

Projet de référencement et de valorisation des développements internes

Réunion du 23 octobre 2007

Ce projet a pour objectif de répondre à un besoin croissant de valorisation de la production scientifique que constituent les logiciels développés dans les laboratoires de recherche.

En effet, de nombreux développements de logiciels sont effectués dans les laboratoires de recherche comme support à la recherche ou aboutissement d'une recherche (pilotage de manipulation, modélisation, simulation, calcul, ...) ou comme outil générique informatique (gestion du laboratoire, administration système, ...). Ils sont développés par des chercheurs ou des ingénieurs, parfois dans le cadre de projets nationaux ou internationaux.

Ces développements sont souvent innovants et peuvent intéresser d'autres entités que le laboratoire. Il est donc nécessaire d'encourager la diffusion et l'échange de ces logiciels dans leur communauté.

Ce problème n'est pas nouveau et quelques initiatives ont déjà été menées pour répondre à ce besoin.

Ce projet est issu d'une réflexion menée au sein d'un groupe de travail constitué de :

- Violaine Louvet¹ (ICJ - CIEL - Mathrice et Groupe Calcul),
- Térésa Gomez-Diaz² (IGM-LabInfo), dont l'étude réalisée au sein de son laboratoire a donné lieu à la rédaction du document [5],
- Geneviève Romier³ (CNRS/UREC - PLUME),
- Stéphane Cordier⁴ (MAPMO, CIEL, département MPPU),
- Jean-Luc Archimbaud⁵ (CNRS/UREC - PLUME).

1 Expériences en valorisation de logiciels

Plusieurs expériences ont été initiées sur la problématique de la valorisation des développements internes. Ces expériences n'ont en général pas abouti aux résultats escomptés.

¹louvet@math.univ-lyon1.fr

²teresa.gomez-diaz@univ-mlv.fr

³genevieve.romier@urec.cnrs.fr

⁴stephane.cordier@math.cnrs.fr

⁵Jean-Luc.Archimbaud@urec.cnrs.fr

Codiciel Codiciel était une Unité Propre de Service du CNRS au service des communautés mécanique et physique dont les objectifs étaient de recenser et mettre à disposition des outils et plate-formes logicielles de modélisation et de simulation numérique dans ces domaines. Elle avait aussi des missions de formation et de documentation relativement aux outils proposés. Cette UPS a été supprimée en 2005.

CIEL Le projet CIEL (Code Informatique En Ligne) a été initié en 2004 avec le soutien du CCSD. Les motivations ayant conduit à ce serveur sont triples :

- *Promouvoir et valoriser les codes de calcul* c'est-à-dire mieux faire connaître les codes de recherche développés dans les laboratoires de recherche et permettre une reconnaissance aux développeurs de ces codes de la même façon qu'un article dans une revue avec comité de lecture.
- *Pérenniser les codes de calcul* pour parer au problème de la perte de savoir-faire due au départ d'un thésard ou d'un chercheur. C'est également l'un des moyens pour faire connaître l'existence de ce patrimoine scientifique dans notre communauté mais aussi dans le milieu industriel.
- *Assurer la reproductibilité des résultats de publication* pour permettre aux personnes intéressées par les articles de disposer d'un outil mettant en œuvre les méthodes proposées et permettant de reproduire les résultats décrits dans l'article.

L'idée principale est que CIEL soit pour les publications de logiciels l'équivalent de ce que HAL est pour les articles, avec une volonté affichée de lier les deux types de productions scientifiques.

Projet du COMI sur le patrimoine logiciel Début 2004, une réflexion a été menée au sein du COMI sur la valorisation du patrimoine logiciel du CNRS suggérant la création d'une UPS (CPL, Coopération des Projets Logiciels) dont les objectifs sont le pilotage opérationnel de projets logiciels, et la diffusion des outils développés afin d'en garantir une certaine pérennité. Cette UPS était proposée sous une forme proche de ce qui existe en Grande-Bretagne (Collaborative Computational Projects - CCPs).

Valorisation logicielle à l'INRIA L'INRIA a une politique de valorisation très forte concernant les développements logiciels. Les logiciels sont recensés sur <http://www.inria.fr/valorisation/logiciels/index.fr.html> de façon succincte mais efficace. Ce recensement contient les informations essentielles permettant d'aller plus loin dans la recherche si le logiciel correspond à ce qu'on cherche. Il faut noter que le fonctionnement par projets favorise la réalisation de plateformes logicielles de grande ampleur et donc naturellement valorisables (développement collaboratif, documentation indispensable, utilisateurs multiples, ...). Leur référencement est donc plus facile que celui de codes de recherche individuels.

2 Analyse des résultats de ces expériences

Codiciel n'existe plus depuis 2005 et le projet du COMI n'a jamais vu le jour. Le projet CIEL périclité depuis deux ans et n'a jamais vraiment pris son envol. La masse critique qui aurait pu générer une utilisation plus forte n'a jamais été atteinte.

Ces initiatives avaient des objectifs différents. Dans le cas de CIEL, plusieurs raisons peuvent être avancées pour comprendre cet échec :

Le partage du travail : il n'est pas toujours facile pour un chercheur de faire la démarche de donner les sources de ses codes, sur lesquels il a passé beaucoup de temps.

Le temps consacré au dépôt : il faut s'investir un minimum pour savoir ce qu'il faut faire, où il faut se connecter, regarder la documentation, obtenir un login, déposer le logiciel et en assurer un minimum de suivi.

La réutilisabilité du logiciel : les codes de recherche ne sont pas développés dans le but d'être utilisés par d'autres. Leur publication suppose donc un gros travail de documentation des sources, voir d'écriture d'interfaces utilisateurs.

La rentabilité de l'investissement : si les chercheurs ne sont pas motivés pour déposer les logiciels, cela peut en partie être dû au fait qu'ils n'en voient pas l'utilité et la rentabilité.

Un dépôt déjà effectué sur une autre plateforme : le logiciel est déjà déposé ailleurs, ou bien dans le forge de développement ou sur la page web du chercheur, ou du laboratoire, et il n'y a donc pas nécessité de le redéposer ailleurs.

Cependant, cette initiative a reçu de nombreux échos positifs. Seul le fonctionnement proposé ne semble donc pas forcément en adéquation avec l'investissement que sont prêts à réaliser les chercheurs. Il faut noter également que malgré quelques évocations (dans le rapport du conseil scientifique du MPPU et dans le rapport du comité de prospective sur le calcul du CNRS), il n'y a pas eu de réel soutien politique à l'intérieur du CNRS.

3 Le Projet PLUME

Le projet PLUME, qui a démarré en 2006, vise à Promouvoir les Logiciels Utiles, Maîtrisés et Economiques dans la communauté de l'enseignement supérieur et de la recherche. Il est porté actuellement par le CNRS à travers l'UREC qui assure la plus grande partie des moyens nécessaires.

Le projet technique consiste à mutualiser les compétences de terrain pour monter un serveur d'information et une organisation pour l'alimenter. Le projet stratégique veut mettre en place un comité de concertation inter-organismes de l'enseignement supérieur et de la recherche. Son objectif sera de promouvoir officiellement ces logiciels économiques, de lancer des préconisations et des actions

majeures de manière concertée (comme une migration à la suite bureautique OpenOffice.org par exemple) et de regrouper les différentes initiatives dans ce sens.

PLUME ne référence que ce qui est utilisé dans la communauté, ce qui est une preuve de l'utilité et de la qualité : ça fonctionne « en production ».

Une autre spécificité de PLUME est de faire connaître les développements internes. En effet, de nombreux développements « libres » ont été faits par des chercheurs, des enseignants ou des ingénieurs sans que cela soit visible comme une production du laboratoire ou de l'université.

Il faut que le logiciel soit déjà diffusé (sur un Web, dans Sourceforge ...) ce qui veut dire que le développement est abouti et possède, en général, une documentation. L'autre contrainte est qu'il soit utilisé par au moins deux autres sites (que le site du concepteur) en production, preuve de l'utilité à d'autres.

4 **Projet de référencement des développements internes**

La plupart des développements internes ne répondent pas à ces contraintes. Si on définit un logiciel de laboratoire comme un programme utile (au sens large) pour faire avancer la recherche, on fait entrer dans cette catégorie beaucoup de développements individuels, plus intéressants par l'expertise de leurs auteurs que par la réutilisation pratique du code. Il est cependant intéressant d'avoir une trace de ces développements, ne serait-ce que pour pouvoir contacter la personne qui en est l'auteur.

Les codes développés dans les laboratoires de recherche servent essentiellement de support aux publications. Ils ne sont donc pas exploitables directement de façon aisée car pas développés pour. Cependant, quand on demande individuellement à quelqu'un d'utiliser son code, ou son algorithme, cette demande est en grande majorité acceptée.

L'idée est de penser CIEL différemment en s'appuyant techniquement sur le serveur de fiches de PLUME pour lequel la plate-forme technique existe déjà. En effet, on peut envisager la mise en service d'un serveur de fiches référençant les développements internes avec un fonctionnement proche de celui du projet technique PLUME, basé sur le même outil technique.

Un référentiel de logiciel interne doit comporter les données suivantes :

- nom du logiciel, avec éventuellement un numéro de version
- auteur(s)
- brève description
- laboratoire(s) dans le(s)quel(s) a(ont) eu lieu le développement
- date de publication ou de version
- licence

On peut rajouter des informations optionnelles :

- référence de publications scientifiques liées au logiciel
- état (en développement, en maintenance, non maintenu)

- site web du logiciel
- exemple d'utilisation

5 Partenaires

Un des points cruciaux est la valorisation de la production scientifique autre que les publications. Si cette idée se développe, les chercheurs seront plus incités à développer des codes de qualité réutilisables facilement. Leur valorisation sera alors plus axée code qu'idée scientifique. Si cette tendance arrivait à se confirmer, le concept initial de CIEL aurait sans doute plus de chance d'aboutir. D'autre part, les laboratoires seront aussi plus enclins à donner des moyens pour aider les développements logiciels.

Les laboratoires ne font en général pas grand chose non plus pour mettre en avant les travaux de développements internes. Les codes développés dans un laboratoire ne sont en général pas répertoriés sur leur site, pas mentionnés dans les rapports du laboratoire.

Les tutelles n'ont pas non plus une politique de valorisation de ces développements claire et cohérente, autre que la rédaction de contrats.

Il est donc très important que ce projet technique soit appuyé par une volonté claire de la part des directions scientifiques du CNRS qui devront informer et inciter les chercheurs à référencer leurs travaux.

Il est important de prendre contact avec les services de valorisation afin que l'initiative soit connue et que l'information soit diffusée au sein des entités également par cette voie.

Un autre point à prendre en compte est la collaboration avec les sociétés savantes, et notamment la SMAI (voir http://smai.emath.fr/article.php3?id_article=62) qui pourrait relayer sur son site et par le biais de ses publications (matapli) l'information relative aux logiciels internes.

6 Etapes d'avancement du projet

L'objectif est donc de s'appuyer sur le serveur PLUME et d'y définir un espace particulier pour les fiches descriptives des développements internes. Ces fiches contiennent les informations minimales listées ci-dessus. Elles sont stockées et traitées sur le serveur PLUME.

Il ne s'agit pas de proposer un environnement de développement. Les logiciels restent sur des forges publiques et/ou sur leurs serveurs Web respectifs (laboratoires).

La mise en place de ce service pourrait être facilitée par la création d'un réseau de "correspondant développement internes" au sein des laboratoires dont le rôle serait similaire à celui des "correspondant formation". Les correspondants développements internes seraient ainsi chargés de l'inventaire des développements réalisés au sein de leur laboratoire et de la rédaction des fiches associées.

Les étapes du projet vont consister à :

- Trouver quelques laboratoires pilotes, susceptibles d’abriter un nombre non négligeable de développements internes, pour commencer un travail de référencement :
 - l’IGM-LaInfo
 - l’ICJ
 - le LAAS
 - le LIG (à solliciter)
 - le laboratoire de Mathématique d’Orsay
- Travailler avec les correspondants développements internes de ces laboratoires pour définir le contenu des fiches, la méthodologie de production et de mise à jour, les développements nécessaires sur la plate-forme PLUME.
- Informer, recueillir le soutien de certaines instances du CNRS, trouver un porteur-sponsor : services valorisation, départements MPPU et ST2I, comité prospective pour le calcul scientifique, ...

Références

- [1] INRIA : <http://www.inria.fr/valorisation/logiciels/index.fr.html>
- [2] CIEL : <http://ciel.ccsd.cnrs.fr/>
- [3] PLUME : <http://www.projet-plume.org/>
- [4] Codiciel : <http://www.codiciel.fr/>
- [5] Autour de la valorisation de logiciels d’un laboratoire de recherche, Teresa Gomez-Diaz,
http://www.mathrice.org/rencontres/octobre.2007/support_logiciels.pdf
- [6] Collaborative Computational Projects : <http://www.ccp.ac.uk/>