

Stirling PDF : La boîte à outils PDF

Dans ce tutoriel, nous allons apprendre à installer et à configurer la solution **Stirling PDF** à l'aide d'un conteneur **Docker**. **Stirling PDF** est une puissante boîte à outils gratuite et open source pour gérer et manipuler vos fichiers PDF localement.

Accessible à partir d'un navigateur Web, cette solution va vous permettre d'effectuer des actions diverses et variées sur vos fichiers PDF, tout en gardant la maîtrise de vos données. En effet, avec Stirling PDF, tout fonctionne entièrement sur votre machine locale, garantissant la confidentialité et le contrôle de vos données. Ceci pourra éviter que vos utilisateurs s'appuient sur des outils en ligne, au risque qu'il y ait une fuite de données par la même occasion.

Voici quelques-unes des fonctionnalités disponibles :

- **Fusionner** ou **diviser** des documents PDF
- **Extraire des pages** d'un document PDF
- **Réorganiser** ou **pivoter** les pages d'un document PDF
- **Convertir** des fichiers sources en PDF (image vers PDF, HTML vers PDF, Markdown vers PDF, etc.)
- **Convertir** un fichier PDF dans un autre format (PDF vers image, PDF vers Word, PDF vers HTML, etc.)
- **Compresser** un document PDF
- **Extraire** les **images** ou **ajouter** une **image**
- **Modifier** les **métadonnées** d'un document PDF
- **Signer** un document PDF (avec un certificat numérique ou une image de signature)
- **Ajouter** ou **supprimer le mot de passe** d'un PDF
- **Ajouter** un **filigrane** sur un document PDF
- **Visionneuse** de documents PDF
- Reconnaissance de caractères (**OCR**)
- Etc...

Toutes ces fonctionnalités sont gratuites puisque cette application est libre et open source. Elle intègre de nombreuses fonctionnalités dont certaines parfois réservées à des outils premium (payants).

image.png

Stirling PDF peut être installé sur une machine **Windows**, en local, car des exécutables sont proposés sur le GitHub officiel. Ceci implique l'**installation de Java sur la machine**. Pour ma part, je préfère m'orienter vers **un déploiement dans un conteneur Docker**, sur un **NAS**

Synology, afin de **mettre à disposition l'application à X utilisateurs**. Ceci est d'autant plus pertinent que l'application est **accessible à partir d'un navigateur**, via l'URL de votre choix.

Installer Stirling PDF sur un NAS Synology

Pour effectuer l'installation sur un NAS Synology, nous allons utiliser l'application "**Container Manager**" (Docker) afin de pouvoir utiliser un fichier de configuration **Docker Compose**. Ceci vous permet d'utiliser cette configuration facilement pour déployer le conteneur sur d'autres plateformes.

Tout d'abord, nous allons créer l'arborescence de dossiers pour accueillir les données de Stirling PDF. Sous le répertoire "docker", voici les répertoires à créer :

- **stirling-pdf**
 - **trainingData**
 - **extraConfigs**

Le répertoire "**trainingData**" est utile uniquement si vous envisagez d'utiliser la fonction liée à l'OCR.

Ce qui donne :

image.png

Ensuite, lancez l'application "**Container Manager**" (Docker) sur votre NAS Synology. Créez un nouveau projet :

- **Projet > Créer**

Commencez par donner un nom au projet, par exemple "**stirling-pdf**". Puis, indiquez le chemin correspondant au dossier précédemment créé, à savoir **"/docker/stirling-pdf"**.

image.png

Puis, sélectionnez "**Créer un fichier docker-compose.yml**" afin de pouvoir personnaliser le déploiement de ce projet basé sur l'image Docker "**frooodle/s-pdf**" dans sa dernière version (vis-à-vis du tag "**latest**"). Voici un aperçu du code de configuration Docker Compose :

image.png

Pour cet exemple, je vais utiliser la configuration Docker Compose suivante :

```
version: '3.3'
services:
  stirling-pdf:
    image: frooodle/s-pdf:latest
```

```

ports:
  - '8080:8080'

volumes:
  - /volume1/docker/stirling-pdf/trainingData:/usr/share/tessdata #Required for extra OCR
languages
  - /volume1/docker/stirling-pdf/extraConfigs:/configs
#  - /location/of/customFiles:/customFiles/
#  - /location/of/logs:/logs/

environment:
  - DOCKER_ENABLE_SECURITY=true
  - SECURITY_ENABLE_LOGIN=true
  - SECURITY_INITIALLOGIN_USERNAME=pdf
  - SECURITY_INITIALLOGIN_PASSWORD=IT-Connect
  - INSTALL_BOOK_AND_ADVANCED_HTML_OPS=false
  - LANGS=fr_FR

```

Quelques explications s'imposent :

- Le conteneur sera accessible sur le port "**8080**" puisqu'il est mappé sur le port "**8080:8080**" (port externe côté NAS : port interne dans le conteneur). Vous pouvez adapter cette valeur (premier numéro de port) si besoin, parce qu'un seul conteneur peut occuper chaque port.
- Sous "**volumes**", indiquez les chemins vers les répertoires "**trainingData**" et "**extraConfigs**" créé précédemment. Les répertoires "**customFiles**" et "**Logs**" sont facultatifs. Ici, ces deux lignes sont commentées.
- La directive "**SECURITY_ENABLE_LOGIN=true**" sert à activer la page de connexion sur Stirling PDF. Autrement dit, il conviendra de s'authentifier avant d'accéder à l'application. Ceci implique aussi de configurer la directive "**DOCKER_ENABLE_SECURITY=true**".
- La directive "**SECURITY_INITIALLOGIN_USERNAME=pdf**" sert à créer un utilisateur par défaut nommé "**pdf**".
- La directive "**SECURITY_INITIALLOGIN_PASSWORD=IT-Connect**" sert à attribuer le mot de passe "**IT-Connect**" à l'utilisateur par défaut.
- La directive "**INSTALL_BOOK_AND_ADVANCED_HTML_OPS=false**" sert à indiquer qu'il ne faut pas télécharger l'application Calibre sur Stirling PDF. Elle permet la conversion de PDF vers/depuis un livre et la conversion HTML avancée. Activez cette option si besoin.
- La directive "**LANGS=fr_FR**" sert à préciser la langue pour la bibliothèque de polices personnalisées à installer pour la conversion de documents. Dans tous les cas, l'interface de l'application sera disponible en plusieurs langues, dont le français.

Une fois que c'est fait, poursuivez jusqu'à la fin de l'assistant pour lancer la création du projet. L'image Docker va être téléchargée et le conteneur configuré puis exécuté.

image.png

L'application Stirling PDF est désormais exécutée au sein du conteneur Docker. Stirling PDF est une application assez gourmande en RAM, probablement à cause de son lien étroit avec Java. Le conteneur consomme entre 400 et 700 Mo de RAM, ce qui n'est pas neutre.

image.png

Découverte de Stirling PDF

Première connexion

L'accès à l'application s'effectue à partir d'un navigateur Web. Il suffit de préciser l'adresse IP ou le nom de domaine du NAS, suivi du port "8080", comme ceci :

```
http://192.168.1.200:8080/
```

Si vous avez activé l'authentification, vous devez vous connecter avec le compte créé par défaut (selon les informations définies dans le Docker Compose).

image.png

Voilà, vous êtes connecté à Stirling PDF ! Vous pouvez profiter de l'ensemble des outils intégrés à cette fabuleuse boîte à outils !

image.png

Gestion des utilisateurs

Avant d'évoquer les outils en eux-mêmes, je vous recommande de changer le mot de passe du compte utilisateur créé par défaut. Vous pouvez aussi créer d'autres utilisateurs. Pour accéder à la gestion de votre compte, cliquez sur "**Paramètres**" en haut à droite (icône en forme de roue crantée) puis cliquez sur "**Paramètres du compte**".

Ici, vous pourrez changer le nom d'utilisateur et le mot de passe du compte.

image.png

Si vous descendez tout en bas de la page, vous pourrez cliquer sur le bouton "**Paramètres d'administration - Voir et ajouter des utilisateurs**". Ceci vous permet de **créer d'autres comptes utilisateurs**, avec différents niveaux de permissions : **administrateur**, **utilisateur**, etc.

image.png

Libre à vous de créer un ou plusieurs comptes. L'intérêt étant principalement de gérer l'accès à l'application. À ma connaissance, Stirling PDF n'a pas vocation à permettre à chaque utilisateur de gérer une bibliothèque de documents PDF.

À tout moment, vous pourrez créer, modifier ou supprimer des utilisateurs.

Utilisation des outils

Nous n'allons pas passer en revue l'ensemble des outils, car l'interface est simple d'utilisation et il y a énormément de possibilités. Il vous suffit de choisir l'outil de votre choix à partir du menu principal, ou de la page d'accueil. Vous pouvez aussi utiliser la fonction de recherche pour gagner du temps, ainsi que mettre certains outils en favoris.

image.png

La fonction "**Fusionner plusieurs PDF**" vous permet, comme son nom l'indique, de fusionner (concaténer) plusieurs documents PDF en un seul fichier.

image.png

Il y a également un outil multifonction qui donne accès à quelques fonctions basiques dans une même interface. Vous chargez un fichier, et ensuite, vous pouvez intervenir des pages, effectuer des rotations de pages, etc... Selon vos besoins.

image.png

À vous d'explorer les différents outils au fur et à mesure que les besoins se présenteront.

Aller plus loin dans la configuration

Est-ce qu'il y a des paramètres de configuration supplémentaires ? La réponse est oui. La configuration de Stirling PDF s'appuie sur un fichier au format YAML nommé "**settings.yml**". Il est situé à cet emplacement :

```
/docker/stirling-pdf/extraConfigs/settings.yml
```

Il permet de configurer l'application plus en profondeur. Nous pouvons **définir un nom personnalisé pour l'application**, mais aussi **configurer l'authentification OAuth2**. Ceci peut s'avérer utile pour s'appuyer sur **un service tiers pour l'authentification des utilisateurs** (Google, GitHub, KeyCloak).

Nous constatons aussi la présence de deux options ayant pour objectif de **protéger l'interface de connexion des attaques par brute force**. En effet, il y a un **verrouillage de comptes** activé par défaut : 5 mauvaises tentatives d'authentification à suivre vont engendrer le verrouillage d'un compte pendant 2 heures.

```
loginAttemptCount: 5 # lock user account after 5 tries
loginResetTimeMinutes: 120 # lock account for 2 hours after x attempts
```

Voici un aperçu de ce fichier de configuration que vous pouvez modifier selon vos besoins :

image.png

Pour aller plus loin, vous pouvez aussi publier l'application sur un nom de domaine à l'aide du reverse proxy de DSM.

Révision #3

Créé 28 décembre 2024 11:13:14 par Angelo

Mis à jour 1 avril 2025 18:42:33 par Angelo