

campusfrance.org et ses sites pays

Dossier d'Architecture Technique (DAT)



AXESS
Digital & Business Experts

Date : décembre 2024

Version : 1.0

Auteurs : AOWS Anne-Sophie Picot

Destinataires : Campus France, Noëlle Kerfant, Coralie Legrand

Préambule

Ce document reprend les principales spécifications techniques liées au socle campusfrance comprenant le site international campusfrance.org ainsi que l'usine à site ayant généré 116 sites pays.

Ce document est une compilation des informations techniques à jour concernant le socle et a été rédigé alors que le socle est déjà en ligne.

1 Environnement technique

1.1 CMS

1.1.1 Version du CMS

Le socle campusfrance.org est propulsé par le CMS Drupal, version 10 dans sa dernière version stable courante, voir pour de plus amples informations : <https://www.drupal.org/project/drupal>

1.1.2 Modules Contribués

Afin de limiter la maintenance du site lors de la phase d'exploitation, les développements des fonctionnalités utilisent au maximum les modules contribués de l'écosystème Drupal.

1.2 Versioning et environnements

L'ensemble des sources du socle campusfrance.org utilise un système de versioning : GIT (cf. <http://git-scm.com/>).

Les 3 branches principales suivantes sont utilisées :

- **main**: branche la plus à jour, correspondant à l'instance de développement principale
- **preprod** : branche correspondant à l'instance de pré-production
- **prod** : branche correspondant à l'instance de production

1.3 Encodage des caractères

L'ensemble des codes sources et données du SGBD du projet utilisent l'encodage UTF-8.

1.4 Infrastructure - hébergement

1.4.1 Serveur web

Nginx : la configuration reprenant la recipe officielle fournie par Nginx :

<https://www.nginx.com/resources/wiki/start/topics/recipes/drupal/> .

Drupal 10.1 demande une petite modification de la configuration Nginx qui est expliqué ici <https://www.drupal.org/node/2888767#nginx-php-fpm>

1.4.2 PHP

L'ensemble des développements (Drupal / modules contribués / développements custom) sont compatibles avec PHP 8.2 (version recommandée à date). Le minimum requis pour un projet Drupal 10 est spécifié sur la page suivante : <https://www.drupal.org/docs/system-requirements/php-requirements>

Les extensions PHP suivantes sont à minima activées :

- | | | |
|---------------------------------|---|-------------|
| • ctype | webp | • openssl |
| • curl | https://www.php.net/manual/en/image.installation.php | • pcrc |
| • date | | • pdo |
| • dom | • json | • pdo_mysql |
| • fileinfo | • pcrc | • readline |
| • filter | • mbstring | • session |
| • hash | • memcached | • sockets |
| • gd (configuré avec un support | • mysqli | • SimpleXML |
| | | • SPL |

- tokenizer
- xml

1.4.3 Base de données

Le SGBD respecte au minimum la configuration requise de la documentation officielle : <https://www.drupal.org/docs/system-requirements/database-server-requirements> ou <https://www.drupal.org/docs/understanding-drupal/how-drupal-9-was-made-and-what-is-included/environment-requirements-of#comment-13741116>.

A titre d'information nous utilisons sur notre serveur de développement MariaDB en version 10.6 (cf. <https://mariadb.org/about/>, [Prérequis drupal](#)).

1.4.4 Syslog

Sur l'instance de production, les logs de Drupal sont configurés pour être écrit en syslog (et non en dblog, dans la base de données) afin d'optimiser les performances. Les logs sont accessible via un outil tiers Graylog.

1.5 Historique du site et volumétries

1.5.1 Volume de contenu

A date de composition de ce document, le site national portait :

- 8274 nodes (tout statut, tout type)
- Dont 2611 actualités (tout statut)
- 794 articles
- 50 dossiers
- 568 espaces
- 118 FAQ
- 1 hub
- 15 pages
- 2882 ressources documentaires
- 7 timelines
- 122 témoignages
- 183 vidéos
- 149 villes
- 19 webforms

- 0 landing
- 751 évènements
- 7971 médias (tout statut, tout type)

1.5.2 Volume de fichiers

A date de composition de ce document, le site national portait :

- 93Go de fichiers (code source + files pour l'ensemble des sites)
- 17Go de fichiers pour le site national uniquement
- 117 bases de données, permettant de créer les 117 sites
- 19Go poids de la base de données du site national

1.5.3 Volume des utilisateurs

A date de composition de ce document, le site national portait : 847 091 comptes utilisateur.

1.6 Sites pays :

A date de composition de ce document 116 sites ont été créés, tous ne sont pas en ligne.

2 Structure des données

Nous distinguons ici 6 types de structures :

- **Les entités**, qui sont des objets avec une/plusieurs propriétés très spécifique à leur utilisation
- **Les nœuds (nodes)**, qui sont des entités ayant un minimum de propriétés et de fonctionnalités associées, adaptées à l'édition de contenu
- **Les taxonomies**, qui sont des entités avec un comportement associé de liaison commune entre les contenus
- **Les médias**, qui sont des entités avec des propriétés à minima de référencement d'un média.

- **Les webforms**, qui sont des entités ayant un minimum de propriétés et de fonctionnalités associées, adaptées à la gestion de formulaire dynamique.
- **Les config pages**, qui sont des entités composées d'un ensemble de propriétés et de fonctionnalités associées, adaptées à l'édition de la page d'accueil, de pages liste ou de contenus répétitifs.

2.1 Les types de contenus (nœuds)

- **Actualité** : page modulable remontant dans la liste des actualités.
- **Article** : page modulable éditoriale.
- **Dossier** : contenu éditorial permettant de regrouper des contenus et de créer de la navigation interne.
- **Espace** : localisation importée depuis un fichier csv
- **FAQ** : contenu de type question/réponse pouvant être remonté sur d'autres pages.
- **Hub** : contenu éditorial utilisé pour créer de la navigation interne au site.
- **Landing** : contenu permettant de créer une page d'arrivée suite à une campagne média (non utilisé).
- **Newsletter** : contenu permettant de générer une newsletter depuis le site et de l'envoyer à sa liste d'abonné
- **Page de base** : page simple, contenus composés de paragraphes.
- **Ressource documentaire** : contenu permettant de présenter un document et de le retrouver dans l'espace documentaire.
- **Timeline** : page éditoriale présentée sous la forme d'une ligne de temps avec étapes.
- **Vidéo** : page mettant en avant un média vidéo.
- **Ville** : page éditoriale présentant une ville, permet la remontée de contenus.
- **Webform** : page permettant d'exposer un formulaire webform
- **Évènement** : contenu éditorial remontant dans l'agenda et disposant de dates.

2.2 Les taxonomies

- **Catégorie espace documentaire** : organisation des ressources documentaires dans une liste à plusieurs niveaux
- **ChatBot** : discussion pré-configurée proposée dans le chatbot intégrée

- **Domaine d'étude** : thématiques d'étude pour étudiants
- **Groupe extranet** : groupes associés au profil utilisateur
- **Guide recherche formation** : pour la partie "Quelle formation recherchez-vous ?"
- **Niveau d'étude** : permet de qualifier les études des utilisateurs
- **Organisme** : Liste proposée sur le profil utilisateur
- **Partenaire** : sont associés à des contenus
- **Pays** : liste de pays pour le profil utilisateur principalement
- **Public cible** : typologie d'utilisateurs
- **Rubrique** : utilisé sur les contenus
- **Région** : liste de région pour géolocalisation
- **Réseaux sociaux** : liste des réseaux sociaux pour composition du menu correspondant
- **Sujets extranet** : grandes thématiques
- **Thème** : utilisé sur les contenus
- **Type d'échange** : utilisé sur les contenus
- **Type newsletter** : permet de créer des newsletters différentes pour un site donné
- **Type évènement** : utilisé sur les contenus
- **Témoignage thème** : utilisé sur les contenus
- **Étiquette** : utilisé sur les contenus

2.3 Gestion des médias

Depuis la version 8.4 de Drupal, le module Media entity a été inclus dans le core : https://www.drupal.org/project/media_entity

Voici la liste des médias qui seront créés :

- **Document** : permet de mettre à disposition un fichier pdf, doc, docx, xls, xlsx, ppt, pptx, zip, odt ou ods
- **Facebook** : permettait de relier une publication facebook, non utilisé à date
- **Iframe** : permet de coller l'url d'une page pour l'afficher dans une iframe sur un contenu
- **Image** : permet de stocker une image sur le site, formats : png, jpeg, jpg et gif.

- **Tweet** : permettait de relier une publication X/Twitter (avant le changement d'API du réseau)
- **Vidéo** : permet d'embarquer une vidéo publiée sur youtube à partir de son url.

2.4 Les webforms

Les webforms sont librement contribuables dans le site et largement utilisés sur les sites pays.

Les webforms sont affichés dans des contenus de type webform.

2.5 Gestion des révisions

Les révisions (core Drupal) seront activées sur la plupart des contenus, afin de garantir une historisation de l'ensemble des modifications sur ces entités.

L'ajout du module diff (<https://drupal.org/project/diff>) permettra d'obtenir une finesse supplémentaire pour consulter quelles ont été les modifications effectuées.

2.6 Liste des différents types de pages

Le site est composé de différents types de page :

- **Page des contenus** : affichage sur une url propre d'un contenu d'un type donné, structure des templates définis en amont, création des pages à la charge des contributeurs.
- **Page de liste** : remontée en liste d'entités publiés par type, avec filtres, pages spécifiées et créées programmatiquement.
- **Recherche** : remontée en liste avec des fonctionnalités de filtres et de tris, depuis un index externe au site.
- **Page spécifique** : interfaces spécifiées en amont et créées programmatiquement

3 Structuration globale du site

3.1 Moteur de recherche

Les fonctionnalités de recherche sont assurées par l'utilisation d'un serveur Apache Solr (cf. <http://lucene.apache.org/solr/>).

3.1.1 Modules utilisés

Les modules utilisés par le moteur de recherche sont les suivants :

- [Search API](#)
- [Search API Solr](#)
- [Facets](#)

3.2 Multilinguisme

Le site campusfrance.org est disponible en 3 langues. Chaque instance site pays peut être configurés en plusieurs langues.

Le back-office nativement disponible en anglais, dispose d'une traduction communautaire en français. Il est possible que certaines chaines ne soient pas toutes disponibles en français, des compléments de traduction pourront être ajoutés et/pi surchargé lors de la vie du site.