

## **l'expression des besoins au calcul des indicateurs**

Dans cette section nous positionnons notre approche dans le cadre général des travaux menés en observation des situations d'apprentissage. Nous nous intéressons aux travaux de conception de scénarios pédagogiques des points de vue processus et moyens disponibles pour produire un scénario pédagogique. Une partie du chapitre sera consacrée à une présentation des travaux sur l'observation des situations d'apprentissage. L'activité d'observation est au cœur de notre approche car elle représente la source des données qui permettent de faire une réingénierie ou une adaptation. Plus spécifiquement, nous nous intéresserons au rôle pouvant être tenu par un enseignant-concepteur lors de la modélisation de l'observation.

### **2.1.1 La conception pédagogique et la scénarisation**

La plupart des activités d'apprentissage nécessitent un ensemble de ressources utilisées pour supporter la transmission et/ou la construction de connaissances, du simple crayon qui sert à schématiser les connaissances sous forme de textes ou de dessins, aux livres scolaires qui contiennent toutes les connaissances destinées à être acquises par les apprenants durant l'année scolaire (le programme annuel). Dans le contexte des situations médiatisées par l'outil informatique, la conception pédagogique peut être vue comme un processus d'organisation et de planification de l'apprentissage, aboutissant à la production d'un scénario pédagogique [El-Kechaï, 2008]. Ce processus est ainsi défini par [Berger et Kam, 1996] : « La conception pédagogique est le développement systématique de spécifications d'enseignement par utilisation des théories de l'apprentis-

sage et de l'enseignement. C'est le processus d'analyse des besoins et des objectifs de mise en place d'une situation d'apprentissage et l'élaboration des artefacts répondant à ces besoins. Il comprend l'élaboration des supports et des activités d'apprentissage ainsi que l'évaluation de toute activité pédagogique ».

### 2.1.1.1 Définitions d'un scénario pédagogique

La notion de scénario pédagogique a suscité l'intérêt de plusieurs équipes de recherche, ce qui conduit à plusieurs définitions de cette notion. [Charlier et Daele, 2002] définissent le scénario pédagogique comme étant le produit d'un processus de conception. Ces auteurs décrivent le contenu d'un scénario pédagogique comme étant composé d'objectifs, d'une planification des activités d'apprentissage, d'un horaire, d'une description des tâches devant être effectuées par les étudiants, des modalités d'évaluation le tout étant défini, agencé et organisé au cours d'un processus de design. Pour Quentin. et al dans [Quintin et al., 2005], le scénario est considéré comme un ensemble structuré et cohérent constitué de deux parties : le scénario d'apprentissage et le scénario d'encadrement. Le rôle du scénario d'apprentissage est de décrire les activités d'apprentissage conçues pour une utilisation par les apprenants ainsi que leur assemblage pour construire une situation d'apprentissage et les productions qui sont attendues. Le scénario d'encadrement, ou scénario de formation tel qu'appelé par [Paquette et al., 1997], précise les modalités d'intervention des enseignants tuteurs telles que conçues afin d'appuyer le scénario d'apprentissage. Le scénario est considéré comme étant une spécification guidant le déroulement de l'activité dans l'environnement informatique pour lequel il a été conçu [Durand, 2006]. Le passage de la simple spécification textuelle décrivant l'activité vers son déroulement s'effectue par une série de transformations qui opérationnalisent le scénario dans l'environnement informatique cible retenu. Pernin et Le Jeune définissent dans [Pernin et Lejeune, 2004] le scénario pédagogique comme étant la description du déroulement d'une situation d'apprentissage en termes de rôles, d'activités et d'environnement nécessaire à sa mise en œuvre, mais aussi en termes de connaissances manipulées. Ces auteurs font la différence entre deux types de scénarios :

- Le scénario prédictif : défini comme établi a priori par un concepteur en vue de la mise en place d'une situation d'apprentissage
- Le scénario descriptif : défini comme un scénario qui décrit a posteriori le déroulement de la situation d'apprentissage en y incluant en particulier les traces de l'activité des acteurs (essentiellement les apprenants) et leurs productions.

### 2.1.1.2 Le processus de conception d'un scénario pédagogique

Dans les travaux menés en conception de scénarios pédagogiques, il est admis que leur description s'inscrit dans un cycle de vie emprunté au domaine du développement logiciel. Ce cycle de vie, notamment adopté dans [Vantroys, 2003] et repris par [Laforcade, 2004], [Choquet, 2007], [El-Kechai, 2008] et [Abdallah, 2009], est composé de six étapes :

1. Expression initiale des besoins. Cette phase a pour objectif d'initialiser ce processus par une expression des besoins à satisfaire indépendamment de la méthode qui sera utilisée pour concevoir la situation d'apprentissage. Il apparaît souhaitable que cette analyse soit menée en amont.

2. Analyse et conception. Cette phase a pour objectif de formaliser les étapes préliminaires de développement de la situation d'apprentissage afin de satisfaire le plus fidèlement possible les besoins énoncés à l'étape précédente. Pour ce faire, le point de départ est un énoncé informel (le besoin tel qu'il est exprimé, complété par des recherches d'informations auprès des experts du domaine pédagogique, tels les enseignants), ainsi que de l'analyse de l'existant éventuel (c'est-à-dire de la manière dont les processus à traiter par le système d'apprentissage se déroulent chez les futurs utilisateurs). La phase d'analyse permet de lister les résultats attendus. La phase de conception permet quant à elle de décrire de manière non ambiguë, le plus souvent en utilisant un langage de modélisation pédagogique, le fonctionnement futur de la situation d'apprentissage, afin d'en faciliter la réalisation. Cette phase amène à produire un scénario pédagogique.

3. Implémentation ou implantation. Cette phase a pour objectif de transformer la conception (par exemple exprimée à l'aide d'un langage de modélisation pédagogique) en code écrit dans un langage donné (XML par exemple). Ce code est interprétable par la machine, le scénario est donc opérationnel.

4. Déploiement. Cette phase a pour objectif de mettre en œuvre le scénario conçu pour qu'il puisse être réalisé par des apprenants. Le déploiement peut, par exemple, se faire sur une plateforme de formation de type LMS (Learning Management System).

5. Tests. Cette phase a pour objectif de tester la validité des spécifications de la conception dans la réalité d'apprentissage concrète. Plus les tests sont concluants et plus ils sont répétés et affinés, plus les spécifications de conception sont jugées stables.

6. Évaluation. L'étape d'évaluation permet de déduire à partir des tests précédents les améliorations pouvant être apportées au système de formation, ce qui peut amener à revoir l'expression des besoins et/ou l'analyse et la conception réalisées au départ.

### 2.1.1.3 La conception des scénarios pédagogiques

Selon [El-Kechaï et Choquet, 2006], il existe deux approches de modélisation d'une situation d'apprentissage. La première approche consiste à utiliser des EMLs génériques tels que IMS-LD [IMS-LD, 2003] ou LDL [Ferraris et al., 2005] [Ferraris et al., 2007]. Les travaux de recherche liés à cette approche tendent à la standardisation des langages de modélisation pédagogique. La deuxième approche préconise d'impliquer davantage les enseignants-concepteurs dans le processus d'élaboration des langages de modélisation pédagogique. Cette approche favorise la production de modèles qui correspondent aux métiers des enseignants-concepteurs. De nombreuses méthodes de conception de scénarios pédagogiques s'inspirent des principes de l'IDM (l'Ingénierie Dirigée par les Modèles) et décrivent une chaîne de transformations de modèles partant de l'idée initiale jusqu'à la mise en place d'un scénario pédagogique sur une plateforme [Villiot-Leclercq et David, 2007]. L'IDM est un paradigme de développement logiciel où les aspects liés aux contingences d'une plateforme d'exécution cible n'apparaissent que tardivement dans le processus [Blanc, 2005]. L'IDM propose un cadre théorique et pratique (des outils et techniques concrets) pouvant guider l'activité de scénarisation [Laforcade et al., 2007]. Cette approche a été adoptée dans plusieurs travaux de recherche traitant de la modélisation des EIAHs en général [Corbière et Choquet, 2005], [Nodenot, 2005] et de la conception des scénarios pédagogiques en particulier [Abdallah, 2009], [El-Kechaï, 2008], [Caron, 2007] et [Laforcade, 2004].

### 2.1.1.4 L'évaluation des scénarios pédagogiques

L'évaluation des scénarios pédagogiques est réalisée en se basant sur les observations de leur utilisation dans des situations réelles [Choquet, 2007]. Selon [Marty et al., 2004], l'activité d'observation des situations d'apprentissage a un rôle important sur la qualité des scénarios pédagogiques. Elle est même considérée dans [Martel et al., 2007] comme étant l'un des piliers de la pédagogie. Cette activité produit des informations qui répondent à différents besoins, comme la ré-ingénierie des scénarios pédagogiques dans notre contexte. Dans ce domaine, un grand nombre de travaux de recherche traitent les problématiques de la modélisation de l'observation.

## 2.1.2 Ingénierie de l'observation dans les EIAH

L'ingénierie de l'observation regroupe l'ensemble des processus, modèles, approches etc. proposés actuellement dans le but de supporter l'activité d'observation des situations pédagogiques.

L'observation est définie dans [Postic et De Ketele 1988] comme un processus dont la fonction première immédiate est de recueillir de l'information sur l'objet pris en considération. L'action de regarder avec attention a pour but de s'intéresser à certaines ou toutes les propriétés de l'objet considéré, de collecter des informations sur cet objet et ses comportements et d'utiliser ces informations afin d'en tirer des conclusions.

Dans notre contexte de recherche, l'observation est utilisée comme un moyen permettant de produire et de remonter des informations aux enseignants-concepteurs. Les résultats d'une observation sont fournis à ces acteurs pour leur permettre de confronter la conception d'une situation d'apprentissage avec son déroulement réel. Cette confrontation a pour but de les amener à affiner leurs conceptions pour les prochaines utilisations de leurs scénarios pédagogiques. Les objectifs d'observation sont donc étroitement liés à

la conception d'une situation d'apprentissage. Les auteurs de [Djouad et al., 2009] considèrent que l'analyse des traces permet, d'une part, de comprendre la dynamique de l'apprentissage, et, d'autre part, d'évaluer efficacement les situations d'apprentissage. Ces traces contiennent des données de bas niveau collectées lors du déroulement d'une activité pédagogique. Elles passent par des étapes de traitement afin d'en extraire des informations d'un niveau plus accessible par les enseignants. La notion d'indicateur est donc introduite dans différents travaux de recherche ; un indicateur est une donnée qui a une signification pédagogique [Choquet, 2007].

Nous présentons et discutons certains travaux de recherche qui traitent différentes problématiques de l'observation. Nous organisons ces travaux en trois catégories. La première catégorie concerne les travaux basés sur une approche fouille de données. Les travaux de la seconde catégorie regroupent l'ensemble des travaux traitant la modélisation des indicateurs. Nous présentons dans la troisième catégorie certains travaux de définition d'indicateurs spécifiques répondant à des besoins particuliers. Cette partie sera conclue par un bilan dans lequel nous discutons l'ensemble des travaux présentés du point de vue de l'implication des enseignants-concepteurs dans la mise en place d'une activité d'observation.

### 2.1.2.1 Travaux sur l'approche fouille de données

La fouille de donnée (en Anglais *Data Mining*) est le processus qui emploie des techniques d'apprentissage automatiques et intelligentes pour analyser et extraire des informations-connaissances à partir de grandes quantités de données [Witten et Frank, 2005]. C'est un ensemble de méthodes et de techniques aidant à la prise de décisions en cherchant à trouver des modèles d'interprétation ou des motifs à partir des données collectées. La définition de la fouille de données la plus utilisée dans la littérature est celle adoptée par [Goebel et Gruenwald, 1999] dans leur étude des outils et techniques qui traitent l'extraction de connaissances à partir des données. La fouille de données est définie comme « le processus de l'extraction de motifs ou de modèles à partir des données observées ». Le recours à ces techniques et méthodes automatiques est dû, d'une part, aux tailles grandissantes des données, et, d'autre part, au besoin de traitement de ces données, parfois hétérogènes, en temps réel. Un exemple de l'utilisation de la fouille de données dans le contexte des entreprises serait par exemple de répondre aux besoins de compréhension des comportements de leur clientèle à partir de ses caractéristiques.

Il y a eu de nombreuses avancées sur les aspects recherche et développement dans le domaine de la fouille de données. De nombreuses techniques d'exploration de données et des systèmes ont été développés [Chen et al., 1996] et sont désormais considérés comme étant mûrs. Dans le domaine de l'éducation médiatisée, plusieurs équipes de recherche se sont intéressées à l'utilisation de la fouille de données. Plusieurs articles ont été publiés autour de l'analyse des travaux de recherche appliquant les techniques de la fouille de données aux données produites par les systèmes éducatifs. On peut citer en guise d'exemple les travaux de [Romero et Ventura, 2007], [Castro et al., 2007], [Baker et Yacef, 2009]. Des événements scientifiques tels que la conférence EDM (Educational Data Mining) sont organisés afin de discuter les travaux sur ce champ. Le site de la communauté de la fouille de données pour l'éducation la définit comme : une discipline émergente, concernée par le développement de méthodes pour explorer les données singulières qui proviennent des environnements éducatifs, et par l'utilisation de méthodes afin de mieux comprendre les apprenants, et le contexte de leur apprentissage.

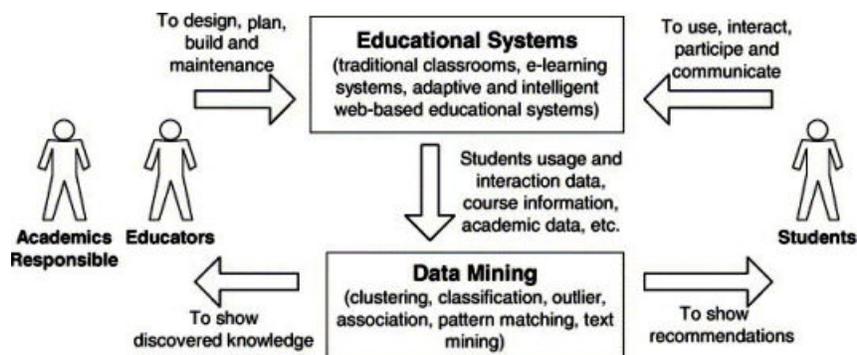


FIGURE 2.1 – Le cycle d’application de la fouille de données dans les systèmes éducatifs (source : [Romero et Ventura, 2007])

Les travaux présentés dans [Merceron et Yacef, 2005] ont pour objectif l’identification des méthodes d’extraction de connaissances ou d’informations qui donnent les résultats les plus pertinents pour les enseignants et les apprenants. Dans l’étude menée par Romero et Ventura [Romero et Ventura, 2007] sur l’utilisation de la fouille de données pour les systèmes éducatifs, les auteurs proposent un cycle (voir figure 2.1) décrivant l’application des techniques de la fouille de données et comment/par qui les résultats sont utilisés. Les connaissances produites par le processus de la fouille de données sont utilisées pour faciliter et améliorer l’ensemble du processus d’apprentissage. Ce cycle est composé de trois étapes :

1. Conception et maintenance : cette phase regroupe les activités de conception de l’environnement d’apprentissage, des plans et supports nécessaires pour le déroulement d’une situation d’apprentissage et la maintenance de ces derniers ;

2. Déroulement d’une situation d’apprentissage : dans cette phase, les apprenants interagissent avec l’environnement d’apprentissage et utilisent les moyens disponibles. Des données sont générées à partir des différentes activités réalisées via l’utilisation de l’environnement d’apprentissage ;

3. Application des techniques de la fouille de données : dans cette phase, les techniques de la fouille de données sont appliquées sur les données générées par le système éducatif afin d’extraire des connaissances. Ces connaissances peuvent être utilisées par les apprenants ou par les concepteurs des artefacts de la situation d’apprentissage afin de procéder à une amélioration de ces derniers.

Il existe une multitude de travaux utilisant la fouille de données pour extraire des informations à destination des apprenants. Les travaux présentés dans [Barnes et al., 2008] ont pour objectif d’automatiser la création de systèmes tuteurs intelligents en se basant sur les techniques de la fouille de données. L’approche présentée repose sur des données générées par le système d’apprentissage, lorsque les apprenants réalisent une activité, afin de leur proposer des recommandations pour les activités suivantes. L’objectif est de rendre un système de recommandation intelligent en se basant sur les connaissances produites par le processus de fouille de données. D’autres travaux partagent cet objectif, par exemple les travaux présentés dans [Mavrikis, 2008] qui servent à prédire si les apprenants peuvent répondre correctement à une question particulière, sans aide supplémentaire, en analysant leurs précédentes interactions avec l’environnement d’apprentissage, ou les travaux d’identification des apprenants qui ont des difficultés dans leurs parcours afin de leur fournir des moyens d’apprentissage adaptés [Antunes, 2008].

Les techniques de la fouille de données sont aussi utilisées pour fournir des informations à destination des enseignants. Romero et al. [Romero et al., 2008b] ont développé un outil de fouille de données spécifique à la plateforme Moodle pour une utilisation par des enseignants concepteurs. Il possède une interface simple facilitant l’exécution des techniques de fouille de données. Cet outil est intégré dans la plateforme Moodle afin de permettre aux enseignants, à la fois de créer et d’entretenir des cours et d’effectuer tous les traitements de fouille de données dans le même environnement. Ceci leur permet de faire des recommandations aux apprenants ou d’apporter des modifications aux cours à partir des informations produites par le processus de fouille de données.

Il existe d'autres travaux cherchant à rendre accessible l'utilisation des techniques de la fouille de données par des enseignants qui ne sont pas forcément experts de ces techniques. On peut évoquer, par exemple, les travaux de visualisation des parcours effectués par les différents apprenants sur une plateforme d'apprentissage basée sur le web [Romero et al., 2008a]. L'outil propose aux enseignants de visualiser les différents parcours effectués par les apprenants afin de découvrir les plus utilisés. Les parcours sont obtenus par application des techniques de la fouille de données sur les données collectées sur les activités des apprenants. Les travaux présentés dans [Ventura et al., 2008] cherchent à aider les enseignants concepteurs à choisir des mesures d'évaluation des règles obtenues par l'application des algorithmes de fouille de données. L'outil proposé intègre un éditeur permettant de créer de nouvelles mesures sous forme d'équations écrites au format Latex.

### **Discussion**

Les travaux sur la fouille de données sont centrés sur l'application des techniques de l'Intelligence Artificielle (analyses statiques, réseaux de neurones, etc.) pour extraire des informations des masses de données brutes. Le travail présenté dans [Merceron et Yacef, 2005] a pour objectif d'étudier les différentes techniques proposées dans ce domaine et de chercher à déterminer comment les combiner pour extraire les informations les plus pertinentes à partir des données brutes. L'étude présentée dans [Castro et al., 2007] présente une quantité importante de travaux qui utilisent la fouille de données dans le domaine de l'éducation. Ils classifient ces besoins en trois catégories : évaluation des performances des apprenants, évaluation des performances des ressources utilisées dans un cours et production de retours aux tuteurs basés sur les comportements des apprenants. Les auteurs choisissent de présenter ces travaux selon les techniques de fouille de données utilisées pour répondre à différents besoins.

La nature de ces études reflète la nature des problématiques auxquelles est confrontée l'utilisation de la fouille de données dans le domaine de l'éducation médiatisée. En partant de certains besoins d'observation, l'objectif de ces travaux est de chercher la ou les méthodes à utiliser pour produire les informations les plus pertinentes. C'est une approche qui reste à caractère technique et nous pensons que l'enseignant concepteur n'est pas en mesure d'utiliser pleinement ces techniques. L'utilisation de ces techniques reste alors réservée aux acteurs qui les maîtrisent. Certains travaux de recherche cherchent à rendre accessible l'utilisation des techniques de la fouille de données par des enseignants-concepteurs, mais la nature de ces techniques impose de concevoir des outils qui implémentent ces techniques et qui proposent des ouvertures, qui restent d'ordre technique, afin de permettre aux enseignants-concepteurs de les prendre en mains. L'exemple des travaux présentés dans [Ventura et al., 2008] montre que cette ouverture reste technique et exige des compétences que nous pensons en dehors des compétences normales des enseignants-concepteurs (nous parlons ici de la définition d'équations sous format Latex). Dans notre contexte de recherche, les besoins d'observation changent d'une situation d'apprentissage à une autre. Plusieurs besoins d'observation peuvent être exprimés pour une même situation d'apprentissage et couvrent différentes parties d'une session d'apprentissage.

Ces besoins d'observations peuvent évoluer ou disparaître à chaque itération d'un cycle de ré-ingénierie. Notre objectif de recherche est de proposer aux enseignants-concepteurs des outils et des techniques qui permettent de les aider à spécifier quelles informations doit produire une activité d'observation.

Bien que cette approche ait fait ses preuves dans divers domaines, y compris dans l'observation des situations médiatisées, nous pensons que son utilisation dans notre contexte n'est pas pertinente car nos objectifs de recherche diffèrent des objectifs des travaux adoptant une approche basée sur la fouille de données.

#### **2.1.2.2 Travaux sur la modélisation des indicateurs**

##### **Les projets ICALTS, IA, CAViCoLA et DPULS**

L'observation et l'analyse des interactions des situations d'apprentissage, collaboratives ou individuelles, ont suscité l'intérêt de plusieurs équipes de recherche. Les actions du Réseau d'Excellence Européen Kaleidoscope sont parmi les travaux de recherche les plus importants menés dans ce champ. Le projet ICALTS (*Interaction & Collaboration Analysis supporting Teachers & Students' Self-regulation*) [Dimitracopoulou

et al., 2004] a débuté en 2004. Le groupe du projet de recherche ICALTS présente, dans son état de l'art, plusieurs indicateurs et outils pour analyser les interactions sur les plateformes collaboratives, comme par exemple les réseaux sociaux qui représentent chaque acteur par un point puis relient les points entre eux selon l'intensité ou la fréquence des interactions. Ce groupe a fourni de nombreux outils d'analyse provenant de plusieurs laboratoires de recherche européens.

Le projet ICALTS a défini l'analyse de l'interaction comme étant dirigée par l'énoncé d'hypothèses sur l'interaction et la volonté de les prouver ou de les rejeter par l'observation. Dans cette optique, les travaux de recherche ont abouti à la formalisation de plusieurs indicateurs. En se basant sur les données de l'interaction des différents acteurs d'une activité collaborative, ces indicateurs sont calculés automatiquement et présentés à ces acteurs afin qu'ils prennent conscience de l'avancement de leurs activités.

Dans [Dimitrakopoulou et al., 2004], une distinction entre trois niveaux d'indicateurs est proposée. Les indicateurs de haut niveau, les indicateurs de niveau intermédiaire et les indicateurs de bas niveau. Les indicateurs de haut niveau sont ceux qui ont une valeur facilement interprétable, souvent calculée par des processus complexes appliqués aux traces. Les indicateurs de niveau intermédiaire sont des indicateurs assez élaborés. Les indicateurs de bas niveau sont obtenus directement à partir des traces brutes. Ces mêmes auteurs précisent que les indicateurs d'un niveau particulier ne sont pas forcément plus importants que ceux des niveaux inférieurs.

Le projet IA [Dimitrakopoulou et al., 2006] a débuté en 2005. Il s'agit de la suite des travaux de recherche menés dans le projet ICALTS. La figure 2.2 schématise l'approche de l'analyse de l'interaction du projet IA. Ce processus commence par le questionnement et la formulation d'hypothèses sur l'interaction, et aboutit à la définition d'indicateurs pédagogiques. La définition des indicateurs conduit à spécifier les données qui doivent être capturées par l'EIAH (flèches en pointillés). Les données brutes, générées par l'EIAH, sont traitées par une méthode d'analyse qui se charge d'établir l'indicateur. Les données brutes conditionnent ainsi la possibilité d'établir ou non l'indicateur. Ce dernier est transmis à un outil pour être utilisé par l'utilisateur final. L'utilisation concrète de l'indicateur répond à une norme ou à une stratégie influencée par la méthode d'analyse, et influence à son tour le choix de l'outil exploitant l'indicateur. Les principales étapes de traitement des données de l'outil d'analyse des interactions proposées dans le projet IA sont illustrées dans la figure 2.3.

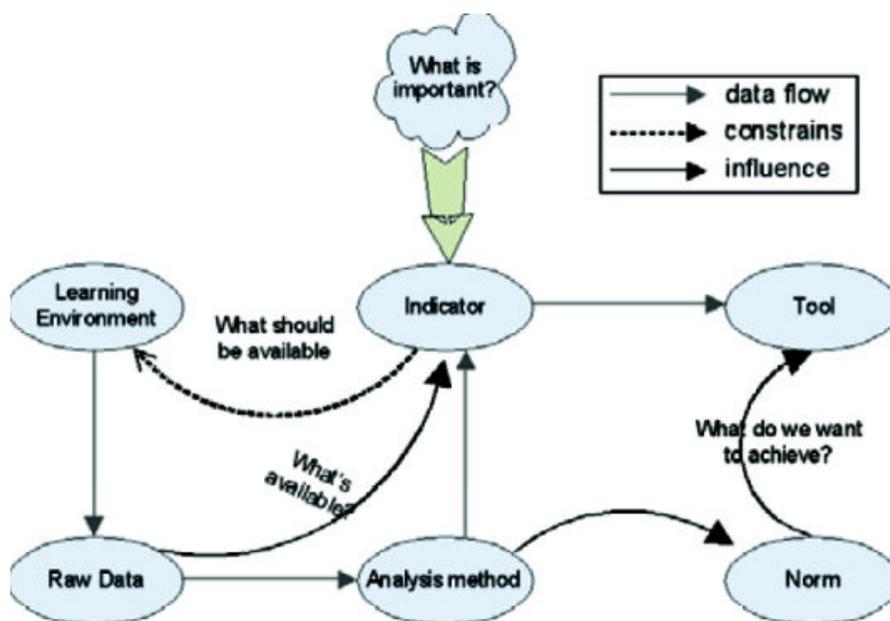


FIGURE 2.2 – Schéma de l'analyse de l'interaction (source : [Harrer et al., 2009])

Le projet CAViCoLA (*Computer based Analysis and Visualization of Collaborative Learning Activities*) est basé sur les résultats des deux projets ICALTS et IA. La modélisation du processus d'analyse de l'interaction dans des environnements d'apprentissage collaboratifs proposée dans ce projet détaille le flot de données depuis la source qui les a générées (l'EIAH) jusqu'à l'outil d'exploitation (les flèches grasses de la figure 2.3). La partie droite de la figure présente le processus d'analyse proposé par CAViCoLA. Ce processus reprend à son compte l'idée de la triangulation des résultats dans une approche multiméthodes : des méthodes quantitatives, des méthodes qualitatives, et un schéma de classification [Spada et al., 2005], [Meier et al., 2007] pour évaluer la qualité du processus de collaboration.

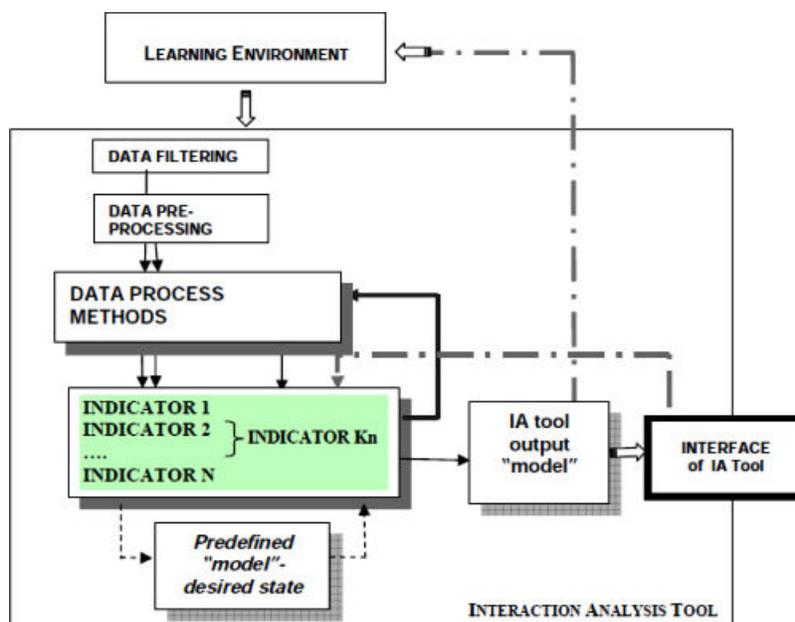


FIGURE 2.3 – Les principales étapes du processus générique de l'analyse des interactions (source : [Dimi-trakopoulou et al., 2006])

Toutes ces méthodes d'analyse suivent la procédure classique [Harrer et al., 2007]. La première étape de cette procédure est la collecte des données brutes. Ces données brutes sont segmentées, c'est-à-dire transformées de manière à isoler des variables significatives. Des pré-traitements, par annotations qualitatives ou mesures quantitatives, sont appliqués sur les données segmentées. Le résultat est analysé selon la méthode considérée. Les données analysées sont visualisées avant d'être interprétées.

DPULS (*Design Patterns for collecting and analysing Usage of Learning Systems*) est une action du Réseau d'Excellence Européen Kaleidoscope qui s'est déroulée en 2005 [Choquet et al., 2005]. Ce projet avait comme objectif de capitaliser le savoir-faire de collecte et d'analyse des usages mis en œuvre pour un EIAH. Ce savoir-faire était capitalisable et réutilisable dans la conception d'autres EIAH. Comme solution, les équipes participant à ce projet ont adopté l'approche de capitalisation à base de patron. Selon [Alexander et Language, 1977] un patron est un modèle qui décrit un problème qui se manifeste dans notre environnement, et donc décrit la solution à ce problème, d'une façon telle que l'on puisse la réutiliser sans jamais le faire deux fois de la même manière. Dans DPULS, une solution est un savoir-faire permettant d'établir des indicateurs mettant en évidence une utilisation particulière d'un EIAH.

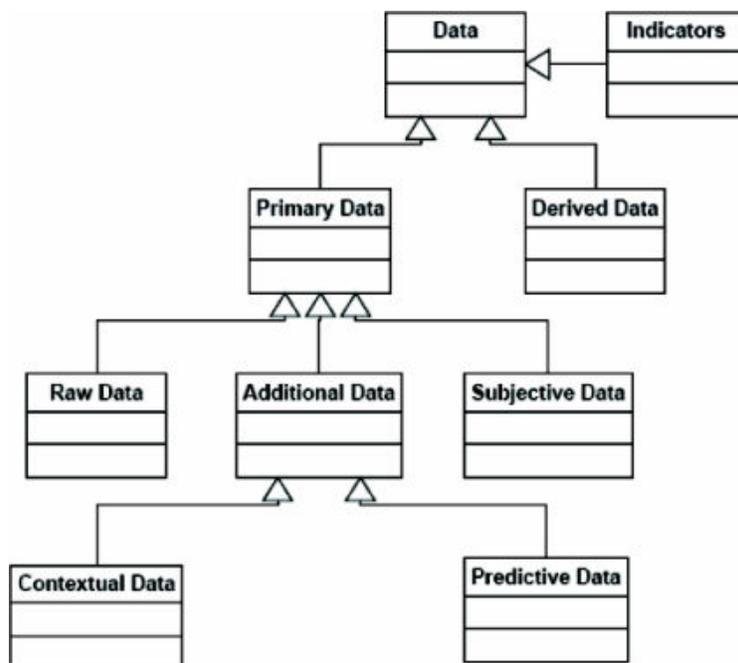


FIGURE 2.4 – Typologie des traces proposée dans le projet DPULS (source : [Choquet et al., 2009])

Suite au projet TRAILS, et dans l'objectif de définir des Patrons d'analyse de trace, le DPULS a proposé une typologie des données d'observation [Pozzi 2005] qui redéfinit le concept d'indicateur introduit par le projet ICALTS. Ces types sont définis en fonction des objectifs d'utilisation et de la provenance des traces. La figure 2.4 montre cette typologie.

Cette typologie permet de différencier une donnée dérivée (derived-datum) qui est une donnée calculée ou établie et une donnée primaire (primary-datum). Cette dernière est décomposée selon trois sous typologies. Elle est considérée comme subjective (subjective-datum) quand elle est établie par l'analyste de la session, et comme brute (raw-datum) quand elle est collectée avant, pendant ou après une session d'apprentissage. Une donnée primaire est également qualifiée d'additionnelle (additional-datum) quand elle est utilisée pour établir une donnée dérivée. Une donnée additionnelle peut être d'ordre contextuel (contextualised-datum) ou prédictive (predictive-datum). Les indicateurs sont des données primaires ou dérivées qui ont un sens au niveau pédagogique.

### L'approche Musette : Modéliser les Utilisations et les Tâches pour Tracer l'Expérience

L'approche MUSETTE [Champin et al., 2003] est développée au sein du laboratoire LIRIS (Laboratoire d'informatique en Image et Systèmes d'information) par l'équipe Silex (*Supporting Interaction and Learning by EXperience*). L'objectif de l'approche Musette est de fournir un cadre pour l'assistance à base de traces des utilisateurs d'un système informatique [Mille et al., 2006]. Le principe de cette approche est illustré dans la figure 2.5. Cette figure se lit de la façon suivante : un utilisateur est en interaction avec un environnement. À partir de ces interactions, un agent observateur génère des traces primitives en respectant un modèle d'utilisation. Ce dernier décrit les objets manipulés par l'utilisateur durant son activité : ce sont les Objets d'Intérêt. Le modèle d'utilisation "décrit les entités, événements et relations à observer pour construire une trace primitive" [Mille et al., 2006].

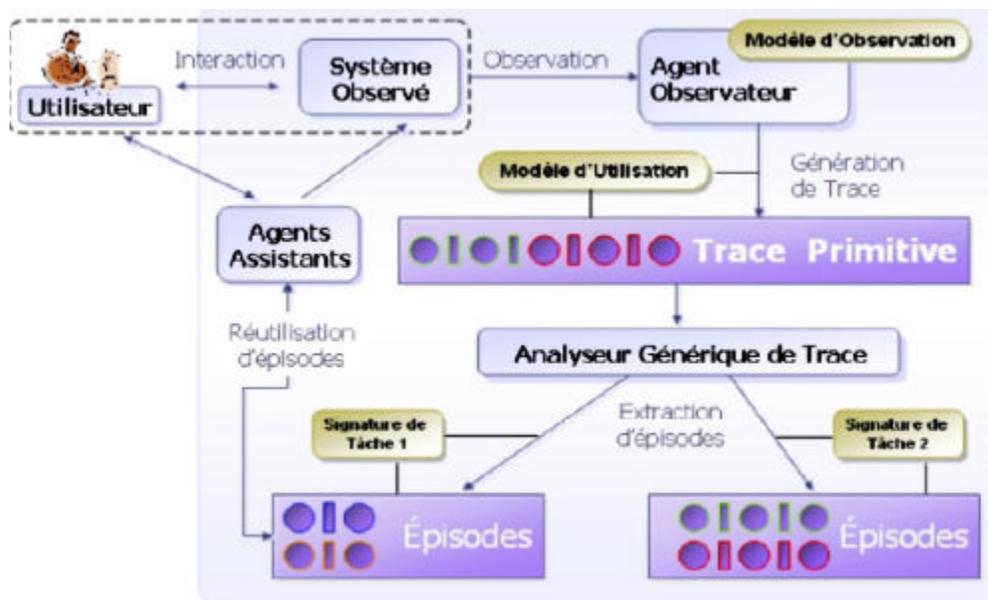


FIGURE 2.5 – L'approche MUsETTE (source : [Mille et al., 2006])

Les traces primitives décrivent les opérations successives d'un utilisateur. Elles sont analysées dans le but de produire des épisodes d'utilisation conformes à des signatures de tâches prédéfinies. Les signatures de tâches permettent d'associer une sémantique aux traces primitives et d'identifier la situation dans laquelle se trouve l'utilisateur de l'environnement informatique. Les épisodes d'utilisation sont utilisés par des agents assistants afin de proposer des aides à l'utilisateur. La construction de la trace, dans l'approche Musette, s'appuie sur une ontologie de haut niveau. La trace primitive est une suite d'états et de transitions contenant respectivement des captures de l'état de l'environnement informatique, etc., et des événements déclenchés par l'appel d'une fonctionnalité par l'utilisateur [Stuber, 2007]. D'une manière générale, la trace primitive est constituée des objets d'intérêt repérés à partir des données brutes, capturées par l'agent observateur [Stuber, 2007], d'interaction de l'utilisateur avec l'environnement informatique.

### Les Systèmes à Base de Traces (SBT)

L'approche des SBTs est un autre résultat des travaux du laboratoire LIRIS. Les SBTs s'inspirent de l'approche MUsETTE, mais en proposant un cadre plus général pour l'exploitation et l'utilisation des traces. Settouti [Settouti et al., 2006] définit le Système à Base de Traces comme « un système informatique permettant et facilitant l'exploitation des traces » [Settouti et al., 2006]. Dans ce cadre on ne parle plus de modèle d'utilisation, ni de modèle d'observation, mais de modèle de trace. Dans ce type de système, les traces sont considérées comme une séquence temporelle d'observés et sont toujours accompagnées de leurs modèles de traces. Ce dernier décrit le contenu de la trace ; il s'agit en fait d'un ensemble de concepts et de relations types. Le principe général d'un système à base de traces est illustré dans la figure 2.6. Ce principe repose sur le fait que les traces collectées, dites traces primitives, peuvent subir plusieurs transformations avant d'être exploitées, ceci afin d'atteindre un niveau d'abstraction significatif pour l'utilisateur. La nécessité de ces transformations est justifiée par certains caractères de ces traces qui rendent leur exploitation difficile. On peut citer, par exemple, le volume important des traces qui peuvent être générées lors de l'observation ou les formats de traces générées par différentes sources de traçage.

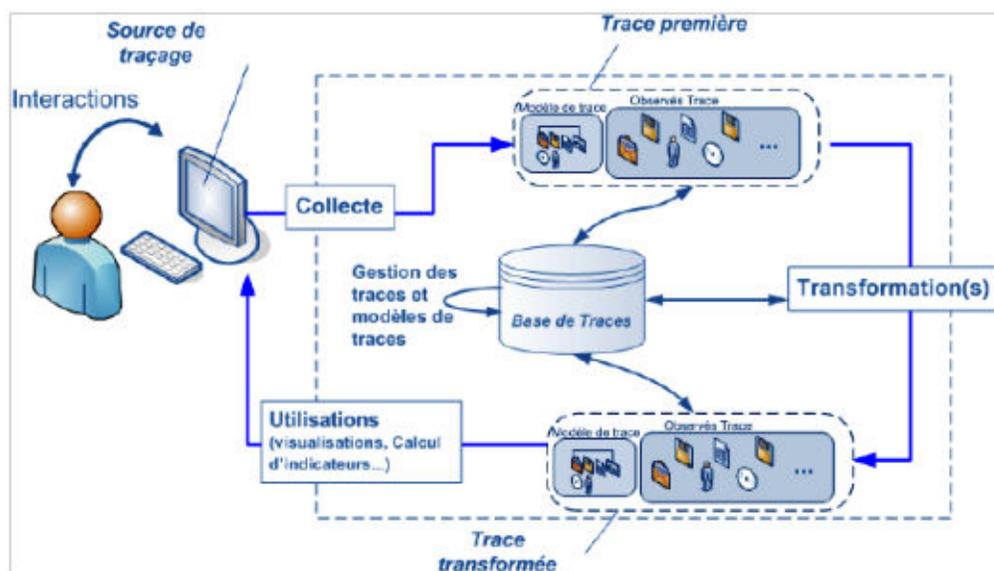


FIGURE 2.6 – Principe général d'un système à base de traces modélisées (source : [Settoui et al., 2007])

Les auteurs de [Settoui et al., 2007] précisent que ces transformations peuvent être réalisées soit d'une manière automatique, soit d'une manière manuelle. Les transformations manuelles sont faites par un utilisateur interagissant avec sa trace. Ces modifications portent sur l'ajout, la modification ou la suppression des observés, changeant ainsi la composition des éléments de la trace. Le SBT se charge de maintenir la conformité des traces avec leurs modèles de traces. Les transformations automatiques, quant à elles, sont faites par le système à base de traces en utilisant des règles de transformations. Trois types de transformations sont définies :

- La sélection : l'objectif d'une telle transformation est d'extraire un sous-ensemble d'observés pertinent, de la trace brute, en respectant des filtres.
- La réécriture de motifs : ce type de transformation permet de changer la trace en remplaçant certains observés par d'autres.
- La fusion temporelle : ce type de transformation permet de fusionner ou de regrouper des traces en provenance de différentes sources.

### Une approche de traçage conduite par les modèles

Broisin et Vidal [Broisin et Vidal, 2007] proposent une approche conduite par les modèles pour la gestion des traces d'activité des utilisateurs. Cette approche vise à assurer l'exploitation des traces d'activité dans des situations d'apprentissage s'appuyant sur des outils hétérogènes instrumentés par les technologies Web. Les propositions scientifiques se présentent sous forme d'un modèle UML générique de traces (la figure 2.7 est un extrait de ce modèle) et d'une architecture distribuée et décentralisée favorisant le partage et la réutilisation des traces collectées par différents outils et services.

La modélisation proposée, qui est une extension du méta-modèle CIM (Common Information Model) et dont l'architecture est conforme au WBEM (Web Based Enterprise Management), s'organise autour de quatre concepts principaux :

- les systèmes d'apprentissage en ligne représentés par le concept EIAH\_System ;
- les ressources intégrées dans ces systèmes représentées par le concept EIAH\_Resource ;
- les activités que ces derniers peuvent réaliser sur les systèmes et ressources représentées par les concepts EIAH\_ActivityOnSystem (activités réalisées sur l'EIAH telles que l'administration du système) et EIAH\_ActivityOnResource (activités réalisées sur une ressource telles que son utilisation par un apprenant) ;

– les utilisateurs. La modélisation de ce dernier concept n’est pas présentée dans la figure 2.7, les auteurs précisent qu’ils ont utilisé les classes du CIM User (le modèle CIM permettant la modélisation de l’utilisateur).

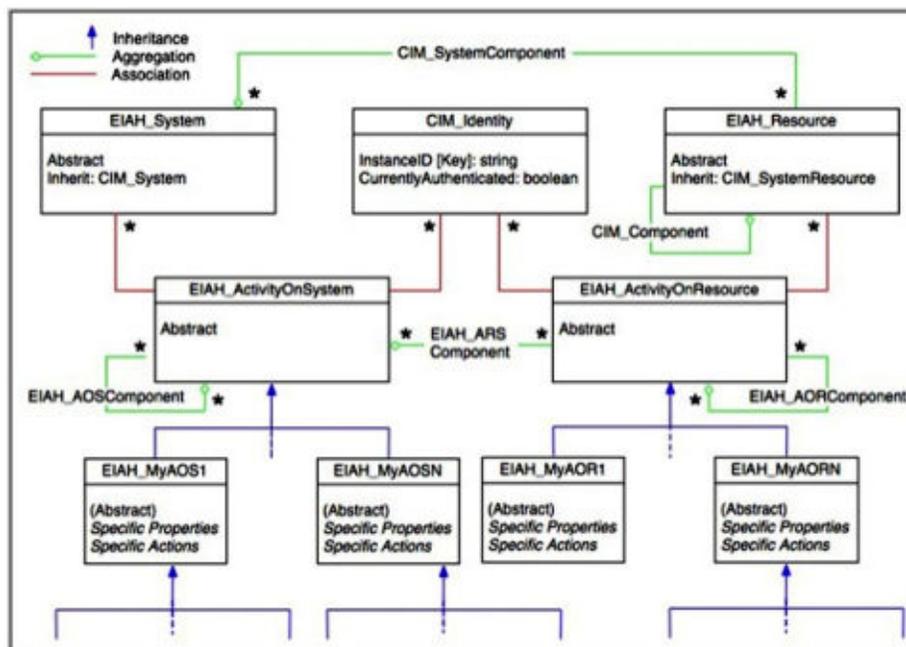


FIGURE 2.7 – Un modèle UML générique de traces pour les activités (source : [Broisin et Vidal, 2007])

Les différentes classes de ce modèle UML permettent la spécification de l’ensemble des traces à collecter sur l’utilisation du système d’apprentissage lui même, sur l’ensemble des ressources mises à disposition des différents acteurs et sur l’ensemble des activités prévues comme devant être réalisées durant une session d’apprentissage. Ces traces sont collectées par des composants logiciels intégrés aux différents systèmes d’apprentissage utilisés dans le contexte d’une situation d’apprentissage. D’autres composants assurent le stockage, l’interrogation et le partage des traces collectées.

### Le méta-langage UTL (Usage Tracking Language)

Le méta-langage UTL [Choquet et Iksal, 2007] est un des résultats du projet de recherche REDiM du laboratoire LIUM. Il a été élaboré dans le but de faciliter l’analyse de traces. La description de données d’analyse peut être réalisée indépendamment d’un langage de scénarisation pédagogique et d’un format de traces particulier.

L’idée générale d’UTL est que chaque trace d’utilisation, prise en considération lors de la modélisation de l’observation, doit concourir à témoigner d’un usage observé, et que se sont ces derniers qui sont à mettre en relation avec le scénario pédagogique prédictif afin de pouvoir les analyser.

La figure 2.8 montre le modèle d’information décrivant un indicateur au sens d’UTL. Les éléments dont la base est colorée en noir sont les éléments accessibles par un enseignant-concepteur. La plupart sont des éléments de type champ de texte. L’élément UTL4DCLFormula (l’élément encadré de la figure 2.9) permet d’exprimer formellement le moyen d’observation par utilisation du langage DCL4UTL (Data Combination Language for UTL) [Ngoc et al., 2009]. Ce langage générique est inspiré des langages de requêtes classiques tels que SQL ou Xquery [Ngoc, 2010]. Son objectif est de combiner les données pour établir de nouvelles données dérivées. Parmi ces données dérivées se trouvent les indicateurs.

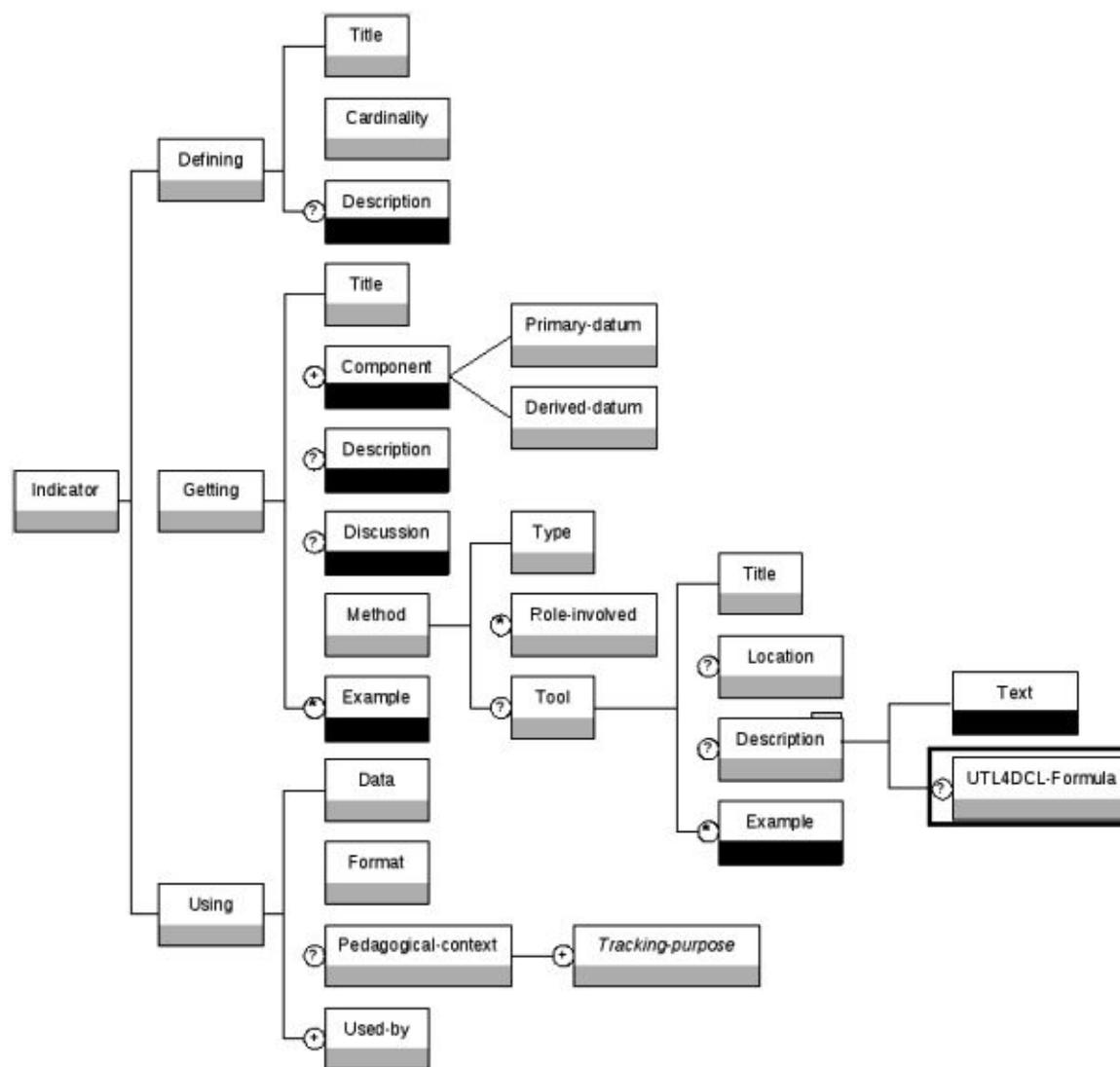


FIGURE 2.8 – Modèle d’information d’un indicateur (source : [Ngoc, 2010])

### Discussion

Les travaux classés dans cette catégorie s’approchent davantage de notre contexte de recherche des points de vue modélisation et préparation de l’observation. Les problématiques des différentes équipes de recherche sont centrées sur le développement de langages, modèles, outils, etc. supports aux processus de modélisation et de traitement des données brutes collectées lors du déroulement d’une session d’apprentissage. Le projet ICALTS a défini l’analyse de l’interaction comme étant dirigée par l’énoncé d’hypothèses sur l’interaction et la volonté de les prouver ou de les rejeter par l’observation. Nous trouvons que ce point de vue est intéressant car l’observation est guidée par un ensemble de questions auxquelles certains acteurs (chercheur, enseignants ; etc.) cherchent à trouver des réponses. Les membres du projet Cavicola ont proposé un format unifié de données d’interaction pour le développement d’outils d’analyse et d’EIAH. La modélisation proposée détaille le flot de données depuis l’EIAH source jusqu’à l’outil d’exploitation. Le projet DPULS propose un modèle structurant les données d’observation, des traces primitives ou premières aux indicateurs. Dans l’ensemble de ces travaux, les propositions scientifiques se sont centrées sur l’analyse et le traitement

des traces.

Les travaux du laboratoire LIRIS, en l'occurrence l'approche Musette et l'approche SBT, proposent des cadres d'utilisation des traces informatiques pour observer l'utilisation des systèmes informatiques. L'objet d'étude principal de ces travaux est la trace informatique. Les épisodes détectés automatiquement grâce aux signatures de tâches expliquées (SiTEx) peuvent être considérés comme des indicateurs de parcours d'utilisateur ou d'utilisation d'un environnement informatique. D'un point de vue d'utilisation des épisodes, le système fournit à l'utilisateur final (l'enseignant ou le tuteur dans le cas des EIAH) des informations dans un niveau d'abstraction facile à interpréter. Les informations extraites et leur quantité dépendent de la définition des signatures de tâches expliquées et du nombre de ces signatures.

La définition du modèle d'utilisation et des signatures de tâches expliquées se fait par une première observation de l'utilisation des outils informatiques [Mille et al., 2006]. Ceci nécessite une connaissance détaillée des fonctionnalités des outils observés. L'exemple d'utilisation de l'outil Word® présenté dans [Mille et al., 2006] montre que pour un même utilisateur, l'observation de l'utilisation d'un outil dépend de la nature des tâches à effectuer. Les signatures de tâches sont donc liées aux différentes utilisations de l'outil informatique utilisé dans la réalisation des tâches concernées par l'observation. Dans notre contexte de recherche, nous pensons que ceci représente un frein car lors de la conception d'une situation d'apprentissage, la planification de l'utilisation d'un outil (un forum ou un tchat par exemple) pendant une activité d'apprentissage est spécifiée dans un but pédagogique, en s'intéressant au service général que rend l'outil informatique et en faisant abstraction des détails de ses fonctionnalités.

Enfin, l'approche Musette est centrée sur l'assistance à la tâche de l'utilisateur en interaction avec l'environnement informatique observé plutôt que sur une approche d'aide à la modélisation ou à la spécification des indicateurs. L'approche SBT [Settoui et al., 2007][ Settoui et al., 2006] offre un cadre d'utilisation des traces informatiques plus général que l'approche Musette. Le calcul des indicateurs est une des trois utilisations possibles dans le contexte des EIAHs. Cette opération se fait par application des transformations sur les traces brutes. La force de cette approche réside dans le cadre qu'elle propose pour l'utilisation des traces pour le suivi des activités médiatisées, mais il nous semble que le fait de considérer le calcul des indicateurs uniquement comme un problème de traitement et d'exploitation des traces est réducteur. L'approche Musette, quant à elle, offre le moyen de définir des signatures de tâches expliquées que nous pouvons considérer comme des modèles d'indicateur potentiels qui répondent à des besoins particuliers pour l'observation d'une tâche (ou activité en général) particulière. Il nous semble regrettable aussi que les rôles ne soient pas clairement définis sur l'utilisation de cette approche. A notre connaissance, elle n'offre aucun moyen pour aider ou pour assister l'utilisateur final, dans notre contexte l'enseignant/concepteur, lors de la définition des indicateurs dont il a besoin pour observer des activités pédagogiques car elle est plutôt centrée sur les traces premières que sur les besoins d'observation.

L'approche de traçage proposée dans [Broisn et Vidal, 2007] permet de modéliser une observation par production d'un modèle UML. Ce dernier peut être perçu comme un méta-modèle de traces défini pour une situation d'apprentissage utilisant des outils hétérogènes. Un des points forts de cette approche est la manière dont les données à collecter sont définies. En définissant l'ensemble des classes décrivant les environnements utilisés, les ressources mises en place et l'ensemble des activités concernées par l'observation, les données collectées sont rattachées à un contexte, facilitant ainsi leur interprétation et leurs traitements. Cette approche nécessite des compétences en modélisation UML que nous pensons hors de portée des enseignants-concepteurs.

Le méta-langage UTL [Choquet et Iksal, 2007] permet de modéliser les données de l'observation d'une session d'apprentissage. Ce méta-langage offre le moyen de modéliser les indicateurs a priori du déroulement d'une session d'apprentissage, les données nécessaires au calcul d'un indicateur ainsi que les moyens nécessaires à la collecte et au traitement des données. La définition de l'indicateur est liée aux éléments du scénario pédagogique. Dans un contexte de ré-ingénierie des scénarios pédagogiques, nous trouvons que ce lien est assez important car il offre le moyen de définir des indicateurs en relation directe avec les éléments du scénario pédagogique dont l'enseignant-concepteur souhaite observer les utilisations ou comportements lors du déroulement réel d'une session d'apprentissage afin de les améliorer.

Le besoin d'observation, source de spécification des indicateurs, est défini d'une manière textuelle dans la facette Defining de l'indicateur. Selon notre point de vue, ceci peut poser un problème car lors de la

spécification de l'indicateur, une interprétation doit être faite de la description textuelle du besoin d'observation. Dans le cas où l'acteur qui spécifie l'indicateur n'est pas le même que celui qui décrit le besoin d'observation, il peut y avoir des différences d'interprétation.

Le méta-langage UTL est centré sur la modélisation et le traitement des données pour la définition et le calcul d'indicateur. Il offre, d'une part des moyens pour une définition des besoins d'observation sous forme textuelle et d'autre part, un moyen de spécification formelle des moyens d'observation par le langage DCL4UTL. La première partie du langage est accessible par des enseignants/concepteurs mais, d'une part, la description du besoin d'observation reste à un niveau textuel, et d'autre part, le langage n'embarque pas de moyens à destination des enseignants-concepteurs pour les guider et les aider dans l'expression et dans la spécification de leurs besoins d'observation. Nous positionnons nos travaux comme un niveau intermédiaire entre la description textuelle du besoin d'observation et sa formalisation par utilisation du langage DCL4UTL. Notre objectif est donc d'aider l'enseignant-concepteur à produire des spécifications de ces propres besoins d'observation tout en restant dans son univers métier. Nos travaux représentent donc une suite des efforts entrepris dans le projet REDiM sur l'implication des enseignants-concepteurs dans le cycle de ré-ingénierie.

### 2.1.2.3 Travaux sur la définition des indicateurs spécifiques

Dans cette partie de notre positionnement scientifique, nous allons présenter divers travaux sur la définition d'indicateurs spécifiques. Nous allons présenter essentiellement trois travaux : les travaux menés dans le cadre de la thèse de F. Diagne [Diagne, 2009] pour la définition et la réutilisation d'indicateurs pour assister un tuteur dans son activité de tutorat ; les travaux de N. Bousbia [Bousbia, 2008] pour la définition et l'utilisation d'indicateurs pour détecter des styles d'apprentissage afin de personnaliser l'apprentissage aux apprenants, et enfin les travaux de M. May [May, 2009] sur l'observation des activités de communication entre les différents acteurs d'une situation d'apprentissage.

#### Des indicateurs pour la détection des styles d'apprentissage

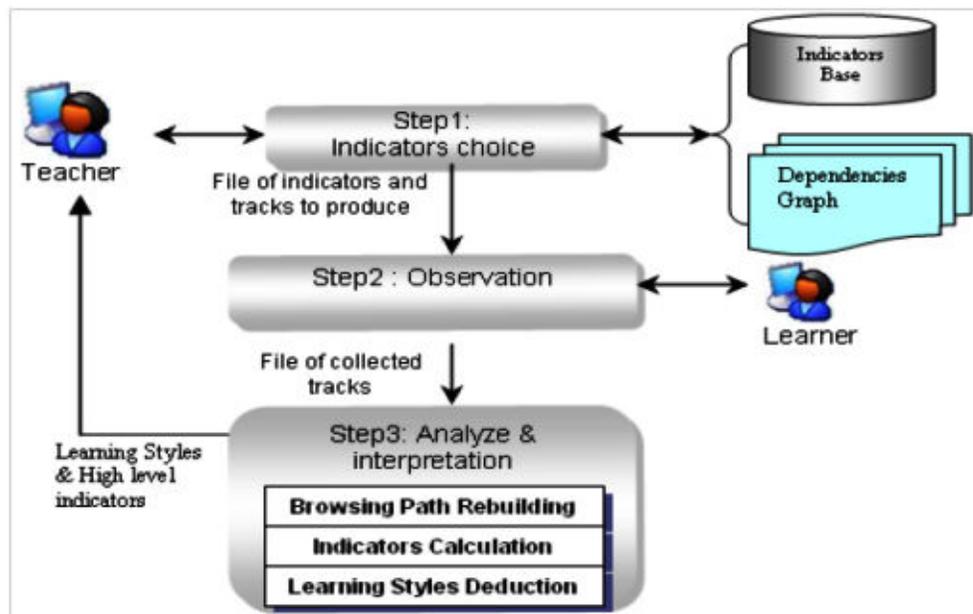


FIGURE 2.9 – Étapes de détection des styles d'apprentissage (source : [Bousbia et al., 2008])

Le processus de détection des styles d'apprentissage présenté dans [Bousbia et al., 2008] est composé de trois étapes (voir figure 2.9) :

- Choix d'indicateur : dans cette première étape, un ensemble d'indicateurs est proposé à l'enseignant. Ce dernier choisit l'indicateur qui correspond à ses attentes sur la base des indicateurs spécifiés par l'équipe de recherche. Ces indicateurs sont considérés comme des indicateurs de haut niveau selon la classification faite dans le projet ICALTS. Cette première étape permet de définir l'ensemble des données à collecter. Le système proposé génère automatiquement les indicateurs intermédiaires nécessaires et permet à l'enseignant de fournir des données additionnelles afin de construire l'indicateur.

- Observation : cette deuxième étape représente la phase de génération et de capture des données sur l'interaction des différents acteurs d'une situation d'apprentissage. Elle est faite par un outil installé sur les ordinateurs des apprenants.

- Analyse et interprétation : c'est l'étape dans laquelle l'indicateur est calculé et les styles d'apprentissage sont détectés. Les résultats sont ensuite présentés à l'enseignant. La liste des indicateurs a été construite en se basant sur l'analyse d'une enquête effectuée auprès des enseignants sur la nature des retours qu'ils souhaitent avoir sur les activités de leurs étudiants [Bousbia, 2008]. Ce questionnaire est divisé en trois parties : le temps de la formation, les interactions de l'apprenant et le parcours pédagogique. Sur la base de cette enquête, les indicateurs les plus signifiants sont identifiés.

### **Réutilisation des indicateurs pour la supervision de l'apprentissage**

Les travaux de thèse de F. Diagne [Diagne, 2009] sont motivés par la complexité de la définition des indicateurs par des experts. Diagne précise que certains indicateurs nécessitent des années de recherche. Elle cite dans [Diagne, 2008] l'exemple de l'indicateur cognitif identifiant les mauvaises conceptions des élèves lors de la résolution d'exercices en électricité qui fait l'objet d'une thèse [Michelet et al., 2007]. La seconde motivation concerne le nombre important d'indicateurs produits par les équipes de recherche durant des années de recherche, comme par exemple, les trente-deux indicateurs définis dans le projet ICALTS. Ces travaux sont donc orientés vers la réutilisation d'indicateurs. Afin de couvrir tous les aspects d'une activité d'apprentissage, [Diagne, 2009] propose le modèle CAS (Cognitif-Activité-Social). Il partage deux types d'indicateurs parmi les trois identifiés dans le projet ICALTS, les types Cognitif et d'Activité. Les indicateurs sont classés en trois types :

- des indicateurs cognitifs qui donnent des informations sur l'acquisition des connaissances lors de leur utilisation dans des activités d'apprentissage ;

- des indicateurs d'activités qui informent sur l'utilisation des ressources pédagogiques et outils informatiques disponibles sur la plateforme d'apprentissage ;

- et des indicateurs sociaux qui portent sur les relations dans un groupe d'apprenants, les positions que les apprenants occupent au sein de ce groupe, leur participation au travail collaboratif et le travail du groupe.

### **Des indicateurs pour la supervision des activités de communication**

Les activités de communication, entre apprenants ou entre apprenants et tuteurs, sont largement utilisées dans le domaine de l'apprentissage à distance. Ce genre d'activité favorise l'apprentissage en proposant des moyens alternatifs pour, d'une part, échanger des idées entre les différents apprenants participant à une situation d'apprentissage, et d'autre part, fournir un espace d'échange entre apprenants et enseignants. Plusieurs travaux de recherche s'intéressent à l'observation ou à l'analyse de ce genre d'activité. Le travail présenté dans [May, 2008] traite deux points. Premièrement, le traçage d'une activité de communication afin de produire des traces riches, et deuxièmement, la proposition de l'architecture d'un système d'observation des activités de communication médiatisées. En ce qui concerne le premier point, toute interaction est tracée par le système proposé. Les interactions sont classées en trois grandes catégories : les interactions homme machine, les interactions entre des humains par l'utilisation de machines et les interactions inter-machines. Le traçage de ces interactions se fait par des outils s'exécutant sur les machines des différents acteurs, et le serveur de communication. Les données générées sont par la suite analysées et peuvent être visualisées par les enseignants/tuteurs ou par les apprenants [May, 2008].

## Discussion

Les travaux que nous avons présentés dans cette catégorie sont orientés vers la définition et le développement de langages, modèles, outils, etc. pour la définition d'indicateurs spécifiques à des besoins d'observation particuliers. Nous nous sommes intéressés plus particulièrement à trois travaux. Ce choix est justifié par la nature de la démarche abordée pour l'identification des indicateurs potentiels pour un contexte d'apprentissage particulier. Dans la première démarche, les indicateurs sont identifiés en interrogeant des enseignants. Nous faisons référence ici aux travaux de détection des styles d'apprentissage proposés par [Bousbia et al., 2008] [Bousbia, 2008]. Cette démarche consiste à travailler avec des enseignants/concepteurs pour identifier leurs besoins. Ces besoins sont donc formalisés et implémentés afin de proposer des indicateurs à destination de ces enseignants concepteurs.

La deuxième démarche est basée sur une étude théorique de l'activité de tutorat. Nous faisons ici référence aux travaux présentés dans [Diagne, 2008] [Diagne, 2009]. Les indicateurs jugés pertinents sont identifiés en étudiant théoriquement une activité pédagogique particulière. Quant à la troisième démarche, les indicateurs sont définis pour fournir des informations sur l'utilisation d'un outil particulier, c'est le cas des indicateurs sur les activités faites sur un forum [May, 2008]. Ces indicateurs peuvent être identifiés en analysant les fonctions ou les services que propose l'outil informatique afin de fournir des informations aux enseignants/tuteurs leur permettant de mieux comprendre le déroulement effectif de la situation d'apprentissage.

Ces démarches peuvent être pertinentes dans un contexte où l'observation est focalisée sur quelques aspects d'une situation d'apprentissage ou sur quelques types d'activité. Cette approche limite la portée de l'observation aux contextes des indicateurs disponibles. Les travaux que nous avons cités ne présentent pas de possibilités de prise en main de ces indicateurs par des enseignants/concepteurs afin de les adapter ou de les faire évoluer selon l'évolution de leurs besoins.

## 2.2 L'exploitation des indicateurs, la perception comme aide à l'analyse

Beaucoup de travaux ont été réalisés dans le cadre de projets autour de l'observation dans un environnement d'apprentissage. Il est nécessaire de les étudier afin de comprendre leur apport et de déceler leurs limites.

Les travaux conduits dans ce sens peuvent être classés selon différents critères (voir Annexe B) :

- Unicité/multitude des plateformes d'apprentissage,
- Moment de l'observation (temps réel, après session),
- Unicité/multitude des mécanismes de perception,
- Destinataire des indicateurs (principal, exclusif),
- Données considérées (EIAH, Système Informatique au sens large).

Afin d'exploiter au mieux les différents critères selon lesquels nous pouvons considérer ces travaux de recherche qui nous intéressent pour la suite de notre travail, nous avons combiné ces différents critères afin de répartir l'ensemble des travaux étudiés en différents sous-ensembles ayant en commun plusieurs de ces critères. La combinaison de ces critères permet d'obtenir différentes classifications de ces travaux de recherche s'intéressant à la visualisation des indicateurs.

### 2.2.1 Perception en temps réel de l'observation d'un scénario sur plateforme unique

Parmi les travaux que nous avons étudiés, certains sont spécifiques à une seule plateforme d'apprentissage. En plus de cette caractéristique, ils permettent de visualiser les indicateurs en temps réel, c'est à dire au moment même du déroulement de la situation d'apprentissage.

Dans [May, 2009] et [May et al., 2011], l'auteur s'intéresse à la visualisation des résultats de l'observation en temps réel. Pour supporter l'observation, une plateforme nommée « TrAVis » (*Tracking Data Analysis*)

and Visualization tools) a été proposée. Des expérimentations de cet outil ont été réalisées sur des sessions implantées dans la plateforme Moodle. L'observation se focalise sur une activité particulière : l'utilisation d'un forum. Dans ce travail, les traces sont générées par l'apprenant lors de la lecture de messages postés dans le forum. Les destinataires des traces sont multiples : l'apprenant, l'enseignant ou le chercheur. La manipulation de la trace est faite en temps réel et il s'agit de calculer de manière automatique le nombre de messages lus par personne ainsi que le temps qui aurait pu être passé pour la lecture d'un message. En effet, il n'est pas possible d'affirmer avec certitude qu'un apprenant a lu un message. Par contre, il est possible d'indiquer si un apprenant donné n'a pas vu un message ou une partie de message. La figure 2.10 montre une visualisation du temps passé par lecteur, sur un message donné.

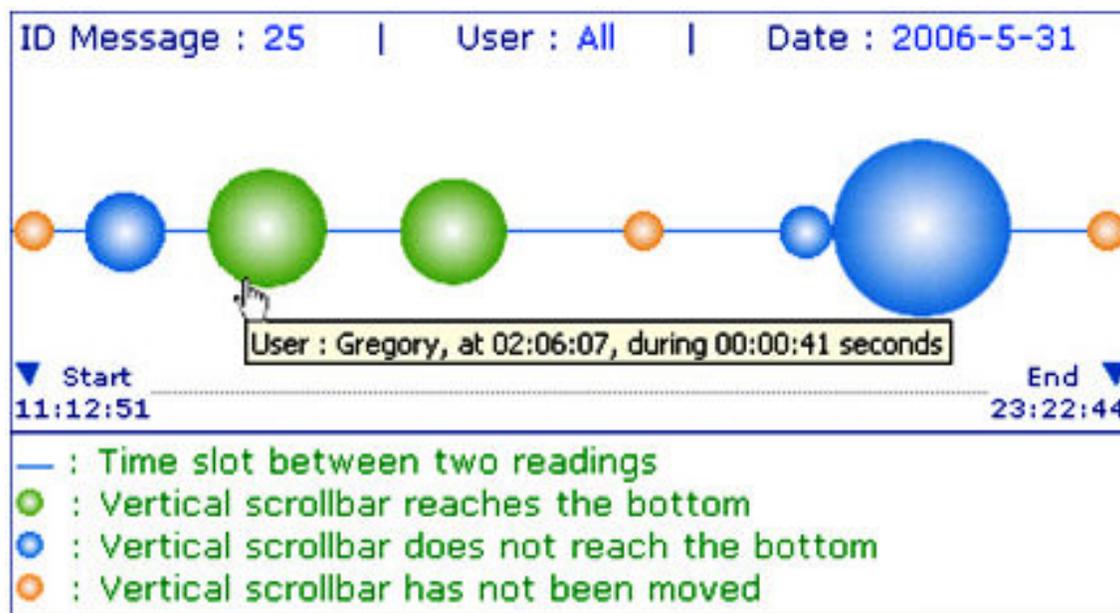


FIGURE 2.10 – Visualisation de l'indicateur « Temps de lecture d'un message » (source : [May et al., 2011])

Comme le montre la figure 2.10, les sphères (chacune correspond à un utilisateur) avec quatre couleurs (chacune correspond à un type d'action faite par l'utilisateur) sont utilisées pour visualiser en temps réel la lecture d'un message ainsi que le temps passé (le diamètre de la sphère) pour la lecture d'un message dans un forum de discussion. La distance entre les sphères correspond au temps écoulé entre deux lectures. Cette visualisation permet une interprétation en détail des différentes actions de lecture faites par les apprenants.

L'apprenant peut utiliser ces informations dans un objectif d'autorégulation : il peut se demander par exemple s'il a lu tous les messages. L'enseignant peut également utiliser ces informations afin de suivre ou de réguler l'interaction : quels sont les apprenants qui ne passent pas suffisamment de temps à lire les messages ? et de quels messages s'agit-il ? Muni d'une telle information, l'enseignant peut intervenir en temps réel.

Pour leur part, les travaux du projet FORMID (FORMation Interactive à Distance) [Guéraud et al., 2004] et [Guéraud et al., 2011] ont conduit au développement d'une plateforme permettant aux formateurs de proposer aux apprenants des activités sur des Objets Pédagogiques Interactifs (par exemple les simulations). La plateforme permet aussi d'assister les formateurs dans leur tâche de supervision et de suivi de la situation d'apprentissage. La plateforme FORMID offre deux interfaces de visualisation différentes : l'interface destinée à l'enseignant (poste formateur) et l'interface destinée à l'apprenant (poste apprenant).

Le poste formateur permet à l'enseignant qui assure le suivi d'une séance de consulter les informations relatives à la séance (vérifier la filière, le groupe, le module et le numéro de séance), de suivre l'avancement

du groupe dans la réalisation des activités de la séance, et ce, en examinant de façon détaillée la progression d'un ensemble d'apprenants ou d'un apprenant particulier et de communiquer avec les apprenants via des messages textuels.

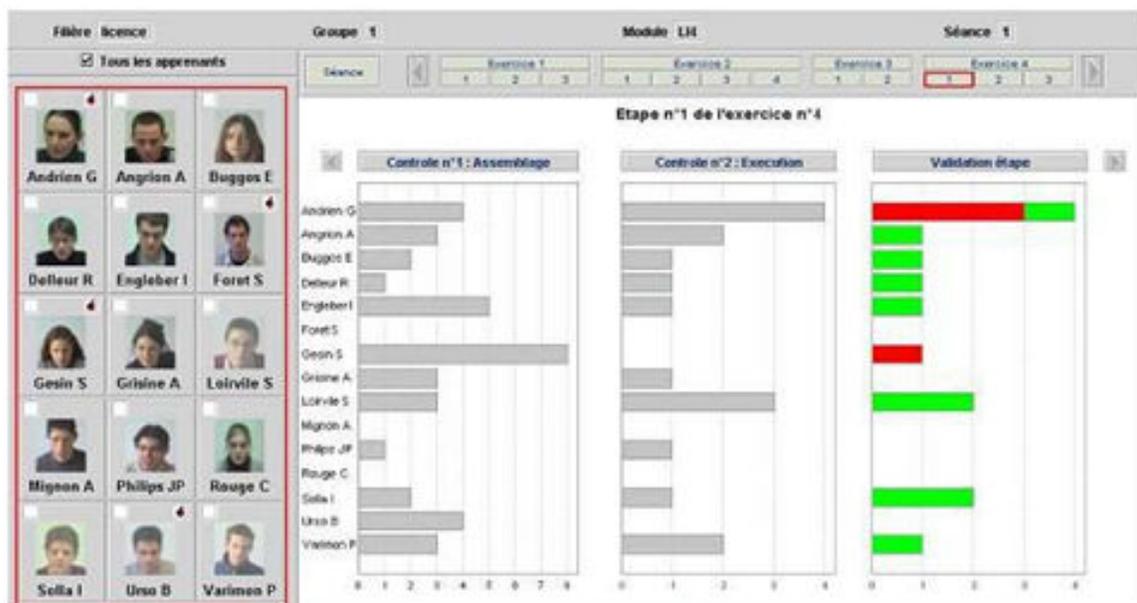


FIGURE 2.11 – Poste formateur, vision de la progression de tous les apprenants sur l'étape 1 de l'exercice 4 de la séance 1 suivie par le groupe 1 dans le module LI4 (source : [Guéraud et al., 2004])

Les apprenants à leur tour sont dotés d'une interface différente leur permettant de sélectionner l'exercice à résoudre, sélectionner l'étape de l'exercice, valider les tâches réalisées et consulter les informations relatives aux exercices sélectionnés.

Par ailleurs, les travaux de Teutsch et al [Teutsch et al., 2004] se sont intéressés à la conception d'outils dédiés à l'activité de tutorat en ligne. L'objectif est de faciliter la perception du tuteur sur la situation de l'apprenant dont il assure le suivi : historique des dernières sessions et de l'activité de l'apprenant (parcours prévu, trajet effectué, ressources sollicitées), inventaire des productions récentes, etc. Les travaux réalisés ont permis l'élaboration d'un environnement de tutorat dédié à l'accompagnement des apprenants utilisant le dispositif de formation en langues « Croisières ».

La page d'accueil présente au tuteur un tableau de bord contenant, entre autres, les événements récents (messages, points de blocage, productions) et la synthèse des trajets des apprenants (Figure 2.12). Les divers modules sont coloriés en fonction de leur taux de fréquentation (en jaune), de la participation aux activités (en orange) et de la réussite à ces activités (en vert).

D'autres auteurs, comme [Teplovs et Scardamalia, 2007] [Schreurs et al., 2013] utilisent les données des fichiers de traces d'un forum pour générer une visualisation interactive des messages aux propriétés similaires (même auteur, même fil de discussion, proximité sémantique). Ils ont développé un outil de visualisation pour un forum dédié à l'apprentissage. Dans la figure 2.13, tous les messages laissés par les apprenants et contenant le mot 'light' sont surlignés.



FIGURE 2.12 – Tableau de bord de l’activité de tutorat du dispositif « Croisières » (source : [Teutsch et al., 2004])

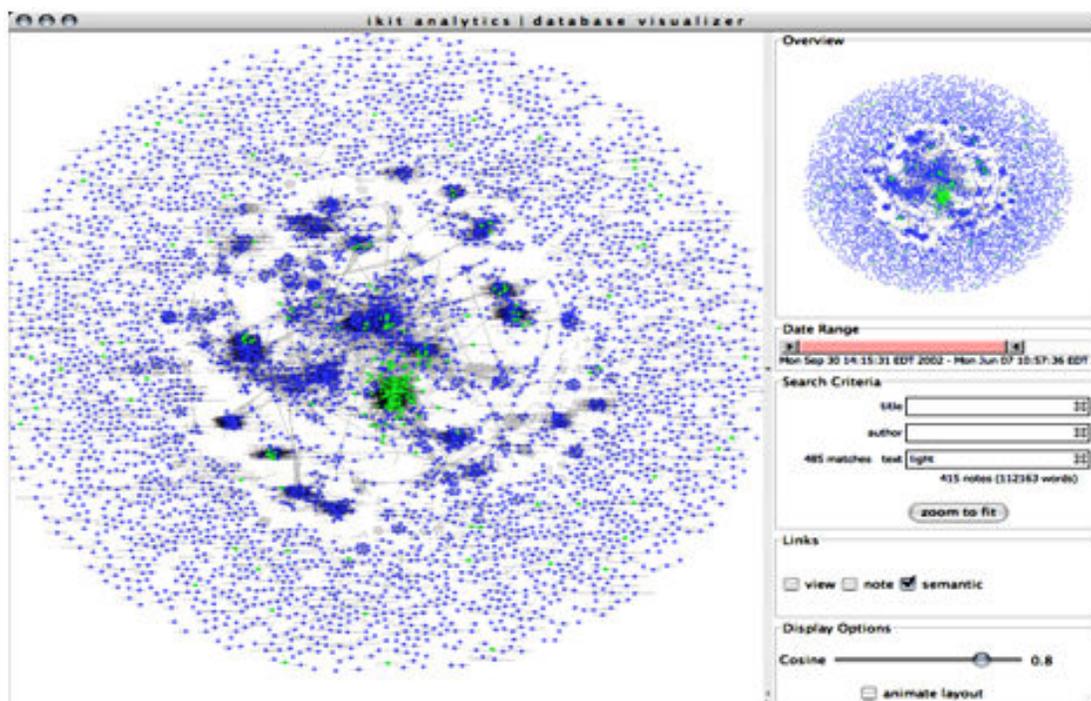


FIGURE 2.13 – Exemple de visualisation de messages contenant des mots similaires (source : [Teplovs et Scardamalia, 2007])

## Discussion

L'outil présenté dans [May et al., 2011] s'intéresse au suivi en temps réel des sessions d'apprentissage en se limitant à l'analyse d'une seule activité : un forum. Cet outil présente l'avantage de suivre les parcours des apprenants en temps réel et permet d'intervenir pour leur apporter des aides personnalisées. Toutefois, l'intervention du tuteur ne peut se faire totalement en temps réel car il faut qu'il y ait suffisamment de données recueillies ce qui amène à un décalage dans le temps. On note aussi, l'usage d'un seul moyen de visualisation pour la restitution des indicateurs.

Pour ce qui est de la plateforme proposée par [Guéraud et al., 2004] et [Guéraud et al., 2011], celle-ci offre plusieurs avantages, dont celui d'assurer le suivi des groupes d'apprenants lors du déroulement des sessions d'apprentissage. Elle offre aussi la possibilité d'interagir à travers l'envoi de messages. On constate néanmoins que cette plateforme n'offre la possibilité de visualisation que de la progression des apprenants dans la réalisation des activités. Elle n'offre pas la possibilité, au tuteur comme au concepteur, d'exprimer des objectifs propre à lui ni de définir de nouveaux indicateurs relatifs aux activités propres à un scénario d'apprentissage particulier. Les interfaces de visualisation dédiées sont développées spécialement pour l'architecture de cette plateforme. Leur intégration dans un tableau de bord plus générique de visualisation nécessite un travail supplémentaire de développement.

Pour sa part, l'outil proposé dans [Teutsch et al., 2004] permet au tuteur d'une situation d'apprentissage en ligne de bien suivre les apprenants et de mieux les accompagner dans la réalisation de leurs activités. Ce travail, étant destiné aux tuteurs, n'offre pas la possibilité aux concepteurs d'exploiter ces informations à des fins de réingénierie.

L'outil proposé dans [Teplovs et Scardamalia, 2007] [Schreurs et al., 2013] est destiné aux chercheurs s'intéressant aux phénomènes de similarité, entre les messages échangés dans un forum, à des fins d'analyses répondant à des objectifs propres. Son exploitation par les enseignants-concepteurs n'est pas envisageable sans travail d'adaptation et d'extension des fonctionnalités. Tout comme le dispositif « TrAVis » évoqué plus haut, l'outil présenté dans cette section s'intéresse aux activités d'un forum. La différence se situe au niveau de l'information recherchée et aussi au niveau des destinataires des résultats de l'observation.

### 2.2.2 Perception après session de l'observation d'un scénario sur plateforme unique

Certains des travaux que nous avons étudiés offrent la possibilité d'une observation après session et permettent ainsi de fournir des indicateurs statistiques sur les sessions d'apprentissage qui se sont déroulées.

L'outil « GISMO » (*Graphical Interactive Student Monitoring System for Moodle*) [Mazza et Milani, 2004] [Mazzola et al., 2010] permet d'extraire les traces des activités des apprenants sur Moodle et les affiche sous la forme d'un graphe représentant le nombre d'accès aux différentes ressources Moodle et les dates de ces accès par les différents apprenants. Les informations visualisées permettent à l'enseignant de mesurer la popularité des ressources mises à la disposition des apprenants et de détecter les styles d'apprentissages (apprenants qui consultent souvent des ressources et d'autres qui ne les consultent que les jours où des activités à réaliser sont demandées). GISMO permet aussi, à l'enseignant, de visualiser le nombre de messages envoyés ou reçus par chaque apprenant ainsi que le nombre de discussions initiées ou commencées par un apprenant.

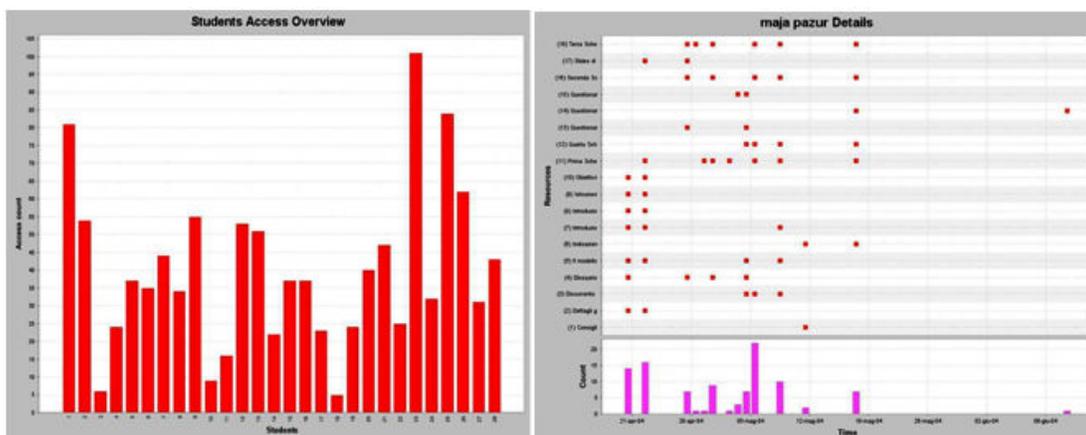


FIGURE 2.14 – Deux exemples d’accès des apprenants aux ressources (source : [Mazzola et al., 2010])

La figure 2.14 montre deux exemples d’accès des apprenants aux différentes ressources mises à leur disposition. L’image de gauche représente le nombre d’accès effectués par les apprenants (axe des abscisses) à toutes les ressources d’un cours (axe des ordonnées). L’image de droite montre en détail les ressources consultées par un apprenant particulier à chaque date du cours. L’histogramme représente le nombre total d’accès réalisés par l’apprenant à toutes les ressources de cours

Les auteurs de [Heraud et al., 2005] et [Marty et al., 2009] étudient les traces laissées par les apprenants pour pouvoir comprendre et interpréter leurs comportements. La représentation des traces se fait suivant un axe temporel horizontal. La figure 2.15 présente les traces de l’observation de la réalisation des tâches d’une activité d’apprentissage, par l’utilisation des barres d’ombres. L’absence de zones d’ombres garantit une faible probabilité d’avoir mal observé l’activité d’apprentissage.

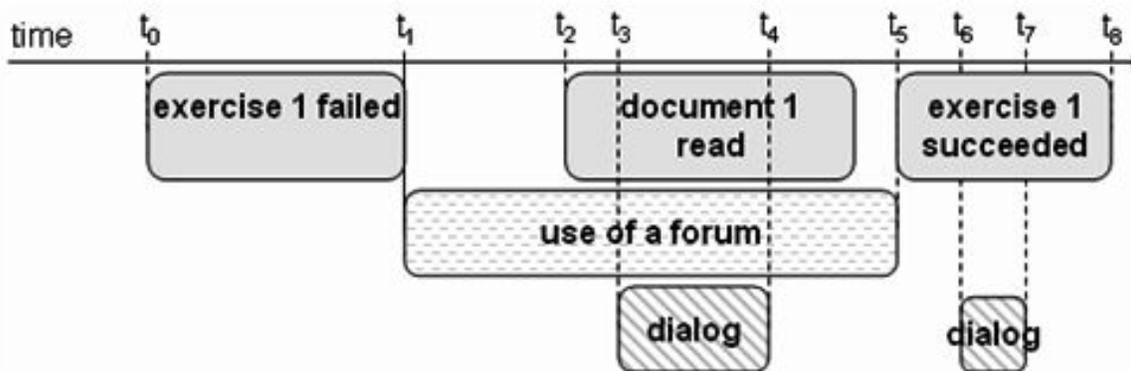


FIGURE 2.15 – Représentation des traces suivant un axe temporel horizontal (source : [Heraud et al., 2005])

Pour ce qui est des travaux sur l’étude des interactions et la collaboration dans l’apprentissage dans le développement des connaissances, Les auteurs [Damsa et Ludvigsen, 2010] [Damsa et al., 2014] ont conçu des outils d’analyse et d’observation intégrés dans le système KPE (*Knowledge Practices Environment*) qui est un environnement web 2.0 offrant un espace virtuel de collaboration entre apprenants. Les traces collectées dans cet environnement ont permis de construire des aperçus généraux sur le nombre et le type de contributions de chaque participant. L’outil développé est doté d’un analyseur visuel des contributions par utilisateur

et d'un analyseur chronologique qui permet de représenter les distributions de types de contribution dans le temps.

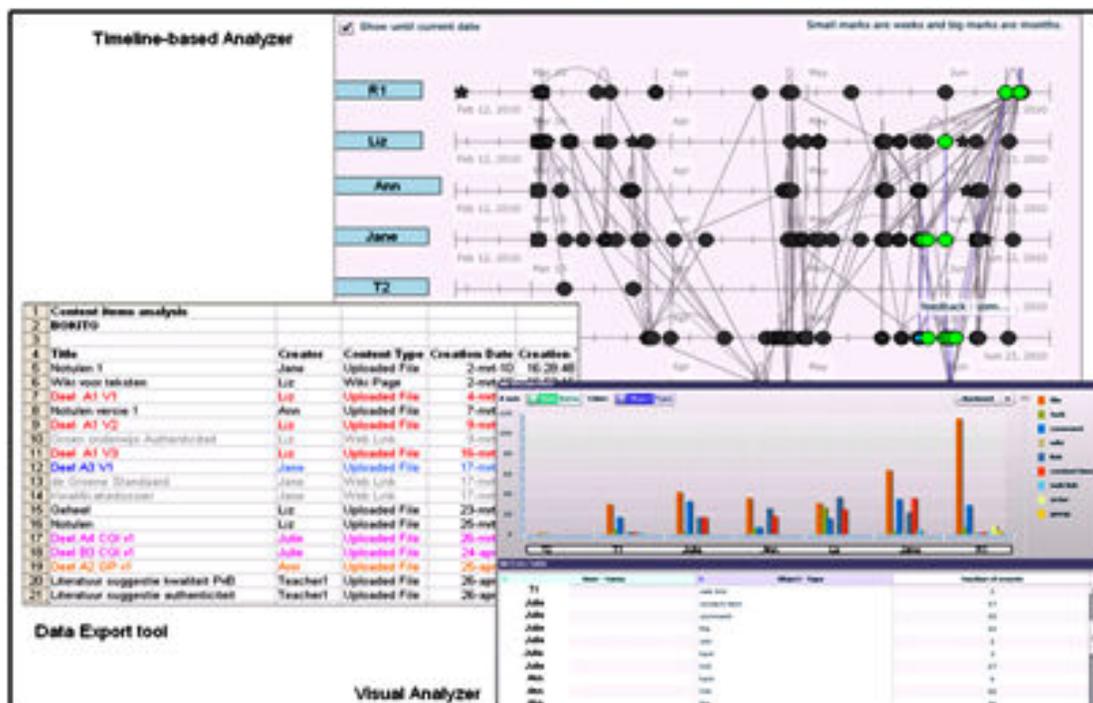


FIGURE 2.16 – analyseur chronologique des contributions de KPE (Knowledge Practices Environment) (source : [Damsa et Ludvigsen, 2010])

Dans [Djouad, 2008] et [Djouad et al, 2011], l'auteur propose un outil générique spécialisé à Moodle pour calculer les indicateurs d'interactions à partir des traces modélisées. Le système proposé est « SBT-IM » (Système à Base de Traces pour calculer les Indicateurs dans Moodle). « SBT-IM » se décompose en trois sous-systèmes : le premier permettant la collecte des données à partir des plateformes d'apprentissages en une trace première, le second gère la transformation des M-Traces, tandis que le troisième permet le calcul des indicateurs à partir des M-Traces transformées. L'interface graphique associée à « SBT-IM » permet aux utilisateurs du système d'interagir avec les services de collecte des données, de transformation des M-Traces, et de calcul de l'indicateur.

La figure 2.17 illustre le calcul d'un l'indicateur. L'interface communique avec la base des M-Traces (Traces spécifiques à Moodle) et appelle : la fonction Compter (Count), le module qui charge les M-traces, ainsi que le module de calcul. Un tableau affiche la liste des M-Traces avec le nombre des instances de chaque M-Trace. Le module de calcul propose deux méthodes pour calculer la valeur de l'indicateur : une valeur numérique avec un calculateur arithmétique, ou un affichage de la valeur de l'indicateur dans un « Secteur » ou un « Histogramme ».

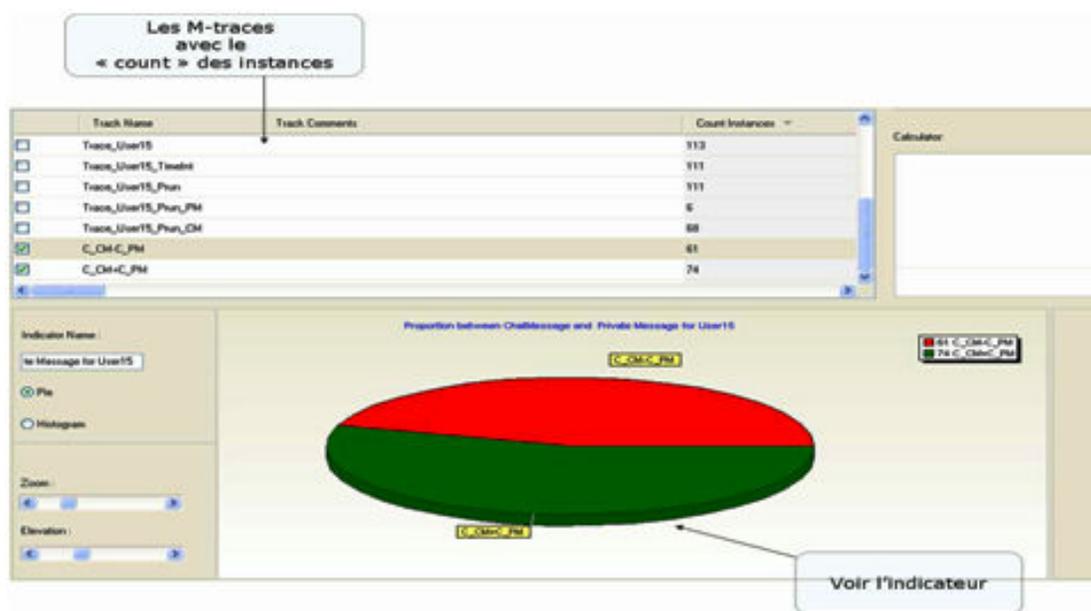


FIGURE 2.17 – exemple de calcul de proportion entre deux types d'indicateurs (source : [Djouad, 2008])

## Discussion

L'outil présenté dans [Mazza et Milani, 2004] et [Mazzola et al., 2010] offre une réelle opportunité pour évaluer la popularité des ressources mise en ligne et pour mesurer les aspects liés à la collaboration entre apprenants par le biais de l'observation des messages échangés mais n'offre pas la possibilité de l'exploiter en temps réel lors du déroulement d'une session d'apprentissage. Nous constatons aussi la non portabilité car il est spécifique à la plateforme Moodle. Il y a aussi le nombre réduit d'indicateurs visualisés par cet outil et son aspect porté principalement sur la visualisation du comportement relationnel entre apprenants (discussions initiées, participants, messages échangés, etc.) et non sur le déroulement des sessions d'apprentissage ni sur les activités réalisées ou les productions faites en résolvant des exercices.

Autrement, le travail présenté dans [Heraud et al., 2005] et [Marty et al., 2009] a pour objectif la détection des comportements des apprenants en analysant les traces générées en fin de session mais n'offre pas la possibilité de suivre la situation d'apprentissage en temps réel et ne permet pas à l'enseignant de choisir ou de définir les indicateurs à visualiser mais se limite à lister, suivant l'axe temporel horizontal, les activités réalisées par les apprenants.

Pour ce qui est du travail présenté dans [Damsa et Ludvigsen, 2010] et [Damsa et al., 2014], ces derniers s'intéresse à la collaboration entre les apprenants et tente de mesurer la part de participation de chacun de ces apprenants à la collaboration. Il est très utile dans la mesure où on essaye d'analyser un travail collaboratif en fonction des apports individuels des participants. Néanmoins, on constate que les informations visualisables se limitent à cet aspect collaboratif sans tenir compte de la nature des activités et du scénario d'apprentissage. Il faut noter aussi que l'analyse intervient en fin de session. Les outils de visualisation proposés dans ce travail seraient d'un apport considérable pour l'amélioration de la compréhension des indicateurs une fois intégrés et exploités par d'autres sources de données.

Quant au travail proposé dans [Djouad, 2008] [Djouad et al, 2011], il offre la possibilité de visualiser sous forme graphique des indicateurs calculés à partir des traces spécifiques à la plateforme Moodle. On constate que l'outil proposé est spécifique à Moodle et nécessite un travail d'adaptation pour pouvoir fonctionner avec les traces des autres plateformes d'apprentissage. Les indicateurs visualisés se limitent à ceux embarqués dans la plateforme et on ne retrouve pas cette notion d'adaptation des indicateurs et des visualisations au contexte de la situation d'apprentissage.

### 2.2.3 Perception en temps réel de l'observation d'un scénario sur des plateformes multiples

Certains travaux que nous avons étudiés permettent une visualisation des indicateurs en temps réel. Ce qui les différencie des travaux présentés précédemment, c'est leur instrumentation à l'aide d'outils supportant les traces issues de plusieurs plateformes d'apprentissage différentes. Ils ne sont donc pas dédiés à un seul EIAH.

L'outil « REFLET » [Després et Coffinet, 2004] permet de visualiser l'évolution du travail d'un apprenant ou d'une promotion d'apprenants en formation ouverte à distance (FOAD). Initialement intégré dans la plateforme WebCT (aujourd'hui Blackboard), REFLET est désormais intégrable dans n'importe quel site web de formation. Les interfaces tuteurs et apprenants sont quasiment les mêmes. Le tuteur peut visualiser l'état d'avancement d'un apprenant dont il a la charge ou comparer l'avancement de tous les apprenants d'un même groupe. Il peut également contre-valider des tâches déjà validées par les apprenants. L'apprenant peut visualiser son état d'avancement dans la réalisation des activités et savoir si une activité validée précédemment a été contre-validée par le tuteur ce qui lui permet de reprendre cette activité dans le cas où le tuteur ne l'a pas validée. Il peut comparer son état d'avancement à celui des autres apprenants qui restent anonymes à son niveau. La figure 2.18 illustre l'état d'avancement d'un apprenant vu du côté de l'interface du tuteur.



FIGURE 2.18 – Visualisation du détail d'une activité pour un apprenant (source : [Després et Coffinet, 2004])

Pour leur part, les travaux réalisés par le laboratoire SysCom dans [France et al., 2006] [France et al., 2007] sur la visualisation des traces permettent au tuteur de voir en temps réel les activités des apprenants, les parcours effectués et les activités réalisées ou en cours de réalisation au sein de l'EIAH, et d'adapter l'activité en interagissant directement sur l'interface de visualisation. La représentation des traces par des outils de visualisation permet à l'enseignant/tuteur de comprendre ce qui se passe réellement, pour pouvoir réagir avec pertinence et suivre l'activité des apprenants en session. La figure 2.19 montre l'outil de visualisation « Classroom » qui utilise les visages de Chernoff où les bulles représentent des activités et les liens représentent les parcours entre activités. Les apprenants sont représentés par des visages.

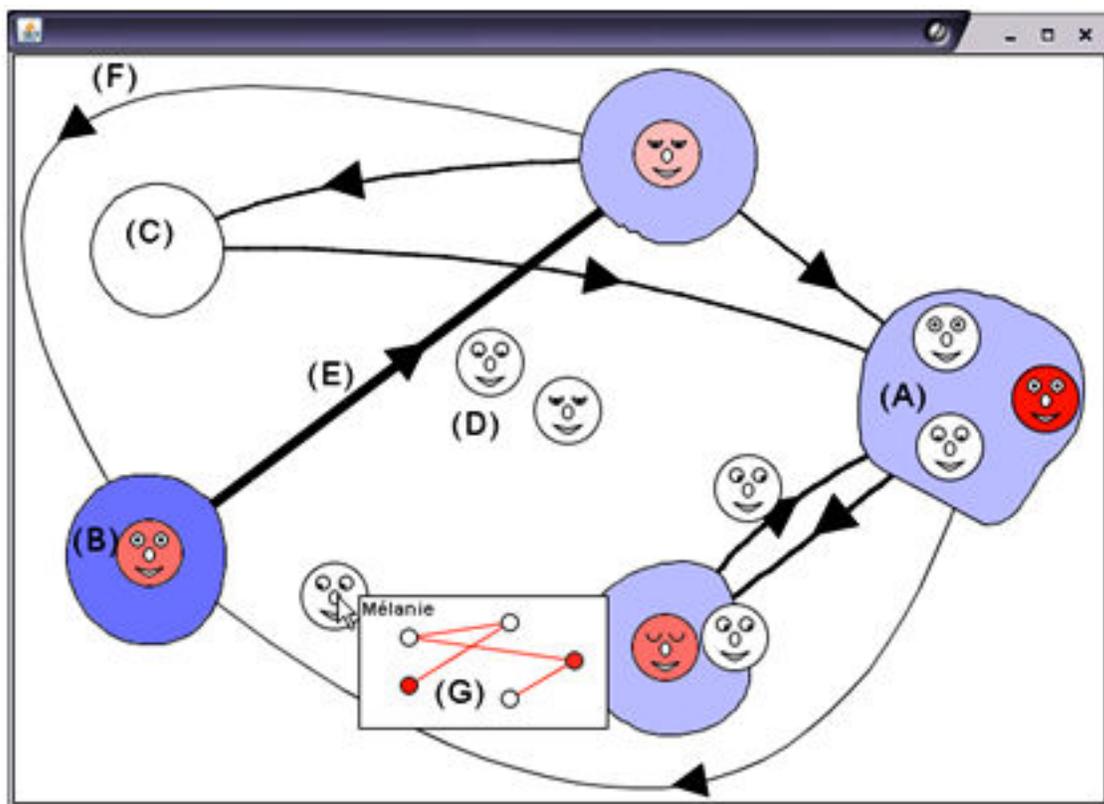


FIGURE 2.19 – Outil de visualisation “Classroom” de l’enseignant basé sur les Chernoff faces (source : [France et al., 2007])

Exemple d’interprétations de cette image : - Mélanie est en train de passer d’une activité à l’autre. Elle est déjà passée dans toutes les activités et on constate (G) qu’elle a déjà passé beaucoup trop de temps dans les activités (A) et (B) qui sont en rouge dans son diagramme. - Les activités en fond bleu montrent qu’en moyenne les apprenants dépassent le temps imparti pour la réalisation de ces activités. - (C) est l’activité compte-rendu et personne n’est en train de la faire actuellement. - (A) est actuellement réalisée par 3 apprenants dont un depuis trop longtemps et un quatrième est en train de passer à cette activité (lien).

Ces différentes visualisations évoluent dynamiquement dans le temps, en fonction des traces laissées par l’apprenant pendant la session. Par exemple, la forme des yeux du « visage » d’un apprenant donné (exemple : fermée, ouverte) correspond à la quantité d’évènements produits par l’apprenant pour une certaine activité. Si un apprenant « dort » dans une activité particulière, le tuteur peut intervenir en lui suggérant de participer plus activement.

Par ailleurs, dans [Avouris et al., 2007] et [Rummel et al., 2011], les auteurs ont étudié les traces des activités en incluant les enregistrements vidéos et audio afin de pouvoir construire une vue complète de l’activité de collaboration et les comportements des apprenants lors d’une situation d’apprentissage, en ne se limitant pas seulement à l’analyse des fichiers de logs. Les auteurs décrivent un outil d’analyse des traces ColAT (Collaborative Analysis Tool). ColAT permet de visualiser les traces sous différentes formes incluant même la vidéo.

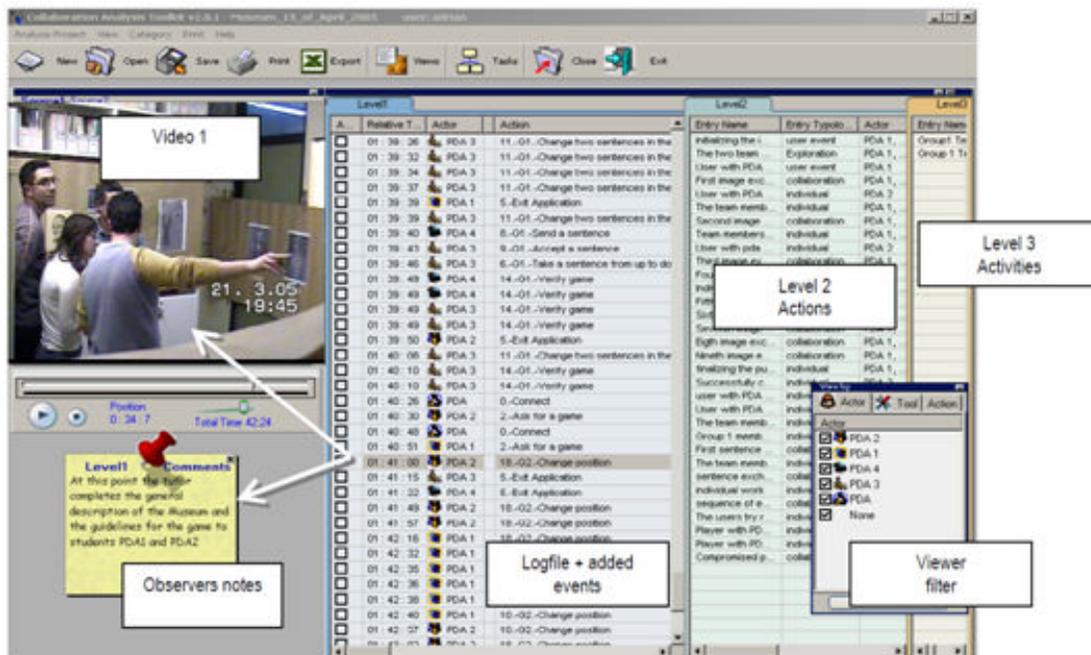


FIGURE 2.20 – L’environnement ColAT : observation de la réalisation d’une activité, (extrait de l’étude des activités d’apprentissage dans un musée) (source : [Avouris et al., 2007])

Dans la figure 2.20 extraite de l’observation des activités d’apprentissage dans un musée, réalisé avec l’outil ColAT, plusieurs types de visualisations sont présentées : les fichiers de logs, les actions effectuées, les activités réalisées ainsi qu’une vidéo reprenant le déroulement de la situation d’apprentissage.

Un autre travail entrant dans cette catégorie : Tatiana (*Trace Analysis Tool for Interaction Analysts*) [Dyke et al., 2010]. C’est un environnement logiciel basé sur le modèle de description des activités d’apprentissage collaboratif médiatisé par ordinateur. Ce modèle est conçu pour aider les chercheurs à gérer, synchroniser, visualiser et analyser leurs données.

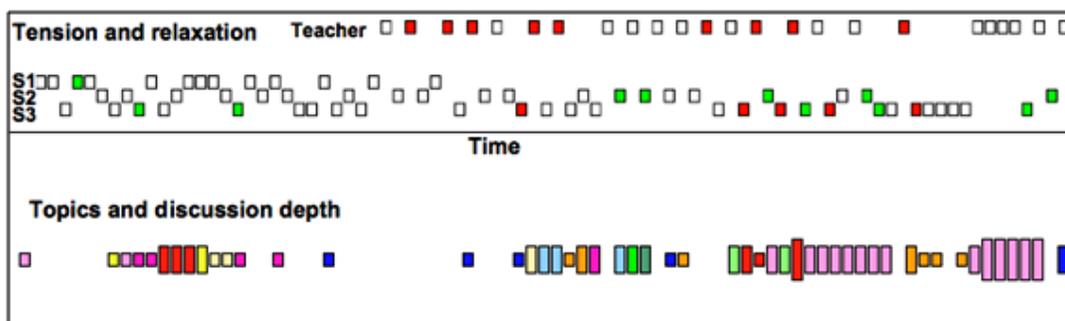


FIGURE 2.21 – Deux visualisations examinées de manière synchronisée au fil du temps. En haut : les rangées correspondent aux intervenants, le vert à la relaxation, le rouge à la tension. En bas : Les topics sont montrés de différentes couleurs, la hauteur des éléments correspond au niveau de profondeur de la discussion. (source : [Dyke et al., 2010])

Dans ce travail, les auteurs ont introduit la notion de « rejouable » qui est un type particulier d’artéfacts, temporellement orienté, qui conserve la notion d’ordre des évènements. Les rejouables peuvent être visua-

lisés dans deux types d'afficheurs : une forme tabulaire qui présente les données dans une forme similaire à Excel et une visualisation temporelle horizontale. Un mécanisme de plugin permettant d'ajouter de nouvelles formes de visualisations a été intégré dans Tatiana et un éditeur graphique permettant de générer les rejouables a été mis à la disposition des chercheurs afin de ne pas les contraindre à maîtriser un langage de requêtes. Tatiana a été intégré à plusieurs travaux de recherche. On peut citer les travaux sur l'analyse des situations de travail en groupes par dialogue oral [Baker et al., 2009] (figure 2.21), l'analyse des données issues de blogs et de Twitter afin de comprendre les structures des communautés [Lomborg, 2009] et la combinaison des compétences de plusieurs chercheurs issus de domaines différents [Prudhomme et al., 2007].

### Discussion

Dans [Després et Coffinet, 2004], l'outil présenté offre une interface conviviale et permet à l'enseignant de suivre le travail des apprenants lors d'une session qui se déroule en formation ouverte à distance et permet de suivre la situation en temps réel. Néanmoins, nous constatons qu'il ne permet de visualiser que peu d'informations (état d'avancement seulement) et limite les opérations d'intervention du tuteur à la contre-validation de certaines tâches réalisées par les apprenants.

Pour sa part, l'outil présenté dans [France et al., 2006] et [France et al., 2007] présente l'avantage de suivre les parcours des apprenants en temps réel et permet d'intervenir pour leur apporter des aides personnalisées. Cet outil est plus destiné à l'activité de tutorat qu'à l'activité de réingénierie car il n'offre pas la possibilité de manipulation par les usagers et les concepteurs de leur propre trace en temps réel. L'intervention du tuteur, aussi, ne peut se faire totalement en temps réel car il faut qu'il y ait suffisamment de données recueillies ce qui amène à un décalage dans le temps. On constate aussi, l'usage d'un seul moyen de visualisation pour la restitution des indicateurs.

Quant à l'outil de visualisation proposé dans le cadre de ce travail [Avouris et al., 2007] [Rummel et al., 2011], il offre l'avantage de combiner un ensemble de moyens de perception pour un même indicateur. Il offre aussi la possibilité de lire des enregistrements audio et vidéo en plus des traces stockées dans les fichiers de Logs. Le son et la vidéo permettent d'avoir une perception plus proche de la réalité de la situation d'apprentissage mais ne suffisent pas à comprendre le déroulement car il est nécessaire de passer à une étape d'analyse et d'étude de ces images et sons.

Pour ce qui est du travail présenté dans [Dyke et al., 2010], qui est destiné essentiellement aux chercheurs du domaine pédagogique. Il leur offre des moyens d'observation et d'analyse des situations d'apprentissage pour répondre à leurs objectifs de recherche. Ce travail a été validé par diverses expérimentations, mais son utilisation n'est pas évidente par les enseignants-concepteurs car l'ergonomie de l'interface n'a pas été l'objectif principal lors de son développement, ce qui est justifié aussi par le public auquel il est destiné.

#### 2.2.4 Perception après session de l'observation d'un scénario sur des plateformes multiples

Certains travaux que nous avons étudiés permettent une visualisation des indicateurs après le déroulement de la session d'apprentissage. Leur particularité est d'être supportés par des outils prenant en charge les données issues de plusieurs plateformes d'apprentissage différentes.

L'outil « CourseVis » [Mazza et Dimitrova, 2003] [Mazza et Dimitrova, 2004] [Mazza, 2009] est utilisé pour la visualisation 3D des traces issues de la plateforme WebCT. WebCT est un spécialiste de la formation en ligne pour universités et grandes écoles. Il édite des solutions de télé-formation permettant aux établissements de proposer à leurs apprenants des cours en ligne. WebCT est distribuée depuis 2006 sous le nom Blackboard. Le but de « CourseVis » est de faire des visualisations à partir de calcul d'indicateurs et des mesures statistiques. Les visualisations peuvent être paramétrées par les utilisateurs. CourseVis permet aux tuteurs qui n'ont pas un niveau de technicité élevé de disposer d'une sorte de tableau de bord leur permettant d'analyser le comportement des apprenants distants et leurs échanges. Cet outil n'est pas spécifique à une seule plateforme d'apprentissage mais il est intégrable dans plusieurs plateformes. La figure 2.22 représente

une visualisation des discussions réalisées par les apprenants et se rapportant aux unités de cours et aux activités réalisées par les groupes d'apprenants.

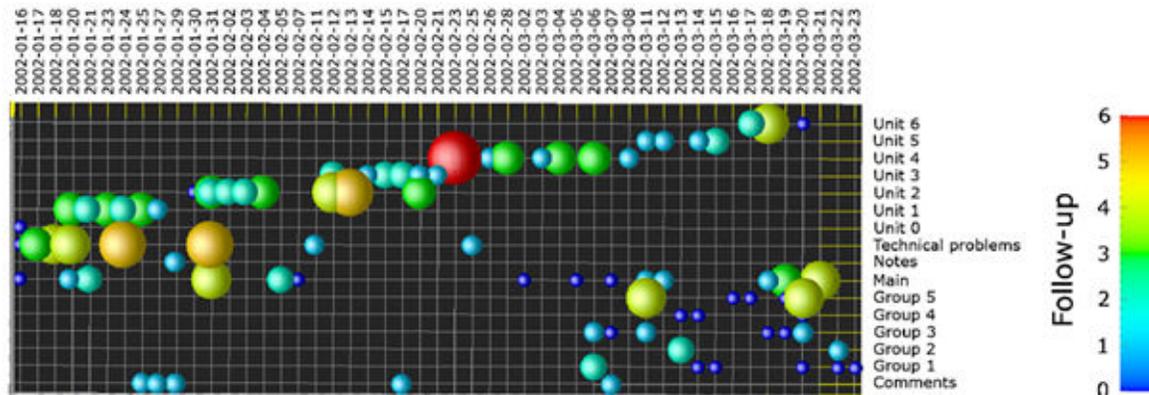


FIGURE 2.22 – Visualisation des discussions relatives aux unités de cours et aux activités de groupe (source : [Mazza et Dimitrova, 2004])

Pour leur part, Les auteurs de [Suthers et al., 2010] [Suthers et Medina, 2011] [Medina et Suthers, 2013] décrivent un cadre analytique pour les interactions médiatisées. Un autre type de visualisation nommé graphes de contingences est présenté dans ce travail où les auteurs ont introduit des annotations relationnelles dans les représentations du type axe temporel horizontal (une contingence étant un évènement imprévisible susceptible de se produire ou non). Les graphes de contingences permettent d'illustrer les relations entre les interactions et l'apprentissage ou la construction de la connaissance. La figure 2.23 illustre un fragment d'un graphe de contingences.

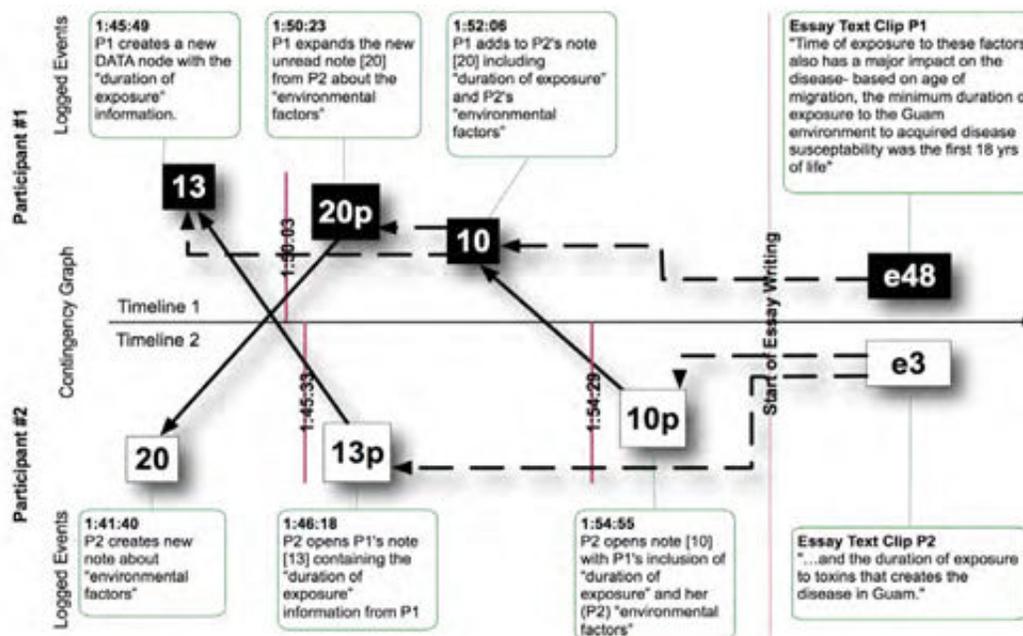


FIGURE 2.23 – Fragment d'un graphe de contingences et des évènements ayant permis de le construire (source : [Suthers et al., 2010])

Cette figure illustre la façon dont les contingences peuvent être mises en évidence par l'édition d'objets média ou par similitude lexicale. A 1 :52 :06, par exemple, P1 a ajouté un commentaire (10) au même objet « note » qu'il vient de lire à 01 :50 :23 (20p) (Un objet « note » peut contenir à la fois une séquence de commentaires des différents participants). Comme la coordination (10) n'aurait pas pu avoir lieu que si l'objet média (20p) existe, il y a synthèse d'une dépendance entre (10) et (20p). Cet exemple illustre aussi des contingences lexicales et temporelles. La coordination (10) utilise l'expression « environmental factors » qui est présente dans les notes accessibles dans (20p), offrant ainsi une contingence épigraphiques de (10) sur (20p). La coordination (10) est aussi dépendante de (13) suivant le chevauchement lexical de « duration of exposure ». (10) a lieu moins de deux minutes après (20p), ce qui peut fournir, par exemple, une preuve circonstancielle par la proximité temporelle que (10) est subordonnée à (20p).

### Discussion

L'outil présenté dans [Mazza et Dimitrova, 2003] et [Mazza, 2009] offre de nombreux avantages à travers l'interface de visualisation dédiée facilitant aux utilisateurs son paramétrage ainsi que par sa compatibilité avec plusieurs plateformes d'apprentissage. Néanmoins, il se limite à la visualisation des indicateurs initialement prédéfinis dans la plateforme qui héberge les ressources pédagogiques en ligne et ne permet pas au tuteur d'exprimer d'autres besoins d'observation spécifiques à une session d'apprentissage particulière et à un contexte de scénario pédagogique particulier.

Quant à l'outil proposé dans [Suthers et al., 2010], [Suthers et Medina, 2011] et [Medina et Suthers, 2013], il s'intéresse à l'analyse des interactions et leurs relations avec la construction de la connaissance. C'est un outil destiné aux chercheurs s'intéressant aux phénomènes d'apprentissage par le biais des interactions, plus qu'aux enseignants-concepteurs ou aux tuteurs. L'enrichissement, par des annotations sémantiques, de la présentation des interactions amène à mieux comprendre l'apport des échanges à l'apprentissage.

## 2.3 L'observation dans d'autres contextes que les EIAH

Lors de l'étude des travaux sur l'observation et la visualisation des indicateurs, nous avons voulu élargir notre champ d'étude et sortir du seul cercle des plateformes d'apprentissage afin de voir comment les autres communautés, et les autres travaux qui nécessitent le calcul et la perception d'indicateurs propres aux domaines concernés, considèrent cet aspect de l'activité de l'observation. Ce qui nous a amenés à l'étude des travaux [Morrison et al., 2006], [Greenhalgh et al., 2007], [Goodman et al., 2006] et [Georgeon et al., 2006] que nous présentons et discutons dans cette section.

Dans [Morrison et al., 2006], les auteurs présentent la plateforme *Replayer* qui offre un ensemble coordonné d'outils de visualisation permettant l'analyse et la compréhension d'informations de différentes natures. *Replayer* combine des outils de visualisation vidéo et des histogrammes ou la représentation d'évènements sur un axe temporel horizontal. La figure 2.24 illustre 5 outils de *Replayer* travaillant en coordination. La figure montre le composant vidéo qui manipule deux flux de données, la série d'évènements qui montre la trace de chaque utilisateur au fil du temps, un histogramme et des séries chronologiques montrant des informations sommaires sur une propriété du système et la carte montrant les positions des utilisateurs mesurées à l'aide d'un GPS.

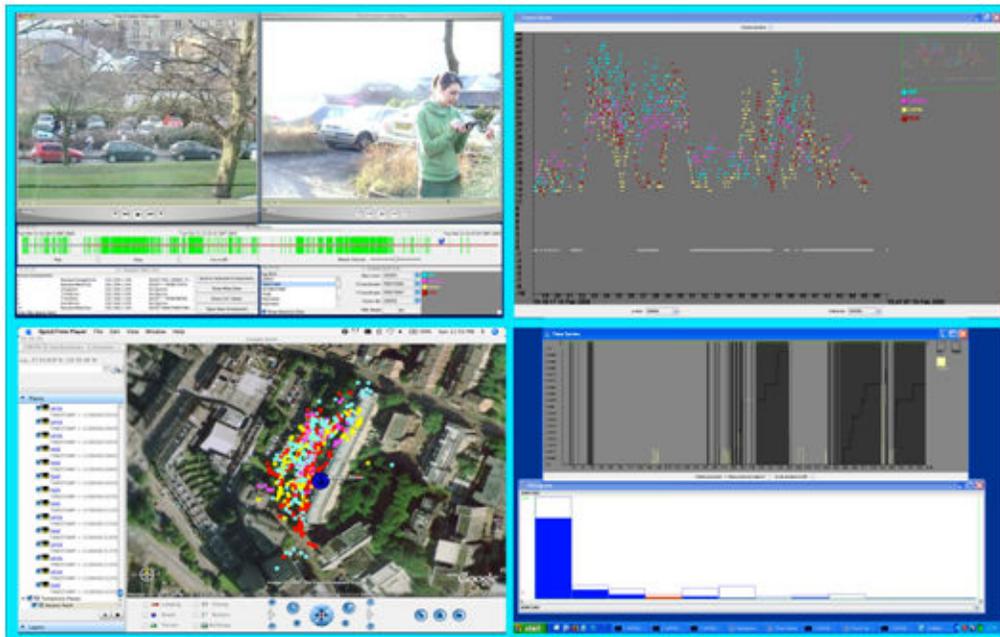


FIGURE 2.24 – Un exemple montrant les outils de visualisation offerts par la plateforme *Replayer* (les données sont tirées d'une application multi-utilisateur mobile) (source : [Morrison et al., 2006])

Par ailleurs, les travaux de Goodman et al. [Goodman et al., 2006], dans le projet *MITRE-Sponsored Research (MSR)* se sont intéressés à l'étude des échanges entre les intervenants dans une équipe de discussion pour la prise de décision sur une question posée. L'objectif étant la synchronisation des interventions afin de comprendre la dynamique de la construction des décisions. Pour mener à bien leurs travaux, les chercheurs ont développé un outil d'analyse et d'observation permettant de visualiser les flux de données synchronisés en les représentant sur un axe temporel horizontal. Cet outil a été nommé le *Synchronized Analysis Workspace (SAW)*. La visualisation des flux de données dans SAW permet d'aider les analystes dans leur processus d'étude et de compréhension des interactions entre les personnes concernées par la prise de décision, et les influences des uns sur les autres dans le cheminement des idées vers un consensus sur les questions sur lesquelles il y a divergence.

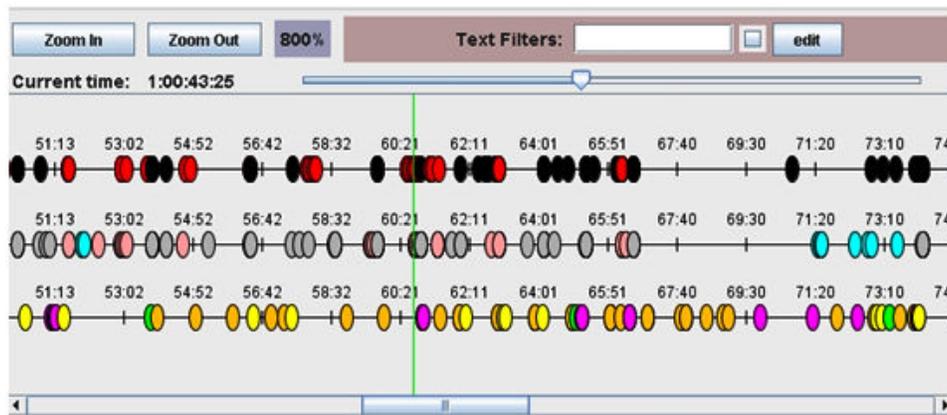


FIGURE 2.25 – Un exemple de visualisation des événements de logs sur un axe temporel horizontal (source : [Morrison et al., 2006])

La figure 2.25 montre la visualisation d'évènements survenus lors d'échanges et de coopérations, l'analyse des évènements permet de déceler les dépendances et les influences des uns sur les autres. Une réaction ou un évènement survenu beaucoup de temps après la lecture d'un message reçu pourrait avoir une signification très différente d'un évènement survenu quelques secondes après la lecture d'un message. SAW permet à l'analyste d'utiliser les représentations pour comprendre le déroulement des évènements sans lire les journaux dans la profondeur, ou pour créer des repères visuels utiles pour naviguer dans le flux de données.

L'outil *Digital Replay System* (DRS) présenté dans [Greenhalgh et al., 2007] et [Brundell et al., 2008] est une application pour la relecture et l'analyse des combinaisons d'enregistrements vidéo, des transcriptions et des fichiers de logs. DRS a été développé au sein du Digital Records pour l'e-Social Science (DRESS) au cœur du Centre national pour l'E-Sciences Sociales(NCESS). Il est actuellement utilisé pour l'observation des sessions d'apprentissage via des simulations, des sessions d'apprentissage de l'anglais parlé et même pour l'analyse de la gestuelle.

La figure 2.26 montre une vue de l'interface de l'application DRS, comprenant des graphiques représentant des traces de capteurs connectés(en haut à gauche), deux enregistrements vidéo (centre gauche), un panneau de contrôle de lecture (en bas à gauche), un aperçu chronologique (centre droit) et la transcription (en bas droite) (ces données sont issues du projet "Thrill", qui explore la nature des sensations fortes en utilisant des promenades dans des parcs à thèmes).

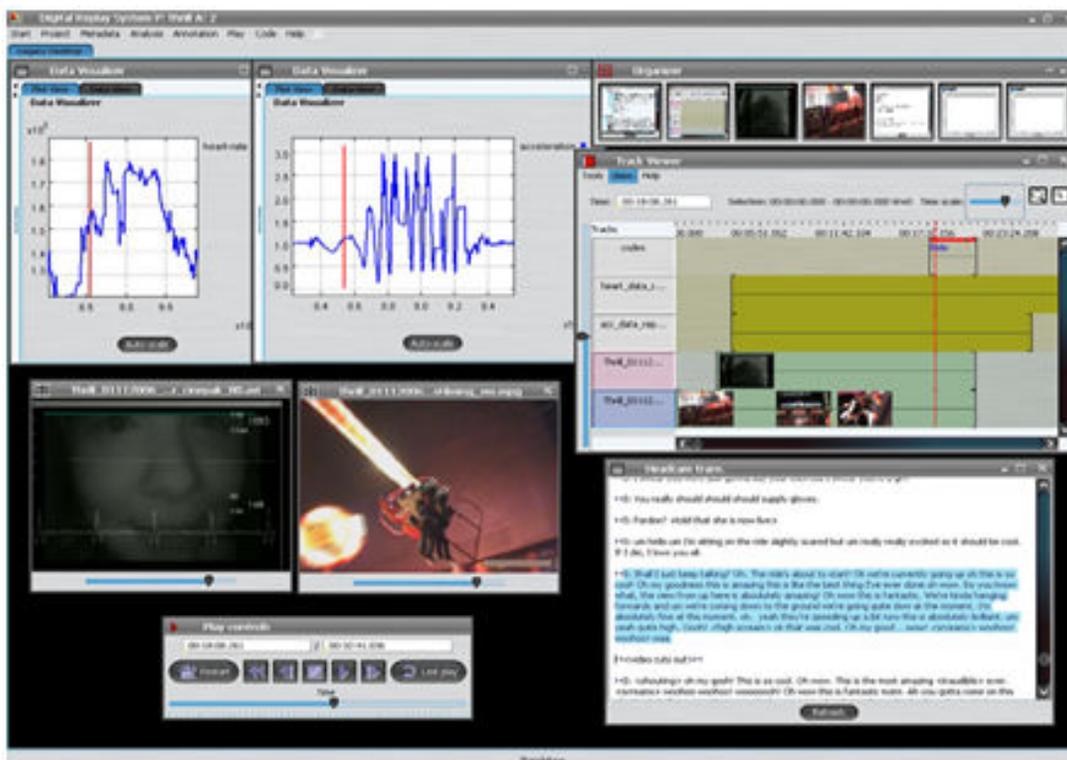


FIGURE 2.26 – Une session de DRS (source : [Greenhalgh et al., 2007])

Un autre travail étudié : *ABSTRACT (Analysis of Behavior and Situation for menTal Representation Assessment and Cognitive acTivity modeling)* [Georgeon et al., 2006][Georgeon et al., 2008] [Olivier et al., 2012] est un outil permettant d'analyser l'activité d'un acteur humain qui interagit avec un dispositif technique complexe. L'activité est observée et les traces collectées sont modélisées sous la forme d'un graphe. Les traces collectées sont essentiellement constituées de descripteurs d'évènements qui se suivent. L'objectif général d'ABSTRACT est de comprendre le comportement d'un utilisateur d'un artefact technique

complexe. Les objectifs sont des objectifs d'ergonomie, dans le cas général d'une activité humaine. Les buts peuvent être multiples : faciliter l'activité, prédire l'activité, prévenir certaines actions indésirables, etc. La figure 2.27 montre une copie d'écran de l'interface utilisateur de l'outil Abstract.

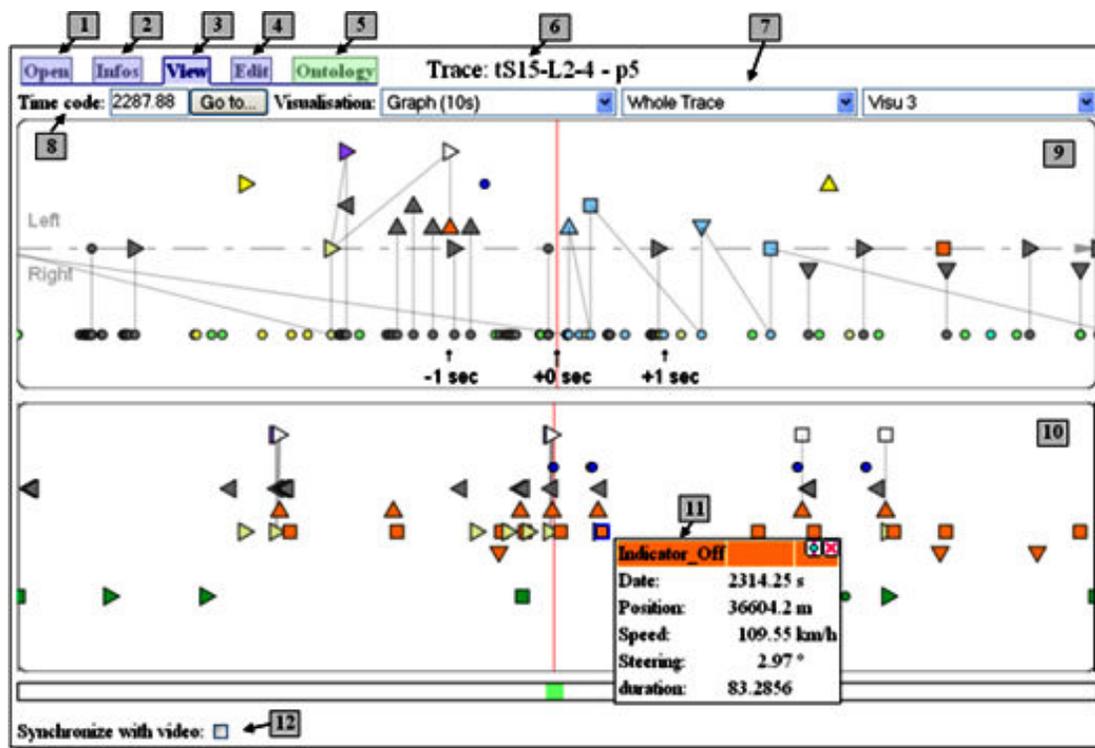


FIGURE 2.27 – Interface utilisateur d'ABSTRACT (source : [Georgon et al., 2008])

Cette interface offre les fonctionnalités suivantes : (1,6) choisir une trace, (2) voir les informations générales de cette trace, (3) visualiser la trace sous différents formats, (4) transformer la trace par l'application de règles de transformation, (5) définir les symboles dans une ontologie, (7) choisir différents modules de visualisation, (8) atteindre un time code particulier. Le module de visualisation (10) montre la trace complète avec les symboles de haut niveau, le module (9) montre un zoom sur un intervalle de 10 secondes. Un clic sur les symboles permet d'afficher une fenêtre d'information (11). L'utilisateur peut faire glisser la trace horizontalement et la synchroniser avec une vidéo (12).

## Discussion

La plateforme proposée dans [Morrison et al., 2006] offre l'avantage de pouvoir suivre des situations mobiles. Elle permet de combiner plusieurs types de visualisations intégrant même la vidéo. Un travail d'adaptation sera toutefois nécessaire pour l'adapter aux situations pédagogiques portées par les dispositifs d'apprentissage. L'un des aspects les plus intéressants est l'exploitation mobile de l'application. Son adaptation à des plateformes d'apprentissage pourrait être bénéfique dans le cas où des scénarios d'apprentissage prévoient des activités réalisables sur le terrain.

Pour sa part, l'outil de [Greenhalgh et al., 2007] et [Brundell et al., 2008] présente un avantage important en offrant une grande combinaison de moyens de visualisation des indicateurs. Il a été expérimenté dans plusieurs situations réelles d'usages mais qui sont pour la plupart des situations d'analyse de gestuelle, d'émotion et de parole. Son usage et son application aux données propres aux scénarios pédagogiques nécessitent des adaptations aux traces des plateformes pédagogiques. Les mécanismes de perception offerts

par ce dispositif pourraient être exploités et adaptés à un tableau de bord d'observation multiplateforme.

Le travail présenté dans [Goodman et al., 2006] offre l'avantage d'analyser les influences dans la construction de la décision relative à un sujet de discussion. C'est un outil destiné aux chercheurs et aux analystes plus qu'aux tuteurs ou aux concepteurs de situations d'apprentissage. Les aspects liés à l'analyse des influences et de la collaboration peuvent être intégrés et exploités dans des plateformes d'apprentissage, après adaptation, afin de superviser les aspects collaboratifs entre apprenants.

Enfin, l'outil proposé dans le cadre du travail présenté dans [Georgeon et al., 2008] et [Olivier et al., 2012] étant destiné à l'analyse des activités d'un acteur humain en interaction avec un dispositif technique complexe, son utilisation en EIAH nécessitera des adaptations et des développements supplémentaires afin qu'il puisse prendre en charge les données propres aux dispositifs d'apprentissage.

## 2.4 Synthèse

Nous avons présenté et discuté l'ensemble des travaux en gardant à l'esprit les deux questions suivantes :

1. Est-ce que ces propositions sont accessibles à des enseignants-concepteurs ?
2. Quel rôle est attribué à l'enseignant-concepteur lors de la préparation de l'observation d'une situation d'apprentissage ?

Bien qu'il existe une multitude de travaux sur l'observation des situations d'apprentissage médiatisées par des outils informatique, nous constatons que la plupart des problématiques traitées sont en relation avec le traitement des traces brutes ou des données collectées lors du déroulement d'une session d'apprentissage. Peu de travaux proposent des aides aux enseignants/concepteurs afin qu'ils puissent exprimer ce qu'ils veulent savoir sur le déroulement réel d'une session d'apprentissage, pourtant ces acteurs sont bien les utilisateurs finaux des résultats de l'observation.

Les enseignants-concepteurs sont impliqués dans tous les travaux de recherches lorsqu'il s'agit de la conception pédagogique mais, paradoxalement, ils ne le sont pas du tout ou pas suffisamment lorsqu'il s'agit de la préparation de l'observation. Nous rejoignons les auteurs de [Martel et al., 2007] et [Corbière et Choquet, 2005] au sujet de la prise en compte de la préparation de l'observation dès la phase de conception des situations d'apprentissage. Dans notre contexte, nous pensons que la préparation de l'observation doit se faire en fournissant des moyens permettant à l'enseignant-concepteur de se poser des questions sur la conception de la situation d'apprentissage qu'il spécifie, de poser des hypothèses sur le déroulement, comme dans le projet IA [Dimitrakopoulou et al., 2006], de la situation en cours de conception.

Plusieurs travaux de recherche axés sur l'observation des traces d'usage de systèmes informatiques, dédié à l'apprentissage pour la plupart, ont été consultés lors de notre étude de la littérature scientifique se rapportant au domaine de l'observation et de la visualisation des indicateurs pédagogiques. Ces travaux ont permis de développer des outils de visualisation des résultats de calcul des indicateurs. Les outils développés permettent de mettre à la disposition des destinataires des résultats de l'observation, un ensemble de moyens leur facilitant la compréhension des indicateurs calculés à partir des traces collectées lors de l'observation des sessions d'apprentissage.

Certains de ces outils permettent le suivi des situations d'apprentissage en temps réel [France et al., 2006] [May, 2009]. D'autres se limitent à la possibilité de visualisation à la fin de la session d'apprentissage [Mazza et Milani, 2004] [Heraud et al., 2005] [Avouris et al., 2007] [Rummel et al., 2011].

Certains outils combinent plusieurs types de visualisation pour le même indicateur [Avouris et al., 2007] [Rummel et al., 2011] [Greenhalgh et al., 2007] [Morrison et al., 2006] tandis que d'autres optent pour la visualisation des indicateurs suivant l'axe temporel horizontal [Goodman et al., 2006] [Van Diggelenet al., 2008], par des graphes de contingence [Heraud et al., 2005] [Marty et al., 2009] ou par les figures de Chernoff [France et al., 2006] [France et al., 2007].

Les outils développés dans [Greenhalgh et al., 2007] [Damsa et Ludvigsen, 2010] [Damsa et al., 2014] [Georgeon et al., 2006] et [Georgeon et al., 2008] montrent les possibilités d'analyse des traces assistées par ordinateur mais sont spécialisés pour certains types de données et nécessitent des développements supplémentaires pour adapter leur fonctionnement aux EIAH.

Certains outils sont destinés plus aux tuteurs afin de leur apporter des éléments pour suivre et superviser une session d'apprentissage qu'au concepteur du scénario pédagogique [Després et Coffinet, 2004] [Heraud et al., 2005] et ne leur offrent que peu de possibilités d'intervention pour réguler la situation d'apprentissage [Després et Coffinet, 2004] sans leur donner les moyens de modifier ou d'adapter leur organisation de l'observation au cours d'une session d'apprentissage. Les indicateurs calculés et visualisés pour la plupart ne concernent que l'activité de communication entre les participants à la session [Mazza et Milani, 2004] [May, 2009] [May et al., 2011].

Certains travaux se sont limités au développement d'outils spécifiques aux traces générées par une seule plateforme d'apprentissage, ce qui ne permet pas leur exploitation sur d'autres plateformes [Mazza et Milani, 2004] [Mazza et Dimitrova, 2004] [Djouad, 2008] [Djouad et al., 2011].

On notera aussi que la plupart des outils ne permettent la visualisation que de peu d'indicateurs qui sont portés sur les échanges de messages dans la majorité des cas. Il y a lieu aussi de noter l'absence de prise en compte du contexte du scénario pédagogique dans les outils cités. Certains travaux [Dyke et al., 2010] ont été validés par un large panel d'expérimentation mais présentent néanmoins des problèmes d'ergonomie et de facilité d'utilisation.

De nouveaux projets sont en cours depuis quelques années en vue d'améliorer la qualité des recherches expérimentales en EIAH. Dans ce sens, les auteurs de [Bouhineau et al., 2013] et [Bouhineau et al., 2013a] ont proposé la plateforme UnderTracks. C'est une plateforme de collecte, conservation, exécution et partage pour les traces d'interaction et les processus de traitement de données expérimentales issues d'EIAH. Cette plateforme se distingue des plateformes existantes par son ouverture à l'ensemble des types d'EIAH et domaines d'enseignements, et de la volonté d'associer données et traitements dans un même environnement. UnderTracks s'intéresse notamment à la question de visualisation interactive pour explorer les données produites dans les EIAH et à comment assurer une utilisation plus facile des outils d'analyse de données.

Un autre projet met en pratique la notion de SGBT (Système de Gestion de Bases de Traces). Il s'agit de kTBS (kernel for Trace Based Systems) [Besnaci et al., 2015] qui est un système informatique, développé au sein de l'équipe SILEX du laboratoire LIRIS. kTBS propose plusieurs fonctionnalités de gestion des traces, à savoir la création, l'interrogation, la transformation et la visualisation. Utilisant les technologies du web, la représentation des traces ainsi que leurs modèles dans kTBS est basée sur le standard RDF (Resource Description Framework) [Schreiber et Raimond, 2014], vu ses possibilités de description structurelles et sémantiques. kTBS présente, cependant, les mêmes problèmes que ceux soulevés en web-sémantiques sur sa difficulté d'utilisation par le grand public [Kong Win Chang et al., 2014]. Ce dernier nécessite la compréhension de graphe RDF et du SQL.

Une autre plateforme propose une architecture adaptée à la mise en place de MOOCs variés basés sur la résolution de problèmes ou sur la démarche par investigation. Il s'agit de SMOOPLE (Semantic, MOOC, Pervasive Learning Environment) [Gilliot et al., 2013]. Cette plateforme permet de suggérer des informations ou des ressources à l'apprenant et d'imposer des activités, avec attendus et échéances, ou libres. Smoople est une plateforme qui intègre des fonctionnalités d'observation prescriptive dans les MOOCs. Ainsi, elle propose des visualisations d'indicateurs, par exemple sous forme de tableaux de bord, indiquant la progression de l'apprenant, de son groupe, ou de la classe. Ces choix permettent ainsi au concepteur du cours de régler l'ouverture de son cours selon cinq dimensions : les objectifs d'apprentissage, le choix des ressources, l'organisation des activités d'apprentissage, l'organisation du travail en groupe et des exercices d'application collaborative.

Un autre projet lancé récemment vise à mutualiser les travaux de recherche français dans le domaine de l'analyse de traces d'e-learning. Il s'agit du projet HUBBLE HUman oBervatory Based-on E-LEarning Traces) [Hubble, 2015] qui regroupe des partenaires académiques et industriels, avec des équipes de recherche en informatique ayant chacune développé leur plateforme d'analyse de traces, et des équipes de recherche en sciences de l'éducation ayant des besoins d'analyse sur des traces de MOOCs, de plateformes d'e-learning ou de jeux sérieux. L'objectif de ce projet est de créer un observatoire national pour la construction et le partage de processus d'analyse de données issues des traces laissées dans des environnements d'apprentissage. Hubble permettra d'analyser et d'expliquer des phénomènes d'enseignement et d'apprentissage avec ces environnements.

