



# **SparkCognition DeepNLP™ Installation Guide**

**A SparkCognition™ Education Document**

**Q3-2019  
v2.1 - 7.2019**

---

---

This document contains copyrighted and proprietary information of SparkCognition and is protected by United States copyright laws and international treaty provisions. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under such laws or with the prior written permission of SparkCognition Inc.

SparkCognition™, the sparkcognition logo, Darwin™, DeepArmor®, DeepNLP™, MindFabric®, SparkSecure® and SparkPredict™, are trademarks of SparkCognition, Inc. and/or its affiliates and may not be used without written permission. All other trademarks are the property of their respective owners.

©SparkCognition, Inc. 2017-2019. All rights reserved.

# SparkCognition DeepNLP™ Installation Guide

## Contents

<b>About this guide</b>	<b>2</b>
<b>Requirements</b>	<b>2</b>
System Requirements . . . . .	2
Hardware Requirements: . . . . .	2
Operating System Requirements: . . . . .	2
System Privileges: . . . . .	2
Virtual Machine Installation Requirements . . . . .	3
<b>Installation &amp; Configuration Process</b>	<b>3</b>
Single-Instance Installation . . . . .	3
Troubleshooting the Installation . . . . .	4
Multi-Instance Installation . . . . .	4
Configure DeepNLP User . . . . .	4
Create a New DeepNLP User . . . . .	4
Create a DeepNLP <i>admin</i> Role . . . . .	4
Assign the <i>admin</i> Role to a User . . . . .	5
<b>Access the DeepNLP Web Interface</b>	<b>5</b>
<b>Manage the DeepNLP Deployment</b>	<b>5</b>
Changing Configuration Parameters . . . . .	5
<b>Contact Support</b>	<b>6</b>
<b>Reference</b>	<b>6</b>
Hardware Requirements . . . . .	6
Operating System/Software Requirements . . . . .	7
Services . . . . .	7
Multi-host Services . . . . .	7
Application Access Requirements . . . . .	7

---

## About this guide

This guide describes the Sparkcognition DeepNLP™ Installation process. This guide is intended for data scientists, software engineers and analysts who want to interact with the DeepNLP system. Note that throughout this document, some long key and token values are truncated - indicated by ellipses (...) - or broken across lines, indicated with a backslash (\).

## Requirements

The general requirements for DeepNLP include:

- Three DeepNLP files - supplied by SparkCognition, either attached as part of the SparkCognition welcome email or available from SparkCognition DeepNLP support:

- deepnlp-docker-bundle-##.##.##.tar.gz  
(where ##.##.## is the DeepNLP version number)
- install\_deepnlp.sh
- versions.env

**Note:** Ensure the filenames are correct (unchanged) following download or copy. Filenames are case sensitive. Because installation scripts call these files, incorrect filenames will result in errors.

- Operating system - Either *CentOS 7* or *RedHat 7*
- Appropriate machine - DeepNLP can be installed on virtual or physical machines running supported operating systems

**Note:** DeepNLP is not supported if the requirements are not met or exceeded

## System Requirements

The following requirements must be met or exceeded to successfully install and run DeepNLP:

### Hardware Requirements:

- Minimum: 4 modern CPU Cores, 32GB memory, 30GB Storage
- Recommended: 16 modern CPU Cores, 64GB memory, 60GB Storage

### Operating System Requirements:

- *CentOS 7* or *RedHat 7*

### System Privileges:

The following system privileges are required for installation and recommended for advanced users:

- ssh -  
provides access to the DeepNLP installation target machine(s) and a control server to run the installer from
- sudo authority -  
enables installing DeepNLP

## Virtual Machine Installation Requirements

Virtual machines (VM) installed with DeepNLP must meet all specified requirements, including those for memory and processor.

DeepNLP requires VM hosts that are compatible with *VMWare Workstation™ 15* and support *VMWare Workstation 15* images.

Contact SparkCognition for additional information concerning support for other VMWare® hosting products or versions of VMWare.

## Installation & Configuration Process

The following sections describe the preparation for installation and configuration of DeepNLP.

### Single-Instance Installation

To install a single instance of DeepNLP, perform the following:

1. Copy the three DeepNLP files to the same directory (your choice)
2. Navigate to the directory with the DeepNLP files
3. Run the following commands as a user with *sudo* authority:

```
sudo chmod +x install_deepnlp.sh
sudo bash install_deepnlp.sh
```

**Note:** The installation creates the *deepnlp* user and copies the *deepnlp app* folder into the */home/deepnlp/deepnlp-docker* directory

4. Verify the installation. Do not proceed without ensuring the installation is successful. Issue the following command and consult the output:

```
./deepnlp.sh status
```

### Example status output

Name	Command	State	Ports
deepnlp-docker_init_1	/bin/bash -c /opt/sparkcog ...	Exit 0	
deepnlp-docker_nginx_1	nginx -g daemon off;	Up	0.0.0.0:443->443/tcp, 0.0.0.0:80->80/tcp
deepnlp-docker_nlpdf_1	/bin/bash -c /opt/sparkcog ...	Up	0.0.0.0:8888->8888/tcp
deepnlp-docker_nlscheduler_1	/bin/bash -c /opt/sparkcog ...	Up	
deepnlp-docker_nlpupload_1	/bin/bash -c /opt/sparkcog ..	Up	0.0.0.0:8180->8180/tcp
deepnlp-docker_nlweb_1	/bin/bash -c /opt/sparkcog ...	Up	0.0.0.0:8080->8080/tcp
deepnlp-docker_postgres_1	docker-entrypoint.sh postgres	Up	0.0.0.0:5432->5432/tcp
deepnlp-docker_solr_1	docker-entrypoint.sh solr ...	Up	0.0.0.0:8983->8983/tcp, 0.0.0.0:9983->9983/tcp
deepnlp-docker_zepplin_1	/bin/bash -c bin/zeppelin.sh	Up	0.0.0.0:9091->8080/tcp

### Notes:

- Verify that *deepnlp-docker\_init\_1* exits with a state of **(Exit 0)**
- Verify that all other docker containers report their state as **Up**

5. Ensure that port **443** of the `deepnlp-docker_nginx_1` container is accessible and not blocked by the firewall. The following example shows the `deepnlp.sh` status output with the appropriate section highlighted:

```
[deepnlp@deepnlp-demo-qa deepnlp-docker]$ ./deepnlp.sh status
~/deepnlp-docker ~/deepnlp-docker
-----
      Name                Command                State                Ports
-----
deepnlp-docker_init_1    /bin/bash -c /opt/sparkcog ...  Exit 0
deepnlp-docker_nginx_1  nginx -g daemon off;      Up                0.0.0.0:443->443/tcp, 0.0.0.0:80->80/tcp
deepnlp-docker_nlpdf_1   /bin/bash -c /opt/sparkcog ...  Up                0.0.0.0:8888->8888/tcp
deepnlp-docker_nlscheduler_1 /bin/bash -c /opt/sparkcog ...  Up
deepnlp-docker_nlpupload_1 /bin/bash -c /opt/sparkcog ...  Up                0.0.0.0:8180->8180/tcp
deepnlp-docker_nlweb_1   /bin/bash -c /opt/sparkcog ...  Up                0.0.0.0:8080->8080/tcp
deepnlp-docker_postgres_1 docker-entrypoint.sh postgres  Up                0.0.0.0:5432->5432/tcp
deepnlp-docker_solr_1    docker-entrypoint.sh solr ...  Up                0.0.0.0:8983->8983/tcp, 0.0.0.0:9983->9983/tcp
deepnlp-docker_zeppelin_1 /bin/bash -c bin/zeppelin.sh    Up                0.0.0.0:9091->8080/tcp
~/deepnlp-docker
[deepnlp@deepnlp-demo-qa deepnlp-docker]$
```

## Troubleshooting the Installation

In the case where the installation fails to meet the verification requirements, perform the following:

1. Verify appropriate disk storage capacity is available
2. Verify the system requirements are met or exceeded
3. Verify the appropriate permissions exist - note that the `deepnlp-docker` directory must be owned by `deepnlp_user`

If all criteria are met and the system continues to fail installation, contact [SparkCognition DeepNLP Support](#).

## Multi-Instance Installation

DeepNLP supports multi-instance installation. Because that procedure is beyond the scope of this document, contact SparkCognition DeepNLP support for assistance to deploy multiple instances.

## Configure DeepNLP User

New user credentials must be created for every DeepNLP user. Note that any DeepNLP user that enriches data must be assigned to the `admin` role.

### Create a New DeepNLP User

To create a new user, issue a `curl` command that includes the `user_id` and `password` for the user. For example:

```
curl -H'Content-type: application/json' -H'Accept: application/json' \
-XPOST localhost:8080/add_user -d '{"user_id":"NEW_USER_NAME",\
"password":"NEW_USER_PASSWORD"}'
```

### Create a DeepNLP `admin` Role

An `admin` role is required for any user that will enrich data.

To define (create) a DeepNLP *admin* role, issue the following command:

```
curl -H'Content-type: application/json' -H'Accept: application/json' \
-XPOST localhost:8080/add_role -d '{"role":"admin"}'
```

### Assign the *admin* Role to a User

To assign the *admin* role privilege to DeepNLP users, issue the following command:

```
curl -H'Content-type: application/json' -H'Accept: application/json' \
-XPOST localhost:8080/add_user_role -d '{"user_id":"USER_ID_TO_ADD" , \
"role":"admin"}'
```

## Access the DeepNLP Web Interface

The DeepNLP web based user interface can be accessed through any modern browser.

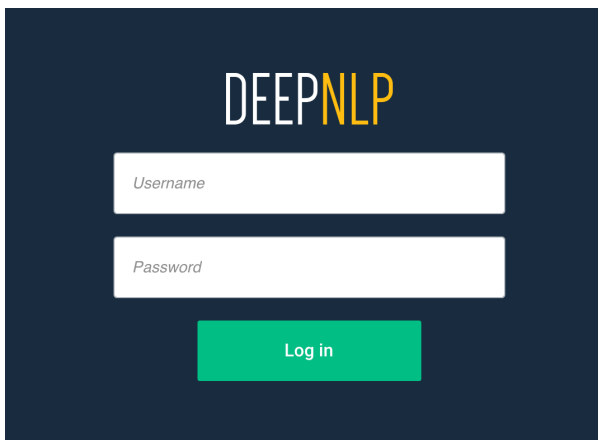
To access the DeepNLP interface:

1. Navigate to `https://<HOSTNAME_OF_YOUR_DEEPNLP_INSTANCE>` to display the DeepNLP web server login screen.
2. Enter the default generic credentials for the DeepNLP web server:

Username = `deepnlp`

Password = `deepnlp`

3. Click **Sign in** to display the DeepNLP application log in screen:



4. Enter the specific user credentials to access the DeepNLP
5. Click **Log In** to display the initial DeepNLP screen.

## Manage the DeepNLP Deployment

### Changing Configuration Parameters

To make changes to the DeepNLP configuration parameters, use the `deepnlp.sh` command. Typing the command without any switches or flags displays the `deepnlp.sh help`:

```
Unknown option <>. Please tell me what to do :/  
Usage:  
  deepnlp.sh start  
    starts all DeepNLP services  
  deepnlp.sh stop  
    stops all DeepNLP services  
  deepnlp.sh restart  
    restarts all DeepNLP services  
  deepnlp.sh clear  
    stops and removes all DeepNLP data  
  deepnlp.sh status  
    shows status of DeepNLP services  
  deepnlp.sh update-services  
    restarts DeepNLP services if their image has changed  
  deepnlp.sh update-password  
    changes the UI password for deepnlp user  
  deepnlp.sh reset-custom-transforms  
    resets custom transforms to its original state  
  deepnlp.sh install-gcs-key <KEY>  
    installs a Google Cloud Storage key  
  deepnlp.sh update-scripts  
    extracts docker-compose.yaml and deepnlp.sh from the image
```

---

## Contact Support

The following resources enable you to research issues, create a support ticket, or contact SparkCognition:

- **FAQ** - [Frequently Asked Questions](#)
- Use the [DeepNLP support portal](#) - Create a [support ticket](#) and log your issue
- **Email Contact** - Send email to [deepnlp\\_support@sparkcognition.com](mailto:deepnlp_support@sparkcognition.com)
- **Call Support** - The DeepNLP support line is +1-512-400-2001

---

## Reference

### Hardware Requirements

- Minimum: 4 modern CPU Cores, 32GB memory, 30GB Storage
- Recommended: 16 modern CPU Cores, 64GB memory, 60GB Storage



## Operating System/Software Requirements

- *CentOS 7 or RedHat 7*

## Services

The two main DeepNLP components are the web frontend and the scheduler. These components expect the following services to be running:

- *Postgres 9.5, or greater*
- *Solr 7.4.0*

## Multi-host Services

For DeepNLP multi-host installations the following are automatically installed, as necessary:

- *Hadoop*
- *Yarn 2.7.x or 2.8.x*

## Application Access Requirements

Although any service can be installed on any machine, consider the following:

- The DeepNLP web front end requires access to both `Postgres` and `Solr`
- The scheduler requires access to the web front end and `Solr`
- The scheduler must be able to access `Yarn` and `Hadoop`, when they are available