

L'authenticité à l'épreuve du numérique – Le cas des cahiers de laboratoire

Lorraine Tosi, Aurélien Bénel

▶ **To cite this version:**

Lorraine Tosi, Aurélien Bénel. L'authenticité à l'épreuve du numérique – Le cas des cahiers de laboratoire. 19ème Colloque International sur le Document Électronique, Nov 2016, Athènes, Grèce. hal-02954835

HAL Id: hal-02954835
https://utt.hal.science/hal-02954835

Submitted on 13 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'authenticité à l'épreuve du numérique : le cas des cahiers de laboratoire

Authenticity in the digital era: the case of electronic notebooks.

Lorraine TOSI (1), Aurélien BÉNEL (2)

(1) Université de Champagne
Université de Technologie de Troyes (UTT) – CNRS, Institut Charles Delaunay (ICD) / Tech-CICO
lorraine.tosi@utt.fr
Société Lex Persona

(2) Université de Champagne
Université de Technologie de Troyes (UTT) – CNRS, Institut Charles Delaunay (ICD) / Tech-CICO
aurelien.benel @utt.fr

Résumé. Établir l'authenticité d'un document numérique ne va pas de soi. La plupart des indices que relevaient les historiens et les philologues à l'intérieur des documents anciens n'existent plus dans les documents créés de manière numérique. Nous proposons un dispositif qui permet de transformer l'ordinateur de n'importe quel innovateur en cahier de laboratoire à valeur de preuve légale, utilisable en travail collaboratif. Les premiers résultats, analysées quantitativement et qualitativement, par des ateliers, des entretiens et du suivi d'usage de l'outil, suggèrent effectivement une adéquation entre les limites soulevées par l'état de l'art, nos hypothèses et les nouvelles pratiques documentaires des utilisateurs travaillant avec le cahier de laboratoire électronique.
Mots-clés. Théorie du document, authenticité, cahier de laboratoire, horodatage, signature électronique, valeur de la preuve, cas juridique.

Abstract. It is not easy to make sure about the authenticity of an electronic document. Most of what historians could check inside those old documents does not exist anymore in electronic documents. We believe that our tool could change every innovator's computer into an electronic notebook that could be accepted in front a court of law. First results were led both in a quantitative and qualitative way, by interviews, workshops, and use study. New documentary practices of electronic notebooks users may actually match our hypothesis and limits spotted in the state of the art.
Keywords. Document theory, authenticity, electronic notebook, trusted timestamping, digital signature, evidential value, case law.

1 Introduction

S'intéresser à la question de l'authenticité des documents numériques revient à s'interroger sur le statut des connaissances de notre époque. Si, à première vue, on pourrait croire que l'épistémologie s'occupe d'abord de vérité, ce serait négliger son rapport à l'authenticité.

Comme chacun sait, il n'y a pas de production scientifique sans références bibliographiques, mais c'est plus particulièrement dans les Sciences de l'Homme et de la Société que le recours au *document* prend toute son importance : le sociologue recueille des témoignages, le linguiste, des formes attestées, le philologue, des testimonia, l'historien, des sources. Tous, à la suite de Dilthey (et dans la tradition herméneutique initiée par Origène), montrent qu'il est possible de développer une pratique scientifique en prenant pour matériaux des discours d'hommes et de femmes, pourtant parfaitement subjectifs, à condition de les considérer non pas comme des descriptions exactes d'une réalité, mais justement comme des discours, sans prétention à la vérité, et dont seule l'authenticité doit nous importer (Bénel et Lejeune, 2009).

Pour autant, établir l'authenticité d'un document numérique ne va pas de soi. La plupart des indices que relevaient les historiens et les philologues à l'intérieur des documents anciens n'existent plus dans les documents créés de manière numérique (Iacovella, 2007). Pour le numérique, la preuve d'origine est externe. L'horodatage certifié et la signature numérique nécessitent un tiers de confiance et l'exécution d'un algorithme. Cette preuve, si elle est incomparablement plus sûre, est aussi beaucoup plus fragile sur le long terme.

La gouvernance de cette confiance envers les tiers, assez oligarchique, est de plus en plus remise en question. La complexité des algorithmes doit régulièrement être revue à la hausse pour faire face à l'augmentation des moyens de calcul. Par ailleurs, contrairement aux indices philologiques qui pouvaient être découverts a posteriori, ces preuves numériques d'origine doivent être mises en place au moment de la production des documents. Or, trop souvent incompris par les producteurs de documents, contraignants à utiliser et coûteux à mettre en place dans une organisation, ces dispositifs ne sont guère utilisés que là où ils représentent des « gains de productivité » (réponses à appel d'offres, factures, déclarations fiscales, bulletins de paye...). L'incapacité de tous nos autres « documents » numériques à devenir des documents au sens fort pourrait nous faire douter que les générations futures puissent écrire l'histoire de notre époque (Iacovella, 2007).

Face à cette question, un peu étourdissante, le sujet des cahiers de laboratoire numérique fournit un champ d'expérimentation, certes beaucoup plus modeste, mais cependant emblématique puisque combinant la constitution de documents authentiques et la pratique scientifique.

Nous établirons tout d'abord l'état de l'art de notre sujet dans les domaines juridique, académique et technique. Ensuite, parce que « tout peut devenir document », nous proposerons une première ébauche de ce qui permettrait de transformer l'ordinateur de n'importe quel innovateur en un cahier de laboratoire à valeur légale. Mais, parce qu'« on ne naît pas document, on le devient », nous présenterons nos hypothèses relatives à la construction de documents au sein d'un cahier de laboratoire numérique. Un document étant par ailleurs une référence partagée, nous présenterons comment, du travail personnel à la collaboration, ce partage peut être un atout pour un cahier de laboratoire. Enfin, ayant matérialisé

toutes ces hypothèses dans un logiciel, nous ferons part des premiers résultats de leur mise à l'épreuve auprès des usagers.

2 État de l'art

2.1 Juridique

Dans le domaine de la propriété intellectuelle, la jurisprudence nous apprend que la corroboration des déclarations de l'accusation ou de la défense peut être fournie autant par des témoins (Reese, 1981) que par des documents. Le principe, par contre, est que ces preuves ne puissent être construites par les parties, ce qui implique, en particulier, que les documents ne soient produits ni par eux ni par une personne travaillant sous leur direction (Mikus, 1976). À condition qu'ils respectent ce principe, les documents peuvent être extrêmement divers dans leur forme (cahiers de laboratoire de l'inventeur, lettres (Reese, 1981), e-mails (Utah, 2015), déclarations d'invention (Mikus, 1976), brouillons sur papier (Everyscape, 2013), fichiers informatiques (Kenexa, 2010)) et dans leur contenu (comptes rendus d'expérience (Reese, 1981), résultats de tests (Mikus, 1976), graphes (Mikus, 1976), photomicrographies (Cooper, 1998), code source (Kenexa, 2010), ou schémas (Goldenbridge, 2013)).

Il est communément admis qu'un cahier de laboratoire (sur papier) doit être constitué de « pages numérotées, cousues entre elles », remplies des « notes quotidiennes » d'une personne, de telle manière que les « entrées [soient] datées [...] les unes après les autres [...] sans laisser de pages blanches », et que les pages soient « tamponnées par un notaire ou relues et signées par un responsable de l'entreprise » (Haber, 1991). Néanmoins, les pratiques réelles sont bien loin de cet idéal et la jurisprudence nous apprend que les situations pouvant disqualifier les cahiers de laboratoire en tant que preuves sont nombreuses :

- des pages signées et datées par l'auteur mais non visées par un tiers (Mikus, 1976),
- des pages « ne suivant pas un ordre numérique dans la numérotation, de nombreuses pages étant même manquantes », même si le propriétaire du cahier expliquait qu'elles n'étaient pas liées au même sujet (Mikus, 1976),
- « le contenu d'une page se prolongeant sur une autre page bien après la page suivante » (Mikus, 1976),
- la trace selon laquelle l'auteur avait « d'abord écrit '1998' et avait réécrit la date en '1988' » (Aptix, 2001),
- la présence de « nombreuses pages blanches ne comportant rien d'autre qu'un large 'X', suggérant que [le tiers] avait signé un cahier vide et que [l'auteur] avait noté le texte et les diagrammes plus tard » (Aptix, 2001),
- des « signatures [des témoins] écrites avec la même encre, en dépit du délai supposé entre les deux de 22 jours » (Aptix, 2001),
- des « preuves scientifiques montrant que les entrées avaient été écrites avec une encre fabriquée pas avant 1994, soit cinq ans après leur date supposée » (Aptix, 2001).

Si certaines de ces anomalies relèvent clairement de la fraude, d'autres peuvent s'expliquer par de la simple négligence, voire même d'une difficulté compréhensible à concilier l'importance du travail réalisé et la lourdeur de la procédure à respecter. L'archivistique reconnaît en effet que :

> « Les dossiers les plus importants [...] sont les moins bien organisés et les moins bien classés. Créés à l'occasion et souvent en plusieurs exemplaires, remplis de papiers inutiles et de duplicata, et mis à part du fonds principal (parce que plus ou moins confidentiels), ces fichiers sont sans aucun ordre [...], alors que les registres et les fichiers correspondant aux procédures routinières sont généralement bien tenus (Pérotin, 1966). »

2.2 Académique

Depuis une quinzaine d'années, la dématérialisation des cahiers de laboratoire intéresse les chercheurs en interactions hommes-machines et en travail coopératif assisté par ordinateur. Ils procèdent d'ordinaire par des observations et des entretiens, auprès des usagers des cahiers de laboratoire sur papier. Par conséquent, les disciplines ciblées sont plutôt la biologie, la chimie, la physique : des disciplines expérimentales connues pour leur usage intensif de cahiers de laboratoire. Pour mettre à l'épreuve leurs hypothèses ou les affiner, ils introduisent, en situation réelle ou dans le cadre d'ateliers, des artefacts divers, tels que des scénarios (Klokmose et Zander, 2010 ; Tabard et al., 2008), des listes de besoins (Klokmose et Zander, 2010), des maquettes vidéos (Tabard et al., 2008), des sondes technologiques (Mackay et al., 2002), des logiciels génériques déjà utilisés sur place (comme un logiciel de note bien connu) mais paramétrés de manière nouvelle (Oleksik et al., 2014), ou des prototypes logiciels conçus et réalisés sur mesure (Schraefel et al., 2004 ; Tabard et al., 2008).

De ces observations et entretiens, on peut retenir en particulier la coexistence de plusieurs types de cahiers : des plus personnels aux plus collectifs, des plus libres (utilisés pour la création) aux plus « disciplinés » (utilisés pour l'archivage et des usages légaux) (Tabard et al., 2008). Même parmi les cahiers les plus formels, on retrouve : des cahiers pour rendre compte des expériences et des cahiers pour protéger les idées ; des cahiers attachés à une personne, et certains autres relatifs à un équipement ou un projet (Tabard et al., 2008). Ces logiques concurrentes d'agrégation de notes impliquent une sorte de duplication au sein de ces cahiers. Pour autant, alors que les contributeurs sont plutôt frileux à dupliquer des éléments uniquement pour les partager (Oleksik et al., 2014), cette duplication s'accompagne souvent d'une réécriture pour les nouveaux destinataires, sorte de « traduction » Latourienne, reconnue comme bénéfique par les acteurs (Tabard et al., 2008).

Sous l'angle de l'implémentation, les sondes technologiques se focalisent plutôt sur l'intégration du numérique et du papier (comme avec une tablette graphique et un PDA (Mackay et al., 2002)), alors que le paramétrage de progiciels et le prototypage se focalisent souvent sur l'aspect collectif de l'information (en particulier avec des mots-clés partagés (Oleksik et al., 2014) et des flux de notification (Tabard et al., 2008)).

Parallèlement, ces études ont mis en valeur un dilemme entre l'organisation chronologique et l'organisation basée sur les projets (Tabard et al., 2008), plus précisément, que l'organisation chronologique était plus utile les deux premières semaines après la création des notes, mais qu'ensuite l'organisation thématique était préférée (Oleksik et al., 2014). Les études ont également montré que l'usage de mots-clés définis librement sans aucune structure était difficile à gérer sur le long terme, lorsque leur nombre explosait, notamment lorsque les prototypes ne permettaient pas de combinaison de mots-clés lors de la recherche de documents (Oleksik et al., 2014).

2.3 Technique

Le problème se pose différemment pour les logiciels commerciaux. Comme le montre une étude très complète (Rubacha et al., 2011), la plupart des logiciels du marché sont des produits « tout-en-un » dans lesquels la rédaction s'effectue dans les champs d'un formulaire ou dans un traitement de texte intégré.

Un indice significatif est que la plupart de ces logiciels ont été conçus pour la chimie, la biologie, ou l'assurance et le contrôle qualité (Rubacha et al., 2011).

Autrement dit, ils sont destinés aux laboratoires, riches en techniciens et en étudiants, qui utilisent quotidiennement les cahiers sur papier pour renforcer le respect des procédures dans l'exécution des expériences.

Un « vrai » cahier de laboratoire numérique, conçu pour la propriété intellectuelle, répond à des enjeux différents. Si on se focalise sur l'invention, il faut en rendre compte dans toutes les étapes qu'elle traverse, notamment les plus informelles, « loin de la paillasse », les mêmes qui sont trop souvent oubliées dans les cahiers de laboratoire existants (Letondal et al., 2009).

Paradoxalement, la plupart de ces étapes laissent déjà des traces numériques (emails, présentations, notes, diagrammes, photos, code, données…). Plutôt que d'imposer un changement de formats, d'organisation ou de processus, il suffirait de munir les traces numériques déjà produites des propriétés leur conférant une valeur légale.

3 Hypothèses

Le cahier de laboratoire idéal qui se dessine « en creux » dans l'état de l'art est très loin d'une simple dématérialisation du cahier de laboratoire papier. Plus que de « singer » sa forme, il s'agit de donner ses propriétés aux fichiers créés quotidiennement par l'innovateur. Cette réinvention ne peut se faire selon nous sans un passage par la « théorie du document » (Briet et al., 1951 ; Buckland, 1997).

3.1 Tout peut être document

Un document est une « preuve » et tout peut devenir document (même une antilope !), du moment qu'il a été sélectionné, qu'on lui a attribué un identifiant et un emplacement dans un « corpus », qu'il a été décrit de manière à être référencé par d'autres documents et à rester accessible et préservé (Briet et al., 1951 ; Buckland, 1997).

Une description naïve du système que nous imaginons pour réinventer le cahier de laboratoire serait de dire qu'il transforme l'ordinateur de chaque inventeur en un cahier de laboratoire à valeur légale.

Transformer tout fichier informatique en un fichier qui ne dépende pas uniquement de soi, c'est à dire « prouver la revendication d'un inventeur d'être en connaissance d'une idée avant une date donnée », n'est désormais plus un problème technique (Haber et Stornetta, 1991). En effet, un protocole informatique peut envoyer une requête à une autorité d'horodatage (selon ses besoins techniques et organisationnels) pour générer un jeton d'horodatage. Ce jeton résulte de la signature électronique de la date courante et de l'empreinte du fichier. Associé au fichier, il permet une vérification. Tout comme des procédures papier bien connues (comme les cahiers de laboratoire signés par des témoins, des lettres envoyées à soi-même et gardées scellées, ou un dépôt chez un notaire), ce protocole peut être utilisé pour s'assurer que ni le contenu ni l'horodatage ne peuvent être altérés hormis par l'autorité d'horodatage (Haber et Stornetta, 1991).

Étonnement, même si de tels formats d'horodatage, de tels protocoles organisationnels ont été définis, implémentés, déployés, ils ont été utilisés principalement pour horodater des signatures électroniques (pour prévenir l'obsolescence des algorithmes). Ces standards n'ont pas autant servi pour l'horodatage de données ou de documents à un sens plus large.

En effet, la signature électronique est souvent considérée comme la seule façon de sceller des documents. Pourtant, elle peut s'avérer être un vrai cauchemar organisationnel au quotidien.

Nous avons donc apporté ces techniques d'horodatage directement sur le bureau (numérique) de nos utilisateurs :

- depuis n'importe quel fichier, le menu contextuel du système propose des actions reliées à l'horodatage,
- un dossier spécifique permet d'accéder à la liste complète des documents horodatés,
- à partir du menu contextuel du document ou d'un pseudo-dossier « Mes documents horodatés », l'horodatage d'un document peut être vérifié d'une manière pédagogique (Tosi et al., 2015).

3.2 On ne naît pas document, on le devient

Même si tout peut être document, cela nécessite un processus. Par exemple, une antilope en cage dans un zoo est un document pour un temps donné. Pour le rester une fois morte, elle devra être empaillée et préservée dans un musée (Briet et al., 1951). Tout comme sa préservation, la description d'un document devrait être considérée comme un processus à part entière. Même si des attributs standards, comme les dates, auteurs et destinataires doivent être renseignés tôt pour rendre toute manipulation plus facile, des attributs heuristiques (comme des tags) peuvent être attachés au fur et à mesure pour aider à l'interprétation des documents (Zacklad, 2006). Ainsi, plutôt que d'espérer créer un document parfait en une seule fois, le défit est plutôt de « documentariser » (Zacklad, 2006) les éléments pas à pas, alors qu'ils prennent de l'importance, afin d'obtenir une preuve de plus en plus solide. Dans beaucoup (trop) de systèmes de gestion de documents, la description demandée à la création d'un document est si contraignante que les utilisateurs préfèrent y inscrire de fausses données. Ils fuient les formulaires interminables aux champs obligatoires.

Cette construction progressive du document n'est finalement pas très différente de la « Théorie des trois âges » en archivistique (Pérotin, 1966). Selon Pérotin, la transformation des documents administratifs, d'une masse informe de dossiers, utiles mais banals, en des archives parfaitement classées, véritables trésors pour les chercheurs, semblerait tenir de la magie, si on n'évoquait pas la difficile tâche d'évaluation des documents par les archivistes. Un juste équilibre doit être trouvé entre accumuler des « déchets » et détruire des perles. L'auteur propose une analogie entre le processus d'archivage et la mémoire humaine : les armoires de classement en tant que subconscient, duquel on peut facilement la retrouver, les piles et les corridors en tant que royaume de l'inconscient, des souvenirs qui graduellement sont oubliés, mais qu'un psychanalyste peut faire émerger. De cette « mémoire administrative », certains éléments de valeur vont être intégrés dans une mémoire historique collective et définitive. Quelques années plus tard, Y. Pérotin utilise une autre métaphore : le processus d'archivage serait semblable à une séquence de bassins de décantation (Caya, 2004). Par exemple, les archives courantes sont conservées sur place pendant 1 à 4 ans, les archives intermédiaires de 20 à 100 ans, sans être consultables par le public pour autant, et enfin les archives

définitives restent dans les archives publiques pour un temps illimité. Dans chaque bassin, seule une petite partie des documents va surnager, selon leur usage administratif ou leur valeur historique, et vont se déverser dans le bassin suivant, et ainsi de suite.

De la même manière, l'objectif de notre système est de transformer graduellement un proto-document en un véritable document, au fur et à mesure de la reconnaissance de sa valeur (voir Figure 1) :

- la surveillance des brouillons en cours d'écriture et leur horodatage à chaque révision (pour obtenir la date la plus ancienne possible),
- l'archivage du contenu des révisions considérées comme importantes par l'auteur,
- le taguage des documents par équipe et projet,
- le nommage d'une invention regroupant plusieurs documents,
- la déclaration d'invention (incluant sa description, les révisions horodatées associées et les signatures des co-inventeurs et éventuellement de témoins).



Figure 1. *L'horodatage au coeur du cahier de laboratoire, pour chaque document*

Comme nous le voyons, à la différence de la simple dématérialisation du cahier de laboratoire papier, l'idée est de différer et de limiter la signature électronique par des humains, connue pour être particulièrement difficile à mettre en place, et de faire par contre un usage intensif de l'horodatage certifié (dont la signature électronique par une organisation tierce est déléguée à une machine), attestant de l'état dans lequel était un document à un moment donné. La signature électronique par des humains, quant à elle, est cantonnée à la déclaration d'invention, c'est-à-dire en fin de processus et uniquement pour les éléments les plus prometteurs.

3.3 Un document est une référence partagée

Un document est une « preuve à l'appui d'un fait » (Briet et al., 1951). Comme l'antilope conservée dans un zoo, qui reste accessible aux défenseurs et aux détracteurs des travaux scientifiques basés sur son existence, un document devrait rester accessible, sous sa forme authentique, à tous ceux qui y ont eu accès à un moment donné et qui ont pu le référencer dans leurs propres documents.

Cette nouvelle propriété du document permet de comprendre pourquoi il était un peu vain d'imaginer le cahier de laboratoire comme un journal personnel, voire intime, et de mieux comprendre pourquoi les entrées sont souvent dupliquées dans des cahiers plus collectifs. Plus encore, ce sont justement les contenus partagés qui

sont appelés à être cités par d'autres documents, à faire « référence ». Ce sont donc ces contenus qu'il est nécessaire de « documentariser ».

Notre idée consiste en quelque sorte à « intercepter » les documents partagés pour les déposer dans notre système et les horodater. Les organisations qui ont tenté d'imposer le dépôt de tous les documents partagés dans un système de gestion de contenu savent qu'il est difficile de lutter contre la facilité et les habitudes de partage de documents par courriel. Par conséquent, l'interface de partage de documents que nous proposons s'inspire du partage par courriel mais en le rendant encore plus facile (voir Figure 2).

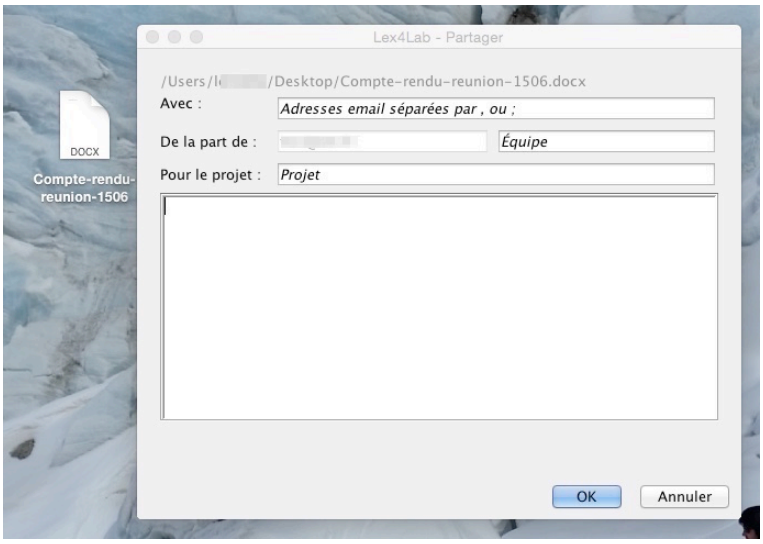


Figure 2. *L'interface de partage accessible en un clic droit sur le document.*

Par ailleurs, l'interface inspirée de celle des courriels nous fournit une occasion d'enrichir la description du document. D'ordinaire l'adresse électronique de l'émetteur et l'objet d'un courriel permettent d'éclaircir la situation de production et diffusion du document : « En tant qu'affilié à X, j'ai produit ce document et vous le transfère parce que vous faites partie du projet Y ». De même, nous proposons dans notre interface de partage d'indiquer l'équipe de la personne à l'origine du document et le projet dans le cadre duquel ce document est partagé.

Enfin cette fonctionnalité de partage alliée à celle de révision vue précédemment permet la co-écriture d'un document. Chaque révision étant associée à son auteur, il devient alors beaucoup plus facile de déterminer la paternité de chaque partie du document.

4 Mise à l'épreuve

4.1 Ateliers

En mars 2015, une fois l'installeur et les mise à jour automatiques implémentés, nous avons organisé deux ateliers, en conditions réelles, avec deux sessions de six utilisateurs cibles : chercheurs, ingénieurs de recherche et doctorants (Tosi et al., 2015).

Questionné à propos de son usage possible d'un cahier de laboratoire numérique, un doctorant en sociologie nous a dit : « L'intérêt d'un carnet de bord est de pouvoir revenir à une idée, la comparer à avant, qu'est-ce qui a évolué ? ». Un membre de cabinet de propriété intellectuelle a déclaré : « Le co-développement, la prestation de service… Il faut voir là où nait le caractère inventif d'un co-développement. Le cahier de labo s'avèrera utile pour déterminer la propriété industrielle ».

Une fois que les utilisateurs pensaient à leur propre discipline, à leurs propres problèmes, ils arrivaient finalement à saisir ce que toute cette technologie a priori complexe pourraient leur apporter.

Notre travail sur le logiciel a pris en compte, pour la conception, des nouvelles remarques des utilisateurs.

4.2 Utilisation sur le temps long (en interne)

Notre logiciel, nommé « Lex4Lab », a été conçu et implémenté en méthode Agile. Des fonctionnalités ont été régulièrement livrées d'avril 2013 à très récemment. Dès les premières livraisons, ses fonctionnalités principales étaient utilisables. Nous pouvions ainsi utiliser l'outil au quotidien, tester et valider les fonctionnalités avant qu'elles ne soient mises à disposition du public. Nous nous sommes vite sentis très concernés par les bugs ou des manques d'intuitivité, qui affectaient nos propres données, et non simplement des données de test. Nous nous sommes donc intéressés à l'historique d'un utilisateur depuis avril 2013 à mai 2016, à la fois sous un angle quantitatif et qualitatif. Cet utilisateur est un membre de l'équipe qui a utilisé Lex4Lab depuis le début. Nous précisons que les données suivantes ne concernent pas les données de test utilisées dans le logiciel. L'historique comporte une grande variété de formats (MS Word, PDF, PNG, Numbers, Graffle, Balsamiq Mockups, Markdown, JPEG). On remarque d'ailleurs une préférence (voir Figure 3) pour le format de type « présentation » (Keynote et Powerpoint). L'utilisateur l'explique par le fait que présenter une idée en réunion, en conférence ou en cours représente une prise de risque en termes de propriété intellectuelle, surtout envers des acteurs externes : entreprises, cabinets, collègues, chercheurs d'autres universités…

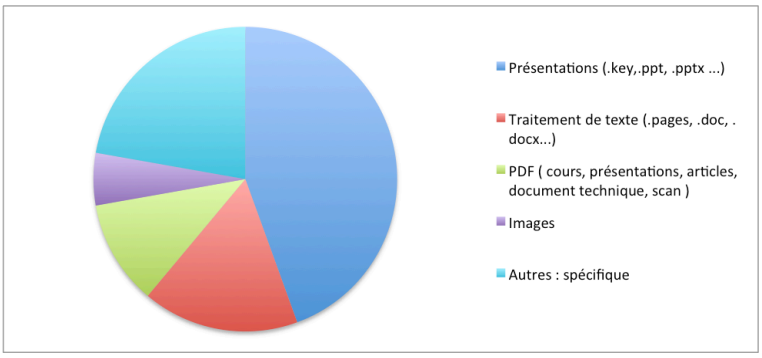


Figure 3. *Proportion par type de document dans l'historique utilisateur d'avril 2013 à mai 2016.*

Cet utilisateur déposait et partageait surtout des versions finales de documents. Il est à noter que certaines d'entre elles provenaient de brouillons qui se sont enrichis peu à peu à travers le cahier de laboratoire. C'est en parcourant l'historique de révisions que l'on peut se rendre compte que la dernière révision compte. Mais c'est l'historique entier, comportant les informations d'horodatage et de partage, qui a de l'importance pour le cahier de laboratoire, représentant la mémoire de ce qu'il s'est passé.

Les tags sont utilisés, comme prévu, pour aider à la recherche des documents de l'utilisateur par filtres (voir Figure 4). Presque tous les documents sont tagués, la majorité avec des tags « Projets ». Nous pouvons ainsi adhérer à notre hypothèse selon laquelle les documents les plus sensibles en terme de propriété intellectuelle sont également les plus partagés et les plus décrits dans le cahier de laboratoire. Même si ce cahier de laboratoire numérique ne serait pas utilisé dans un cas de litige, l'utilisateur ne peut pas savoir si un jour quelque chose pourra être utile pour protéger ses intérêts, et encore moins savoir quel document sera concerné. Il peut uniquement compter sur le fait que un de ses documents, partagé, taggué, révisé, le tout en collaboration entre des acteurs d'un projet (voir Figure 5), pourra vraiment aider en cas de litige, constituant un dossier de preuve. De façon générale, pour cet utilisateur comme pour d'autres, nous avons observé une transition, entre la première année (ou l'outil était surtout utilisé pour le projet en cours et tester le logiciel) et le moment où l'outil a commencé à être utilisé dans d'autre projets d'enseignement et de recherche.

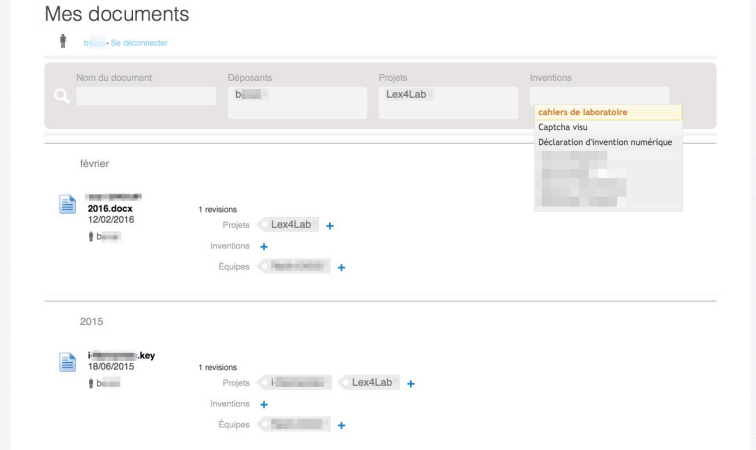


Figure 4. *Liste des documents horodatés filtrés grâce aux tags*

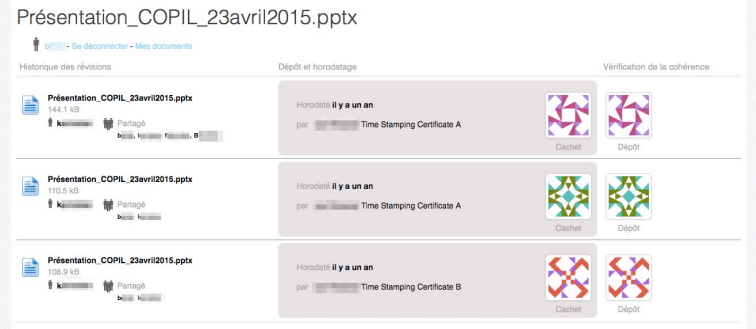


Figure 5. *Historique d'un document partagé et révisé entre plusieurs acteurs*

Nous continuons à guetter les résultats de notre étude sur le temps long, menée dans une université avec davantage de chercheurs de toutes disciplines.

4.3 Des (futurs) utilisateurs enthousiastes

Nous souhaiterions à présent tester notre logiciel sur la durée d'un projet en entier. Pour commencer, nous avons identifié quinze utilisateurs potentiels, et mené des entretiens semi-directifs de novembre 2015 à juin 2016 dans une université française.

Vis à vis de la propriété intellectuelle, la responsabilité juridique ressortait beaucoup des témoignages, surtout en ce qui concernait la réplication des données. Plusieurs des chercheurs interrogés ont déjà utilisé un cahier de laboratoire précédemment. « Regardez, j'ai gardé tous mes cahiers depuis mon premier labo [...] Au pire on fait des photocopies [...] On peut me prendre mon disque dur, ils n'ont pas le cahier pour comprendre tout mon travail » a expliqué un chercheur en nanotechnologies. Un collègue évoquait la complexité de faire signer un témoin pour ce genre de cahier de laboratoire : « plus qu'un cahier de labo, on devrait dire cahier de chercheur ». Selon la plupart des autres interviewés, un cahier de

laboratoire, s'il est utilisé dans une structure, devrait représenter le coeur de leur travail. Un chercheur en mécanique des solides a raconté qu' « il est arrivé deux ou trois fois qu'un reviewer m'appelle pour m'avertir à propos d'un papier déposé par un ancien collègue ou un doctorant qui utilisait directement des données de travaux réalisés pendant son doctorat ». Cependant, aucun des chercheurs ne voudrait abandonner une collaboration à cause de simples risques de propriété intellectuelle. En France, ils sont en effet beaucoup évalués sur leurs partenariats et le nombre de publications.

Nous attendons d'autres utilisateurs pour compléter notre étude, afin de mieux comprendre leurs habitudes de collaboration, et leurs besoins en terme de protection de la propriété intellectuelle. Le suivi se manifestera par un entretien d'auto-confrontation (Theureau, 2010) sur l'historique de chacun.

4.4 Au niveau légal

Un avocat a écrit un avis juridique sur le logiciel. Les fichiers émanant de Lex4Lab pourraient être reconnus comme preuve électronique dans un litige, selon la loi française. Il va bientôt compléter ses déclarations pour les ajuster au règlement européen eIDAS.

5 Conclusion et perspectives

Au cours de cet article, nous avons montré qu'il était possible de transposer à un environnement numérique les propriétés du cahier de laboratoire sur papier sans pour autant mettre en danger la flexibilité nécessaire au travail scientifique quotidien. Les documents au sens fort ainsi produits, munis d'éléments de preuve cryptographiques et de contexte particulièrement riches, satisfont à la double injonction d'être conservés sous leur forme authentique et de rester accessibles à leurs auteurs et destinataires.

Nous ne cherchons pas uniquement à fournir un homologue numérique au cahier de laboratoire, mais vraiment à faire hériter les innovateurs des bonnes pratiques d'un cahier de laboratoire, même s'ils n'en n'ont jamais utilisé auparavant. D'abord, une étude de l'état de l'art sous plusieurs angles nous a permis de déterminer les principaux critères utilisés pour prouver l'antériorité de conception d'un document, ainsi que les limites des travaux réalisés autour des cahiers de laboratoire. L'étude des éléments juridiques relatifs aux cahiers de laboratoire suggère également que la construction des documents doit englober la question de l'authenticité, et ce dès la création. L'état de l'art des cahiers de laboratoire, notamment numériques, met en évidence un manque dans les solutions proposées. Cette lacune se situe bien au carrefour entre un cahier de laboratoire numérique et un cahier de laboratoire collaboratif. Pourtant, aujourd'hui, les éléments techniques relatifs à de telles fonctionnalités ne constituent plus une barrière pour en assurer la faisabilité et la pérennité.

L'idée de proposer l'horodatage cryptographique aux innovateurs est basée sur trois hypothèses tirées de la théorie du document, par lesquelles notre logiciel se distingue des cahiers de laboratoire existants.

Plutôt que d'obliger à remplir des formulaires ou des formats de dépôt précis, car tout peut être document, nous proposons l'horodatage pour tout format de fichier choisi par les innovateurs. De plus, nous avons préféré proposer les fonctionnalités comme une extension de leur espace numérique de gestion de document favori (bureau, explorateur de fichiers), plutôt que de les envoyer sur une autre plateforme (web par exemple).

Plutôt que de chercher à créer un document parfait en une fois (avec les bonnes métadonnées et des signatures), et parce qu'un document se construit au fil d'un processus, nous offrons la possibilité de passer par différentes transitions, tout comme par différents bassins de décantation : de brouillons, les révisions se transforment peu à peu en documents les plus complets et les plus renseignés.

Plutôt que de concevoir le cahier de laboratoire comme un coffre fort, car tout document est une référence partagée, nous préférons suivre la construction et le partage des documents, au fil de la collaboration, en permettant la co-écriture et la traçabilité.

Nous imaginons observer des phénomènes similaires dans une étude commencée à l'université, et, nous l'espérons, auprès d'un public plus large, l'outil étant distribué par un éditeur de logiciel. Les premiers résultats ont été récoltés de manière quantitative et qualitative, par des ateliers, des entretiens et du suivi d'usage de l'outil. Ils suggèrent effectivement une adéquation entre les limites soulevées par l'état de l'art, nos hypothèses et les nouvelles pratiques documentaires des utilisateurs travaillant avec le cahier de laboratoire électronique. Nous espérons surtout que ce type de cahier de laboratoire électronique pourra aider les utilisateurs à se protéger et à éviter les cas de litiges, tout en gardant trace de documents de qualité.

Bibliographie

Aptix Corporation v. Quickturn Design Sys. (2001). Court of Appeals, Federal Circuit, No. 00-1468, 2001.

Briet S., Day R.E., Martinet L., and Anghelescu H.G.B. (2006). *What is documentation? : English translation of the classic French text.* Scarecrow Press, Lanham, Md. Disponible à : http://ella.slis.indiana.edu/~roday/briet.htm

Bénel A., Lejeune C. (2009), Partager des corpus et leurs analyses à l'heure du Web 2.0. In *Degrés, revue de synthèse à orientation sémiologique*, 136-137, section "m".

Buckland M.K. (1997). What Is a 'Document' ? JASIS , 48, 9, 1997, 804–809. DOI: http://dx.doi.org/10.1002/(SICI)1097-4571(199709)48:9<804::AID-ASI5>3.0.CO;2-V

Caya M. (2004). La théorie des trois âges en archivistique. En avons-nous toujours besoin? In *ELEC, Conférence de l'Ecole des chartes*, Vol. 2.

Everyscape, Inc. v. Adobe Systems Inc. (2013). Dist. Court, D. Massachusetts, Civil Action No. 10-11597-RGS, 2013.

Golden Bridge Technology Inc. v. Apple Inc. (2013). United States District Court, D. Delaware, Civ. No. 10-428- SLR, 2013.

Haber S. and Stornetta W.S. (1991). How to Time-Stamp a Digital Document. J. *Cryptology* 3, 2, 1991, 99–111. DOI: http://dx.doi.org/10.1007/BF00196791

Iacovella A. (2007), *Éléments pour une historiographie digitale.* Brossaud, Claire ; Reber Bernard (Éds.) Humanités numériques 1 – Nouvelles technologies cognitives et épistémologie. Hermes-Lavoisier.

Kenexa Brassing Inc. v. Taleo Corporation. (2010). United States District Court, D. Delaware, Civ. No. 07- 521-SLR, 2010.

Klokmose C.N. and Zander P. (2010). Rethinking laboratory notebooks. In *Proceedings of the 9th international conference on the design of cooperative systems,* COOP 2010, 119–140.

Letondal C., Tabard A., and Mackay W.E. (2009). *Quand tu l'as écrit, tu l'as lu !* Rapport de recherche 1526. Laboratoire de Recherche en Informatique (LRI). Disponible à : https://www.lri.fr/~bibli/Rapports-internes/2009/RR1526.pdf

Mackay W.E., Pothier G., Letondal C., Bøegh K., and Sørensen H.E. (2002). The missing link: augmenting biology laboratory notebooks. In *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology*, Paris, France, October 27-30, 2002, Michel Beaudouin-Lafon (Éd.). ACM, 41–50.

Mikus and Shaffer v. Anselm Wachtel (1976). United States Court of Customs and Patent Appeals, Patent Appeal No. 76-594, 1976.

Oleksik G., Milic-Frayling N., and Jones R. (2014). Study of electronic lab notebook design and practices that emerged in a collaborative scientific environment. In *Computer Supported Cooperative Work*, CSCW'14, Baltimore, MD, USA, February 15-19, 2014, Susan R. Fussell, Wayne G. Lutters, Meredith Ringel Morris, and Madhu Reddy (Eds.). ACM, 120–133. DOI: http://dx.doi.org/10.1145/2531602.2531709

Pérotin Y. (1966). Administration and the 'Three Ages' of Archives. *The American Archivist,* 29, 3 (1966), 363–369.

Reese and Katz v. Hurst and Crouse v. Wiewiorowski and Miller (1981). United States Court of Customs and Patent Appeals, Appeal Nos. 81-525 to 81-527 1981.

Rubacha M., Rattan A.K., and Hosselet S.C. (2011). A review of electronic laboratory notebooks available in the market today. *Journal of the Association for Laboratory Automation*, 16, 1 (2011), 90–98.

Schraefel M.C., Hughes G.V., Mills H.R., Smith G., Payne T.R., and Frey J. (2004). Breaking the Book: Translating the Chemistry Lab Book into a Pervasive Computing Lab Environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, ACM, New York, NY, USA, 25–32.

Tabard A., Mackay W.E., and Eastmond E. (2008). From individual to collaborative: the evolution of prism, a hybrid laboratory notebook. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, CSCW 2008, San Diego, CA, USA, November 8-12, 2008, Bo Begole and David W. McDonald (Eds.). ACM, 569–578. DOI: http://dx.doi.org/10.1145/1460563.1460653

Theureau J. (2010). Les entretiens d'autoconfrontation et de remise en situation parles traces matérielles et le programme de recherche 'cours d'action'. *Revue d'anthropologie des connaissances* 4, 2, 2010, 287–322.

Tosi L., Bénel A., Devoret F. (2015). Améliorer l'expérience utilisateur pour encourager la confiance entre collaborateurs. Rencontres doctorales et poster*, IHM 2015,* Toulouse, octobre.

University of Utah v. Max-Planck-Gesellschaft Zur Fo'Erderungderwissenschaften e.V. (2015). United States District Court, D. Massachusetts, Civil Action No. 11-10484-PBS, 2015.

Zacklad M. (2006). Documentarisation processes in Documents for Action (DofA): the status of annotations and associated cooperation technologies. *Computer Supported Cooperative Work*, CSCW15, 2-3, 2006, 205–228.