



Portail math www.portail-math.fr point d'avancement sept 2013

Jean-Luc Archimbaud Mathdoc Jean-Luc.Archimbaud@ujf-grenoble.fr

Journées mathrice - sept 2013 V1.0 – 1 oct 2013





PLAN



- Le projet
 - Page : http://www.mathdoc.fr/projet_portail_math
 - Présentation 'Portail math en mai 2013 : une vue de l'intérieur' ANF RNBM mai 2013 : http://www.mathdoc.fr/publis/2013_05_portail_maths_inside_v1.1.pdf
- Le contenu

Drupal

La suite : V2...



Projet: objectif du portail



- L'objectif du projet portail math est d'aboutir à la mise en place d'un portail à façon satisfaisant les besoins professionnels des membres de la communauté enseignement supérieur et recherche en mathématiques installés en France en offrant un guichet d'accès simplifié :
 - À la documentation scientifique et aux services associés, en accès libre ou contrôlé, avec des fonctions de recherche... évoluées;
 - Aux services facilitant le travail nomade & collaboratif;
 - Aux informations institutionnelles et professionnelles.
- 4+1 partenaires : INSMI, RNBM, Mathrice, Mathdoc + UJF
- Conception V1: sept 2012 sept 2013 puis V2, V3...



Projet: planning prévu



- NSMI : première version 'Le plus tôt possible' → V1 en sept 2013
- 4 étapes
 - Définir : organisation du projet, services à rendre (quoi ? pour qui ?), matériels et logiciels et moyens humains nécessaires : 9/2012-1/2013.
 - Mettre en place une maquette basée sur Drupal : 10/2012-8/2013 contenant des descriptifs précis et indexés de ressources et services disponibles (situés ailleurs que sur le portail).
 - 2 utilités : état de lieux et un contenu minimal mais utile pour des utilisateurs Ouvrir la maquette en accès public V1 du portail en septembre 2013. Cela remplacera entre autres l'actuel Portail Documentaire Mathématique de Mathdoc.
 - → Faire des choix techniques pour les services avancés : 3/2013-12/2013 en particulier sur les référentiels, contrôles d'accès, recherches, indexations, formats...
 - ➤ Faire évoluer l'existant de la communauté et installer des services avancés (contrôle d'accès, recherches..) sur le portail : 7/2013-... Ceci donnant lieu à de nouvelles versions du portail : V2-V3-... La version V2 du portail, en janvier 2014, intégrera des fonctions d'authentification pour accéder au moins aux ressources documentaires actuellement accessibles via la PLM.



Projet : planning : état le 1/10



- Serveurs, drupal, contenu : tout est prêt
- Graphisme : sociétés Kitchenette et Negative Network
 - Logo, favicon, choix des couleurs : fait
 - Thème : en test (encore quelques bugs à corriger) par le comité éxécutif et l'équipe technique
- Reste à faire :
 - Tests par équipe portail
 - Tests par testeurs externes
- Ouverture semaine prochaine si comité de pilotage est d'accord



Projet: organisation



- Responsable du projet : Thierry Bouche Coordinateur technique : Jean-Luc Archimbaud
- Comité de pilotage
 - Animateur : Guy Métivier
 - Bureau exécutif : Thierry Bouche et Jean-Luc Archimbaud
 - Autres membres :
 - RNBM : Francesca Leinardi, Frédéric Hélein
 - Mathrice : Jacquelin Charbonnel
 - UJF : Hervé Gaussier
 - Communauté recherche maths : Stéphane Cordier, Fabrice Planchon
 - (Thierry Bouche représente aussi Mathdoc dans ce comité)
- Comité exécutif
 - Thierry Bouche (Mathdoc), Francesca Leinardi (RNBM), Jacquelin Charbonnel (Mathrice), Jean-Luc Archimbaud (coord technique)



Projet: concepteurs



- Jean-Luc Archimbaud Mathdoc 60 % depuis juillet 2012 : animateur
- Nadège Arnaud LMV et RNBM 10 % depuis janv 2013
- Dominique Barrère Institut de Mathématique de Toulouse (IMT) et RNBM 10 % depuis nov 2012
- Brigitte Bidegaray-Fesquet LJK et MI2S 5 % depuis mai 2013
- Thierry Bouche Mathdoc 15 % depuis juillet 2012
- Jacquelin Charbonnel Mathrice 10 % depuis sept 2012
- Simon Chevance Mathdoc 10 % depuis oct 2012
- Philippe Depouilly Institut de Mathématique de Bordeaux (IMB) et Mathrice 10 % depuis oct 2012
- Nicolas Houillon Mathdoc de l'ordre de 50 % (% pas vraiment fixé) depuis fév 2013
- Myriam Laurens Mathdoc 5 % depuis avril 2013
- Francesca Leinardi Institut Fourier et RNBM 10 % depuis sept 2012
- Franck Lontin Mathdoc 15 % depuis oct 2012
- Henri Massias XLIM et Mathrice 5 % depuis mai 2013
- Alexandra Miric Bibliothèques MIR et RNBM 10 % depuis déc 2012
- Alain Royer Département de Mathématiques d'Orsay 5 % depuis fév 2013
- Gregory Thureau IRMA, Strasbourg 5% depuis mars 2013
- Anna Wojciechowska LATP, FRUMAM et RNBM 10 % depuis oct 2012

→ En 2012 : 1.6 ETP – actuellement : 2.4 ETP mais très répartis



Projet : responsables éditoriaux - relecteurs



- Responsables éditoriaux de site : Thierry Bouche et Jean-Luc Archimbaud.
- Responsables éditoriaux de rubrique :
 - Ensemble de documents : Anna Wojciechowska et Dominique Barrère.
 - Point d'accès : Alexandra Miric et Anna Wojciechowska.
 - Base bibliographique : Francesca Leinardi (et Thierry Bouche de manière provisoire).
 - Licence/contrat : Gregory Thureau et Francesca Leinardi.
 - Documents utiles : Jean-Luc Archimbaud (recherche de binome)
 - Service informatique : Jacquelin Charbonnel et Henri Massias.
 - Site utile: Thierry Bouche et Brigitte Bidegaray-Fesquet.
- Relecteurs syntaxique :
 - Nadège Arnaud, Myriam Laurens



Projet : groupe référent utilisateurs



- Chercheurs ou Enseignants chercheurs : Brigitte Bidegaray-Fesquet (LJK), Benoit Claudon (IHC), François Ducrot (LAREMA), Benoit Kloeckner (Institut Fourier), Boris Kolev (LATP), Sébastien Loustau (LAREMA), Hussein Mourtada (Institut de Mathématiques de Jussieu), Geoffrey Powell (LAREMA), Karim Ramdani (Institut Élie Cartan), Bernard Teissier (IMJ)
- Doctorants : Marie-Angela Cornelie (Institut Fourier), Guenaëlle De Julis (Institut Fourier), Kevin Layat (Institut Fourier), Danièle Turchetti (LMV)
- Informaticiens : Gérard Grancher (LMRS), Henri Massias (XLIM), Benoit Métrot (LMA)
- Documentalistes: Odile Luguern (ENS DMA), Julie Janody (Bibliothèques MIR), Grégory Thureau (IRMA)



Projet : méthodes de travail



- Tous les concepteurs ont participé à la conception et continueront
 - Selon leurs responsabilités, compétences, temps disponible
- Depuis début 2013, travail sur une plate-forme opérationnelle (serveur Web drupal) qui s'est construite au fil du temps (fonctionnalités et contenu)
 - Aucune mise en forme graphique au départ
 - Menus colonne de gauche : futurs menus utilisateurs
 - Menus colonne de droite : menus de travail
- Minimum de réunions, utilisation de listes de diffusion...
- Documentation maximale
- Essai d'avoir pour chaque tâche un binôme
- Une fois le site ouvert, la publication de nouveaux contenus continuera au fil de l'eau ainsi que l'ajout de fonctionnalités



Projet: risques



Listés en mai 2013 - par poids décroissant (à mon avis)

- Portail spécifique pour les chercheurs en math : inutile
 - Il y a déjà des outils puissants : google, Zentralblatt, HAL, PLM...
- Projet trop spécifique INSMI, indépendant des projets INIST...: mal intégré dans plans IST et DSI CNRS
- → Perte du soutien fort d'une instance (ou d'une personne influente) : INSMI actuellement
- Objectifs trop ambitieux par rapport aux moyens humains
 - Actuellement participation trop éparpillée et trop faible (en particulier côté documentalistes)
- Faible participation des membres des partenaires RNBM, Mathrice, Mathdoc
- Frein important voire refus d'évolution de l'existant chez RNBM, Mathrice, Mathdoc
 - Le portail va empiéter sur les périmètres de chacun
 - Et il va demander de faire évoluer l'existant de chacun
- Mauvais choix techniques et modes de création du contenu
 - Coeur technique : choix Drupal : ne fait pas l'unanimité à Mathdoc, ni à Mathrice par exemple. Autres choix : SPIP, tout développer, reprendre un développement existant...
 - Contenu : choix descriptions rédigées. Autres choix : aspiration de contenu, moteurs de recherche...
- Mauvais choix en terme de contenu



Contenu: outil: glossaire



- Besoin d'un langage commun, compréhensible par tous et de référence
 - Pour les concepteurs
 - Pour les utilisateurs
- Création d'un glossaire : module Lexicon de Drupal
 - Lorsqu'un terme est défini dans le glossaire, lors de sa première itération sur une page affichée il est souligné et quand on passe la souris audessus, sa définition s'affiche.
 - Les utilisateurs pourront comprendre, les concepteurs peuvent parler de la même chose.
 - Ne sont définis que les termes qu'on emploie avec une signification particulière : communauté math, ressource documentaire, point d'accès...
 - Ajout de nouveaux termes au fil de l'eau.
 - Coordinateurs : Francesca Leinardi, Dominique Barrère, JL Archimbaud



Contenu: quelles informations?



NB : En italiques : définitions extraites du glossaire interne au portail

- Ressources documentaires (document ou outil (catalogue, index...) qui gère des documents)
 - Ensembles de documents (regroupement identifiable de documents, ayant les mêmes conditions d'accès, présenté comme un seul objet)
 - Bases de données bibliographiques (ensemble organisé et structuré de références de documents)
 - Points d'accès (page Web point d'entrée à une ressource documentaire (un ensemble de documents ou une base de données bibliographiques) ou un service informatique)
 - Licences/contrats (contrat : document contractuel qui permet à un ensemble de personnes (appelées bénéficiaires) d'accéder à un ensemble de ressources documentaires qui ne sont pas en libre accès, selon les termes d'une licence d'utilisation)
 - Documents utiles (document utile à la communauté math, en lien avec les thématiques de ce portail et créé sur ce portail)
- Services informatiques (service, accessible par internet, offrant des fonctions d'organisation (annuaire, agenda...) ou de travail collaboratif (espaces partagés...) ou de calcul ou de communication (listes de diffusion, visioconf...)...)
- ◆ Sites utiles (site web (institutionnel, scientifique...) utile à la communauté math)

En V1 : des descriptifs



Contenu: descriptifs

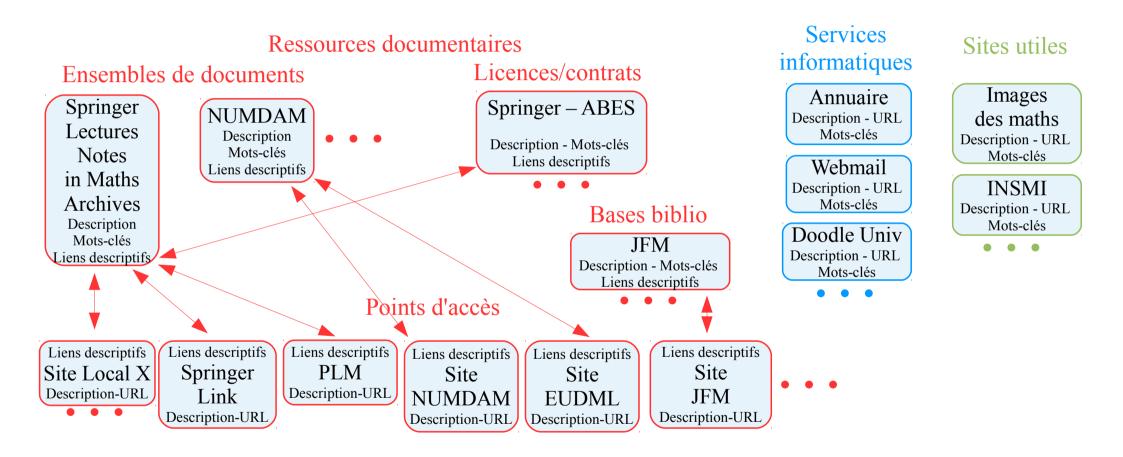


- Les ensembles d'information sont sous forme de descriptifs
- Chaque descriptif contient
 - Description de l'objet : suite de champs
 - Liens vers d'autres descriptifs internes
 - Liens vers des sites extérieurs
 - Mots-clés



Contenu: ensembles d'info (V1)







Contenu : champs des descriptifs Ensembles de documents



- Description courte
- Points d'accès : descriptifs sur ce portail URL
- Précisions sur les points d'accès
- Conditions d'accès : accessible
 - A tous
 - A la communauté Enseignement Supérieur -Recherche française
 - Aux laboratoires INSMI
 - Aux laboratoires membres du RNBM
 - A un ensemble de laboratoires ou d'établissements
- Précisions sur les conditions d'accès
- Description longue
- Description longue en ligne : URL
- Titres des documents
- Titres des documents en ligne : URL

- Types de documents
 - Revues
 - Livres
 - Actes de congrès
 - Thèses
 - Autres
- Précisions sur les types de documents
- Bases bibliographiques : URL
- Existence de licence non libre ? (O/N)
- Licences/contrats : descriptifs sur ce portail
- Hébergeur
- Contact hébergeur
- Autres informations
- Auteur du descriptif
- Date relecture complète
- Mot-clés : gestionnaire/hébergeur, éditeur



Contenu : champs des descriptifs Points d'accès



- URL
- Description courte
- Description longue
- Description longue en ligne : URL
- Descriptifs des ensembles de documents sur ce portail
- Autres ensembles de documents
- Descriptifs des bases bibliographiques sur ce portail
- Autres bases bibliographiques
- Descriptif en ligne des contenus accessibles : URL
- Descriptifs des licences/contrats sur ce portail
- Autres licences/contrats
- Mécanismes de contrôle d'accès
 - Aucun
 - Filtre d'adresses IP
 - Logins-Mots de passe sur le point d'accès
 - Fédération d'identité
 - Certificats électroniques
 - Autre

- Types de recherche
 - Plein texte
 - Mots clés
 - Metadonnées
 - Listes
 - Facettes
 - Autres
- Administrateur
- Contacts techniques
- Statistiques d'utilisation
- Modalités pour récupérer les statistiques d'utilisation
- Autres informations
- Auteur descriptif
- Date relecture complète
- Mots-clés : gestionnaire/hébergeur, éditeur



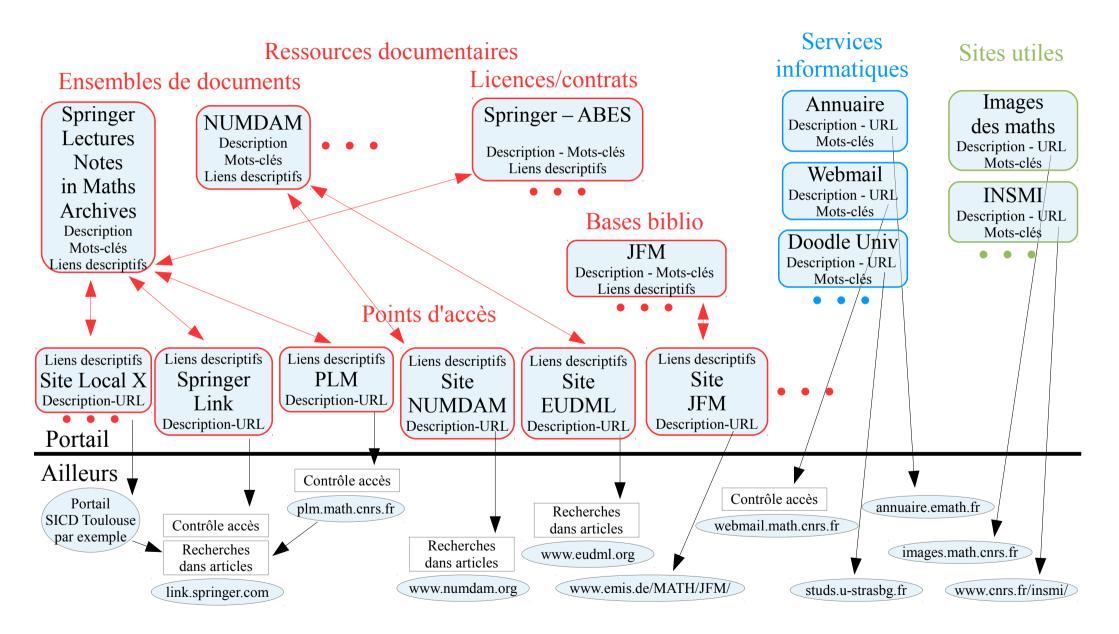
Champs: service informatique



- Description courte
- URL pour accéder directement à ce service
- Description longue
- Description en ligne de ce service
- URL pour avoir plus d'informations sur le service
- Gestionnaire de ce service
- Pour contacter le gestionnaire de ce service
- Etat de ce service Opérationnel / Opérationnel avec améliorations en cours / En projet / En développement / En test
- ◆ Conditions d'accès à ce service Accessible à tous / Enseignement Supérieur -Recherche française / Laboratoires INSMI / Laboratoires membres du RNBM / Personnes inscrites dans La PLM /à un ensemble de labos ou établissements
- Précisions sur les conditions d'accès
- Mécanismes de contrôle d'accès à ce service Aucun / Filtre d'adresses IP / Logins-Mots de passe sur le point d'accès / Fédération d'identité / Certificats électroniques / Autre

- Précisions sur les mécanismes de contrôle d'accès
- Coût d'accès à ce service
 Totalement gratuit / Gratuit pour l'utilisateur / Payant pour l'utilisateur / Payé par une organisation (laboratoire/université/...)
- Précisions sur le coût d'accès à ce service
- Informations quantitatives sur ce service
- Utilisateurs réguliers de ce service
- Services similaires
- Langues de l'interface utilisateur du service Français / Anglais / Allemand / Espagnol / Russe /Autre
- Descriptifs sur ce portail en lien avec ce service
- Mots-clés en vrac
- Auteur de ce descriptif
- Date création / relecture complète de ce descriptif
- Mots-clés Gestionnaire/hébergeur,
 Fonctionnalité informatique

Portail math V1: architecture fonctionnelle (sans utilisateurs)

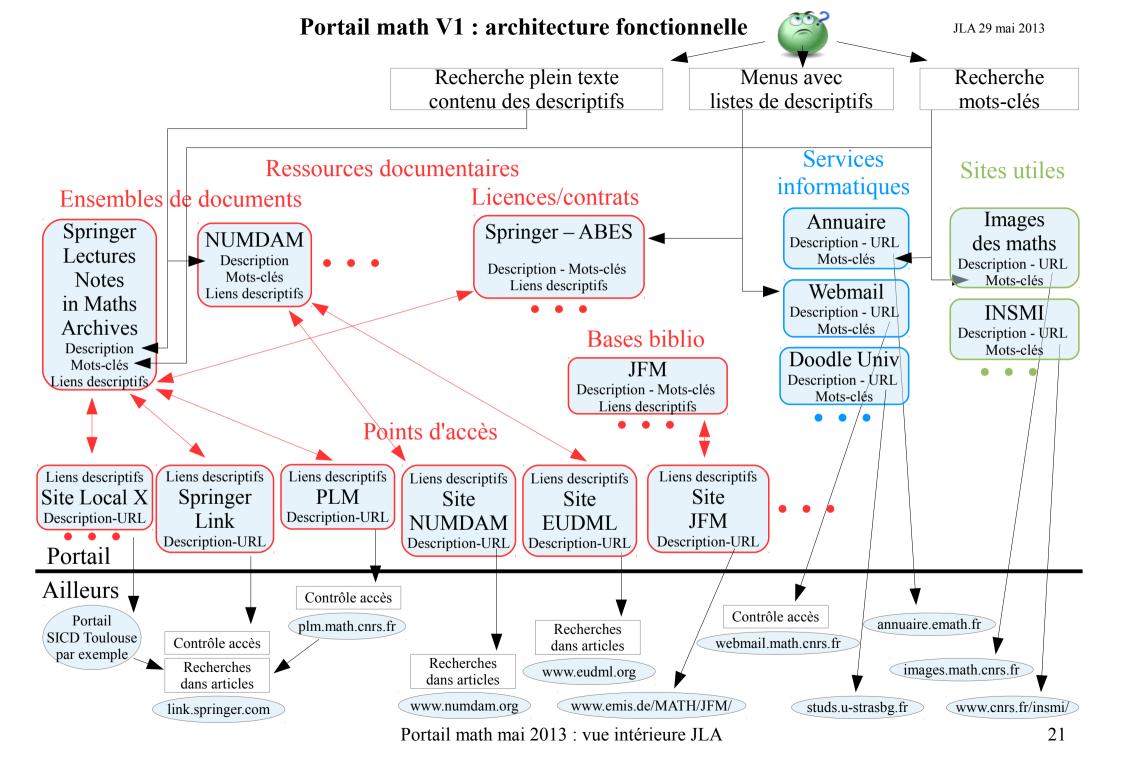




Contenu: recherches possibles



- Trois modes de recherche sont possibles
 - Recherche texte intégral sur le contenu du portail, c'est à dire sur les pages et sur les descriptions de ressources documentaires, services informatiques et sites utiles
 - Via une arborescence de menus avec des listes
 - Par mots-clés affectés aux descriptifs et une recherche par facette (progressive) sur ces mots-clé
- Dans la version V1, ces possibilités de recherches pour les utilisateurs portent uniquement sur les descriptifs présents sur le portail, pas sur le contenu des objets décrits comme les articles.





Contenu: workflow de publication



Les différents états des descriptifs dans le processus de rédaction-publication

- A faire : le descriptif est créé
- En rédaction ou En rédaction collaborative : un/des rédacteurs saisissent en ligne
- En relecture du contenu (c'est-à-dire scientifique et technique): vérification du contenu par un des responsables de rubrique concerné
- En normalisation technique : ajout de liens, mots-clés... par Jean-Luc Archimbaud et Nicolas Houillon
- En relecture syntaxique : Myriam Laurens et Nadège Arnaud
- En relecture par l'auteur
- Bon à publier (donné par l'auteur)
- Publié (par un des responsables éditoriaux de site)

Un message électronique est envoyé à chaque changement d'état pour avertir l'intervenant suivant



Contenu : descriptifs le 1/10



- Ressources documentaire8
 - Ensembles de documents : 47 descriptifs : 27 publiés, 21 en cours de création.
 - Bases de données bibliographiques: 7 descriptifs : 1 publié, 6 en cours de création.
 - Points d'accès: 44 descriptifs : 28 publiés, 16 en cours de création.
 - Licences/contrats: 25 descriptifs : 13 publiés, 12 en cours de création.
 - Documents utiles : 23 descriptifs : 20 publiés, 3 en cours de création.
- Services informatiques : 6 descriptifs : 5 publiés, 1 en cours de création.
- Sites utiles : 21 descriptifs : 13 publiés, 8 en cours de création.



Drupal: serveurs - rôles



- Rappel Mathrice : liste de diffusion et page Drupal
- Drupal 7 Ubuntu Apache MySQL
- 3 serveurs : prod, test et dév
 - Contenu mis à jour sur prod → test / dév
- Rôles
 - Administrateur (4)
 - Responsable éditorial du site (2)
 - Responsable éditorial de rubrique et/ou relecteur (8)
 - Concepteur portail (15)
 - Rédacteur (20)
 - Utilisateur authentifié (1 actuellement → 5000)



Drupal: types de contenu



- Page publique (9)
- Ressources documentaires :
 - ◆ Ensemble de documents (47 actuellement → plusieurs centaines ?)
 - Base de données bibliographiques (7)
 - Point d'accès (42)
 - Licence/contrat (25)
 - Document utile (23)
- Service informatique (6)
- Site utile (21)
- News (59)
 - Fil RSS à venir
- Document de travail (54)

et Fichiers attachés (49) et Vues (15 environ)



Drupal: conception - graphisme



- Conception progressive
- Thème Bartik (le thème par défaut ?) avec 2 colonnes de menus
 - A gauche : menus qui seront définitifs
 - A droite : menus pour travailler
 - Pas de couleur, pas de logo, graphisme minimal
 - Mais certains CSS modifiés pour présentation claire, ajout d'info...
 - Présentation des résultats de recherche plein texte spécifique
 - Janv 2013 → Sept 2013
- Appel d'offre --> société externe : logo graphisme thème
 - Livraison : thème (responsive)
- ◆ Actuellement : possibilité de basculer dynamiquement ancien ↔ nouveau thème pour chaque concepteur
 - Pour la phase de tests



Drupal : modules un peu spécifiques utilisés



- Menu attribute
 - Plus de souplesse dans l'affichage des menus
- Diff
 - Pour afficher les différences entre versions (on conserve les versions de tous les contenus)
- Lexicon
 - Lexique créé progressivement : certains mots définis
- Views
 - Pour créer des vues souvent sous forme de tableaux listes
- Apache Solr
 - Moteur d'indexation : Solr
- Apache Solr search attachments
 - Pour indexer fichiers attachés (PDF, texte...)



Drupal : modules un peu spécifiques utilisés



Workflow

- Tous les types de contenus ont un workflow
- Permissions différentes / rôle /état, envoi de mail à chaque chgt d'état

Rules

- Pour créer des événements (envoi de mails) quand chgt de workflow
- Node Reference
 - Champ d'un type de contenu X → (champ d')un type de contenu Y
- Pathologic
 - Problème de chemin relatif quand on copie sur site de test/dev
- Metatag
 - Configuration de meta balise HTML pour meilleure indexation google



Drupal: modules développés PHP (Nicolas Houillon Mathdoc)



- Customization portail
 - Objet : Diverses modifications qui ne justifient pas un module chacune. Pour l'instant, ajout de l'état de workflow au début de l'affichage d'un noeud et ajout du type d'affichage "Titre avec lien et URL" pour affichage d'infos de noeuds référencés.
 - Environ 30 lignes.
- Ajouts recherche portail
 - Objet : Ajoute le contenu du glossaire dans un noeud pour indexation
 - Environ 30 lignes.
- Fields portail:
 - Objet : Fourni des pages listant les champs de chaque type de contenu, rempli le contenu de noeuds de type document local avec la description des champs (pour indexation).
 - Environ 100 lignes.
- Tokens portail
 - Objet : Ajout de tokens pour nb de descriptifs de tel type publié...
 - Environ 200 lignes.
- User profile portail
 - Objet : Dans visualisation d'un compte utilisateur : ajout de champs (rôles et date de dernière connexion)
 - Environ 30 lignes.



Démonstration (petite)



- http://www.portail-math.fr
- Recherche
 - Conférence
 - PLM
 - Annales de l'ENS
 - Champs service informatique



La suite: V2 – début 2014



- Identifier et authentifier les utilisateurs de la communauté math
 - Référentiels à utiliser
 - Laboratoires et tutelles : annuaire des laboratoires géré par Mathdoc
 - Personnes : annuaire des personnes géré par Mathrice (synchronisé avec le précédent)
 - Authentification : fournisseurs d'identités des universités et JANUS pour CNRS (fédération d'identité RENATER) + fournisseur d'identités Mathrice : outils Shibboleth
 - Méthode 1 pour portail-math Drupal
 - Utiliser des modules Drupal standards : modules DRUPAL-LDAP et modules DRUPAL-Shibboleth
 - Non accepté par PLM Team de Mathrice → utilisation du service AuthPLM développé par Mathrice
 - Methode 2 pour portail-math Drupal : sera mise en place
 - Drupal interroge le service AuthPLM : qui authentifie l'utilisateur et donne comme informations son/ses labos, son/ses tutelles
 - Développement d'un module spécifique Drupal pour interface avec AuthPLM



La suite: V2 – début 2014



- Identifier et authentifier les utilisateurs de la communauté math (suite)
 - Sur le portail (Drupal) chaque utilisateur appartiendra à plusieurs groupes de type
 - Laboratoires (80-90 laboratoires de maths)
 - Universités (X universités tutelles des labos)
 - Organismes/structures (5-10 organismes/structures : NRBM, Couperin...)
 - Cette appartenance sera construite automatiquement (d'après les infos fournis par AuthPLM)
 - Affichage spécifique des données pour chaque utilisateur (progressivement, pas tout en V2). Cette méthode est à valider :
 - Un descriptif d'ens de documents est lié à un descriptif de licence. A ce descriptif de licence est affecté les mots-clés labos, univ, structure. Ces mots-clés correspondent à la liste des ayant droits pour cette licence.
 - Ainsi le portail (Drupal) pourra, suivant le profil de l'utilisateur authentifié, afficher :
 - Uniquement ce à quoi l'utilisateur peut accéder (si on opte pour cette solution)
 - Les meilleurs points d'accès pour son profil
 - ***** . . .
- Intégration des accès aux ressources documentaires actuellement accessibles par la PLM
 - ◆ Dans Drupal ou simplement par lien Drupal → l'interface PLM actuel ou PLM-V2 ou mixte ?



La suite: V3



- Remarque : c'est un projet modeste (ce n'est pas EuDML : 13,5 ETP pendant 3 ans)
- Le comité de pilotage et le comité exécutif décideront
- Faut il continuer avec Drupal et dans quelles proportions ?
 - Autres solutions : basculer sur un autre CMS : SPIP ?, tout développer ou développer une grande partie ?
 - Avis personnel : il faudrait continuer avec Drupal et ses modules en logiciels libres. Mais si d'autres veulent faire mieux...
- Idées personnelles en vrac
 - Généraliser les fonctionnalités V2 : remplacer (supprimer) les contrôles d'accès externes (schéma page 21)
 - Négocier avec les éditeurs de nouveaux contrats pour ce portail (et son @IP) : portail math = authentification de confiance pour des éditeurs (ce qu'est l'INIST pour certains éditeurs)
 - > Portail pour plate-forme ISTEX : assure l'authentification et le portail utilisateur
 - Intégrer des services ou bases d'information existants dans le portail (chez Mathdoc, Mathrice-PLM...) ou créer de nouveaux services/bases
 - Recherches sur titres des revues → recherche sur contenus des articles (remplacer les rectangles externes 'Recherche dans articles' du schéma page 20)
 - Recherche sur métadonnées
 - Recherche sur contenu
 - → ---> aspiration de contenu ou API vers recherches sur les points d'accès ou...