

Autour de la valorisation de logiciels d'un laboratoire de recherche

Teresa Gomez-Diaz

Laboratoire d'informatique de l'institut Gaspard-Monge
CNRS-Université Paris-Est

Journées MATHRICE

IHP - Paris, 9-11 octobre 2007



Plan

1 Présentation

- Buts
- Mission

2 Cadre de travail

- IGM-LabInfo
- Définition de logiciel
- Les ressources linguistiques

3 Élaboration du plan de logiciels

- Étapes dans l'élaboration
- Contacter les tutelles
- Questionnaires direction
- Enquête développeurs

4 Conclusions

- Conclusions enquête développeurs
- Conclusion 1 : référence logiciel
- Conclusion 2 : Rôle d'un correspondant de logiciels
- Conclusions - Fin

5 Si on a le temps

Plan

1 Présentation

- Buts
- Mission

2 Cadre de travail

- IGM-LabInfo
- Définition de logiciel
- Les ressources linguistiques

3 Élaboration du plan de logiciels

- Étapes dans l'élaboration
- Contacter les tutelles
- Questionnaires direction
- Enquête développeurs

4 Conclusions

- Conclusions enquête développeurs
- Conclusion 1 : référence logiciel
- Conclusion 2 : Rôle d'un correspondant de logiciels
- Conclusions - Fin

5 Si on a le temps

Pourquoi je suis venue vous voir ?

- Diffusion (1 an de travail)
- Intérêt ?
- Réflexion
- Collaborations
 - ▶ pour faire évoluer

Mission

Mission

Le directeur de mon unité (UMR 8049 IGM-Labinfo) m'a demandé d'*étudier la mise en place de services* (suivi de versions, publication, etc.) associés au développement logiciel au sein du laboratoire, avec pour **objectif** de favoriser la visibilité de ces développements.

Mission

Mission

Le directeur de mon unité (UMR 8049 IGM-Labinfo) m'a demandé d'*étudier la mise en place de services* (suivi de versions, publication, etc.) associés au développement logiciel au sein du laboratoire, avec pour **objectif** de favoriser la visibilité de ces développements.

Propositions

- Étude == Plan de logiciels(*)
- Mon rôle == Correspondante de logiciels(*)

(*) Terminologie volée au réseaux de correspondants de formation du CNRS.

Plan

1 Présentation

- Buts
- Mission

2 Cadre de travail

- IGM-LabInfo
- Définition de logiciel
- Les ressources linguistiques

3 Élaboration du plan de logiciels

- Étapes dans l'élaboration
- Contacter les tutelles
- Questionnaires direction
- Enquête développeurs

4 Conclusions

- Conclusions enquête développeurs
- Conclusion 1 : référence logiciel
- Conclusion 2 : Rôle d'un correspondant de logiciels
- Conclusions - Fin

5 Si on a le temps

Le laboratoire

C'est un laboratoire d'informatique avec 5 équipes de recherche :

- Algorithmique
 - ▶ Combinatoire des mots, automates
 - ▶ Programmation générique
 - ▶ Bio-informatique
 - ▶ Réseaux
- Combinatoire algébrique et calcul symbolique
- Informatique linguistique
- Géométrie discrète et imagerie
- Signal et communications

Le laboratoire

C'est un laboratoire d'informatique avec 5 équipes de recherche :

- Algorithmique
 - ▶ Combinatoire des mots, automates
 - ▶ Programmation générique
 - ▶ Bio-informatique
 - ▶ Réseaux
- Combinatoire algébrique et calcul symbolique
- Informatique linguistique
- Géométrie discrète et imagerie
- Signal et communications

Forte composante “mathématiques” :

- au niveau thématique
- dans l'utilisation de systèmes de calcul formel
Maple, Matlab, MuPAD

Mais qu'est-ce que c'est un logiciel ?

Définition ?

J'entends par **logiciel** du laboratoire tout programme

- utile (**au sens large**) pour faire avancer la recherche
- au moins un membre du laboratoire a participé à son développement.

Mais cela dépend un peu du cadre !

Mais qu'est-ce que c'est un logiciel ?

Définition ?

J'entends par **logiciel** du laboratoire tout programme

- utile (**au sens large**) pour faire avancer la recherche
- au moins un membre du laboratoire a participé à son développement.

Mais cela dépend un peu du cadre !

- je me place dans un laboratoire de recherche
 - ▶ des **publications associées** (si possible)
- dans l'équipe de linguistique informatique on développe
 - ▶ des **ressources linguistiques** :
tables de lexique-grammaire, dictionnaires électroniques, grammaires locales

Mais qu'est-ce que c'est un logiciel ? (2)

Les ressources linguistiques ont des **points communs** avec les logiciels :

- ce ne sont pas de publications,
- il y a des références associées (articles, thèses, livres, ...)
- il y a des contrats associés
- elles font partie active de la recherche
- elles ont des licences associées

Elles sont aussi **accompagnées par des logiciels** développés dans l'équipe pour leur manipulation et leur visualisation.

Mais qu'est-ce que c'est un logiciel ?

Définition ?

J'entends par **logiciel** du laboratoire tout programme

- utile (**au sens large**) pour faire avancer la recherche
- au moins un membre du laboratoire a participé à son développement.

Mais cela dépend un peu du cadre !

- Dans mon cadre, **logiciel** contient aussi les ressources linguistiques
- **valorisation** de la production scientifique **autre** que les publications

Plan

1 Présentation

- Buts
- Mission

2 Cadre de travail

- IGM-LabInfo
- Définition de logiciel
- Les ressources linguistiques

3 Élaboration du plan de logiciels

- Étapes dans l'élaboration
- Contacter les tutelles
- Questionnaires direction
- Enquête développeurs

4 Conclusions

- Conclusions enquête développeurs
- Conclusion 1 : référence logiciel
- Conclusion 2 : Rôle d'un correspondant de logiciels
- Conclusions - Fin

5 Si on a le temps

Élaboration du plan de logiciels

- élaboration questionnaires pour chefs/direction
- élaboration enquête pour développeurs
- contacter les tutelles + voir INRIA
- premier recensement des logiciels du laboratoire (octobre 2006)
- lancement enquête (9 mars-2 oct.), entretiens chefs (mars-mai 2007)
- enquêtes Mathrice et Calcul
- rédaction plan de logiciels (mai-septembre 2007)

Élaboration du plan de logiciels

- élaboration questionnaires pour chefs/direction
- élaboration enquête pour développeurs
- contacter les tutelles + voir INRIA
- premier recensement des logiciels du laboratoire (octobre 2006)
- lancement enquête (9 mars-2 oct.), entretiens chefs (mars-mai 2007)
- enquêtes Mathrice et Calcul
- rédaction plan de logiciels (mai-septembre 2007)

Avec les objectifs :

- bien saisir la situation actuelle au niveau
 - ▶ du laboratoire
 - ▶ des tutelles
 - ▶ ailleurs
- savoir où les gens veulent aller (si possible)

Les tutelles et autres

- SAIC UMLV (activités industrielles)

questionnaires, rôle d'un correspondant de logiciels

- ▶ licences et droits d'auteur
- ▶ préparation et gestion de contrats, de projets, ...

- CRI UMLV

questionnaires, services à mettre en place, ∃ cluster

- ▶ vision trop "enseignement" encore, peu "recherche"
- ▶ mais manque de personnel !

Les tutelles et autres

- CNRS : CALCUL (petite enquête)
- CNRS : Mathrice (petite enquête, licences Maple)
- CNRS : CIEL (1 logiciel IGM déposé en 2005), 23 up today
- CNRS : PLUME (2 fiches IGM dont 1 pas de relecture)
 - ▶ 3 log. “scientifiques” : FreeFem++, MuPAD-Combinat, Populations

Les tutelles et autres

- CNRS : CALCUL (petite enquête)
- CNRS : Mathrice (petite enquête, licences Maple)
- CNRS : CIEL (1 logiciel IGM déposé en 2005), 23 up today
- CNRS : PLUME (2 fiches IGM dont 1 pas de relecture)
 - ▶ 3 log. “scientifiques” : FreeFem++, MuPAD-Combinat, Populations
- listes : calcul, suplibre
- INRIA : inria.fr/valorisation/logiciels/, gforge.inria.fr
- Pôle System@tic, groupe thématique Ouverture : 2 logiciels IGM
- Sourceforge

Des projets (PLUME, INRIA, IGM, ...) ont démarré vers mi-2006.

Questionnaires chefs d'équipe - direction du laboratoire

Voici les points les plus importants :

- Évaluation situation actuelle
- Utilité : avancer la recherche, contrats, ...
- Direction : Où va-t-on ?
- Besoins : humains, matériels, logiciels, court/long terme
- Que demande-t-on aux développeurs ? + Critères de qualité
- Tutelles, offre, ce qu'on nous demande, ce qu'on doit donner
 - ▶ CIEL
 - ▶ PLUME
 - ▶ SAIC - droits d'auteur
 - ▶ Mathrice - Maple
- Classification logiciels par thème

Enquête développeurs

- Renseignements personnels (nom, directeur de thèse, ...)
- Renseignements logiciel (nom, dates, objectifs, \exists doc, \exists page web)
- Renseignements techniques (licence, diffusion, exemples, CIEL, PLUME)
- Renseignements pour l'évaluation du laboratoire
(contrats ou subventions, collaborations, publications, utilisateurs)
- Services au développement attendus au laboratoire
(cvs/subversion, accès externe/restreint, gestion bugs, conseil licences, moyens calcul,...)
- Formation
- Problèmes rencontrés, remarques

Plan

1 Présentation

- Buts
- Mission

2 Cadre de travail

- IGM-LabInfo
- Définition de logiciel
- Les ressources linguistiques

3 Élaboration du plan de logiciels

- Étapes dans l'élaboration
- Contacter les tutelles
- Questionnaires direction
- Enquête développeurs

4 Conclusions

- Conclusions enquête développeurs
- Conclusion 1 : référence logiciel
- Conclusion 2 : Rôle d'un correspondant de logiciels
- Conclusions - Fin

5 Si on a le temps

Conclusions enquête développeurs

● Remarques

- ▶ Seul champ obligatoire : le nom
- ▶ Trop longue car complète
- ▶ Donner la possibilité d'exprimer les besoins, les problèmes, ... (doctorants)
- ▶ L'enquête n'était pas "bien adaptée" aux ressources linguistiques
- ▶ **Ce n'est pas une enquête** dont on a besoin (dans le sens réponse volontaire)

● Résultats

- ▶ Total questionnaires : 42 (dont 2 ress. ling.)
- ▶ Total logiciels au labo : **43** (sans compter ress. ling.)
 - ★ dont 5 sans questionnaire
 - ★ sans licence (connue) : au moins 22, 17 avec LGPL
 - ★ avec référence : au moins 29
 - ★ sans page web : 5

Référence logiciel

Définition ?

Une *référence logiciel* **doit être définie et utilisée** systématiquement, avec

- nom du logiciel, avec version (par exemple nom-1.0.0)
- auteur(s) ou le(s) responsables du projet
avec mention explicite des membres du laboratoire impliqués
- date (de la dernière version)
- licence

Optionnel :

- état (en développement, en maintenance, non maintenu)
- site web du logiciel

Étudier le cas des ressources linguistiques ...

Rôle d'un correspondant de logiciels

- Élaboration du plan, recueil et analyse des besoins, suivi du plan
- Point de contact pour des tutelles : SAIC, CRI, ...
- Suivi des critères de qualité : licence, doc, page web, ...
- Gestion des listes de logiciels du laboratoire
 - ▶ interne, avec recensement complet
 - ▶ externe, pour favoriser la visibilité
- Communication : élaboration rapport, pages web, plaquettes, ...
- Formation : être à l'écoute des besoins
- Valorisation : rédaction projets concernant les logiciels
- Suivi des démos des logiciels (pérennisation du code)

Conclusions

Qu'est-ce qu'on fait avec des logiciels de recherche ?

- au niveau chercheur :
 - ▶ publications
 - ▶ pages web, en général personnelles
- au niveau laboratoire :
 - ▶ ne sont pas “systématiquement” mentionnés aux rapports du labo
 - ▶ parfois sont perdus/confondus/ignorés dans la relation des résultats scientifiques
 - ▶ les logiciels des thésards “sont oubliés” après leur départ
 - ▶ peu (ou pas) de page web logiciels
 - ▶ contrats/projets : les logiciels sont +/- mentionnés
- au niveau tutelles :
 - ▶ leur diffusion encourage “la recherche reproductible”
 - ▶ leur diffusion fait augmenter le nb de citations
 - ▶ CIEL, PLUME, INRIA

Plan

1 Présentation

- Buts
- Mission

2 Cadre de travail

- IGM-LabInfo
- Définition de logiciel
- Les ressources linguistiques

3 Élaboration du plan de logiciels

- Étapes dans l'élaboration
- Contacter les tutelles
- Questionnaires direction
- Enquête développeurs

4 Conclusions

- Conclusions enquête développeurs
- Conclusion 1 : référence logiciel
- Conclusion 2 : Rôle d'un correspondant de logiciels
- Conclusions - Fin

5 Si on a le temps

Droit d'auteur

- auteur == développeur(s) du code
- travail salarié, l'employeur est titulaire de ces droits (sauf accord) et l'auteur conserve ses droits moraux
- protection automatique (sans rien faire !)
- pendant **70 ans** (après la mort de l'auteur)
- condition : originalité
- ce qui n'est pas protégé :
 - les fonctionnalités, les algorithmes, les interfaces, les lang. de programmation
- il est illicite de :
 - exécuter, copier (excep. sauvegarde), modifier, distribuer

Droit d'auteur et licences

- auteur == développeur(s) du code
- travail salarié, l'employeur est titulaire de ces droits (sauf accord) et l'auteur conserve ses droits moraux
- protection automatique (sans rien faire !)
- pendant **70 ans** (après la mort de l'auteur)
- condition : originalité
- ce qui n'est pas protégé :
 - les fonctionnalités, les algorithmes, les interfaces, les lang. de programmation
- il est illicite de :
 - exécuter, copier (excep. sauvegarde), modifier, distribuer

La **licence sert** à mieux protéger :

- les auteurs
- les utilisateurs
- les éventuels collaborateurs au développement

Droit d'auteur et licences

- auteur == développeur(s) du code
- travail salarié, l'employeur est titulaire de ces droits (sauf accord) et l'auteur conserve ses droits moraux
- protection automatique (sans rien faire !)
- pendant **70 ans** (après la mort de l'auteur)
- condition : originalité
- ce qui n'est pas protégé :
 - les fonctionnalités, les algorithmes, les interfaces, les lang. de programmation
- il est illicite de :
 - exécuter, copier (excep. sauvegarde), modifier, distribuer

Il faut bien établir : **date, utilité et licence.**

Enquête Calcul et Mathrice (1)

- ① Y-a-t-il des développeurs de logiciels pour la recherche dans vos laboratoires (permanents ou thésards) ?
 - ▶ 10 rép., 9 oui, 1 non
- ② Est-ce que votre laboratoire/institution fait quelque chose avec les logiciels ?
 - ▶ 10 rép., 7 oui (recherche (4), distribution (2), projets (2), enseign. (1)), 2 non, 1 il n'y a pas
- ③ Est-ce que les logiciels sont mentionnés dans le rapport d'évaluation (tous ?, comment ?) ?
 - ▶ 10 rép., 6 oui, 2 non (dont 1 dit projets), 1 pas tous, 1 pas de rapport
- ④ Est-ce qu'il y a une page web pour les logiciels développés par équipe/laboratoire/institution ? Lesquels ?
 - ▶ 10 rép., 4 oui (labo), 1 oui (équipe), 1 non (page perso.), 2 non, 1 il n'y a pas

Enquête Calcul et Mathrice (2)

- 1 Est-ce que les logiciels développés par les thésards sont mentionnés quelque part ?
 - ▶ 8 rép., 2 oui, 2 thèse, 2 non, 1 je ne sais pas, 1 il n'y a pas
- 2 Existe-t-il des collaborations en matière de développement avec d'autres institutions ?
 - ▶ 10 rép., 8 oui, 2 non
- 3 Est-ce qu'il y a une personne et/ou service/entité qui "est en charge" (cela peut être un simple point de contact) de la diffusion/promotion ou valorisation des logiciels de recherche produits au laboratoire ?
 - ▶ 10 rép., 6 non, 1 oui, 3 indécis (un peu, en cours d'étude, 1 groupe de travail)
- 4 Est-ce que des chercheurs de votre institution ont déposé un de ses logiciels dans CIEL ?
 - ▶ 10 rép., 6 non (dont 1 APP), 2 oui, 2 incertain (1 pas moi, 1 je ne sais pas)

Enquête Calcul et Mathrice (3)

Seulement 10 réponses (6 Calcul + 4 Mathrice) mais 143 + 63 consultations de la page web !

<http://igm.univ-mlv.fr/~teresa/Calcul-et-Mathrice/>