LexPredict ContraxSuite Documentation

Technical FAQ

Release 1.0.1 - September 1, 2017

How to Get Support

For support or request for other information, please contact support@contraxsuite.com.

Frequently Asked Questions (FAQ)

1. What operating systems does ContraxSuite support?

ContraxSuite is designed to run on any operating systems that support Python and Java. In practice, this means that ContraxSuite can be run on Windows, Linux, Mac, or similar operating systems. Furthermore, as a multi-tier application, organizations can deploy different tiers of ContraxSuite on different operating systems. This means that you can run the web and application tiers on Linux while using a Windows server to run SQL Server.

From a support and development perspective, we primarily develop, test, and deploy ContraxSuite on Linux. Our roadmap includes automated testing and deployment for Windows.

2. What languages is ContraxSuite written in?

ContraxSuite is a three-tier application, primarily written in Python. Knowledge of Java may also be useful in the installation, configuration, maintenance, and customization of ContraxSuite.

3. What kind of security does ContraxSuite support?

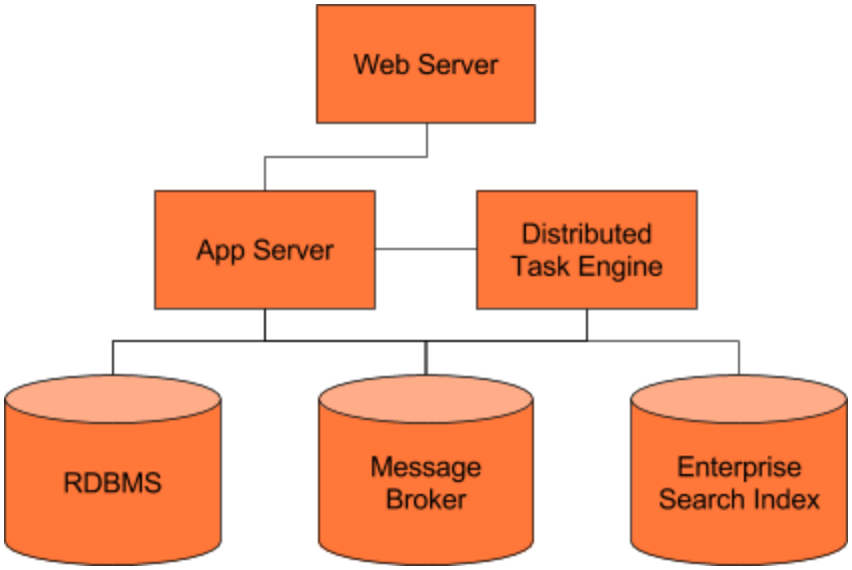
We have written an entire FAQ dedicated to Security, including topics such as encryption, two-factor authentication, and on-premises or dedicated hosting. You can find this Security FAQ in the public documentation.

4. What dependencies does ContraxSuite rely on?

We have written a complete document dedicated to Software and Data Dependencies, including information about their installation, customization, and licensing. You can find this Software and Data Dependency document in the public documentation.

5. How is ContraxSuite architected?

ContraxSuite is architected as a multi-tier application, separating components for the web server, application server, distributed task engine, relational database (RDBMS) server, enterprise search server, and message broker. These components roughly align with the common three-tier, presentation-application-data layer model. A simplified architecture diagram can be found below.



6. How can I customize ContraxSuite?

As discussed in (2) above, ContraxSuite is primarily written in Python. The web tier and application tier of ContraxSuite are built using the Django and celery frameworks. The natural language processing and machine learning code rely primarily on NLTK and scikit-learn.

ContraxSuite is designed to allow for modular additions, modifications, and removals. For example, users can add or replace the OCR engine, add new file formats for document ingestion, and develop their own locators, analyzers, extractors, or reports.

Our Roadmap includes the development of documentation to clearly explain, with code samples, how to customize and extend ContraxSuite. Look for more information on this topic in Q3 and Q4 2017.

7. Does ContraxSuite support cloud or virtualization?

Yes, we have run ContraxSuite in both virtualized and physical installations. ContraxSuite has been tested on IaaS and PaaS providers like Amazon Web Services and Digital Ocean. To our knowledge, any virtualization or paravirtualization technology that supports Linux, Windows, Java, and Python should be possible.

8. I'm a data scientist but not a developer. What can I do?

ContraxSuite ships with support to run Jupyter notebooks inside of the application. Jupyter notebooks, which can run over 40 different languages including Python, R, Julia and Scala, provide a highly-interactive, graphical experience for data science and rapid prototyping. Our consultants and data scientists frequently use Jupyter notebooks to improve and customize ContraxSuite, debug issues, or develop new reports.

Our Roadmap includes the development of sample notebooks to clearly explain, with code samples, how to use Jupyter notebook with ContraxSuite. Look for more information on this topic in Q3 and Q4 2017.

9. How can I install ContraxSuite for test or development purposes?

We have developed a "one-line" install script for single-machine, development purposes. This simple installation automation allows developers and users to quickly deploy the core ContraxSuite components and application on a Linux server. Contributors who are interested in helping support Windows can reach out to us at support@lexpredict.com or send PRs on Github.

We can also help through our services team. For support or help in setting up the application, please contact support@contraxsuite.com.

Note: The result of this installation process is not likely to meet real production usage needs. For example, most productions installations require more security configuration or RAM than is commonly available on a single laptop or workstation.

10. How can I install ContraxSuite for production purposes?

We have written an Installation and Configuration Guide for ContraxSuite, which can be obtained with the source code repository. This installation guide is currently written to support Linux installations, and assumes basic Linux IT administration or operations knowledge.

Our Roadmap includes the improvement of existing documentation and development of new documentation. We intend to release a Windows installation and configuration guide by Q1 2018. Contributors who are interested in helping support Windows can reach out to us at support@contraxsuite.com or send PRs on Github.

We can also help through our services team. For support or help in setting up the application, please contact support@contraxsuite.com.

11. Do you have unit tests for languages in place?

Public unit tests are currently under development for all core natural language, machine learning, and extraction algorithms. Currently, we have implemented over 100 tests with over 400 test cases on real contract and document data. English language unit tests are scheduled for release in early Q4 2017.

We are actively working on developing unit tests in multiple languages that will be released with accuracy, recall, and runtime performance data. Please refer to our Roadmap for more details.