

# Charte éthique pour la recherche en Learning Analytics

## CONTEXTE, OBJECTIFS ET USAGES DE CETTE CHARTE

Dans le contexte du projet ANR HUBBLE, le lot 6 (Éthique) est chargé de proposer une charte servant de guide au comité éthique constitué par le projet HUBBLE au moins pour la durée du projet. Cette charte a vocation à s'appliquer aux projets de recherche ressortissant des méthodes étudiées dans le cadre du projet Hubble, donc possède une finalité plus large que de s'appliquer dans le projet. Cette charte permet d'explicitier le cadre et les limites dans lesquelles pourront se faire les actions de recherche incluant l'utilisation de dispositifs numériques dans l'enseignement/apprentissage pouvant observer les activités d'apprentissages constituant ainsi des traces à des fins d'analyse. Le comité d'éthique devra être particulièrement attentif :

- aux protocoles d'observations,
- à la nature des données recueillies (anonymes, personnelles, sensibles...),
- aux fins d'analyses de ces données,
- aux méthodes de traitements de ces données,
- aux risques encourus par la collecte, le traitement, le stockage et le partage de ces données.

Le comité devra évaluer ces risques et s'assurer qu'ils ne dépassent pas l'intérêt du résultat escompté et qu'ils ont été clairement présentés aux participants qui y ont librement consenti.

L'objectif principal de cette charte est

**l'«articulation des libertés individuelles, des vies privées et du progrès social, scientifique et technique»**

(Victoria Pérès-Labourdette Lembé, 2014).

Cette charte se doit de rappeler l'existence des textes de loi notamment française (Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, [Version du 28 janvier 2016](#)) et européennes ([Directive 95/46/CE sur la protection des données personnelles](#)). Elle est considérée par le consortium de recherche du projet HUBBLE comme un document vivant qui pourra être amendé et précisé pendant toute la durée du projet.

Le comité d'éthique ne se substitue pas à la CNIL (en France) ou aux autorités compétentes sur d'autres territoires (Europe, autres continents), mais vient s'assurer que la loi est connue et appliquée dans la communauté scientifique. Il sert aussi et surtout à alimenter la réflexion, à partir des cas concrets de recherche, sur les questions d'éthique en relation avec le traitement des données potentiellement massives issues des écosystèmes numériques d'apprentissage.

Les principaux aspects qui concernent la réflexion de ce comité pour soutenir l'éthique dans la recherche en learning analytics seraient :

- préciser les limites dans lesquelles peuvent se trouver les questions de recherche soutenables d'un point de vue éthique
- aider les chercheurs à exercer leur activité dans la légalité
- aider les chercheurs à mieux protéger les données personnelles des participants

- aider les chercheurs à construire les mécanismes les plus appropriés pour
  - recueillir le consentement éclairé des participants,
  - organiser la traçabilité des ensembles de données
  - mieux définir les traitements appliqués aux données
  - Éclairer les pratiques d’anonymisation des données
  - partager les données tout en maintenant les droits de retrait ou de modification des données personnelles aux participants
- donner des solutions (cahiers de laboratoire?) aux chercheurs pour protéger leurs droits (propriété intellectuelle) et diffuser les connaissances qu’ils produisent (licences d’utilisations appropriées)

### **Comment intégrer la protection des données personnelles et des libertés individuelles dans les dispositifs numériques de formation et de recherche ?**

#### **TRAITEMENT ETHIQUE DES DONNEES**

##### **Collecte de données :**

Définition Concernant les dispositifs d’enseignement et d’apprentissage (incluant des outils numériques ou non), nous appelons “collecte de données” l’ensemble des opérations (et pré-traitements : filtres, requêtes, etc.) prévues dans le dispositif pédagogique et/ou de recherche, et qui permettent de recueillir des données ou des observables, de les organiser et de les stocker.

Notons que la collecte est toujours orientée vers un objectif qui gouverne la pertinence des données recueillies. Le comité d’éthique veillera à la cohérence entre les données recueillies et les objectifs des analyses en particulier si le recueil inclut des données personnelles.

Dans le cadre du projet HUBBLE, pour chaque collecte de données doit être identifié un “responsable de traitement des données à caractère personnel”. Le comité d’éthique pourra alors se référer à lui pour :

- Compléter la demande formulée au comité d’éthique pour le mise en oeuvre d’une expérimentation (cas d’usage sur le terrain) ;
- Lui rappeler ses obligations légales ;
- Répondre à toute question dans le cadre d’un éventuel recours (participants ou tiers)

#### **LES OBLIGATIONS DU RESPONSABLE DE TRAITEMENT DE DONNEES A CARACTERE PERSONNEL**

Cette partie est détaillée et simplifiée ici mais les membres du comité d’éthiques sont vivement invités à consulter le texte vivant de la loi n° 78-17 du 6 janvier 1978 dite “informatique et liberté” disponible et à jour [sur le site Legifrance](#).

##### **La sécurité des fichiers**

Tout responsable de traitement informatique de données personnelles doit adopter des mesures de

sécurité physiques (sécurité des locaux), logiques (sécurité des systèmes d'information) et adaptées à la nature des données et aux risques présentés par le traitement.

### **La confidentialité des données**

Seules les personnes autorisées peuvent accéder aux données personnelles (telles qu'identifiées dans la directive européenne [95/46/CE](#)) contenues dans un fichier. Il s'agit des destinataires (chercheurs) explicitement désignés (par contrat spécifique dûment signé par l'autorité compétente : directeur de laboratoire, chef d'établissement) pour en obtenir communication et des «tiers autorisés» ayant qualité pour les recevoir de façon ponctuelle et motivée (ex. : la police, le fisc).

Les usages des données personnelles ont une date de péremption. Le responsable d'un fichier fixe une durée de conservation raisonnable en fonction de l'objectif du fichier.

### **L'information des personnes**

Le responsable d'un fichier doit permettre aux personnes concernées par des informations qu'il détient d'exercer pleinement leurs droits. Pour cela, il doit leur communiquer : son identité, la finalité de son traitement, le caractère obligatoire ou facultatif des réponses, les destinataires des informations, l'existence de droits, les transmissions envisagées.

### **L'autorisation de la CNIL**

Les traitements informatiques de données personnelles qui présentent des risques particuliers d'atteinte aux droits et aux libertés doivent, avant leur mise en œuvre, être soumis à l'autorisation de la CNIL.

### **Le CIL correspondant informatique et libertés**

La désignation d'un correspondant informatique et libertés, ou «CIL», permet au responsable de traitement d'alléger ses obligations de déclaration auprès de la CNIL. Au sein de sa structure, ou en externe pour les petites structures, il désigne une personne qui sera chargée :

- de tenir un registre des traitements mis en œuvre au sein de l'organisme (HUBBLE?)
- de veiller au respect des dispositions de la loi « informatique et libertés » au sein de l'organisme.

En contrepartie, l'organisme qui a désigné un CIL n'a plus à effectuer les formalités pour les traitements qui relèvent du régime de la déclaration.

## **PARTAGE DES DONNEES ET DES RESPONSABILITES**

Dès lors que des données, recueillies dans un contexte (pour lequel un responsable des traitements des données personnelles a nécessairement été identifié), sont susceptibles d'être partagées, une demande doit être adressée au comité d'éthique et un contrat de confidentialité (nominatif et exclusif) et de droit à publication à partir de ces données doit être rédigé et signé par les parties : autorité administrative de l'entité possédant les données d'une part et autorité administrative de l'entité les recevant pour traitements d'autre part.

Dans le cas où les acteurs (chercheurs ayant réalisé la collecte, participants à la collecte, et partenaires impliqués dans la construction d'un corpus des données de recherche) sont d'accords entre-eux et que le comité d'éthique ne s'y oppose pas, de tels corpus peuvent être mis à disposition de la communauté scientifique et plus largement de la société civile en utilisant des [licences pré-formatées](#) (ex : Creative Commons, Licences Ouvertes, Copyleft, etc.) d'usages des ressources mises à disposition de tiers. Mais ces licences ne peuvent en aucun cas déroger aux clauses spécifiées dans le formulaire de consentement éclairé préalablement établi avec les participants de

la situation d'enseignement/apprentissage, sans leur accord explicite.

### **Composition du dossier – Délai ??? - pour que ce soit fonctionnel ???**

En aucun cas des données personnelles ne pourront être communiquées à des chercheurs sans qu'ils aient (eux-même, et leur autorité administrative) accepté les conditions de confidentialité et de publication dans un contrat tel que mentionné ci-dessus.

Et si c'est le cas, les chercheurs doivent s'engager (a minima) : à respecter les mêmes clauses que celles du responsable des traitements des données personnelles et y compris à informer les participants des intentions de traitement, conditions et durées de stockage des données, transmissions, etc. et à traiter tout recours d'un participant avec le comité d'éthique; à informer le responsable initial de tout manquement à l'une de ces clauses;

### **Transparence des algorithmes utilisés**

Inspiré par la contribution (Paul, 2015) à la consultation sur la République numérique, ne serait-il pas essentiel de demander à tout nouveau traitement d'être documenté (au moins en version anglaise) à destination des participants pour qu'ils puissent faire valoir leur droit s'ils s'opposaient à de tels traitements sur leur données personnelles ? Cette documentation doit préciser les rapprochements opérés et les risques encourus pour le participant dans le type d'informations qui pourraient être inférées.

## **QUESTIONS DE RECHERCHE SOUTENABLES**

L'objectif premier des algorithmes et analyses développés sur les cas d'usage à l'intérieur du projet HUBBLE est-il seulement ?

**L'exploitation des traces numériques en contexte pédagogique a pour objectif de favoriser le développement de pratiques éducatives plus individualisées et plus réflexives, afin de rendre l'apprentissage plus efficace .**

L'ensemble des chercheurs du projet HUBBLE doit se prononcer les questions majeures qui serviront ensuite de guide et cadre au comité éthique pour donner son aval ou ses conseils dans le cadre d'une demande qui lui sera adressée.

Reste à définir un protocole d'interaction entre les chercheurs et le comité éthique du projet HUBBLE. (Dossier, délais, Documents annexes, rapporteurs, échanges, délibération, décision, recours, etc.)

## **PUBLICS USAGERS DES LEARNINGS ANALYTICS**

### **Apprenants**

Accompagne la personnalisation de la formation médiée par les systèmes numériques, en développant en particulier le tutorat automatique ou des tableaux de bords à destination des tuteurs en vue de soutenir l'apprentissage des apprenants

### **Enseignants/formateurs/tuteurs**

Met à disposition des outils d'analyse de l'apprentissage et du comportement des étudiants

Permet d'améliorer les processus de décision (justifier les objectifs premiers et prendre garde aux détournements possibles)

### **Administrateur (de direction ou technique)**

Aide à déterminer la meilleure façon d'organiser et d'utiliser des ressources et des espaces connectés

## **REFERENCES**

Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés : [Version consolidée au 20 janvier 2016](#), page consultée sur LegiFrance le 20/01/2016.

Directive européenne relative à la protection des données personnelles : [Directive 95/46/CE](#).

Consultation République Numérique (2015). "République Numérique – Projet de loi pour une République numérique". Consultation nationale du 26/09/2015 au 18/10/2015. disponible en ligne, <https://www.republique-numerique.fr/>, Consulté le 17/01/2016.

Victoria Pérès-Labourdette Lembé (2014), "La protection des données personnelles et des libertés individuelles dans des dispositifs numériques de formation : une discipline à construire", MOOC Enseigner et Former Avec le Numérique, 2014.

Paul (2015). "Article 11bis – Transparence des algorithmes utilisés par les services publics et renforcement de la sécurité en ligne", dans Consultation République Numérique (2015), [Disponible en ligne](#), Consulté le 17/01/2016