

Infegy Linguistics

Where Unstructured Text Meets Big Data

The world's volumes of unstructured digital text continues to grow into the zettabytes. While a wealth of valuable information lies within these volumes of natural language, computers and software have struggled to effectively make use of it at scale.



Introducing Infegy Linguistics

A system that is finally able to deeply analyze enormous collections of text, spanning petabytes in size, at incredible speed.

Based on more than seven years of research and development, Infegy Linguistics is built from the ground-up to handle massive datasets and high-volume streams of unstructured textual data. The service is designed to pull valuable insight from a variety of sources including customer feedback, call center transcripts, news articles, social media, HTML and more.

Data that was previously difficult to work with is made valuable. Quickly and easily extract insights such as aggregate topics, highly-accurate sentiment and polarity, emotion analysis, topical themes, structural analysis and our unique automatically-generated headlines feature.

Whether you have a large archive of content, a high-volume stream of new data or both, Infegy Linguistics can create structured insights from your unstructured data fast.



www.infegy.com +1 816-494-1650 frontdesk@infegy.com
4151 N Mulberry Dr., Suite 240, Kansas City, Missouri 64116

Infegy Linguistics **Functionality**

Functionality **Overview**

Infegy Linguistics is a service accessed either over the internet, behind your firewall with leased hardware installations, or integrated into your own solutions.

Simply connect with the easy-to-use API and begin streaming your data. As you stream, the service will respond with document-level results at incredible throughput and remarkably low latency. When finished, summary results are provided with a variety of collection-level information including topics and generated headlines.

The API is fully-featured, and even includes a checkpointing system that makes it easy to pause work and resume later. This enables jobs to be recovered even after power outages.

Highlight Features

Sentiment Analysis

Industry-leading 93% accuracy and 97% recall (0.952 F-Score), benchmarked using 1.9 million customer reviews spanning 81 categories, from food to video games.

Emotion Detection

Measure frequency and intensity of the eight major emotions. Are your customers and fans happy? Do they trust you? Emotion Detection can tell you.

Topic Extraction

Extract key topical phrases and words from large collections of text, scored with a powerful relevance system gauging their topical importance, and each with their own subject-specific sentiment information.

Headlines

Understand what's really going on in your text. Headlines builds natural text sentences highlighting the important elements within your text in an immediately-understandable natural language format.



Infegy Linguistics **Flexibility**

Infegy Linguistics offers a variety of options for connectivity, ensuring we can meet your needs for performance, security and scalability, no matter the requirements and workload.

Streaming Service via the Internet

Our internet streaming service is the quickest way to get started, most cost-effective and offers terrific performance. Send documents at blazing speeds up to 1Gbps (128 megabytes per second), with processing latencies under 10 milliseconds. Our service provides enough performance to process even the largest datasets quickly, or the busiest streams of data live. Additionally, our streaming service has no base costs, ensuring you only pay for data processed.

On-Site Behind Your Firewall

When security is a primary concern, sending your data over the internet may not be an option. For these cases, we offer leased hardware capable of the same features as our streaming service. These machines can exist completely off the internet, administrated by your company, and contain no hard drives or mass-storage medium, ensuring your data is kept completely under your control. A single machine (1U rack or medium tower) can handle data at up to and beyond 2 Gbps (with supported network infrastructure) and processing latencies under 5ms. Need more? Our system can scale over multiple clustered machines ready for any performance demand.

In The Cloud On Your Service

Hosting your data on a cloud provider? Similar to our on-site option, we can build and manage a dedicated processing machine or cluster for you on most popular platforms including Amazon EC2, Rackspace Cloud and Microsoft Azure. Combining these platforms and our distributed processing, performance is easily scalable to your needs.



Infegy Linguistics Summary

Feature List

Full unicode support, language detection, rich tokenization with robust HTML support, document statistics (such as word and sentence counts), in-depth sentiment (general and subject-specific), readability score, emotion detection, customizable topical themes with sentiment (such as purchase intent or taste), collection-level topic extraction and system-generated collection headlines.

Tech Specs

Specs below apply to the Internet Service offering. On-site or cloud solutions have differing performance.

Throughput	Up to 1 gbps (upgrade options available)
Processing Latency	As low as < 1 ms
API Data Format	JSON / UTF-8
Maximum Dataset Size	16,384 petabytes
Supported Languages	Detection of 68 languages, sentiment for English, French, German, Spanish, Portuguese, Financial English

About Infegy

Founded in 2006 in Kansas City, Missouri, USA, Infegy was born to create the future of market research and consumer insight. From day one, Infegy's goal has been to leverage the exponential growth of content shared online to better understand consumers through deep analysis of natural communication and dialog. Infegy's Atlas service leverages the Infegy Linguistics platform in combination with an in-house 40-billion document real-time dataset of online content to provide clients with incredible insight into their brands, consumers and marketing. Now, Infegy Linguistics makes Infegy Atlas' core analysis technologies available to everyone.



www.infegy.com +1 816-494-1650 frontdesk@infegy.com
4151 N Mulberry Dr., Suite 240, Kansas City, Missouri 64116