

**QUAND TU L'AS ECRIT TU L'AS LU !**

LETONDAL C / TABARD A / MACKAY W E

Unité Mixte de Recherche 8623  
CNRS-Université Paris Sud – LRI

09/2009

**Rapport de Recherche N° 1526**

**CNRS – Université de Paris Sud**  
Centre d'Orsay  
**LABORATOIRE DE RECHERCHE EN INFORMATIQUE**  
Bâtiment 490  
91405 ORSAY Cedex (France)

# Quand tu l'as écrit tu l'as lu !

*Catherine Letondal*

INRIA In|Situ LRI  
Univ. Paris-Sud  
letondal @ lri.fr

*Aurelien Tabard*

ENST  
75013 Paris  
aurelien @ tabard.fr

*Wendy E. Mackay*

INRIA In|Situ LRI  
Univ. Paris-Sud  
mackay @ inria.fr

## RESUME

Cet article étudie l'espace de spécification et de conception des cahiers de laboratoire mixtes sous l'angle de l'écriture. Nous montrons, à travers des observations de terrain, que la rédaction du cahier de laboratoire se place dans un espace complexe d'artefacts et d'activités informatiques et intellectuelles qu'il faut comprendre pour aborder la conception d'un cahier de laboratoire intégrant les différents supports d'écriture. Ce travail se situe dans la prolongation d'une approche de sondes technologiques exploratoires qui, à travers le développement de systèmes s'appuyant sur des hypothèses différentes, permet de définir l'espace de conception de ces cahiers.

**MOTS CLES :** Support de l'activité, système de gestion de l'information personnelle, outils réflexifs, co-adaptation, interfaces utilisateur papier-stylo, écriture mixte.

**CATEGORIES AND SUBJECT DESCRIPTORS:** H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

**GENERAL TERMS:** Documentation.

## INTRODUCTION

L'introduction, dans les laboratoires, des technologies numériques de communication et de gestion de l'information a changé radicalement la façon de travailler des biologistes [12]. Leurs activités se partagent maintenant entre le physique (expériences, cours, réunions) et le virtuel (analyses, recherche d'outils et d'articles, échanges d'information). Contrairement à certaines prédictions [4], ils continuent à utiliser le papier sous diverses formes, de brouillons aux cahiers de laboratoire. Nous avons ainsi pu observer qu'aujourd'hui les biologistes continuent à utiliser les cahiers papiers, ce qui ne les empêche pas d'explorer des alternatives numériques.

Cet article pose la question du, ou des, rôle de l'écriture dans l'utilisation des cahiers de laboratoire. Pour poursuivre cette exploration, nous nous appuyons sur notre expérience avec différents types de cahiers :

- des cahiers papiers augmentés de couches numériques,
- des cahiers numériques intégrant le papier, et
- plus récemment des expériences avec des cahiers augmentés d'audio et des blogs.

Dans chacun des cas, nous avons pu nous rendre compte que le support influençait le type d'écriture, mais aussi que selon le type de rédaction, les utilisateurs se destinaient plus à un support ou à un autre. Nous nous proposons de caractériser ces différences dans les façons d'écrire et explicitons les questions qui se posent pour la conception de cahiers numériques à écriture mixte.

## TRAVAUX ANTERIEURS

Les cahiers de laboratoire numériques sont développés avec des hypothèses sous-jacentes concernant le type d'écriture qu'ils sont destinés à recevoir. Ces hypothèses influent sur la forme des cahiers. Ainsi les cahiers utilisés dans l'industrie (pharmaceutique par exemple) sont principalement destinés à capturer le travail de laboratoire et à prouver l'antériorité d'un développement: ils contraignent fortement les usages, suivent des processus et contrôlent strictement les dates d'édition.

Dans le monde de la recherche, plusieurs hypothèses technologiques ont exploré des modes d'écriture différents. myTea [5] se concentre sur la transposition des affordances des cahiers papier dans le monde numérique. myTea propose d'utiliser un tablet PC ajoutant à la prise de note des possibilités de calcul et de communication. Dans ce cadre, l'écriture est vue comme un moyen de capturer et de lier l'information, parmi d'autres (logs, bookmarks). Avec ButterflyNet [13], Yeh et al. se sont intéressés aux biologistes de terrain. Pour ces derniers les données capturées sont intrinsèquement physiques. Aussi plutôt que de prendre un parti tout numérique, ButterflyNet permet aux biologistes de lier leurs données capturées (notes, échantillons, photos, relevé GPS) à des données numériques, et ainsi de les organiser et de naviguer entre elles sur l'ordinateur. Ici aussi, l'écriture dans les cahiers établi des faits, elle est considérée comme

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IHM 2009, 13-16 Octobre 2009, Grenoble, France

Copyright 2009 ACM 978-1-60558-461-4/09/10 ...\$5.00.

définitive (“They take their notebooks everywhere, using them as the definitive record”).

Nous avons travaillé par le passé sur différents cahiers augmentés [6] permettant de lier cahiers papier et numériques. Un premier prototype permettait de numériser l’écriture dans un cahier grâce à une tablette graphique. Un autre prototype, doté d’un PDA agissant comme une lentille ou fenêtre numérique, permettait lui de considérer les notes non seulement comme un moyen d’archivage de l’information mais aussi comme un moyen d’interagir avec des données en créant un lien bidirectionnel entre le papier et l’ordinateur.

Prism [11] s’est intéressé à l’intégration du cahier de laboratoire avec les différents flux de l’activité virtuelle des utilisateurs, mêlant cahier de laboratoire papier, cahier de laboratoire électronique et différents flux d’activités numérique (pages web, documents, emails, flux RSS). Une fois l’information agrégée dans Prism nous avons pu observer que les utilisateurs avaient recours à un cahier pour lier et réfléchir sur leur données. L’écriture dans le cahier prend alors un rôle narratif et de réflexion important.

## METHODE

L’étude, que nous présentons ici, utilise une approche de *sonde technique* [1]. Nous avons exploré avec deux systèmes, Prism initialement puis un blog basé sur Wordpress, l’association de cahiers numériques et papiers (normaux ou Anoto). Ces systèmes ont trois objectifs:

- mieux comprendre l’utilisation des cahiers dans le contexte des laboratoires;
- tester différentes technologies de cahier et leur robustesse sur le terrain;
- inspirer les biologistes et les chercheurs quant aux possibilités et aux implications de l’utilisations de cahiers hybride en les ancrants dans des cas concrets d’utilisation.

Les aspects que nous présentons dans l’étude concernent les artefacts et les activités relatifs à l’utilisation des cahiers de laboratoire. Ces usages illustrent l’importance du support technologique de la rédaction, les situations rédactionnelles, le processus d’écriture et son cadre intellectuel tout comme son déroulement dynamique dans le temps.

## Outils

**Cahiers et stylos Anoto** La technologie Anoto<sup>1</sup> permet de capturer les traits laissés par un stylo sur une feuille de papier. Le stylo comprend une caméra qui lit en permanence une grille de points imprimée sur le papier et définissant des coordonnées. Pour Prism, les stylos utilisés étaient des Logitech associés au Paper Toolkit [14] permettant de manipuler la version brute de l’écriture.

	A	B	C
1	1	2	3
2	4	5	6
3	11	12	13
4			

Figure 1 : Entrée de nombres manuscrits.

Le stylo Livescribe Pulse<sup>2</sup> utilisé dans la suite de l’étude est un stylo Anoto programmable qui possède également des fonctions d’enregistrement audio. Cependant il n’existe pas encore de boîte à outil permettant de manipuler ce qui a été écrit sur les cahiers à posteriori. Nous avons pu cependant obtenir quelques informations à travers le système de log du stylo, permettant par exemple de charger un tableau de nombres manuscrits sur l’ordinateur, ou de repérer des marqueurs permettant de diviser le texte saisi en catégories (figure 1 et 19).

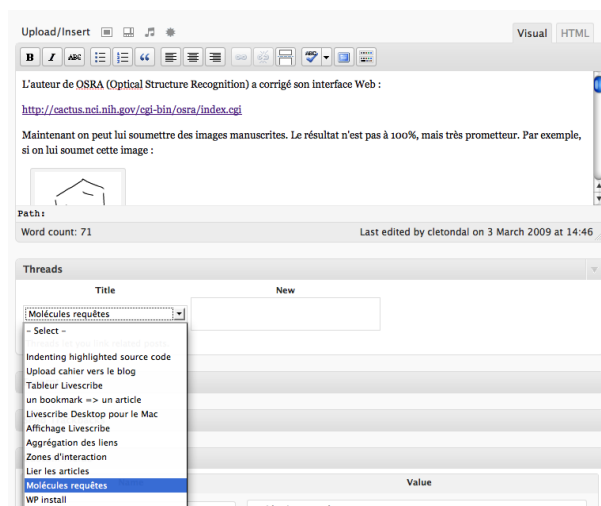


Figure 2 : blog : Une interface de rattachement d’un article à une *thread*. Sur le blog, la liste des articles de la thread est affichée en bas de chaque article.

**Blog** Dans la continuité de l’approche de Prism, nous avons commencé à adapter un outil de blog, Wordpress, pour en faire un cahier numérique augmenté. Tout comme le cahier de laboratoire, un blog repose en effet sur une infrastructure chronologique linéaire pré-déterminée, nécessaire pour refléter le processus intellectuel avec ses évolutions et ses transformations. Il bénéficie d’autre part d’une importante communauté d’utilisateurs et de développeurs et d’une très grande flexibilité d’adaptation.

Dans le cadre de cette étude, nous développons et nous proposons un certain nombre d’extensions Wordpress. Par exemple, une expérimentation de biologie pouvant s’étaler sur plusieurs jours, et aussi pour apporter un support à l’activité réflexive des scientifiques, nous avons proposé un outil pour créer des fils de discussion ou “threads” (figure 2). Nous avons aussi adapté les modèles d’affichage du blog pour permettre une organisation par

<sup>1</sup><http://www.anoto.com/>

<sup>2</sup><http://www.livescribe.com/smartpen/>

date plutôt que par article. Parmi les extensions proposées, et qui feront sans doute l'objet d'adaptations de notre part, citons un outil de visualisation de type timeline pour repérer des articles sur une échelle chronologique ; un outil de planification de tâches permettant d'intégrer les "choses à faire" dans le flux des activités reportées par le blog ; une fonction de soumission de bookmark depuis un navigateur, qui permet au biologiste d'intégrer des liens dans le contexte chronologique de son activité.

Depuis le début du projet, nous maintenons une sorte de blog réflexif, appelons-le meta-blog, dont l'objectif est de discuter les différents aspects du projet, de proposer des fonctionnalités ou de démontrer des extensions qui nous semblent utiles, et qui sont parfois proposées par les participants eux-mêmes, démontrant le potentiel co-adaptatif de ce type d'outils pour une étude participative. Ce blog fournit également un mode d'emploi des extensions développées et comporte une page de "choses à faire" pour la suite du projet. Plusieurs des participants sont inscrits comme lecteurs ou comme auteurs sur ce blog, y font des commentaires, et s'abonnent au flux RSS.

### Participants

Les observations que nous présentons ont été effectuées durant la période d'utilisation du blog et du Livescribe, d'octobre 2008 à avril 2009. Nous avons travaillé avec deux biologistes de l'Institut Pasteur, CM et CS, travaillant dans un même laboratoire qui étudie la génétique des interactions macromoléculaires. Tous deux sont à la fois biologistes expérimentaux et programmeurs. CM est aussi le bioinformaticien du laboratoire, chargé des développements de bases de données. CS maintient le blog-intranet du laboratoire. L'un des auteurs de l'article connaît CM et CS depuis plusieurs années et a travaillé avec eux lors d'un cours d'informatique. A l'INSERM, nous avons mené une série d'interviews ainsi qu'un atelier de prototypage vidéo. Les participants sont des chimioinformaticiens (OS et DL), bioinformaticien structuraliste (JB) et une biologiste (AV), également enseignante en bioinformatique. Nous avons également proposé à une biochimiste et ex-bioinformaticienne de l'Institut Pasteur (KS), actuellement dans une société de bioinformatique, de participer à l'étude. Trois des participants (AV, CM et KS) ont pu avoir un cahier Livescribe. Certains des biologistes participant à l'étude utilisaient déjà un blog, parfois comme cahier de laboratoire où ils consignent leur activité - souvent plus bioinformatique que biologique, parfois aussi comme gestionnaire de contenu pour un site intranet. Durant l'étude, quatre participants (AV, DL, OS, CM) ont progressivement adopté l'idée d'utiliser un blog.

### RESULTATS

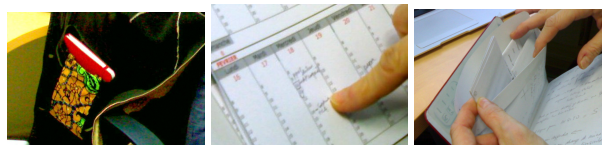
Nous décrivons ici en quoi la dimension temporelle de l'écriture (encodage chronologique), les différents modes de production de l'information (enregistrement brut et synthèse), ainsi que les déplacements qui s'opèrent entre

les niveaux d'information éclairent la crise actuelle des cahiers de laboratoire.

### Des cahiers de laboratoire en crise

La rédaction, qu'il s'agisse du blog, du cahier, ou de toute information concernant l'activité du biologiste a besoin d'être située dans un espace reconnu comme une structure convenue qui peut *accueillir* l'information. Cela garantit qu'on saura la retrouver, que cette structure rassemble des informations qui vont ensemble pour une raison bien déterminée - pas nécessairement parce qu'elles sont de même type d'ailleurs. AV, depuis qu'elle travaille avec des bioinformaticiens, perd parfois des éléments essentiels des discussions, parce qu'aucune structure n'a pu être décidée pour accueillir les étapes scientifiques d'un projet : "Je me suis donné du mal, je n'ai pas trouvé de structure pour réussir à garder les résultats des cinq personnes qui chacun avait un sujet différent là-dedans."

Le support a ainsi une grande importance, il "supporte" non seulement l'activité de rédaction du cahier, mais il induit une sémantique de son contenu. Il y a le cahier de séminaire, le cahier thématique (OS : "j'ai un fichier par cible thérapeutique : je ne veux pas tout mélanger"), le cahier fourre-tout, sorte de cahier-vie ou de "main courante" toujours dans le sac, souvent "celui qu'on ne pourrait pas perdre" et dont, quand on dit "tout est dedans", on affirme le fait que ce tout n'est pas tout et n'importe quoi mais "le tout de ce qui existe" (figure 3). Il y a le cahier-blog ou le cahier Livescribe qui intriguent et déstabilisent par leurs multiples possibilités, comme l'explique DL : "Je pense que ce qui est frustrant c'est de ne pas réussir à imaginer tout ce qu'on peut faire avec. D'être freiné par son imagination".



**Figure 3 :** Le cahier-vie: le cahier doit être disponible à tout moment, comme l'illustre la confection d'une poche interne pour le cahier, le collage d'un agenda à la fin du cahier, et le bricolage d'une pochette de documents en attachant deux feuilles du cahier.

Les biologistes sont ainsi en recherche active d'affectation d'outils pour assurer le traçage de leur activité et leur réflexion scientifique. Les systèmes qu'ils essaient vont des outils de cartes heuristiques (mind-mapping), aux logiciels de planifications de tâches comme ProjectManager en passant par Word, wiki, les modèles de pages HTML. Certains utilisent des blogs, soit comme moyen de maintenir un cahier de laboratoire, soit comme outil de gestion de contenu de site web (CMS). KS utilise toute une série de supports différents : agenda, cahier libre pour cartes heuristiques, cahier fourre-tout, cahier de laboratoire laissé sur le lieu de travail, dossiers plastifiés des documents courants, annotés sur le plastique, wiki, ... Cette

multiplicité nous a frappé, mais la raison principale en est que cette bioinformaticienne est à la recherche du système idéal et essaye plusieurs solutions, forcément redondantes (figure 4).



**Figure 4 :** Multiplication des supports pour essayer des alternatives.

### Aspects temporels

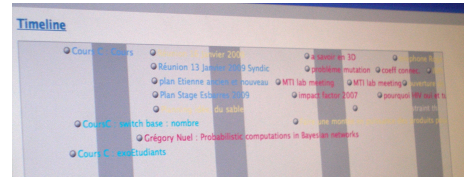
Dans cette section, nous relevons les occurrences des aspects rédactionnels liés au temps, telles que la datation des articles ou à l’encodage du processus scientifique dans l’objet temporel qu’est le cahier [10] [3].

**Datation** Qu’il s’agisse des cahiers manuscrits ou des articles de blog, les entrées sont datées. Un des participant utilise par exemple la fonction de date de word pour insérer la date d’un ajout dans un fichier de projet. Sur les cahiers, la date n’est pas systématiquement reportée : le caractère linéaire et chronologique du cahier semble suffire pour s’y retrouver en cas d’oubli ou d’omission volontaire. Les outils de blog associent une date automatique, qui est en général la date de publication de l’article. Les biologistes de l’étude ont parfois dû jongler avec les mécanismes de datation du blog, notamment souvent pour voir figurer un article non finalisé en haut de la liste pour le retrouver plus vite. Les stratégies techniques varient pour obtenir ce type d’accès direct :

- en modifiant le squelette Wordpress pour qu’il affiche les articles par date de modification,
- en éditant la date de l’article,
- en écrivant d’abord une page statique pour la transformer ensuite en article,
- en utilisant le mécanisme d’article “sticky”.

Le calendrier pourrait aussi constituer un moyen pratique d’accéder aux articles, mais il ne semble pas utilisé de façon systématique et peu de participants le mentionnent. Une des raisons est qu’il n’affiche les titres des articles que sur passage du pointeur, contrairement à la timeline qui les affiche directement (figure 5). Ce qu’apporte surtout la timeline et qui est très apprécié, c’est la possibilité de visualiser les occurrences de publications sur une échelle temporelle globale (annuelle, mensuelle ou hebdomadaire).

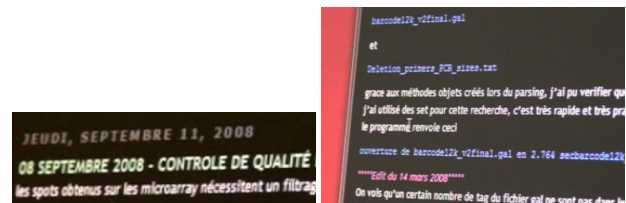
C’est plutôt pour affecter une datation réaliste à l’article, et encoder correctement les étapes de leur processus de travail, que certains participants préfixent le titre de l’article de la date de l’expérience relatée (figure 6.a). A



**Figure 5 :** Visualisation globale des articles sur une échelle temporelle.

noter que CM par exemple souhaite aussi conserver la date automatique et “objective” de Wordpress .

**Fil d’activité** Dans l’extension que nous fournissons aux participants, nous proposons un moyen supplémentaire d’afficher le lien chronologique entre les articles à travers la création de “threads”, qu’on pourrait appeler “fils de discussions” si ce n’est qu’il s’agit plutôt de discussions avec soi-même (figure 2). Cette fonction a été développée suite à une discussion avec un biologiste qui observait qu’il était difficile, avec les outils fournis par les blogs, de reporter une expérience se déroulant sur plusieurs jours. Son utilité est également observable chez CM qui, pour matérialiser le fait qu’un article relatait deux étapes distinctes mais relatives d’une même analyse, avait annoté les passages concernés d’une marque “edit” (figure 6.b).

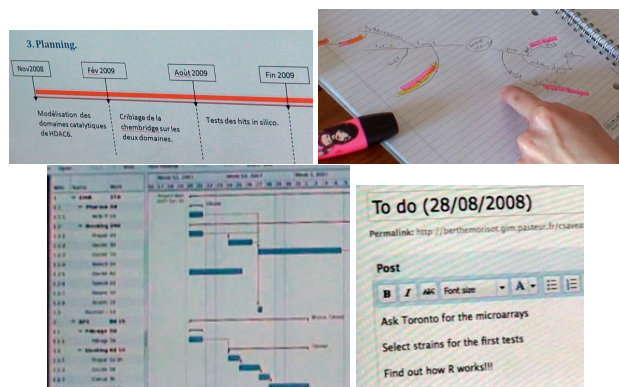


**Figure 6 :** (a) Décalage entre la date d’écriture et la date de rédaction. (b) Mentions explicites des éditions d’un article et de leur date.

**Gestion du temps** Le cahier de laboratoire et les documents associés sont fréquemment le support d’outil de gestion du temps - de “feuilles de temps” pour reprendre l’expression de KS (figure 7). Les entrées de type “choses à faire” ou “TODO”, parfois barrées une fois accomplies, sont soit mêlées au flux des articles soit reportées dans un cahier à part. L’outil de timeline suscite ainsi parfois des suggestions d’y incorporer des informations d’échéances, de durée d’accomplissement d’une tâche. CM reporte très souvent des informations de temps dans son blog pour prendre en conscience du temps que lui prennent les choses. Ce type d’information répond à un besoin de réassurance chez certains biologistes, particulièrement ceux, souvent bio-informaticiens, qui ont plusieurs projets à gérer et pour qui la mesure du travail accompli et la planification des tâches futures est un peu plus difficile.

La gestion du temps se manifeste aussi par la mise en place d’un rendez-vous quotidien du biologiste avec son cahier : certains ont une entrée “pour demain”, que ce soit

dans le cahier, après le rapport de la journée, ou dans le blog, sous la forme par exemple d'un article "sticky".



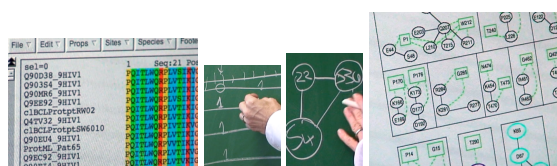
**Figure 7 :** Gestion du temps : (a) un planning de projet sous Word. (b) un carte heuristique pour planifier la semaine. (c) Utilisation de ProjectManager et (d) écriture d'un article pour les tâches à planifier.

**Historiques** Wordpress fournit un certain nombre d'outils pour connaître la date des versions successives des éléments d'information. Ainsi, un historique des éditions est disponible, mais nous avons constaté que peu de biologistes s'en servent ou l'ont même remarqué. C'est plutôt un versionnement implicite que nous avons vu utiliser, associé à la traçabilité. Les bioinformaticiens changent souvent leurs scripts pour les adapter aux évolutions des données et apprécient de pouvoir conserver le code d'analyse associé à chaque résultat. Le script associé est donc attaché à l'article et chargé sur le serveur, qui maintient ainsi une archive des versions non pas d'édition mais d'exécution des scripts. Cet historique devient ainsi un log des scripts exécutés.

**Vitesse, temps passé** Sur un autre plan, qui serait plutôt celui du processus intellectuel, nous avons observé que ce n'est pas la vitesse qui est recherchée à tout prix. Le temps nécessaire à la maturation d'une idée, voire la lenteur imposée par un dispositif, sont prisés pour leur efficacité en termes de production d'idées.

### Penser

La question de la rédaction d'un cahier de laboratoire est bien sûr aussi celle des étapes de la réflexion scientifique et de ses modalités pratiques. Nous en retiendrons essentiellement trois aspects.



**Figure 8 :** Histoire de données et d'analyses : séquences de protéase alignées (a), encodage des positions corrélées (b), graphe de cet encodage au tableau, (c) puis sous GraphViz

**Histoires** Les biologistes ont un penchant naturel pour les histoires. Leurs objets d'investigation - virus HIV, complexes protéiques, gènes - sont des objets qui se racontent, à qui il arrive des choses, qui ont une vie qu'on peut suivre à travers les étapes d'un processus. De même, les données qu'ils analysent, en particulier en bio-informatique, subissent des transformations au fil des explorations : de séquence nucléique encodée dans une chaîne de caractères, elles deviennent tableau de booléens, puis graphe ou représentation 3D (figure 8). Les biologistes que nous avons rencontrés font souvent émerger un sens de ces données par une mise en situation narrative [9], qui n'est d'ailleurs pas limitée aux données. Ainsi, AV, lorsqu'elle réfléchit à sa question de recherche sur la co-occurrence négative de mutations dans la protéase, semble-t-elle se glisser dans la machine et faire les calculs à sa place, non pas seulement pour nous expliquer sa problématique, mais pour en réfléchir la logique et la cohérence :

- "les matrices de score ne reflètent pas obligatoirement sa façon de muter... et du coup je me dis que je vais rajouter un cache supplémentaire sur quelque chose qui va peut-être pas me raconter la bonne histoire..."
- "Et alors, juste pour vous dire la petite histoire qui est drôlement intéressante scientifiquement : la figure la plus grosse des résultats, c'est des acides aminés qui se trouvent là..."
- "... et en fait, on revient à l'histoire du départ, de dire : si ces trucs ne mutent jamais c'est sûrement qu'ils ont une fonction ... qu'ils font une fonction ensemble"
- "et là j'étais dans cette histoire de : est-ce que ça pouvait se coincer ou pas"

De même, JB nous explique pourquoi elle stocke les traces des discussions dans des pochettes plastique : "...quand ça m'intéresse, quand j'ai un truc qui est sorti de ça... pour regarder le film de ce que j'ai fait".

**Secret et liberté** L'exploration libre, la concentration, tout comme la rédaction du cahier, nécessitent un espace approprié, mais privé cette fois : "Alors ça c'est mon cahier mais vous voyez c'est mon cahier secret ... C'est mon cahier où il se passe des trucs. (...) Et je me suis aperçue que les trucs vraiment importants qui se passent, il faut pas que les gens soient au courant... Tu vois tu as besoin d'avoir une liberté absolue, ... pour avoir une espèce d'esprit créatif maximum."

Nous avons aussi généralement constaté une réaction non pas de méfiance, mais toujours une question en proposant l'idée de blog : les participants s'inquiétaient du fait que le blog soit une publication, hébergé sur un site serveur, et non un espace d'exploration personnel.

**Appropriation, recopie, rituel** A l'instar de l'écrivain, l'appropriation d'un problème, la nécessité d'obtenir une concentration suffisante demandent parfois une action qu'on pourrait presque qualifier d'incantatoire : préparer

du matériel de rédaction, amorcer l'écriture, vecteur de la pensée : "j'ai besoin de me concentrer énormément sinon je ne dépasse pas le taux d'énergie nécessaire à ... voila... Il me faut un rituel obligatoire : le bureau proprissime - sinon trop d'interférences, ce n'est pas possible". Plusieurs des participants ont ainsi besoin de recopier leurs données ou du moins une partie de l'ordinateur sur le papier pour amorcer une réflexion. Par exemple JB ou AV recopient des tableaux de chiffres, des listes de noms de gènes, des valeurs calculées, des graphes, ... pour "rentrer dans les données" (figure 9.a). Comme l'exprime JB : "Quand tu l'as écrit, tu l'as lu!". Il n'est pas rare de rencontrer souvent dans les cahiers du code source, en SQL par exemple, ou en python, ou bien sûr des algorithmes en pseudo-code (figure 9.b), toujours par besoin qu'on les biologistes de s'approprier l'information, de s'en imprégner.

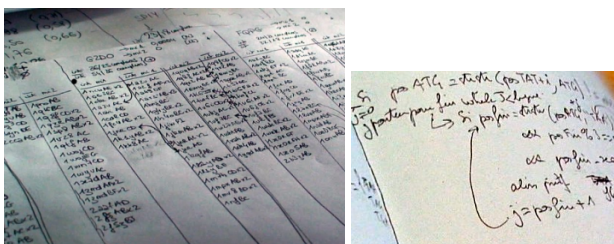


Figure 9 : Recopie de données, code rédigé ou recopié dans le cahier.

### Promotion, élaboration de l'information

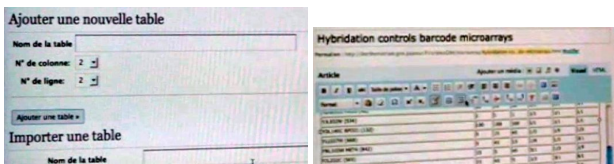


Figure 10 : Données tabulaires : essais d'extensions wordpress

L'information, les idées, l'expression des problèmes et de leur solution traversent des niveaux de diffusion de la note de brouillon temporaire à l'article international. Ainsi, dans le laboratoire de CS et CM, les biologistes consignent leurs manipulations de paillasse dans un cahier de laboratoire tant que ces expériences sont encore en attente de validation. Une fois la décision prise d'en faire une thématique du groupe, elles sont promues et figurent dans le blog du projet, puis dans le blog de l'équipe mais sous une forme plus synthétique, puis une procédure automatique extrait l'information appropriée sur le site public de l'Institut Pasteur. Le passage sur le blog pose en réalité un vrai problème technique : la mise en forme des données tabulaires, pour lesquelles CS, mainteneur du blog, a peine à trouver les extensions adéquates (figure 10). Cette promotion est unidirectionnelle : une thématique, une fois passée dans le blog à des fins de mise en commun de données et de partage d'information, ne repasse jamais dans le cahier de laboratoire, même si elle nécessite des expérimentations sur la paillasse. Ces différents passages,

ces "mises au propre", induisent évidemment une certaine redondance, et encore une fois de la recopie au moins partielle, accompagnée de filtrage et de synthèse (figure 11).

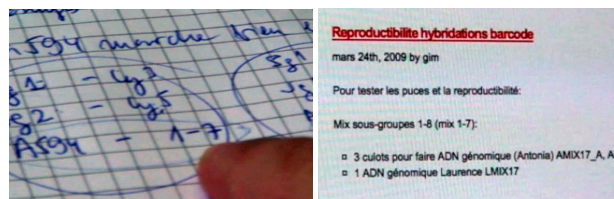


Figure 11 : Mise au "propre" : sur le blog, les informations sont moins elliptiques.

Ce mouvement de promotion correspond aussi au passage d'une information d'un état dynamique à un état plus statique. Pour KS, mais aussi pour CS, ce qui figure dans le wiki ou le blog - utilisé comme un site intranet - est plus définitif que ce qui figure dans le cahier, qui en progrès.

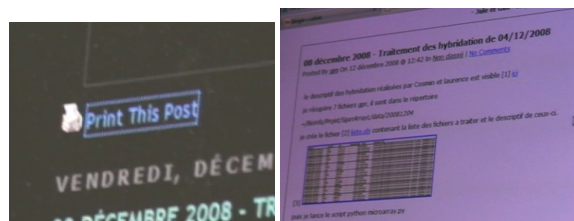


Figure 12 : Imprimer un rapport d'activité à partir du blog.

Enfin, une autre forme de promotion ou de "mise au propre" est la possibilité de produire un rapport d'activité à partir du blog, mentionnée par l'un des participants de l'INSERM. CM avait ainsi prévu, par l'installation d'une extension, cette possibilité d'imprimer un article pour rendre compte de son travail dans une réunion (figure 12).

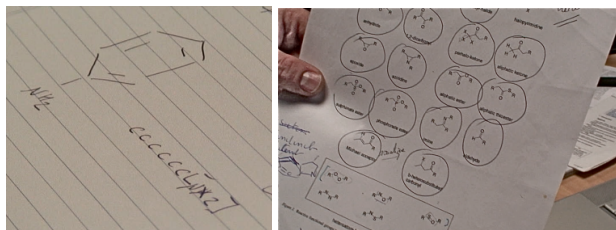
### Volatilité : interactions, discussions

La volatilité aussi bien des interactions avec l'ordinateur que des discussions scientifiques avec les collègues est le problème central auquel la pratique du cahier de laboratoire est censé faire face. CM, quand il travaille seul, prend soin de retracer les principales étapes de ses interactions avec l'ordinateur. Mais quand, lors d'une discussion avec le responsable du laboratoire, il doit essayer à sa suggestion un certain nombre d'outils, il ne peut pas se souvenir de tout. Comme l'exprime OS, "l'informatique ça s'évapore", et il faut garder des logs de ce qui se passe, des commandes qui ont servi à produire tel et tel résultat. Le défaut actuel de structure d'accueil a pour conséquence directe cette perte d'information. Comme le dit AV à propos d'une discussion cruciale pour la suite de son projet : "La conclusion était carrément inexistante. Sur le moment elle l'a été et puis on l'a stocké nulle part".

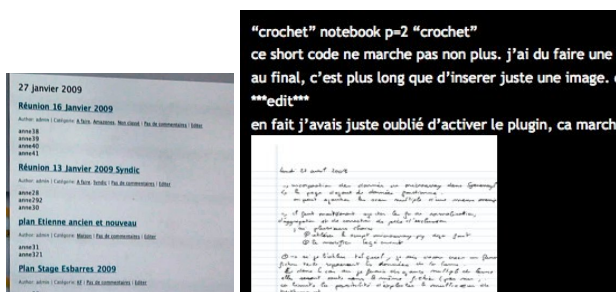
### blog et cahier Livescribe

L'introduction auprès de deux participants, l'un à l'INSERM, l'autre à l'Institut Pasteur, de la possibilité d'utiliser conjointement le blog et le cahier Livescribe nous a apporté

des premiers éléments de réflexion. La mise en relation entre le cahier électronique manuscrit et le blog a inspiré des idées autour de la reconnaissance de caractères et la reconnaissance de schémas, notamment de molécules chimiques (figure 13).



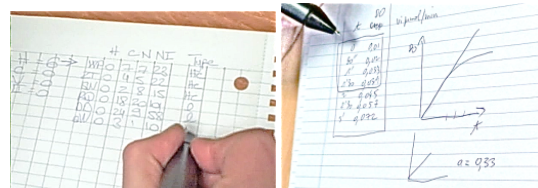
**Figure 13 :** Reconnaissance de schémas de molécules dessinées sur le cahier Livescribe. Une fois reconnues et transformées en code SMART, ces schémas pourraient servir de requêtes pour des calculs ou des recherches dans des bases de schémas.



**Figure 14 :** (a) Le blog comme index du cahier Livescribe. (b) Recopie complète.

Pour AV, l'introduction du cahier Livescribe a été comprise comme l'équivalent d'une saisie manuscrite du blog. Elle s'attendait à de bonnes performances des outils de recherche de texte fournis avec le stylo et pensait qu'il serait aisé de répartir automatiquement les pages du cahier sur le blog en fonction des catégories déclarées au stylo sur le papier. Ces mécanismes étant loin d'être fonctionnels, elle a progressivement orienté sa rédaction de cahier vers le blog, dont elle est satisfaite. Elle utilise le cahier, mais le critique comme un système informatique qui ne marche pas : une information qu'on ne peut pas retrouver est perdue. Elle organise d'ailleurs le cahier de façon analogue au blog, avec les mêmes catégories qui sont aussi celles de son système de fichiers, et déclare que sur le blog, "une catégorie c'est comme un petit cahier". Le blog devient donc progressivement le document maître, et sert à l'occasion d'index dans le cahier manuscrit (figure 14.a).

Pour CM, qui n'avait qu'une expérience très brève du cahier Livescribe quand nous l'avons interviewé, le cahier manuscrit convient mieux que le blog, plus fastidieux à maintenir. Il apprécie la possibilité de pouvoir disposer le cahier à côté de l'ordinateur plutôt que d'avoir en permanence une fenêtre sur le blog affichée à l'écran. Selon lui, qui fait pourtant plus de bioinformatique que AV, c'est certainement le cahier qui deviendra le document maître.



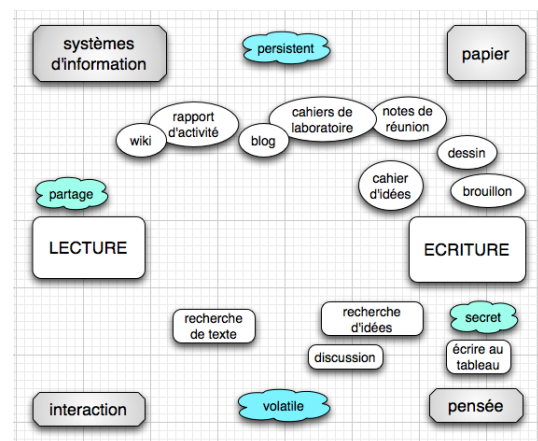
**Figure 15 :** Conception participative d'un tableau pour le Livescribe.

Il compte charger régulièrement les pages sur son blog, en les complétant par des adresses de fichiers ou de sites web, et par des copier-coller de résultats informatiques (figure 14.b).

### Interaction avec le stylo Livescribe

Les possibilités d'interaction avec le stylo ont fait l'objet d'un atelier. Plutôt que de s'intéresser aux aspects d'enregistrement audio, qui suscitent l'intérêt mais sont déjà très opérationnels, la thématique s'est concentrée sur les tableaux (figure 15 a et b), structure de données essentielle pour les biologistes (figure 10). Cet atelier a donné lieu à un développement de prototype, permettant de saisir un tableau interactif de  $n$  colonnes sur le cahier, en vue de le charger sur le blog. Un brainstorming a aussi porté sur un outil de définitions de fonctions simples de calcul, comme une analyse cinétique, utilisant des données numériques d'un tableau (figure 15 c et d). Enfin, une courte session de brainstorming avec CM a permis d'identifier les interactions avec une feuille livescribe qui permettrait de relier les données informatiques au cahier.

### DISCUSSION



**Figure 16 :** Cartographie de l'écriture axée sur deux pôles : écriture-pour-la-lecture et écriture-pour-l'écriture.

Dans cette section, notre objectif est d'expliquer les observations rapportées dans la section précédente. Nous avons constaté, à travers nos observations, le fait que les biologistes, et avec eux la plupart des *knowledge workers*, écrivent à la fois dans le but de produire de l'information stockée et structurée dans un système d'information numérique et par besoin de récolter et de travailler leurs



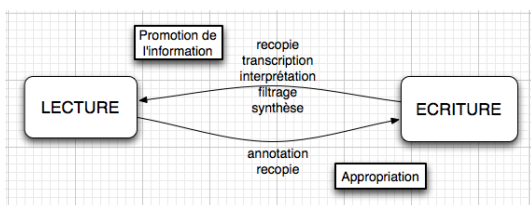
idées sous une forme plus discursive. Le diagramme cartographique de la figure 16, par une prise en compte de l'intentionnalité sous-jacente à l'acte d'écrire, permet de modéliser l'écriture autour de deux intentions, la lecture et l'écriture :

- **l'écriture-pour-la-lecture** représente tout ce qui a trait à la production d'un système d'information à but de communication, de partage et d'activité réflexive ;
- **l'écriture-pour-l'écriture** décrit l'acte d'écrire lui-même en tant que vecteur de la pensée et en tant qu'il effectue en même temps un enregistrement de cette dernière dans un support externe.

Les étiquettes disposées aux quatre coins du schéma orientent également les notions de cette carte :

- **systèmes d'information** : ce qui à la fois est destiné à être conservé, partagé, et réside sur un support informatique ;
- **interaction** : les activités et artefacts ayant lieu sur un ordinateur, mais sans nécessairement laisser de trace ou donner lieu à un enregistrement ;
- **papier** : les activités et artefacts intellectuels qui utilisent le papier comme outil d'inscription ou d'archivage ;
- **pensée** : les activités et artefacts intellectuels volatiles, n'utilisant ni ordinateur ni papier.

Autour du pôle "lecture", se trouvent tous les artefacts et activités relatifs à l'écriture destinée à une lecture future, à la production d'information partagée : l'objectif est d'organiser l'information et de la rendre accessible. On doit pouvoir effectuer des recherches de texte, proposer différents vecteurs de navigation, comme des liens hypertexte, une classification thématique. L'information est accessible à la fois de manière globale et directe grâce aux techniques d'indexation et aux algorithmes efficaces de recherche de texte. La production collective d'information permettant une meilleure communication est favorisée, par exemple à travers des outils comme le wiki. L'information produite est de nature plutôt définitive et compilée : on y trouve par exemple des résultats scientifiques, des compte-rendus de réunions ou de conférences, des documentations logicielles, des protocoles de laboratoire, des dossiers soumis pour des financements, la description d'une équipe, etc...

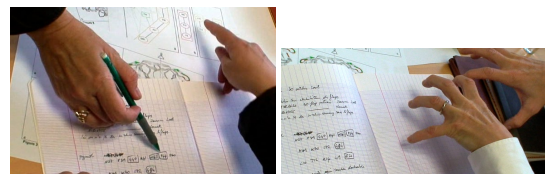


**Figure 17** : Transferts d'information : promotion et appropriation.

Le pôle "écriture" décrit ce qui a trait à la dynamique de la pensée et à son enregistrement : la recherche cette

fois non pas d'information mais d'idées, la création intellectuelle ou esthétique. De ce côté, il s'agit non pas de rendre l'information partageable, mais de lui donner corps pour l'empêcher de s'envoler, en se donnant un espace de notation disponible et temporaire dans lequel rien n'est décidé définitivement, et où règne une grande liberté d'expression personnelle. Pour cette raison, contrairement au côté lecture qui est celui de la communication collective, on a parfois besoin ici d'un espace privé, favorable à l'absence de censure, et où l'on s'autorise des formulations non définitives, des questionnements qui encombreraient inutilement les pages d'un site Web de laboratoire. Cet espace doit en réalité être autant protégé de sa propre évolution vers l'espace "lecture" que du regard extérieur. Ce qu'exprime AV lorsqu'elle déclare, à propos du Livescribe, : "Comment veux-tu avoir des bonnes idées sur un truc qui est en train de *tout* garder ?", c'est bien cette crainte de penser dans un espace d'informations trop définitives, à but de communication.

Le diagramme est également orienté du bas vers le haut, au long d'un axe qui part du volatile vers le persistant. On trouve en haut d'un côté par exemple les bases de données, rapports d'activité, sites Web externes, et de l'autre les cahiers de laboratoire manuscrits, trace définitive d'un processus scientifique. Ce qui est caractéristique de cette partie du diagramme est l'utilisation du support papier comme support d'archivage [2]. Vers le bas, ce sont du côté lecture les interactions avec un système d'information : recherches, visites de pages Web, du moins quand il n'en reste aucune trace sous forme de log ; de l'autre côté, on trouve par exemple les schémas au tableau pendant une discussion.



**Figure 18** : (a) Recopie de données sur le papier pour discuter. (b) "Moi je gratte".

Les deux pôles sont rattachés par des opérations de transfert (figure 17). Du pôle écriture vers le pôle lecture, on trouve la recopie pure et simple, la transcription, ainsi que les diverses techniques d'interprétation comme la reconnaissance d'écriture ou de diagrammes dessinés. Ces opérations s'accompagnent souvent de filtrage, de classement ou de synthèse. Un texte abrégé pour une utilisation personnelle peut aussi être complété afin d'être plus compréhensible. Les déclencheurs de ces différents types de transfert sont multiples : ils correspondent en général à un besoin de promotion de l'information, de mise en forme plus définitive et communicable, de "mise au propre" et d'élaboration scientifique et intellectuelle. Dans l'autre sens, les informations numériques sont aussi très souvent associées et incorporés dans les

documents manuscrits, par impression, découpage et collage. Nous avons souvent observé chez les biologistes que les transferts par recopie de données (figure 9 ou 18.a) correspondent à un besoin d'appropriation, recyclage de l'information compilée dans un compost propice à la réflexion et à l'émergence d'idées nouvelles ("moi je gratte avec mes doigts" dit explicitement AV, figure 18.b).

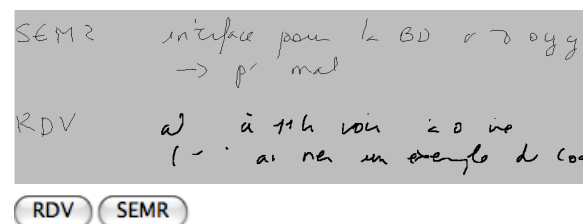
## IMPLICATIONS POUR LA CONCEPTION

Cette modélisation nous permet de mieux appréhender l'espace de conception de cahiers de laboratoire mixtes et nous guide dans nos choix technologiques. Le cahier de laboratoire des biologistes se situe manifestement à la frontière des deux espaces : il est à la fois personnel - une seule personne écrit dedans - et collectif puisque toute l'équipe peut venir le consulter dans le bureau du scientifique. Il est très souvent manuscrit sur un support papier, pour des raisons simplement pratiques : cela le rend utilisable dans des parties du laboratoire où il n'y a pas d'ordinateur, la saisie de valeurs affichées sur un appareil de mesure étant plus aisée que la saisie au clavier, mais il fait partie du système d'information définitif du laboratoire, et peut être conservé pour une durée indéterminée, notamment par le service des archives ou des brevets. Enfin il comporte à la fois des éléments de type dynamique : trace du processus intellectuel sous la forme d'hypothèses personnelles, ou d'expérimentations, et de type statique : relevé de résultats directement reportés dans un article international. Par son inscription dans un support tangible, le cahier papier offre des capacités de localisation rapide, mais ne constitue pas un espace de recherche à accès direct. Cependant, par sa structure généralement chronologique, il permet une recherche d'informations localisables par leur date ou leur co-occurrence avec d'autres éléments datés. Comme on l'a vu, le développement récent de l'information numérique en biologie rend cet outil de plus en plus impraticable.

L'outil de blog, pour sa part, se situe aussi à la frontière des deux espaces [7]. Il est en effet à la fois personnel et public, et a souvent pour objectif - parfois même explicite - de favoriser la recherche d'idées et le travail sur la formulation de la pensée [8]. Il est linéaire et chronologique, autorisant par exemple la répétition avec variation dans le temps. Il présente souvent la même dissymétrie entre auteur et lecteurs que le cahier de laboratoire. Mais on l'a vu dans l'étude, le blog est souvent utilisé comme gestionnaire de contenu et se situe donc clairement pour cette composante-là du côté "lecture" du schéma. Contrairement aux cahiers de laboratoires manuscrits il permet une recherche de texte et une organisation non chronologique par définition de catégories et de tags. Enfin, tout comme le cahier de laboratoire, déjà bien installés dans la culture collective, les blogs constituent un espace de référence de l'expression personnelle. Cet aspect est particulièrement important pour les biologistes à la recherche d'un "contenant", d'un réceptacle légitime,

reconnu et technologiquement praticable de leur travail intellectuel (voir page 3).

Notre étude permet de reposer la question des *rattachements* entre le cahier manuscrit et le cahier électronique ainsi que la question du rattachement des articles du blog entre eux. Prism propose un mécanisme automatique d'assemblage, dont l'intérêt est de fournir à chaque élément d'information un contexte chronologique. Cet attachement pose d'abord une question : les thèmes abordés au niveau de l'écriture du cahier sont-ils contextuellement pertinents par rapport aux thèmes plus travaillés reportés dans le blog ? S'ils sont produits, matériellement parlant, au même moment, s'agit-il pour autant des mêmes activités et des mêmes questions ? du même type de données ? On l'a vu, les données des cahiers sont parfois plus liées aux expérimentations de laboratoire, celles du blog aux manipulations et analyse de données. Le lien d'un type de donnée à l'autre ou d'une étape du processus scientifique à l'autre est-il un lien pertinent à l'échelle quotidienne ? Mais, comme nous l'avons constaté, il existe aussi une grande différence de nature dans les activités de rédaction, de réflexion, de production de systèmes d'information, d'interaction avec un ordinateur et de discussions entre collègues. Si KS nous déclare : "Je n'aurais jamais un seul cahier, ça je le sais déjà, ça ne marchera pas", c'est que le transfert d'information d'une activité à l'autre ou d'un artefact à l'autre est problématique, et suppose de respecter une granularité assez fine, tenant compte de la complexité de cet espace. Par exemple, les difficultés que les participants de l'étude ont manifesté à l'idée d'utiliser un blog, c'est-à-dire un outil de publication, et surtout d'y associer automatiquement des pages de leur cahier, suggèrent qu'une étude fine des mécanismes de gestion des droits d'accès doit être menée et qu'un outil de transfert du cahier Livescribe vers le blog permette de choisir entre un transfert automatique et un transfert sélectif.



**Figure 19 :** Catégorisation manuscrite (les mot-clés sont SEMR pour séminaire et RDV) et interface de sélection.

S'il faut bien comprendre les différences entre les deux pôles de l'écriture et les soutenir techniquement en tant que tels par des outils différents que sont par exemple le cahier manuscrit et le blog, il reste bien sûr tout fait utile de permettre des déplacements et des rattachements d'information entre les deux. C'est l'objectif notamment d'un outil de transfert de tableaux manuscrits vers l'ordinateur (figure 1) ; nous avons également commencé

à prototyper des transferts sélectifs d'extraits manuscrits repérés par des catégories vers le blog (figure 19). Sur le plan du rattachement des articles entre eux, nous avons proposé un support des threads (page 4) ; nous envisageons aussi un meilleur support des références croisées et un outil de "mashup" d'articles, permettant de composer une visualisation sélective.

Un outil comme la timeline (figure 5), outre la co-visualisation globale des articles et des pages manuscrites affichées sous forme de vignettes, permet un déplacement de la réflexion entre l'espace de l'écriture-enregistrement et l'espace de l'écriture-lecture. Mise en présence de cet outil qui lui affichait toute son activité des derniers mois, AV, a commencé à analyser son activité par une re-lecture, ce qu'elle ne faisait ni avec le cahier ni avec le blog. Cette expérience nous a confirmé l'existence d'une importante distinction entre la *réflexion* qui s'inscrit dans l'écrit, opérant une décharge cognitive par l'enregistrement de la pensée sur un support, et la *réflexion sur l'activité*, retour sur les traces de la réflexion par la lecture, et qui opère plutôt, selon l'expression de notre participante, une "recharge" cognitive.

Enfin, dans le cas d'une discussion avec essais techniques autour d'un schéma manuscrit (voir page 6), les données issues du stylo peuvent être rattachées chronologiquement aux interactions sur l'ordinateur : un outil de défilement temporel peut permettre de suivre d'un côté les étapes manuscrites et de l'autre un historique d'interaction.

## CONCLUSION

Dans cet article nous apportons des éléments à la fois empiriques et théoriques pour comprendre le problème de la rédaction des cahiers de laboratoire et de l'écriture scientifique. Nous distinguons ainsi deux types d'écritures. L'une qui se destine à une lecture future, par soi-même ou par d'autres : *l'écriture-pour-la-lecture* permet la communication avec soi (processus réflexif) ou d'autres (partage). L'autre tient sa valeur du processus d'écriture lui-même : *l'écriture-pour-l'écriture* permet le dépôt et la construction d'une pensée, via un enregistrement explicite de l'activité de recherche. Cette distinction des activités d'écriture permet d'envisager des outils permettant et intégrant ces modes : avec une composante papier, et des outils de synthèse, de suivi dans le temps et de vision globale, le blog, qui comme le cahier de laboratoire est à mi-chemin entre l'écriture pour la lecture et celle pour l'écriture, peut devenir un vrai outil de recherche.

## BIBLIOGRAPHIE

1. Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., Hansen, H., Roussel, N., and Eiderbäck, B. Technology probes: inspiring design for and with families. In *CHI '03*, New York, NY, USA, 2003. ACM.
2. Letondal, C., Mackay, W. E., and Donin, N. Paper-oles et musique. In *Proceedings of IHM 2007*, 2007.
3. Letondal, C., and Tabard, A. Temporal data and data temporality: Time is change, not only order. In *Proceedings of CHI 2009 workshop: Interacting with temporal data*, 2009.
4. Lysakowski, R. Comparing paper and electronic laboratory notebooks. Scientific Computing and Automation, 1997.
5. m. c. schraefel, Hughes, G. V., Mills, H. R., Smith, G., Payne, T. R., and Frey, J. Breaking the book: translating the chemistry lab book into a pervasive computing lab environment. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 25–32, New York, NY, USA, 2004. ACM.
6. Mackay, W., Letondal, C., Pothier, G., Boegh, K., and Sorensen, H. The missing link: augmenting biology laboratory notebooks. In *UIST '02: Proceedings of the 15th annual ACM symposium on User interface software and technology*, pages 41–50, New York, NY, USA, 2002. ACM Press.
7. Mortensen, T., and Walker, J. Blogging thoughts: personal publication as an online research tool. In *Researching CTs in context*, May 2007.
8. Nardi, B. A., Schiano, D. J., Gumbrecht, M., and Swartz, L. Why we blog. *CACM*, 47(12):41–46, 2004.
9. O'Day, V. L., Adler, A., Kuchinsky, A., and Bouch, A. When worlds collide: Molecular biology as interdisciplinary collaboration. In *Proceedings of ECSCW'01*, pages 399–418, 2001.
10. Stiegler, B. *La technique et le temps. Tome 1: La faute d'Epiméthée*. Editions Galilée, 1994. 396 pages.
11. Tabard, A., Mackay, W. E., and Eastmond, E. From individual to collaborative: the evolution of prism, a hybrid laboratory notebook. In *CSCW '08: Proceedings of the ACM 2008 conference on Computer supported cooperative work*, pages 569–578, New York, NY, USA, 2008. ACM.
12. Wouters, P. Public proofs, science, technology and democracy: What is the matter with e-science? In *4S & EASST Conference*, 2004.
13. Yeh, R., Liao, C., Klemmer, S., Guimbretière, F., Lee, B., Kakaradov, B., Stamberger, J., and Paepcke, A. Butterflynet: a mobile capture and access system for field biology research. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 571–580, New York, NY, USA, 2006. ACM.

14. Yeh, R. B., Paepcke, A., and Klemmer, S. R. Iterative design and evaluation of an event architecture for pen-and-paper interfaces. In *UIST '08: Proceed-*

*ings of the 21st annual ACM symposium on User interface software and technology*, pages 111–120, New York, NY, USA, 2008. ACM.