis

Monitoring and analysis of unstructured data is an important solution to consider. There are companies such as [AlchemyAPI](#) and [IBM Watson](#) that provide APIs that can be used to enable real-time social media monitoring, sentiment analysis, competitive intelligence and extraction capabilities (text, keyword, entity, relation, etc). There are data visualization and analytics platforms that provide APIs such as [GoodData](#), [Keen io](#) and [Tableau](#).

Dashboard platforms are another popular way to monitor and analyze data in real time. Many dashboard platforms such as [Cyfe](#), [Ducksboard](#), [Geckoboard](#), [Leftronic](#) and [Klipfolio](#) provide APIs. It should be noted that data monitoring and analysis platforms Bottlenose, Brandwatch, DataSift, Engagor, Gnip, Nexalogy, and Viralheat all provide APIs as well.

Conclusion

Before any of the above solutions can be considered, an organization must figure out their priorities and goals for unstructured data. Which unstructured data problems are the most important and need to be solved first? What types of unstructured data hold the most value and will provide important business insights? What information is critical to the organization’s success?

Many organizations have already answered these questions and are now using unstructured data to gain a competitive edge in the marketplace. If your organization is not already looking for solutions to solve its unstructured data problems and to gain key business insights, then it should start looking now. Your competitors may already be using unstructured data to their advantage, don’t wait.

About AlchemyAPI, An IBM Company

AlchemyAPI’s mission is to power smart applications that understand human language and vision by making breakthroughs in deep learning-based artificial intelligence available to everyone. AlchemyAPI is used by more than 40,000 developers across 36 countries and a wide variety of industries to process over 3 billion texts and images every month. For more information, visit our website at alchemyapi.com.

Looking For a Partner in Data Analysis?

Ask the Right Questions

- ✓ Technology
- ✓ Vendor Positioning
- ✓ Customer Sentiment

[Get the Buyer’s Checklist](#)

