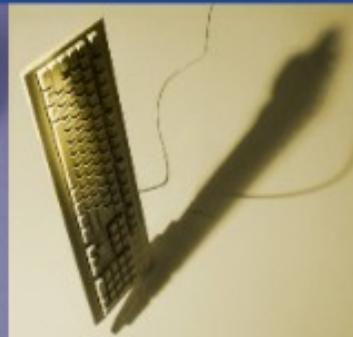




# ORI-OAI

Outil de Référencement et d'Indexation  
Les thèses électroniques

Colloque international « l'Université à l'ère du numérique »  
Bordeaux, 10 décembre 2008





- Le contexte
- TEF
- STAR
- ORI-OAI
- ORI-OAI : thèse
- Calendrier
- Retour d'expérience



# Le contexte

## Le contexte juridique

- **Circulaire du Bulletin officiel du Ministère (2000-149 du 21/9/2000) >>**  
*« Les thèses soutenues dans les universités et les autres établissements d'enseignement supérieur constituent des **documents de première valeur** dont il convient, tant dans l'intérêt des jeunes docteurs et des établissements que pour une **meilleure visibilité internationale de la recherche française, d'assurer la promotion.** »*
- **Circulaire du 29 mars 2005 sur le Dépôt, signalement, diffusion et archivage des thèses sous forme électronique >>**  
(Cette circulaire a été abrogée par l'arrêté du 7 août 2006 relatif au dépôt)
- **Arrêté du 7 août 2006 relatif aux modalités de dépôt, de signalement, de reproduction, de diffusion et de conservation des thèses ou des travaux présentés en soutenance en vue du doctorat >>**

### Objectifs :

Valoriser la recherche

Favoriser l'insertion des jeunes docteurs

Diffuser, recenser, signaler et conserver la production scientifique

## Le contexte juridique

- **Arrêté du 7 août 2006**

L'objectif du circuit des thèses est d'assurer **quatre fonctions essentielles** :

- Dépôt
- Signalement
- Diffusion
- Archivage

### **Choix du mode de dépôt : dépôt papier OU dépôt électronique**

→ **Deux circuits :**

- **Thèses papier**

Le circuit traditionnel reste en vigueur pour les établissements n'ayant pas encore adopté le dispositif de diffusion électronique des thèses.

- **Thèses électroniques**

Dépôt dans l'établissement de la thèse sous forme numérique, accompagnée d'un bordereau électronique.

Mise en œuvre d'un **outil national** d'enregistrement du dépôt de la thèse et de ses métadonnées, en vue de son signalement et de son archivage : **STAR**

## Le passage au tout électronique : impacts et enjeux

### Enjeux :

- Améliorer les services en termes : d'exhaustivité, d'accessibilité, d'authentification, de pérennisation et de rapidité de traitement.
- **Mise en oeuvre d'un dispositif** assurant des **fonctions complexes** pour des **acteurs multiples dans les établissements** et au **niveau national** au **moyen d'outils ergonomiques et d'usage simple**.

### Impacts : bouleversement dans les habitudes de travail des universités

- Définition d'une politique de diffusion de la production scientifique dans les établissements
- Mise en place de nouvelles pratiques (circuit administratif, mode de production de la thèse, modalités de dépôt, etc.)
- Mise en oeuvre de nouveaux outils :
  - outil de gestion local,
  - plateforme de diffusion,
  - système d'archivage pérenne,
  - jeu de métadonnées.



**TEF**

## Un jeu de métadonnées spécifique

La recommandation TEF (Thèses Électroniques Françaises) - groupe de travail AFNOR CG46/CN357/GE5 – 2006

« *La recommandation TEF (Thèses Électroniques Françaises) définit un jeu de métadonnées pour les thèses électroniques soutenues en France. Son objectif est **d'organiser de manière cohérente des métadonnées de thèse riches et normalisées, afin de faciliter leur échange et leur diffusion, au niveau national ou international.** Ces métadonnées recouvrent des **métadonnées descriptives (bibliographiques)** et des **métadonnées de gestion (administration, droits, conservation).** »*

**TEF 2.0**

<http://www.abes.fr/abes/documents/tef/recommandation/index.html>

## Un jeu de métadonnées spécifique

- **une modélisation : une articulation modulaire des 4 types de métadonnées des thèses**
  - **métadonnées descriptives (DC)** : description bibliographique
  - **métadonnées administratives (DC étendu)** : suivi administratif de la thèse
  - **métadonnées de droit (METS Rights)** : respect des conditions juridiques associées au document (droits de propriété intellectuelle notamment)
  - **métadonnées de conservation** (définies avec le CINES) : relatives à l'archivage pérenne
  
- **un format XML** : format d'échange et outil de validation.  
Ce format XML s'appuie sur le schéma METS (MetadataEncoding and Transmission Standard) qui permet d'articuler les différents types de métadonnées.



# **STAR**

**(Signalement des thèses électroniques,  
archivage et recherche)**

## **STAR (Signalement des thèses électroniques, archivage et recherche)**

Application nationale développée par l'Abes pour le dépôt légal des thèses électroniques en France :

- 12-10-06 : lancement de STAR
- 02-07 : formation des premiers établissements utilisateurs
- 10-07 : formation 2e vague
- 01-08 : retours d'expérience
- 09-08 : formation 3e vague

*« [...] **point central dans le nouveau circuit national des thèses.** STAR sert d'interface entre les établissements de soutenance et les autres acteurs du paysage français intéressés par la valorisation des thèses. STAR est le guichet unique par lequel l'établissement de soutenance peut donner à ses thèses de nouveaux débouchés, qu'il s'agisse de débouchés incontournables (archivage, signalement) ou optionnels (serveurs de diffusion extérieurs à l'établissement, indexation du texte intégral). »*

## Les fonctions de STAR

- **Enregistrement du dépôt** de la thèse, version d'archivage et version de diffusion ainsi que de ses **métadonnées**
- **Signalement automatique dans le Sudoc.**
- A la demande de l'établissement, **envoi des métadonnées et/ou de l'édition de diffusion de la thèse** vers les sites de diffusion désignés par celui-ci.
- **Envoi de l'édition d'archivage** -y compris dans le cas d'une thèse non diffusable- au CINES.
- Attribution d'un identifiant pérenne

## STAR est un intermédiaire

### En entrée :

- Saisie ou import des données administratives, juridiques et bibliographiques
- Dépôt de l'édition d'archivage (XML ou PDF)
- Dépôt de l'édition de diffusion de la thèse (HTML ou PDF)

### En sortie :

- Signalement automatique dans le catalogue Sudoc
- Dépôt au CINES à des fins de conservation
- Création d'une URL pérenne
- Entrepôts de diffusion local et éventuellement national (TEL, PASTEL...)
- Serveur OAI-PMH de l'ABES

### STAR n'est pas :

- Un catalogue Sudoc des thèses
- Une chaîne de transformation
- Un outil de diffusion



# ORI-OAI



- **Gérer et publier les ressources numériques de l'établissement**
  - Référentiel unique de ressources connecté au SI
  - Production de métadonnées et diffusion des ressources produites par l'établissement
- **Valoriser la production de l'établissement par une indexation de qualité**
  - Indexation avec des normes et standards
  - Utilisation de classifications paramétrables articulés sur des standards comme Dewey



- **Partager les ressources numériques avec d'autres établissements**
  - Exposition des fiches de métadonnées de l'établissement
  - Agrégation de fiches de métadonnées
- **Accéder aux ressources numériques à distance en fonction des droits d'accès**
  - Système de recherche avancée et thématique
  - Définition des droits



- **ORI-OAI** : **O**util de **R**éférencement et d'**I**ndexation
    - mettre en réseau des ressources numériques
  - **+ Protocole OAI/PMH**
    - échanger des données
  - **+ .ORG**
    - développé en logiciel libre
    - pour les UNT, UNR, les établissements universitaires, ...
- = ORI-OAI.ORG**

## **Les acteurs du projet**

- 1 Équipe de pilotage
- 1 Équipe fonctionnelle
- 1 Équipe technique



# **ORI-OAI : groupe fonctionnel thèse**



## **ORI-OAI appliqué aux thèses électroniques**

- **Outil de valorisation** des thèses : diffusion web et OAI
- **Outil de gestion** du cycle de vie du traitement des thèses
- **Outil de dépôt** vers l'application STAR et HAL



# ORI-OAI thèse : groupe de travail fonctionnel

- **Un groupe de travail fonctionnel** composé de différents gestionnaires de thèses d'établissements français :
  - INP Toulouse, INSA Lyon, Université de Clermont-Ferrand, Université de Rennes 1 et 2, INPL, ...
- **Objectif : rédiger un cahier des charges ORI**
  - en respectant la législation française
  - en profitant de l'expérience des établissements



## Le groupe de travail a défini :

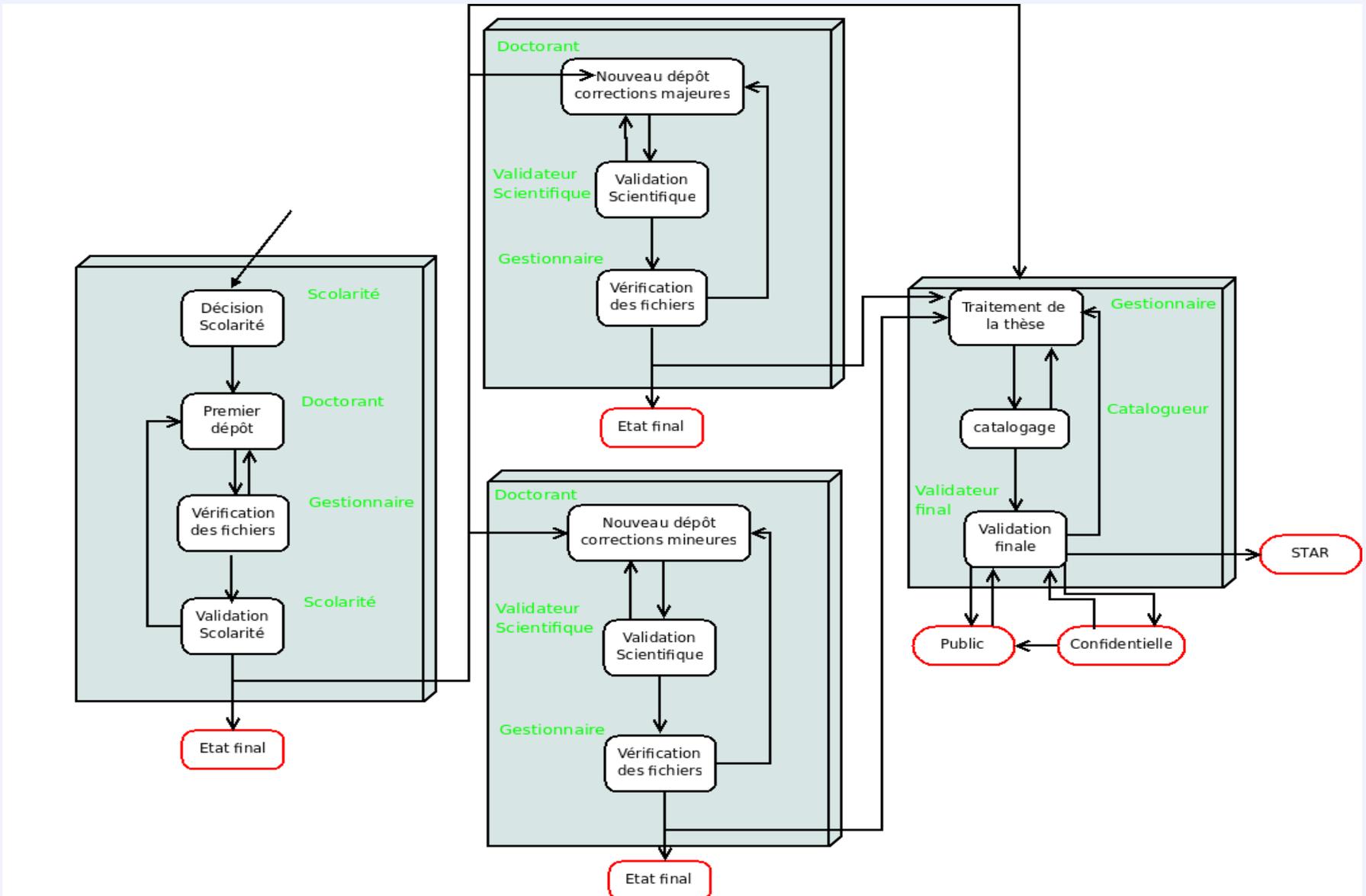
- les différents **acteurs**,
- les différentes **étapes** du workflow,
- les **échanges** entre acteurs,
- **l'interface** des éditeurs de métadonnées,
- les **délais** entre les étapes,
- les **services**,
- ...



## Les acteurs du workflow

- Doctorant
- Service scolarité (SEVU)
- Service Commun de Documentation (SCD)
- Valideur scientifique
- Valideur final

## Workflow des thèses électroniques





## Les services associés au workflow

- **Vue éditeur de métadonnées différentes** suivant les acteurs et en fonction des étapes du workflow
- Autocomplétion (éditeur)
- Profils avec des **droits différents**
- Trace et **historisation** des actions
- Accès aux **webservices** de l'ABES (autorités) et au SI (LDAP)



# ORI-OAI thèse : droits et versions

- **Les droits**

*en fonction des droits déclarés dans la TEF*

- accès intranet ou internet
- embargo
- suivant le cas : diffusion métadonnées et/ou texte intégral

- **Les versions**

*s'appuie sur le modèle FRBR de la TEF*

- plusieurs versions d'une même thèse pourront être gérées et diffusées



## **Le suivi du cycle de vie du traitement des thèses**

- données complémentaires confidentielles
- données de gestion (temps de traitement, outils utilisés, etc)
- échanges entre les acteurs (mail, fil RSS)



- **Diffusion**

- plateforme **locale** de l'établissement
- diffusion **OAI-PMH**
- **transfert** éventuel vers HAL

- **Services**

- fil **RSS**
- **import/export** de métadonnées

- **Archivage**

- transfert et validation des thèses (métadonnées et fichiers) vers **STAR**
- attribution à venir d'un **URL pérenne**



- Version 1.4.0
  - Workflow
    - Protocole de communication avec d'autres plates-formes
    - LOMFR / SupLOMFR
  - Moteur d'indexation
    - Indexation en texte intégral
  - Moteur de recherche
    - Nouveaux modes de recherche
  - Gestionnaire de vocabulaires
    - Gérer des vocabulaires distants et centralisés
  - Évolutions générales du projet
    - Compatible avec outils de statistiques



- Prochaines versions
  - 1.5.0 Intégration du stockage à la version 1.4.0
  - Référencement et publication dans l'archive de l'établissement
    - des thèses avec le format TEF
    - des publications de la recherche
  - ORI-OAI / STAR
    - Export de métadonnées TEF vers l'archive nationale STAR
  - ORI-OAI / HAL
    - Échange des métadonnées et des ressources de l'archive de l'établissement avec l'archive nationale HAL
- CEPIA
  - Référencement dans ORI-OAI depuis une plate-forme LMS et un outil de chaîne éditoriale



- Site du projet
  - <http://www.ori-oai.org>

- Contacts

Liste de discussion **ori-oai-utilisateurs**

<http://listes.univ-rennes1.fr/wws/info/ori-oai-utilisateurs>

Liste d'information **ori-oai-info**

<http://listes.univ-rennes1.fr/wws/info/ori-oai-info>



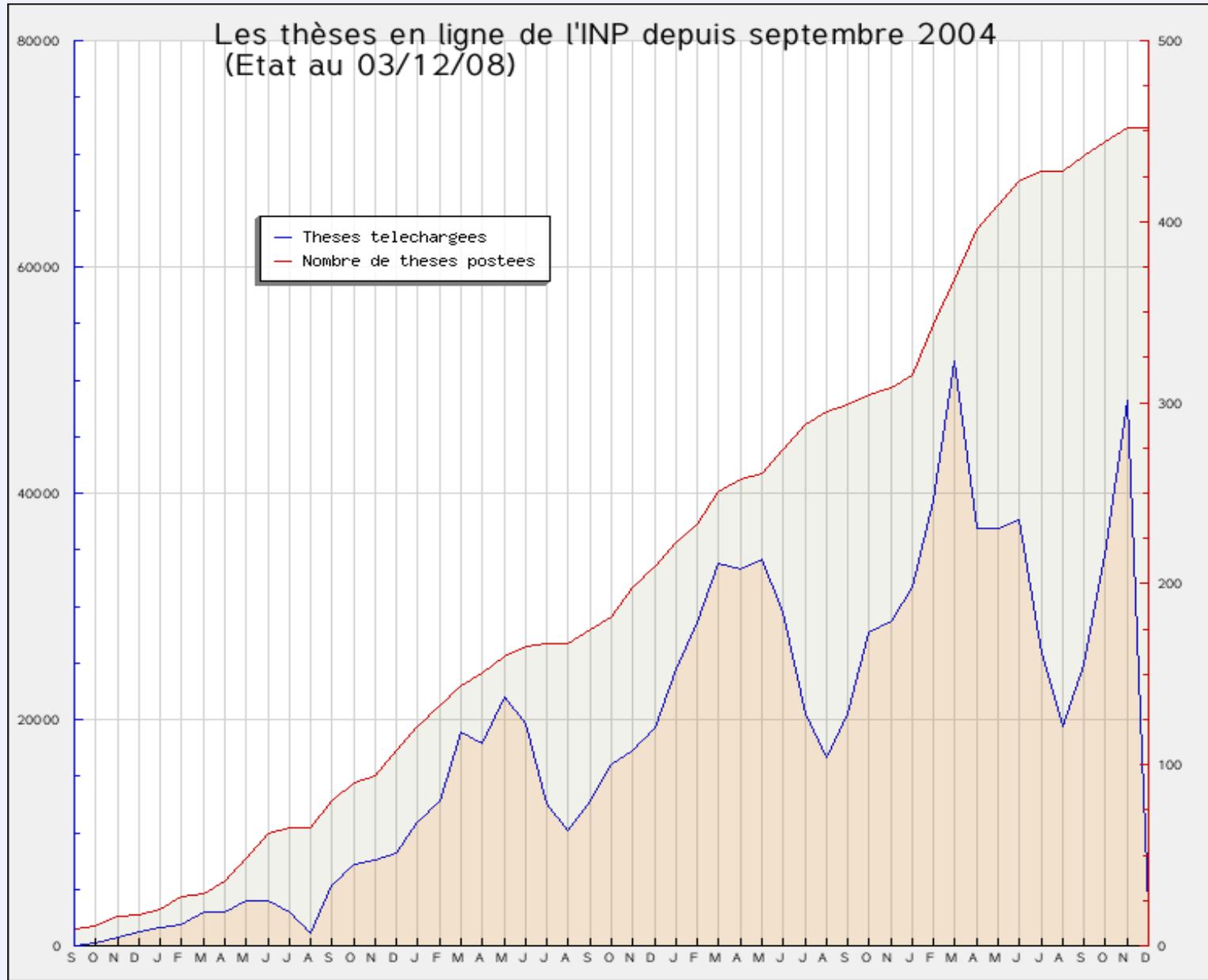
# Retour d'expérience



## INP Toulouse

- En chiffres :
  - 120 thèses/an en moyenne
  - 450 thèses diffusées depuis septembre 2005
- Plateforme utilisée actuellement : Eprints
- Passage au tout électronique :
  - Vote au CA de l'établissement : juin 2008
  - Début : janvier 2009
- Migration des données vers ORI programmée courant 2009

# ORI-OAI thèse : expérience des établissements



Ce travail est mis à disposition sous une licence Creative Commons

### **Vous êtes libres**

De reproduire, distribuer et communiquer cette création au public

De modifier cette création



Cette création est mise à disposition selon le Contrat Paternité-NonCommercial-ShareAlike 2.5 disponible en ligne <http://creativecommons.org/licenses/by-nc-sa/2.5/>

Emilie Romand-Monnier (BCIU - Clermont-Ferrand)  
Jean-Marie Le Behec (INP Toulouse)  
Dalila Boudia (INSA Lyon)  
Sébastien Erhel (Université de Rennes 1)  
Magali Jacquot (INPL)  
Mathilde Thiriet (INPL)  
Laurence Leroux (Université de Rennes 2)  
Andry Rahobisoa (SICD Toulouse)

Rosa-Maria Gomez (INSA de Lyon)  
Yohan Colmant (Université de Valenciennes)  
Raymond Bourges (Université de Rennes 1)  
Monique Joly (INSA de Lyon)  
Alain Mayeur (Université de Valenciennes)  
François Jannin (ENSEEIHT – INP de Toulouse)  
Nolwen Huet (INSA de Lyon)