

# GROUPE WEBHOSTING

## 1. INTRODUCTION

Une offre Web complémentaire de celle de la DSI et plus en adéquation avec les besoins exprimés par le département MathNum, a été proposée et mise en place en 2008, avec le Cati 1ère génération CIAM.

Le groupe de travail WebHosting propose et assure le suivi du CMS (content management system) Drupal, dans la création de sites Web au sein d'un hébergement sécurisé.

Les temps forts du groupe WebHosting pour l'année 2022 sont **l'arrivée** dans le groupe de **Sylvain Lanoe** de l'unité BioSP à Avignon et **le départ** du groupe de **Christian Poirier** de l'unité MaIAGE à Jouy en Josas

## 2. ETAT DES LIEUX

### Présentation de l'organisation actuelle

Le groupe est composé de quatre agents de différentes unités du département MathNum.

#### Mission principale :

**Valérie Vidal**, pour le Département MathNum (0,5 ETP) est l'animatrice du groupe WebHosting et occupe les fonctions de Webmaster du Portail MathInfo (site institutionnel du département), d'administratrice CMS Drupal et Webmaster, de la communication auprès des utilisateurs, aide, à la conception de site, du développement et d'accompagner les utilisateurs (conseils, support etc.)

#### Mission transversale :

**Mikael Grialou**, pour le Département Mathnum (0,2 ETP) occupe les fonctions d'administrateur LAMP (Linux, Apache, MySQL, PHP), administrateur CMS Drupal, développements spécifiques éventuels et l'accompagnement des utilisateurs (conseils, support etc.)

**Sylvain Lanoe**, pour le Département Mathnum (0,2 ETP) occupe les fonctions d'administrateur LAMP (Linux, Apache, MySQL, PHP), administrateur CMS Drupal, l'accompagnement des utilisateurs (conseils, support etc.)

**Maurice Baudry**, pour le Département Mathnum (0,05 ETP) occupe les fonctions de support à l'administration système et conseil et veille technique

### Description de l'hébergement

En 2021, 43 sites étaient hébergés. En 2022, 47 sites sont hébergés par la plateforme sans compter les sites statiques. Le nombre de site est quasi constant sur les six dernières années.

#### Différentes catégories de sites sont hébergées

- Les réseaux métiers (ex RESSTE ...)
- Les évènements (ex colloque StatMathAppli)
- Des Unités internet/intranet (ex MaIAGE)
- Des outils de gestion (ex ColDem MaIAGE)
- Les CATI & PEPI (ex IMOTEP, IBIS)
- Les projets & les groupes de travail (ex MSE, Stat4Plant)
- Des plateformes (ex Migale).

- Des portails (ex Workflow4Metabolomics, MathInfo).
- Des écoles chercheurs (ex EPICASA)
- Les clusters (ex BioSP-Cluster)

## Tâches de fond réalisées au niveau du groupe

### Mise à jour mensuelle du core (noyau), des modules et des thèmes de Drupal avec

- Planification de la tâche à accomplir.
- L'envoi d'un mail d'informations aux gestionnaires de sites.
- Réalisation de la tâche (script Rundeck).
- Vérification et validation de la tâche.
- L'envoi d'un mail d'informations aux gestionnaires de sites.

### Création d'un site

- Réception d'une demande de site par mail via un formulaire ou par voie de connaissance.
- Échange avec la personne des besoins exprimés :
  - *Choix du nom du site avec ou sans DNS.*
  - *Établir le ticket auprès de la DSI (le demandeur ou nous) si DNS.*
- Validation avec l'utilisateur des besoins :
  - *Par exemple le module Webform pour les colloques.*
  - *Mail applicatif.*
- Création, installation et paramétrage du site.
  - *Création d'une base de données MySQL.*
  - *Création et configuration d'une instance Drupal.*
  - *Configuration du fichier setting.php de Drupal et Apache.*
  - *Installation des modules que nous conseillons et d'un choix de thèmes stables.*
- Rédaction et affichage des contenus obligatoires : Mentions légales, CGU et Gestion des cookies.
- Création d'un compte administrateur pour l'utilisateur.
- L'envoi d'un mail le prévenant et lui rappelant les points importants :
  - *Le mail du support.*
  - *Son login et mot de passe.*
  - *Ce qui a été fait, selon les besoins exprimés.*
- Demande de certificat de sécurité à la DSI.

### Aide et accompagnement pour les sites déjà créés

- *Installation de module spécifique à la demande des webmasters de sites.*
- *Aide, au paramétrage de ces modules*
- *Accompagnement à la migration de sites dans les différentes versions du CMS*
- *Transfert de site sans évolution dans un mode dit « site statique » (page HTML et CSS) dans un but d'archivage avec visibilité.*

### Rôle de Webmaster pour une dizaine de sites

- Design du site
- Installation modules pour l'ergonomie du site
- Mise en ligne du contenu

### Veille et évolution des outils et services :

- *Mise en place de certificats de sécurité générique pour faciliter l'exploitation.*
  - *Type : gen-SSL-wildcard pour les sites en \*.mathnum.inrae.fr*
  - *Type : gen-SSL-AlterName pour les sites avec un DNS dédié*

## Infrastructure

Espace NetAPP : Les données de webh représentent 16Go sur les 3 To loué à la DSI (mutualisé avec la ForgeMIA)

Machines virtuelles :

- Serveur - vm-pkg-32 - 4 VCPU - 8go de RAM - Actuellement inutilisée, nous la conservons, car une prochaine migration vers Drupal 10 est annoncée pour 2023.
- Serveur Dupal 9 - vm-pkg-61 - 4 VCPU - 8go de RAM - Drupal 9.
- Serveur d'administration - admin-webh - 1 VCPU - 2 Go de RAM - 20 Go disque
- Serveur Pré-Production MATHINFO - vm-pkg-27 - 1 VCPU - 2go de RAM - 5,4 Go disque - Est amené, à disparaître fin octobre 2023
- Serveur Production MATHINFO - vm-pkg-28 - 1 VCPU - 2go de RAM - 5,4 Go disque - Est amené, à disparaître fin octobre 2023

Taux d'occupation des sites de l'infrastructure des VMs			
<b>vm-pkg-61</b>			
46 sites sous Drupal 9 (taille - nom du site)			
289M : AG-MathNum2022 829M : applibugs 269M : biblio-maiage 383M : biobaye 289M : bioinfomics 1,2G : biosp 235M : biosp-cluster 378M : ingenum 418M : coldem-maiage 381M : coldem-mia-ps 290 M : dynstoch 579M : imotep	1,5G : maiage 238 M : metasimfood 1,2G : mexico 367 M : mia-ps 426M : migale 479M : mistea 660M : modstatsap 269M : mse 491M : ofe2021 258M : omer 314M : peakforest 1,2G : pepi-ibis	243M : pherosensor 302M : phis 259M : pole_sci 273M : portail-mathinfo 341M : prodmia 240M : pse 260M : reseauA2 523M : reseau-cisstats 538M : reseau-payote 1,2G : reseau-resste 282M : resiste 346M : rh-maiage	295M : sesame 265M : smpgd 260M : ssb8 289M : stat4plant 387M : statmathappli 650M : sysmics 370M : trombi-maiage 305M : usemetabo 326M : webh 586M : workflow4metabolomics
10 sites statiques en HTML (taille - nom du site)			
556K : cbme.ssbgroup.fr 4,6M : ciam 696K : nemo.ssbgroup.fr	481M reseau-m3d.fr 2,3M : projet-peerless 48M : workshop-ssiab9	74M : swg2014 61M : cascisdi.inrae.fr 203M : ag2019.mia.inrae.fr	255M : epicasa
2 sites sous Drupal 7 (taille - nom du site)			
<b>vm-pkg-28 - Prod MathInfo</b>		<b>vm-pkg-27 - Pré-prod MathInfo</b>	
288M mathinfo		281M mathinfo	

## Stack technique mise en œuvre

**CMS Drupal** est un système de gestion de contenu libre et open-source, publié sous la licence publique générale et écrit en PHP.

**Matomo** est un logiciel libre et open source de mesure de statistique web, conçu pour être une alternative libre à Google Analytics. Matomo fonctionne sur les serveurs web PHP/MySQL et s'appuie sur l'analyse des logs du serveur web pour fournir des indicateurs de fréquentation.

**Rundeck** est un logiciel libre permettant l'automatisation d'administration de serveurs via la création de jobs ou tâches. L'interface de Rundeck permet de créer des jobs ou tâches applicatives, qui peuvent s'exécuter de façon périodique ou manuelle, sur un ou plusieurs serveurs.

**KeePass** est un gestionnaire de mot de passe utilisant deux algorithmes de chiffrement, permettant de sauvegarder un ensemble de mots de passe dans une base de données chiffrée sous la forme d'un fichier. Ce gestionnaire de mot de passe s'ouvre avec un mot de passe maître chiffré.

**Uptime Robot** est un service de surveillance en ligne de la disponibilité des sites, notre choix s'est porté sur le forfait gratuit qui propose la surveillance du HTTP(S) sans les erreurs SSL, du temps de réponse, des pings, des ports, des mots clés et une rétention de trois mois des journaux.

**NextCloud** est un portail d'accès, de partage et de coédition de document, il s'agit d'une offre de la DSI, auprès de qui nous avons ouvert un compte.

## Identifications des risques encourus et moyens de contrôles mis en place

Le risque majeur est la sécurité des serveurs, des logiciels et des sites. Pour assurer cette sécurité, le collectif a mis en place une veille technologique afin d'être réactif en fonction des alertes nous pouvons intervenir sous 24 heures :

- **Serveurs virtuels packagés (VM)** : il s'agit d'une offre d'environnement logiciel de virtualisation hébergé au data center de Toulouse. La sécurité et les problèmes relatifs à ses différentes couches sont gérées et administrées par la DSI. Nous interagissons avec eux via des tickets Ariane.
- **Une veille technologique du CMS Drupal** :
  - Réception de mails quotidiennement, sur une adresse mail dédiée (groupe-webhmia@inrae.fr) via les instances Drupal, des mises à jour de sécurité du core, des modules et des thèmes qui doivent être faits.
  - Compte individuel sur le site de drupal.org pour recevoir la lettre d'information hebdomadaire de Drupal, le bulletin mensuel de l'Association Drupal et des mails ad hoc occasionnels avec des mises à jour importantes.
- **Une veille technologique de Matomo** : Compte commun au groupe WebHosting donnant accès aux données des sites utilisant ces statistiques et au blog de Matomo où les informations concernant les dernières mises à jour disponibles.
- **Une veille technologique de Rundeck** : Compte individuel sur l'interface en ligne donnant accès aux scripts de gestion des projets et aux news de la communauté Rundeck.

### 3. PERSPECTIVES

- **Audit sécurité** : Faire réaliser un audit de sécurité.
- **Architecture système** : Réfléchir aux évolutions de l'architecture du matériel pour suivre les besoins et collaboration avec les acteurs de la DSI.
- **Le groupe WebHosting** : Élargir le groupe dans la consolidation et la perspective de d'évolution du nombre de site hébergé.

### 4. LA FEUILLE DE ROUTE

#### Travaux réalisés en 2021

Nous nous étions donnés comme objectifs de réaliser ces deux actions étroitement liées ce qui a été fait.

**La migration vers la version 9 du CMS Drupal** avant le 2 novembre 2022, date de fin du support de la version 8 et 7 pour cela :

- Le provisionnement et la configuration d'une nouvelle VM a été nécessaire, pour accueillir Drupal 9.
- Une récolte des souhaits pour chaque site a été faite, auprès des gestionnaires (les choix étaient de rendre le site statique, l'archiver, le supprimer, le migrer vers la nouvelle version de Drupal ou from scratch).
- Une action a ensuite été effectuée, selon les vœux exprimés et le travail exécuté a été validé par les gestionnaires concernés.

L'évolution du département MIA vers MathNum dans le nouveau contexte de INRAE, a nécessité **le changement du nom de domaine** de `informatique-mia.inrae.fr/NomSite` pour `NomSite/mathnum.inrae.fr` puis, une règle de redirection apache a été mise en place pour les sites concernés.

#### Travaux en cours de réalisation en 2022

Nous avons été contactés courant août, par Karine Daubas (Responsable de l'équipe Access Management), dans le cadre du projet IAM, un travail est en cours sur les accès LDAP à l'annuaire INRA et de nombreux accès à la vm-pkg-61 via le LDAP INRA ont été constatés.

C'est la raison qui nous pousse à réviser notre mode de connexion au LDAP INRA, car il est voué à disparaître en 2023 et les cnx (connexions) directes pour l'authentification Web seront proscrites sur le nouveau LDAP.

Il est donc nécessaire d'étudier la possibilité de migrer vers la solution nationale authentification.inrae.fr avec une intégration dans la fédération Renater.

Concernant les connexions via le CAS ce protocole étant compatible avec la nouvelle solution, il suffit d'établir un ticket Ariane afin qu'on déclare le site sur la nouvelle solution, ainsi la migration pourra se faire à la date souhaitée.

Nous allons mettre en place le SAML (Security Assertion Markup Language) pour cela des up grade et mises à jour seront nécessaires :

- Up grade de PHP 7 vers PHP 8 (fait par Mikael),
- up grade de composer (fait par Mikael),
- tests de mise en œuvre (en cours – action de groupe).

## Prévision pour l'année 2023

Migration vers la version 10 du CMS Drupal : la communauté Drupal a annoncé le 17 juillet 2020 que la version de Drupal 9 ne sera plus supportée du fait de la sortie de la version de Drupal 10 le 10 novembre 2023, plus d'information sur le site officiel de [Drupal](#).

A prévoir courant 2023 :

- Récolte des souhaits de chaque gestionnaire de site puis évolution vers l'état souhaité.
- Vérification via un module de la comptabilité des modules installés sur chaque site avant migration
- Test de migration des sites et migration des sites après validation.

## 5. FORMATION

Nous réfléchissons à ce style de formation autour des outils que nous utilisons quotidiennement, afin d'améliorer notre prestation, auprès des utilisateurs comme :

- **Drupal** administrateur et développeur
- Initiation à **Matomo** et **Symfony**

Nous ne formulons aucune demande particulière auprès du département pour l'année à venir.

## 6. BILAN

Analyse des forces et les faiblesses de notre organisation, ainsi que les opportunités et les menaces de l'environnement.

### Notre force :

- La cohésion du groupe qui est stable depuis 2016,
- la mise en place d'une Visio conférence hebdomadaire favorisant les échanges sur l'avancement d'un projet, collaborer pour répondre rapidement à une demande de service, etc. ,
- l'élaboration de tutoriels disponibles à tous les membres du groupe (partage de connaissance),
- l'amélioration des outils d'administration pour faciliter la gestion,
- le respect du RGPD et l'affichage sur tous les sites des Mentions Légales, CGU et Gestion des cookies, ainsi que le suivi des nouveautés/modifications sur la protection des données personnelles à INRAE.

**Notre principale difficulté :** nous avons signalé en 2021 que la taille du groupe représentait un risque « résidant dans la taille du groupe et en particulier sur le fait établi, qu'il ne s'appuyait que sur la disponibilité de deux personnes, consacrant du temps sur les demandes de sites et les interventions » C'est aujourd'hui révolu avec l'arrivée Sylvain Lanoe.

**Les opportunités :** chacune des onze Unités du département représente un potentiel important de création de site pour des réseaux, évènements, Unités internet/intranet, CATI & PEPI, projets & les groupes de travail, plateformes, portails, écoles-chercheurs, etc.

**La menace :** manque de visibilité sur l'offre DSI d'un hébergement dans l'état de l'art comme l'offre du Data Center de Toulouse (offre VM et stockage DSI).

En conclusion, le soutien sans faille du département MathNum reste pour nous la meilleure des motivations.