Ingénierie des traces numériques d'interaction comme inscriptions de connaissances

Julien Laflaquière[2, 1], Yannick Prié[1], Alain Mille[1]

[1] LIRIS -UMR 5205- Bât Nautibus - UFR Informatique
Université Claude Bernard Lyon 1 / F-69622 Villeurbanne,
{yprie,amille}@liris.cnrs.fr
[2] ICD/Tech-CICO FRE CNRS 2848, Université de Technologie de Troyes
12 rue Marie Curie, 10000 Troyes,
Soutenu par le Conseil Régional de Champagne-Ardenne et le FSE
julien.laflaquiere@utt.fr

Résumé : Cet article s'intéresse aux traces numériques d'interaction, et en particulier à ce que nous définissons comme les traces modélisées. Contrairement à une majorité d'approches se basant sur les traces, nous considérons trace et modèle de trace explicite comme indissociables. Cet article aborde deux questions en particulier : la première, théorique, concernant le statut d'inscription de connaissances des traces modélisées ; la seconde, méthodologique, concernant une démarche de modélisation de trace que nous considérons comme une « ingénierie de la trace ». Nous y considérons l'exploitation des traces dans la dynamique d'une activité toujours susceptible d'évoluer et à laquelle le modèle de trace doit rester adapté.

Mots-clés : Trace d'interaction, Inscription de connaissance, Modèle de trace, Modélisation, Ingénierie des connaissances.

1 Introduction

Considérer le caractère dynamique des connaissances est un enjeu essentiel en ingénierie des connaissances. En effet, quelle que soit la qualité de la modélisation et du processus de recueil de connaissances, les inscriptions de connaissance qui en résultent seront mobilisées dans des contextes jamais complètement prévisibles et alors que ces inscriptions rendront compte d'une réalité qui aura évolué par elle-même. Pour tenter d'apporter des éléments de réponse à ce défi de la dynamicité et de la plasticité de la connaissance, l'équipe Silex du LIRIS mène depuis plusieurs années des travaux sur les traces numériques d'interaction avec l'idée de les utiliser comme support à l'inscription de *l'expérience d'utilisation* d'environnements numériques (Mille *et al*, 2006a). Ces travaux nous ont conduits à développer les concepts de *trace d'utilisation* (modèle *Musette* en 2002), puis de *trace modélisée* (notion de *Système à Base de Traces* en 2006), dont il sera question dans cet article. Ces concepts sont le fruit d'une généralisation et d'une formalisation générique de ce que sont les traces numériques d'interaction, ce que les travaux de la littérature ne se sont pas proposés de faire jusqu'ici (section 2). Au-delà des applications que nous avons pu mettre en

place, nous souhaitons dans cet article, aborder deux questions : (a) celle du *statut* des traces (d'interactions) modélisées, sur le plan théorique (Section 3), et (b) celle de la *modélisation* de ces traces en termes cette fois méthodologiques (section 4). Nous verrons à travers les propositions que nous formulerons, combien nos travaux s'inscrivent dans et s'inspirent du domaine de l'Ingénierie des Connaissances, ce que souligneront également la discussion (section 5) et la conclusion de cet article (section 6). Précisons d'entrée de jeu que le traitement de tous les aspects nécessaires à une description complète de notre approche ne se prête pas au format de cet article. Aussi serons-nous amenés, au cours de nos explications à renvoyer le lecteur vers d'autres articles, et à synthétiser parfois un peu radicalement notre exposé.

2 Trace d'interaction, trace d'utilisation, trace modélisée

La possibilité théorique qu'offre le numérique d'enregistrer toutes les interactions entre un utilisateur et son environnement, a depuis longtemps, et abondamment, suscité l'intérêt des chercheurs. Ces enregistrements, que l'on peut appeler de manière générale *trace numérique d'interaction*, ont été utilisés dans quasiment tous les domaines de recherche concernés par l'exploitation d'environnement numérique.

2.1 Les traces d'interaction dans la littérature

Le nombre et surtout la diversité des finalités attribuées aux traces d'interaction dans les approches en question sont un obstacle à toute tentative de classification claire. Pour n'en citer que quelques-uns : analyse d'utilisabilité, d'utilité ou de l'ergonomie des interfaces (IHM), modélisation utilisateur, caractérisation des comportements d'opérateurs (Psychologie Cognitive), caractérisation des parcours sur le *Web* (Sciences de l'information), assistance à l'utilisateur, personnalisation d'interface (Informatique), création et validation de scenarii pédagogiques, ou encore réflexivité des activités d'apprentissage (EIAH), *etc*. Tout juste peut-on distinguer, avec A. Wexelblat (1998), les approches où les traces d'interaction sont exploitées *en dehors* de l'activité observée (dans une posture « d'analyse »), de celles où les traces le sont *au sein même* de ladite activité (dans une posture de « réflexivité »). Dans ce second cas, la trace de l'interaction est rendue disponible d'une façon ou d'une autre à l'utilisateur de l'environnement, qui est lui-même à l'origine de l'interaction en question[1].

Si l'on trouve trace des traces d'interactions dans de nombreux domaines, certains constituent des pôles de recherche particulièrement actifs. C'est le cas des recherches liées au *Web* (tant en termes d'analyse que de réflexivité) que nous prendrons comme illustration ici. Du côté de l'analyse d'abord, on retrouve par exemple de manière commune chez tous les hébergeurs de sites (ou de *blogs*) des fonctionnalités de traçage de l'activité des visiteurs à l'image de ce que fournit *Google Analytics*[2]. Mais

[1] Les deux rôles pouvant être endossés par une seule et même personne.

[2] http://www.google.com/analytics/

on trouve également sur le sujet des travaux exploratoires visant à caractériser les parcours des internautes sur le Web (Beauvisage, 2004), et plus généralement les travaux dits de *Web Mining* ou *Web Usage Mining* (Brinkman *et al*, 2006), qui se basent sur des analyses statistiques des *logs-files*[3]. Du côté de la réflexivité cette fois, on trouve de plus en plus de fonctionnalités additionnelles de navigateurs (Firefox le plus souvent) permettant à un internaute de tracer sa navigation un peu à l'image de *Google Web History*. Des recherches plus poussées encore sont également menées avec des technologies similaires, à l'image du *NavTracer*[4], ou en y ajoutant des couches logicielles supplémentaires afin de fournir des *assistants* à la navigation, c'est le cas *Letizia* (Lieberman, 2001), ou d'étudier la *visualisation* de parcours ou de données temporelles (Magnusson *et al*, 2000).

Ces différents travaux exploitent le fait que des traces d'interactions « existent déjà », si l'on peut dire, à travers le fonctionnement des navigateurs. Dans le domaine des EIAH (Environnements Informatiques d'Apprentissage Humain), autre pôle très actif, on retrouve pareillement des démarches d'analyse des interactions (Plaisant *et al*, 1999), mais également (et de manière plus sophistiquée) des tentatives d'assistance à l'apprenant s'appuyant sur ses propres traces d'interaction (Ollagnier, 2006). Le trait particulier de certains de ces travaux est de ne pas se baser uniquement sur des *logs* existants comme précédemment, mais de parfois créer de toutes pièces un environnement dans lequel les interactions ont été déterminées à l'avance (Soller *et al*, 2005) délimitant ainsi de fait le champ d'observation pour les traces.

Bien que dans certains cas les traces d'interactions semblent faire l'objet d'une définition précise, elles ne font pas l'objet d'une *modélisation explicite* en tant que telle. Quel que soit le domaine ou le niveau d'abstraction auquel sont exploitées les traces d'interaction, les approches restent *ad hoc*. Les traces elles-mêmes ne bénéficient pas (ou peu) dans ces travaux d'une réflexion théorique, ni même d'une tentative de généralisation, s'effaçant souvent en tant qu'objet de recherche derrière les objets qu'elles aident à caractériser ou à découvrir.

2.2 Une autre approche des traces numériques

Les travaux menés par notre équipe au LIRIS se sont justement distingués en tenant dès le départ de prendre le contre-pied de cette tendance, en tenant de considérer les traces d'interaction comme importantes à modéliser pour les considérer *d'emblée* comme des inscriptions de connaissances issues de l'expérience. Nous avons commencé par considérer les traces numériques d'interaction, comme le support potentiel d'une *expérience d'utilisation* (Mille, 2006b). L'objectif était alors de donner à un utilisateur la possibilité d'accéder, à travers une trace d'utilisation à la *singularité* d'une telle expérience. Nous avons alors généralisé (approche Musette[5]), et formalisé (approche des Systèmes à Base de Traces) notre approche pour travailler aujourd'hui sur la notion de *trace modélisée*, qui est définie comme *l'association*

[3] En informatique, fichier récoltant l'historique de certains évènements machine.

[4] http://navtracer.mozdev.org/

[5] Détails disponibles dans d'autres contributions (Champin *et al*, 2004) par exemple.

d'une collection d'observés temporellement situés structurée par leurs relations et d'un modèle explicite de cette collection. Dans cette définition un *observé* est entendu comme tout objet informatique décrivant un élément issu de l'observation de l'activité d'utilisation d'un environnement informatique, une *relation* entre observés permet d'exprimer des liens entre observés autres que temporels, et le modèle de trace est quant à lui le « vocabulaire » de la trace permettant la compréhension et la manipulation de la trace en en décrivant abstraitement les éléments, c'est-à-dire en typant les observés et les relations.

Nous aurons l'occasion, dans la suite de cet article à revenir sur ces concepts[6]. Cette formalisation a elle-même rendu possible la création des *Systèmes à Base de Traces* (SBT) qui sert de cadre conceptuel pour penser tout système informatique exploitant des traces modélisées, et permet d'encadrer l'implémentation des phases nécessaires à la gestion de traces (collecte, transformation, et visualisation de traces) (Cram *et al*, 2007). Le principe général des SBT est illustré ci-dessous (figure 1).

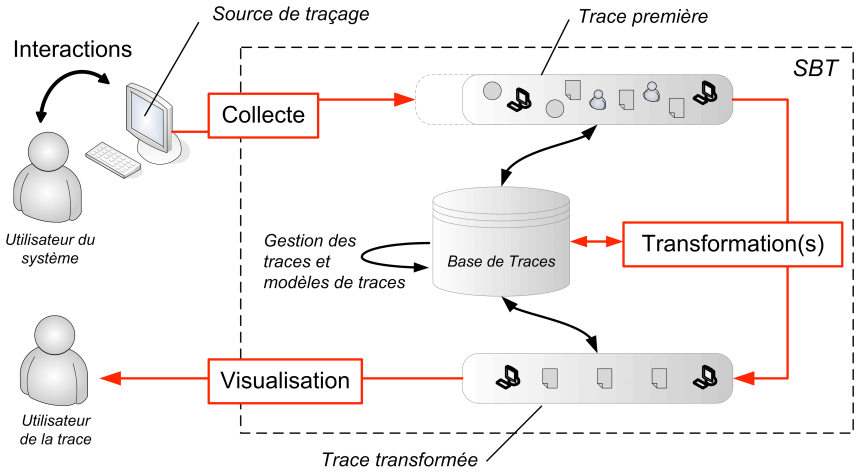


Figure 1 : Principe d'un Système à Base de Traces

Ce principe est ou a été mis en œuvre dans plusieurs domaines d'applications : appropriation d'EIAH (Ollagnier, 2006), plateforme de elearning (eLycée, projet ANR Ithaca), conduite automobile (Georgeon *& al*, 2006), Gestion des connaissances (projet ANR Procogec), *etc*. Le *framework* générique Open Source ATER (Atelier pour la gestion de Traces, leurs Exploitations et Représentations) permet de gérer modèles de traces (OWL), traces (graphes RDF) ainsi que transformations, utilisateurs, *etc*.

Les grandes lignes de notre approche ayant été posées comme préalable, nous pouvons aborder la première des deux questions que cet article se propose d'aborder, celle du *statut* attribuable aux *traces modélisées* du point de vue du domaine de l'ingénierie des connaissances.

[6] Le lecteur intéressé pourra trouver des descriptions complètes dans (Settouti *et al*, 2006).

3 Statut d'inscription de connaissance pour les traces modélisées

« Une connaissance est la capacité d'exercer une action pour atteindre un but » (Bachimont, 2004a). Cette capacité n'étant par nature ni formalisable, ni réifiable dans quelque modèle que ce soit, B. Bachimont (Bachimont, 2004b) fait habilement remarquer que l'Ingénierie des Connaissances porte bien mal son nom puisque ce ne sont pas des *connaissances* en tant que telles qu'elle manipule, mais des *inscriptions de connaissances* qui permettent elles-mêmes « l'expression, la transmission et l'appropriation d'un contenu ». Pour cela, une inscription de connaissance doit être *mobilisable*, et se prêter à *l'interprétation critique*, afin de (re)construire la connaissance dont elle est l'inscription. Nous défendrons ici l'idée est que les *traces modélisées* peuvent prétendre au statut d'inscription de connaissance.

Considérons le cheminement suivant. Prenons une trace générée[7] par l'activité effective d'un utilisateur exploitant son environnement numérique, pendant une période de temps et une activité données et en respect d'un modèle explicite. Cette trace est composée d'observés (éléments mobilisés dans l'interaction, des entités manipulées ou bien des évènements, ou actions, réalisés) liés entre eux. Ce lien, cette relation est par construction au minimum temporelle[8], mais le modèle de trace peut spécifier/décrire d'autres relations correspondant à ce que l'activité elle-même créée en son sein comme liens entre les objets qu'elle manipule[9].

Il en découle que la trace générée est par construction une *organisation*, une *structuration singulière* d'un ensemble d'observés. Singulière parce que cette structuration a pour logique celle du déploiement de l'activité dans l'environnement numérique, sur une période de temps cohérente par rapport à une activité donnée (durée d'une session d'utilisation de la machine, ou durée d'un projet par exemple). Dans la mesure où les observés ont été déterminés pour être pertinent relativement à l'activité observée, cette trace structurée est *mobilisable* par un observateur comme une description structurée de l'activité du point de vue de l'interaction (ce que seul un environnement numérique peut proposer). De plus, cette trace n'est pas uniquement une description de l'activité pour elle-même. Elle est également porteuse de sens pour les observés dont elle se compose. Elle leur offre un *contexte d'interprétation* en leur donnant place dans une représentation de leur utilisation dans l'activité. L'effet de cette contextualisation est particulièrement sensible quand les observés sont eux-mêmes des inscriptions de connaissances (Bachimont, 2004b). Leur *contextualisation* modifie les conditions d'interprétation dédites inscriptions, et donc les possibilités de (re)construire des connaissances[10].

---

[7] Générée au sens d'enregistrement automatique des observés durant l'utilisation de l'environnement.

[8] De part la séquentialité des interactions avec un environnement numérique.

[9] Exemple : dans un environnement donné, les « fichiers de texte » et les « destinataires de mails » sont des observés. L'utilisation du client mail pour envoyer le fichier A à un destinataire B permet d'instancier une relation entre les observés A et B, relation exploitable plus tard dans l'utilisation de l'environnement. Pour accéder par exemple à des informations du type « liste personnes ayant travaillé sur ce fichier » sans avoir à passer par une procédure de recherche exploratoire.

[10] Elle est porteuse de sens comme toute représentation structurée d'inscriptions de connaissances peut l'être (une classification, une lecture) à la seule différence près que cette structuration n'est pas volontaire, déclarative, mais induite par l'utilisation des inscriptions à travers celle de l'environnement numérique.

3.1 Le rôle du modèle de trace

On vient de dire que des traces générées par une activité sont *mobilisables* pour (re)construire des connaissances relative à l'activité elle-même mais également à ses composants (observés) – connaissances permettant de ré-agir dans la conduite de l'activité, dans sa compréhension pour un observateur, ou dans le rapport qu'entretien un utilisateur aux objets manipulés. Le tout n'est pas de comprendre l'intérêt d'une trace mobilisable. La trace modélisée doit se plier à *l'interprétation critique* de celui ou celle qui tentera de l'exploiter, et ce qui garantit cette possibilité c'est le *modèle de trace explicite* qui accompagne nécessairement une *trace modélisée*. Non seulement il permet de donner des *clefs de l'interprétation* d'une trace générée singulière, mais mobilisable lui-même, ce modèle est capable d'être modifié, indépendamment du système observé, pour s'adapter au regard *critique* d'un observateur. L'association de la trace et de son *modèle explicite* garantit la possibilité de l'interprétation critique d'une trace générée singulière.

Pour insister sur l'importance du modèle de trace, et notamment de son caractère explicite, dans la notion de trace modélisée, répondons à la question suivante : pourquoi tout enregistrement d'interaction ne peut-il être considéré comme une inscription de connaissance[11] ? Prenons l'exemple un peu trivial, mais clair, de *l'historique des navigateurs Web*. Il s'agit d'une succession temporelle des pages visitées (avec parfois des fonctionnalités de tri), sensée représenter la navigation d'un internaute. Cet historique, à première vue, correspond à la définition que nous avons donnée des traces modélisées. S'agissant du modèle de trace, nous pouvons en effet postuler qu'il existe et que les « pages visitées » sont des observés mis en relation de succession temporelle. Le modèle de trace dont il est question ici est *implicite*, trivialement lié au modèle de conception du navigateur (dont le but est l'affichage de pages *Web*). Pourquoi alors « l'historique en tant que trace » ne pourrait-il pas prétendre au statut d'inscription de connaissance ?

Premièrement, les conditions d'interprétation ne sont réunies que parce que le modèle implicite peut être « déduit » par l'utilisateur. Si le modèle de conception empêchait (c'est un exemple arbitraire) que les pages dont l'URL contient la lettre « a » apparaissent dans la liste des pages visitées, l'interprétation de l'historique en serait rendue pour le moins difficile. Deuxièmement, pages et URL sont au cœur du modèle de conception d'un navigateur, et sont donc traçables facilement. À y regarder de plus près cependant, l'historique ne correspond pas à une trace de *l'activité de navigation* elle-même, mais au fonctionnement du navigateur lui-même. La navigation hypertextuelle est une activité complexe (Bachimont, 1999), une « lecture active », dont il n'est pas possible de rendre compte, pour un utilisateur, sur la simple mise en relation temporelle des pages visitées. Les historiques des navigateurs Web semblent d'ailleurs « sous-utilisés » à en croire par exemple les études qui les compare à la fonction « retour-en-arrière » (Tauscher & Greenberg, 1997).

[11] La question est justifiée puisque dans un environnement numérique, si toute interaction est théoriquement enregistrable, c'est qu'il existe au minimum un modèle sous-jacent (de conception) qui crée les conditions d'existence de cette interaction, et de l'enregistrement lui-même.

Pour résumer le raisonnement précédent, une *trace modélisée* est inscription de connaissance parce que le modèle de trace est à même de garantir son interprétation critique et par là même la (re)construction de connaissances relatives (a) au déploiement structuré de l'activité dans le temps et/ou (b) à l'organisation des observés entre eux au sein de la trace. En s'ouvrant à l'interprétation d'un observateur, les traces modélisées ne sont pas seulement une représentation de l'activité pour elle-même, elles deviennent potentiellement une ressource de cette activité au sein de l'environnement observé. Pour que ce raisonnement reste cohérent, il faut considérer que la mobilisation et l'interprétation critique d'une trace ne peuvent être menées concrètement que si celles-ci sont *instrumentées*, c'est-à-dire dans la mesure où on se donne les moyens *d'exploiter la trace* en rendant possible *transformations*, *visualisations* et *manipulations* de la trace en question[12].

3.2 Finalité des inscriptions de connaissances tracées

Il nous reste pour clore cette section à évoquer, la finalité des traces modélisées, i.e. les possibilités d'exploiter les connaissances dont elles sont l'inscription. De manière très générale nous avons envisagé dès le départ une finalité de *facilitation* de l'activité médiée (Mille, 2006). Cette facilitation prend deux formes principales selon qu'une trace modélisée vise à *enrichir les possibilités d'interaction* entre l'utilisateur et les objets qu'il manipule dans son activité (facilitation directe), ou que la trace modélisée, représentation de l'activité pour un utilisateur, lui donne des prises cognitives[13] sur celle-ci (facilitation indirecte).

Nous retiendrons de cette troisième partie, que non seulement les traces modélisées sont assimilables à des inscriptions de connaissances mais qu'elles peuvent également constituer une instrumentation facilitatrice de l'activité médiée. Ces assertions reposent essentiellement sur le rôle pivot du *modèle de trace*, dont nous allons maintenant décrire la démarche de modélisation, seconde question abordée par cet article.

4 Une ingénierie de la trace

Dans la partie précédente, nous avons ancré notre vision théorique des traces dans le champ de l'Ingénierie des Connaissances. Notre démarche de modélisation de trace s'y inscrit également sur un plan *méthodologique*. En effet, si cette démarche de modélisation possède naturellement des spécificités dues à son objet, elle se fonde principalement sur un processus similaire à la modélisation de connaissances au sens classique, lui empruntant outils et techniques.

---

[12] Ce que se propose d'encadrer l'architecture des SBT (c.f. section 2).

[13] Processus métacognitifs, prise de recul nécessaire à une analyse (même sommaire) de sa propre activité, prise de conscience de savoir-faire ou de stratégies d'ordinaire occultées, et ce en s'appuyant sur des *expériences toujours singulières* que la trace modélisée permet de réutiliser (Laflaquière, 2005).

4.1 Un processus de modélisation dynamique

Notre idée de la modélisation de trace est qu'il s'agit d'un processus dynamique, et que puisque les concepts et l'architecture de notre approche SBT nous le permettent, il est nécessaire de mettre en place un cadre de modélisation qui tienne compte du déploiement, puis de l'exploitation des traces modélisées dans le temps. Issu de l'expérience accumulée par les diverses applications que nous avons pu développer, le principe de la démarche, illustré dans la figure ci-dessous (figure 2) est celui d'une ingénierie laissant la place à des ajustements réguliers et à la possibilité de retour arrière. Ce principe constitue pour nous le cadre de la modélisation de trace proprement dite et s'articule de la façon suivante : **(a)** un modélisateur détermine quels sont les éléments constituants de la trace dans une phase d'*amorçage* ; **(b)** un SBT est mis en place (c.f. partie 2) ; **(c)** le modélisateur (ou bien l'utilisateur lui-même) peut analyser la trace pour en évaluer la pertinence et des *ajustements successifs* sont possibles dans un *phase d'évolution* (pour que la trace modélisée évolue en fonction de l'activité).

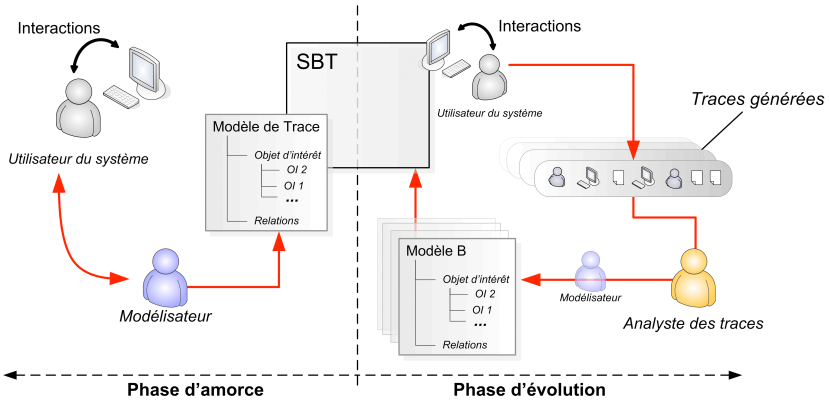


Figure 2 : Démarche de modélisation de trace

Nous avons choisi de focaliser notre présentation sur la phase d'*amorce* qu'est la première modélisation de la trace, puis sur les possibilités d'évolution ou plutôt *d'évolution conjointe* de l'activité et de l'exploitation des traces. Dans la discussion, nous évoquerons les modalités d'ingénierie de la connaissance qui seraient adaptées aux phases d'évolution.

4.2 Première phase de modélisation de trace : la phase d'amorce

À l'instar d'un processus de modélisation de connaissances « classique », la modélisation de trace doit composer avec deux mouvements de modélisation opposés (Bachimont 2004b). Le premier est *descendant*, partant d'un *modèle de l'activité* pour atteindre un modèle de trace capable de décrire l'utilisation de l'environnement dans le cadre d'une activité que l'on veut observable. Le second est *ascendant*, partant du *modèle de conception* de l'environnement observé (et des éléments techniquement

enregistrables que ce dernier définit) pour se hisser à un niveau de description auquel la trace fera sens pour un utilisateur relativement à son activité. L'objectif sera donc pour le modélisateur de trouver en termes de description, le point d'équilibre de ces deux mouvements.

4.2.1 Déterminer le système observé

Le processus de modélisation de trace commence par la détermination du système observé, qui constitue une première illustration de la double contrainte que nous venons d'évoquer. Cette détermination repose sur *une analyse de l'activité effective* des utilisateurs d'un environnement numérique, sachant que celle-ci ne se limite généralement pas à l'exploitation de ce dernier. De plus, il faut tenir compte du fait **(a)** qu'on ne peut tracer que des logiciels qu'il nous est possible d'instrumenter pour cela, et **(b)** que la multiplication des sources de traçage rend la tâche techniquement plus délicate. La recherche de compromis dès la première étape de la démarche que nous décrivons n'a évidemment de sens que dans la mesure où les objectifs d'exploitation de la trace peuvent être servis. Ce sont bien ces *objectifs d'exploitation envisagée de la trace qui doivent prioritairement guider la démarche et non les possibilités techniques d'enregistrement des interactions des outils impliqués*. Une fois le système observé délimité, peut commencer le cœur du travail de modélisation, celui qui consiste à définir les « *Objets d'Intérêts* », i.e. les éléments d'interaction qui seront choisis comme *constituants* de la trace.

4.2.2 Déterminer des Objets d'Intérêt qui seront observés

Les Objets d'Intérêt, choisis comme éléments « saillants » de l'interaction d'un utilisateur avec son système dans le cadre d'une activité donnée, doivent faire sens une fois instanciés et mis en relation dans une trace, ce qui implique finalement une modélisation « indirecte » de la trace.

Modélisation « descendante »

Il s'agit d'une « analyse de l'activité » qui combine diverses approches de modélisation, et a pour objectif fondamental de donner les moyens de comprendre *comment s'inscrit l'utilisation du système observé dans le reste de l'activité*[14]. De manière très générale, les questions principales à se poser sont : quels sont les objectifs de l'activité de l'utilisateur ? Quelles ressources et quels outils mobilise-t-il ? Quelles sont les caractéristiques de cette activité (homogénéité, échelle temporelle, dépendance à d'autres activités) ?

Le modélisateur trouvera dans le domaine de l'Ergonomie, de la Conception, des Sciences Cognitives et évidemment dans le domaine de l'IC des outils pour mener ce travail à bien. Dans ce dernier cas, la définition, même grossière d'un *modèle du domaine* nous paraît importante pour éclairer les autres éléments d'analyse[15].

[14] Notons qu'une analyse complète de l'activité n'est effet pas un objectif en soi, il s'agit bien de cerner comment l'utilisation de l'environnement s'inscrit dans l'activité.

[15] Dans la mesure de l'expertise du modélisateur.

Notons que l'analyse doit être adaptée au regard des particularités de l'activité considérée. Par exemple lorsqu'il s'agit d'une *activité collaborative*, elle doit alors inclure des dimensions socio-organisationnelles : quelle organisation, quelle structuration de l'activité, quels moyens de coordinations mobilisés, *etc*.

Modélisation « ascendante »

De manière symétrique, si on peut dire, la modélisation de trace est aussi *ascendante*. Il s'agit cette fois de prendre comme point de départ le système lui-même : *modèle de conception* du système observé ou tout au moins *modèle fonctionnel*. Plus que l'architecture logicielle des différentes fonctionnalités, on insiste ici sur les *objets* et les *actions mobilisables* par l'utilisateur et qui sont *accessibles au traçage*. Le but est d'arriver à se faire une idée assez précise de l'ensemble des enregistrements possibles des interactions (qu'on pourra appeler « trace brute ») directement à la source de traçage. Cette modélisation ascendante peut s'appuyer sur des outils et techniques existants : *UML*, les *modèles de tâches*, *scenarii d'utilisation*.

Ajustements et convergence

Une fois les deux mouvements de modélisation engagés, doit se mettre en place un processus visant à les faire *converger*, par *ajustements successifs*, vers le niveau d'abstraction recherché. On peut réaliser un « ajustement descendant ». Par exemple, l'ensemble des interactions enregistrables n'étant pas forcément complet par rapport à la description de l'utilisation de l'environnement que l'on veut en obtenir, il est possible de devoir ajouter les moyens de récupérer des « *logs* » supplémentaires[16]. Il est également possible de réaliser des « ajustements ascendants » permettant d'arriver à ce que nous appelons « trace première »[17], qui consistent à mener des processus d'abstraction sur des enregistrements bruts. Par exemple le fait qu'un fichier soit ouvert par un « clic droit » ou par un « double clic » peut ne constituer qu'un seul observé : « ouverture de fichier ».

Dans ce contexte, comment faire concrètement apparaître les éléments saillants pour l'utilisateur dans son activité médiée ? De notre point de vue, le meilleur moyen est de confronter les descriptions que peut produire un utilisateur de sa propre activité, aux différents modèles que nous avons évoqués plus haut. Un moyen de procéder qui nous a semblé pertinent a été de combiner différents types d'entretiens[18] : **(a)** Un scénario d'activité est discuté en face à face avec un utilisateur du système observé. Le modélisateur lui demande de décrire la manière dont il procèderait et comment il utiliserait son système pour mener son activité à bien. **(b)** L'utilisateur face à sa machine est cette fois invité à réaliser le scénario prévu en (a) et de verbaliser son activité (avec l'appui éventuel du modélisateur). **(c)** Un troisième entretien, réalisé *a posteriori*, a pour objet d'une part de déterminer ce que l'utilisateur retient comme

[16] Pour reprendre l'exemple de la trace d'une navigation sur le *Web,* on peut ainsi être conduits à enregistrer toutes les saisies de champs de formulaire, ce qui *a priori* n'est pas traçable d'emblée.

[17] Première trace qui correspond complètement à un modèle, de fait manipulable dans un SBT.

[18] Qui peuvent être appuyés par un enregistrement audio et/ou vidéo.

remarquable dans l'activité qu'il vient de mener (notamment du point de vue des éléments potentiellement réutilisables, par lui-même, dans une activité similaire), et d'autre part un retour sur la façon dont la description de (a) a été mise en action en (b). Les résultats de ce genre d'investigations peuvent conduire à mettre en évidence des éléments saillants pour l'utilisateur dans son activité mais *absents du modèle de conception* du système observé et qui ne sont donc pas traçables « directement ».

Nous disposons au terme de la phase d'amorce un modèle permettant de mettre en place un SBT capable de délivrer une trace exploitable. Voyons maintenant comment nous envisageons les possibilités d'évolution des traces modélisées.

4.3 Évolution conjointe de l'activité et du modèle de trace

Nous venons de voir à travers la phase d'amorce, par quels ajustements passe la modélisation de trace. Dans le contexte d'une activité en perpétuelle évolution (sous l'influence de facteurs externes à l'utilisation du système, ou bien encore en considérant l'impact progressif des traces sur l'exploitation dans le système observé), nous considérons que la modélisation n'est jamais totalement aboutie. Selon les besoins de l'utilisateur, ou d'un analyste des traces, avec l'appui éventuel du modélisateur lui-même (figure 2), il est souhaitable de faire évoluer les modèles de traces sans remettre en cause le SBT lui-même.

5 Discussion

Nous voulons insister sur la plasticité qu'offre notre approche des traces modélisées relativement à l'évolution de l'activité. Un utilisateur ne fait pas évoluer le SBT en lui-même, mais propose des manières pertinentes de « parler » de son activité, que ce soit pour des raisons intrinsèques à l'activité, sur un changement de l'utilisateur dans sa pratique, ou encore suite à une modification du système observé. Des ajustements, voire la création de nouveaux modèles, sont donc réalisés dans la phase d'évolution de la démarche. Pour répondre à des besoins spécifiques de l'activité il est possible (premier niveau d'évolution) de modifier simplement des modèles de traces existants pour changer les conditions d'accès à la trace (sa visualisation). Pour prendre en compte une nouvelle fonctionnalité du système observé en revanche, il sera nécessaire de modifier le modèle de trace lui-même (second niveau d'évolution). Cette possibilité est offerte grâce à *l'indépendance du modèle de trace* relativement au système observé.

Ce sont donc les modèles de traces qui évoluent, permettant par la même occasion la mise au point de nouvelles « interprétations » de la trace et des éléments tracés. Cette émergence peut être concrétisée par la réification de ce que nous appelons des « signatures de tâches », motifs remarquables dans les traces qui peuvent être expliqués, réutilisés ou partagés à la condition qu'ils soient exprimés au bon niveau d'abstraction. L'évolution du (ou des) modèle(s) de trace permet de répondre, indépendamment de l'environnement observé, à ces besoins. La plasticité qu'elle permet en termes de méthodologie est le point fort de notre approche. D'autres travaux ont déployé de gros efforts de modélisation pour créer des systèmes

sophistiqués exploitant des traces d'interaction. Ainsi (Plaisant *et al*, 1999), au sein d'un EIAH, a proposé une modélisation des interactions afin de permettre aux apprenants d'être confrontés à une trace de leur activité représentant leur avancement dans les activités pédagogiques. Ce type d'approche, qui reste *ad hoc*, ne permet d'envisager ni l'évolution de l'exploitation des traces, ni la mise en place d'un processus encadrant cette évolution.

Notre volonté de laisser la possibilité aux modèles de trace d'évoluer par ajustements successifs conduit à la question : faut-il tracer l'exploitation de la trace pour en rendre compte lors des ajustements en question ? Nous ne pensons pas que l'intégration de la trace à l'environnement dans lequel se déploie l'activité, doive se faire en ces termes. Toutefois l'exploitation des traces manipulables par un utilisateur lui-même (posture de réflexivité des traces), doit être traçable et tracée, sans pour autant tomber dans le piège d'une mise en abîme d'une trace de l'utilisation d'une trace, *etc*. La question sera mise à l'épreuve des faits dans des situations d'exploitations réelles sur de relativement longues périodes de temps.

6 Conclusion

Nous avons débuté cet article en replaçant la notion de trace numérique d'interaction dans le contexte de ses diverses utilisations, puis défini les traces modélisées. Deux questions ont ensuite été traitées : l'une théorique concernant le statut *d'inscription de connaissances* des traces modélisées, l'autre méthodologique, concernant une démarche de modélisation de trace que nous considérons comme une *ingénierie de la trace*. Nous y avons considéré l'exploitation des traces dans la dynamique d'une activité toujours susceptible d'évoluer, et explicité la démarche de modélisation que nous avons mise au point.

Les travaux sur les traces modélisées se poursuivent en exploitant plusieurs pistes, en gardant pour objectif la facilitation des activités médiées. Notre attention se porte notamment sur les questions soulevées par la modélisation et l'exploitation de traces dans le cadre d'activités *collaboratives*. Deux terrains d'applications, une plateforme de classe virtuelle pour l'apprentissage collaboratif et une plateforme de production collaborative de contenus de formation professionnelle, sont aujourd'hui investis. Sur le plan théorique, le travail décrit dans cet article est quant à lui toujours en cours d'approfondissement. Plusieurs voies restent à explorer, tout spécialement dans le domaine de l'IC dans lequel notre travail répond, à sa manière, à une préoccupation centrale : « L'IC n'aura donc pas d'expérience scientifique, mais des *expériences humaines d'utilisation dont il faudra garder la mémoire* » (Bachimont 2004b).

7 Références

BACHIMONT B. (2004a). *Arts et sciences du numérique : ingénierie des connaissances et critique de la raison computationnelle*. Mémoire de HDR, UTC. Compiègne.

BACHIMONT B. (2004b). Pourquoi n'y a-t-il pas d'expérience en ingénierie des connaissances ? Actes de la conférence « *Ingénierie des connaissances (IC2004)* ». p.55-64. Lyon.

BACHIMONT B. (1999). De l'hypertexte à l'hypotexte : les parcours de la mémoire documentaire. *Technologies, Idéologies, Pratiques.* Mémoire de la technique et techniques de la mémoire, sous la direction de Charles Lenay et Véronique Havelange. p.195-225.

BEAUVISAGE T. (2004). Sémantique des parcours des utilisateurs sur le Web. *Thèse Sciences du langage*. Paris : Université Paris X – Nanterre, 361p.

BRINKMAN W.P., GRAY P., RENAUD K. (2006). Computer-Assisted Recording, Pre-Processing and Analysis of User Interaction Data, *actes de HCI 2006*, Brinkman W.P. (eds), Londres.

CHAMPIN P.A., PRIE Y., MILLE A. (2004). Musette : a framework for knowledge capture from experience. *Extraction et Gestion des Connaissances EGC'04*. Clermont Ferrand.

CRAM D., JOUVIN D., MILLE A. (2007). Visualisation interactive de traces et réflexivité : application à l'EIAH collaboratif synchrone eMédiathèque. Numéro spécial *« Analyses des traces d'utilisation dans les EIAH »*. STICEF (à paraître).

GEORGEON O., MILLE A., BELLET T. (2006). Analyzing behavioral data for refining cognitive models of operator. Dans *Philosophies and Methodologies for Knowledge Discovery*, 7th International Workshop on Database and Expert Systems Applications, IEEE Computer Society (ed). p.588-592. Krakow, Poland.

LAFLAQUIERE J., CHAMPIN P-A., PRIE Y., MILLE A. (2005). Approche de modélisation de l'expérience : utilisation de systèmes complexes pour l'assistance aux tâches de veille informatiquement médiées. Dans *ISKO-France 2005*, PUN. p. 209-230. Nancy.

LIEBERMAN H. (2001). Interfaces that Give and Take Advice, dans Carroll J. (Ed). *Human-Computer Interaction for the New Millenium,* ACM Press/Addison-Wesley, p.475-485.

MAGNUSSON M. S. (2000). Discovering Hidden Time Patterns in Behavior: T-Patterns and their Detection. *Behavior Research Methods, Instruments and Computers*, 32(1), p.93-110.

MILLE A., CAPLAT G., PHILIPPON M. (2006a). Faciliter les activités des utilisateurs d'environnements informatiques : quoi, quand, comment ? Dans *Intellectica* 2006/2, n° 44: *Systèmes d'aide: Enjeux pour les technologies cognitives*. p.121-143.

MILLE A. (2006b). Raisonner à Partir de l'Expérience Tracée. Dans *Le storytelling : concepts, outils et applications*, Eddie Soulier (dir), Traité IC2, Série Informatique et SI, Hermes.

OLLAGNIER-BELDAME M. (2006). Traces d'interactions et processus cognitifs en activité conjointe : Le cas d'une co-rédaction médiée par un artefact numérique. *Thèse en Sciences Cognitives* de l'Université Claude Bernard (Lyon 1), 247p.

PLAISANT C., ROSE A., RUBLOFF G., SALTER R., SHNEIDERMAN B. (1999). The Design of History Mechanism and Their Use in Collaborative Educational Simulations, Actes de *Computer Support for Collaborative Learning*, p.348-359. Palo-Alto.

SETTOUTI L.S., PRIE Y., MILLE A., MARTY J-C. (2006). Système à base de traces pour l'apprentissage humain. Actes du *Colloque international en «Technologies de l'Information et de la Communication dans l'Enseignement Supérieur et l'Entreprise »,* Toulouse.

SOLLER A., MARTINEZ A., JERMANN P., MUEHLENBROCK M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal of Artificial Intelligence in Education*, 15, p.261-290.

TAUSCHER L., GREENBERG S. (1997). How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems. *International Journal of Human Computer Studies, Special issue on World Wide Web Usability*, 47(1). p.97-138.

WEXELBLAT A. (1998). History-rich tools for social navigation. In proceeding *CHI'98 conference summary on Human factors in computing systems*, California. p. 359-360.