



HAL
open science

Projet METAL : Plan de collecte de données

Azim Roussanaly, Thomas Toulotte, Laura Infante Blanco, Anne Boyer,
Armelle Brun, Geoffray Bonnin

► To cite this version:

Azim Roussanaly, Thomas Toulotte, Laura Infante Blanco, Anne Boyer, Armelle Brun, et al.. Projet METAL : Plan de collecte de données. [Rapport de recherche] Université de Lorraine (Nancy); CNRS. 2017. hal-01801890

HAL Id: hal-01801890

<https://inria.hal.science/hal-01801890v1>

Submitted on 28 May 2018

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

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



Projet METAL

Modèles Et Traces au Service de l'Apprentissage des Langues

Plan de collecte de données



Azim Roussanaly, Thomas Toulotte, Laura Infante Blanco,
Anne Boyer, Armelle Brun, Geoffray Bonnin

Équipe KIWI, LORIA, Université de Lorraine

20/12/2017 - v. 1,2

- Objectifs..... 3
- Identification des données..... 3
 - Acteurs du projet..... 4
 - Nature des informations..... 5
 - Formation en cinquième..... 7
- Modèle de données..... 8
 - Méthodologie..... 8
 - SCORM..... 8
 - xAPI ou TinCan..... 9
 - OneRoster..... 11
 - Le modèle conceptuel..... 11
- Conclusion..... 15
 - Précision sur l’approche..... 16
 - Interdépendance des actions..... 16
 - Culture du standard et de l’analytics..... 16
 - Compromis entre collecte “raisonnée” et recherche..... 17
- Enjeux juridiques et éthiques..... 17
 - Les enjeux juridiques des Learning Analytics liés à la nature des données..... 17
 - La responsabilisation des responsables de traitement..... 18
 - Consentement..... 18
 - Privacy by design..... 19
 - Accountability et transparence algorithmique..... 19
 - L'empowerment des utilisateurs..... 19
 - Droit à la portabilité (Art 20 RGPD)..... 19
 - Droit d'accès (Art 15 RGPD) (Art 40 LRN)..... 20
 - Droit de rectification (Art 16 RGPD) (Art 40 LRN)..... 20
 - Droit à l'effacement ou droit à l'oubli (Art 17 RGPD) (Art 40 LRN)..... 20
 - Droit à la limitation de traitements (Art 18 RGPD)(Art 40 LRN)..... 20
 - Le cas de la recherche scientifique..... 20
 - Enjeux éthiques particuliers..... 21

Objectifs

Les technologies du *Learning Analytics* visent à analyser automatiquement les données issues de l'interaction des apprenants avec leur environnement numérique de travail. Le but est d'établir des modèles de comportements permettant d'améliorer l'apprentissage des apprenants ainsi que le contexte dans lequel il se produit. L'identification des informations et des facteurs qui permettent de modéliser le comportement des apprenants est une tâche complexe, notamment en raison de la grande quantité de variables susceptibles de déterminer ces modèles.

Ce document a pour objectif de présenter le plan de collecte des données sur lesquelles s'appuieront les algorithmes de modélisation qui constituent un préalable aux applications reposant sur les *Learning Analytics* dans le cadre du projet METAL¹.

Il comporte deux axes majeurs : 1) l'identification des données à travers leur localisation et leur nature au sein des différents systèmes d'information des organismes susceptibles de contribuer à leur mise à disposition pour le projet METAL et 2) la caractérisation de ces données sous la forme d'un modèle conceptuel destiné à implémenter à terme un entrepôt de données diverses alimenté à partir de sources également diverses.

Identification des données

Une des activités préalables à cette tâche de collecte de données consiste à bien comprendre le fonctionnement et les procédures du système éducatif, de manière à recenser les applications numériques mises au service des élèves et des enseignants afin d'identifier les informations sur lesquelles s'appuient ces applications ainsi que les traces d'usage qui seraient potentiellement exploitables par les outils de *Learning Analytics* à concevoir dans le cadre du projet METAL. Cette activité résulte de nombreux échanges avec les acteurs concernés par le projet.

¹ <http://metal.loria.fr/>

Acteurs du projet

Les principaux interlocuteurs avec lesquels nous avons échangé sont (voir Figure 1):

- la DANE (Délégation Académique du Numérique Éducatif) à travers Christine François, déléguée académique au numérique, Michèle Drechsler, Déléguée académique au numérique adjointe, Laurence Langlois, chargée de mission ENT Place, Emmanuel Coste, IA-IPR, Inspecteur d'académie - inspecteur pédagogique régional de langue Allemande, et Sophie Rolée, chargée de mission ENT Oze.
- le Rectorat de Nancy-Metz : à travers Maxime Mazzini, responsable de la Direction du Système d'information,
- le Conseil Départemental de Meurthe et Moselle : à travers Franck di Santolo, responsable de la Direction des Système d'information,
- le Conseil Départemental de Moselle : à travers Raphael Régnier, responsable de la Direction de l'Organisation et des systèmes d'information,
- la Société ITOP, gestionnaire du logiciel ENT PLACE, qui est l'actuel environnement numérique de travail déployé dans l'Académie, à travers Olivier Tardella, responsable du projet chez ladite entreprise.

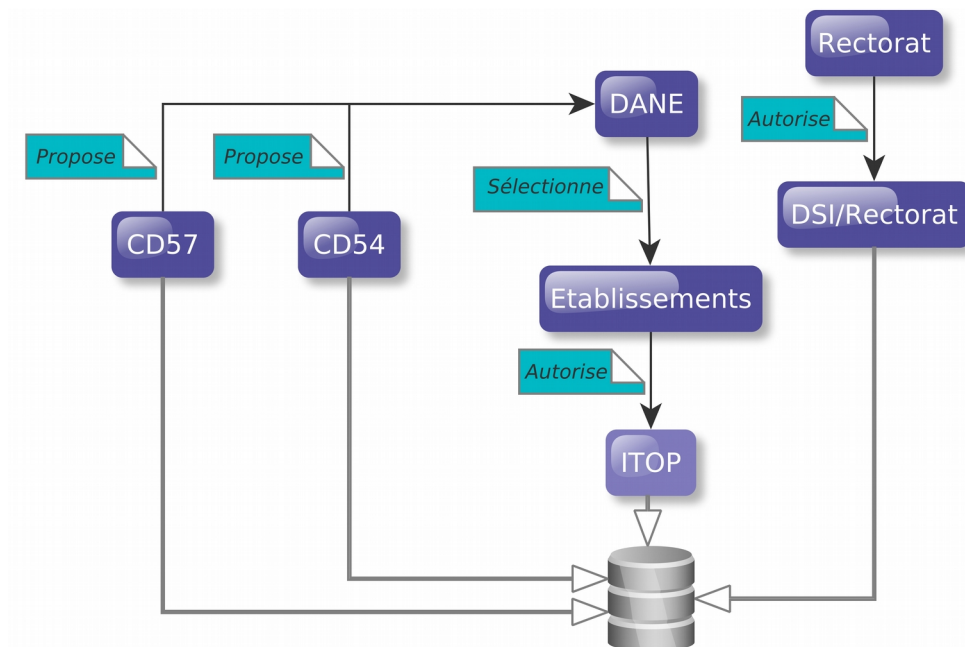


Figure 1: les acteurs pour les données

La DANE est en charge de la détermination de la liste des classes et des établissements qui vont participer au projet. L'objectif est de sélectionner au sein de l'Académie, un groupe d'établissements volontaires pour participer aux expérimentations liées au projet. Le niveau retenu est la classe de cinquième (Collège). Ce travail est, à l'heure actuelle, encore en cours.

Les informations concernant ces futurs établissements sont regroupées dans un système d'information opéré par la Société ITOP (partenaire du projet) à travers l'ENT PLACE, qui constitue l'Espace numérique de travail utilisé communément dans l'ensemble des établissements. La responsabilité de rendre ces données disponibles pour le projet incombe aux chefs des établissements sélectionnés.

D'autres sources de données ont également été identifiées. Il s'agit, d'une part, de données provenant des services numériques des Conseils de département (CD54 et CD57) en charge des accès par les collèges aux infrastructures numériques et au réseau, et, d'autre part, des données stockées par le Rectorat concernant le parcours de chaque élève.

Nature des informations

La Figure 2 fournit une vue synthétique des informations potentiellement intéressantes issues des différentes applications et des systèmes d'informations des organismes-acteurs précédemment cités.

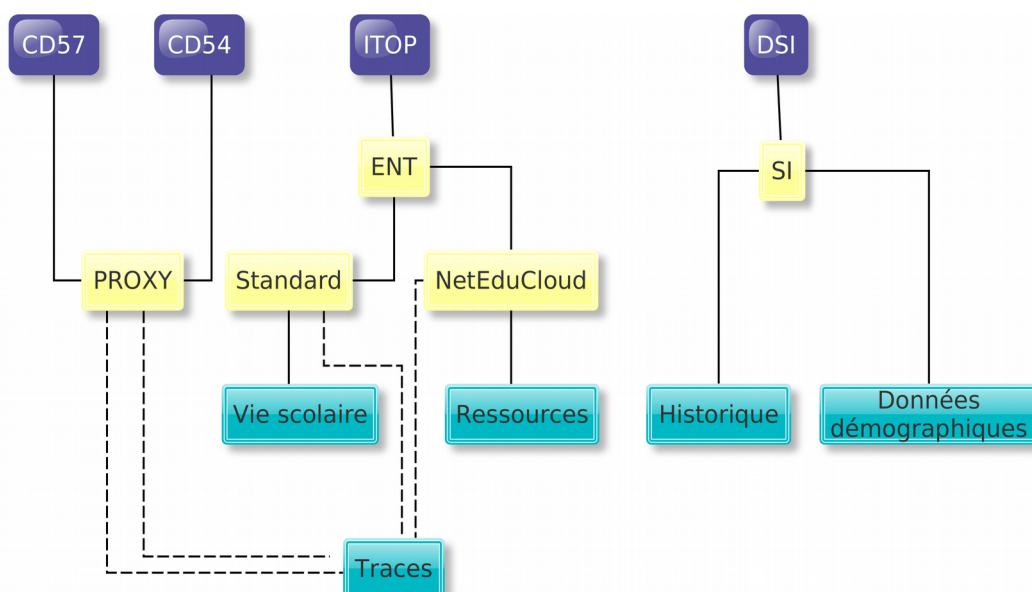


Figure 2: nature des informations

La source de données majeure concerne l'ENT, c'est-à-dire la plateforme hébergeant une variété de services destinés aux enseignants, aux élèves et à leurs parents, et accessibles à travers un portail en ligne unique. Les principales données identifiées sont relatives à la vie scolaire, aux ressources pédagogiques et aux traces d'usage des services. La solution actuelle, basée sur l'ENT PLACE, appelée à évoluer dans un futur proche, se présente comme une plateforme offrant de manière standard des services liés à la vie scolaire. Cette plateforme peut accueillir, selon les établissements, des applications complémentaires permettant un travail collaboratif entre enseignants et élèves, et fournissant des accès à diverses ressources pédagogiques. L'extension *NetEduCloud* en est un exemple.

Par ailleurs, les accès aux services sont enregistrés dans un journal par le système (traces d'usage). Certaines applications ont la possibilité d'aller plus loin et de tracer plus finement les usages pédagogiques des ressources selon la norme *xAPI* (sur laquelle nous allons revenir un peu plus loin).

D'autres traces d'usages peuvent être fournies par les Conseils départementaux par l'intermédiaire des journaux des PROXYS mis en place qui constituent un point de passage de toutes les connexions au réseau par les établissements situés dans chacun des départements. Nos contacts se sont limités pour le moment aux seuls départements de la Meurthe-et-Moselle et de la Moselle, mais il est envisagé d'élargir la coopération à tous les départements des établissements qui vont être sélectionnés pour l'expérimentation. Les premiers exemples de données qui nous sont parvenus restent encore insuffisamment précis, mais nous espérons obtenir de meilleurs résultats grâce à une adaptation de la journalisation des proxys.

Et enfin, les données complémentaires concernant le parcours sur plusieurs années de chaque élève ainsi que les informations sociodémographiques de leurs parents peuvent être extraites du système d'information du Rectorat. Un contact en cours avec la DSI du Rectorat vise à définir de manière précise les caractéristiques de ces informations.

Formation en cinquième

Notre découverte de l'information éducative s'appuie sur un travail de documentation sur le système éducatif et, plus spécifiquement, sur ce qui concerne la classe de cinquième au collège, classe expérimentale ciblée par le projet METAL.

La cinquième appartient au cycle pédagogique 4 (cycle des approfondissements) et se compose de 26 heures d'enseignements hebdomadaires obligatoires dont 4 heures correspondent à des enseignements complémentaires (enseignements pratiques interdisciplinaires - EPI - et accompagnement personnalisé - AP).

Les enseignements obligatoires en cinquième sont résumés dans le tableau suivant :

Matière	Heures/semaine
Français	4h30
Mathématiques	3h30
Langue vivante 1	3h
Histoire - géographie	3h

Enseignement moral et civique	
Langue vivante 2	2h30
Sciences de la vie et de la Terre	1h30
Sciences physiques	1h30
Technologie	1h30
Éducation physique et sportive	3h
Arts plastiques	1h
Éducation musicale	1h

Source: www.education.gouv.fr

Depuis 2017, les établissements ont également la possibilité de proposer des enseignements facultatifs comme :

- Les langues et cultures de l'antiquité
- Les langues et cultures régionales
- Les langues et cultures européennes

Modèle de données

Méthodologie

Depuis l'apparition des LMS modernes (Learning Management Systems), des efforts pour définir des mécanismes permettant la formalisation et standardisation des processus et contenus éducatifs ont été fournis par divers organismes internationaux d'éducation numérique. Si ces standards ne sont pas généralement intégrés dans la solution de l'ENT actuelle, nous travaillons néanmoins avec l'hypothèse que les prochaines solutions s'y appuieront de manière plus systématique.

Dans notre étude, nous nous proposons d'examiner sur les standards les plus émergents.

SCORM

Le premier d'entre eux, est le standard *SCORM*² (Sharable Content Object Reference Model), créé en janvier 2001 par l'Advanced Distributed Learning (ADL), une division du Département de Défense des Etats Unis).

SCORM s'appuie sur d'autres standards comme le standard *LOM*³ (Learning Object Metadata) de l'IEEE Standard Association qui fournit une spécification en XML pour la description de la structure du contenu des ressources éducatives.

Outre les propriétés d'importation et d'exportation des ressources pédagogiques facilitant leur comptabilité avec les plateformes *LMS*, la version la plus récente (*SCORM* 2004) introduit des spécifications pour la définition de parcours d'activités et pour la capture, transfert et partage de traces d'usage des apprenants lors de leur utilisation des ressources éducatives

De nombreux éditeurs de *LMS* utilisent le standard *SCORM* dont la célèbre plateforme libre *MOODLE* adoptée par la majorité des universités françaises.

Actuellement, et malgré son utilisation répandue, *SCORM* est jugé obsolète et la tendance est de le remplacer par *Experience API* (*xAPI* ou *Tin Can API*)

xAPI ou TinCan

Le projet *TinCan*⁴ est né en 2011 à la suite d'une étude de modernisation du standard *SCORM*, confiée la Société américaine *Rustici Software*⁵ qui se fixe pour objectif d'améliorer les spécifications de *SCORM* en vue de faciliter son utilisation.

Les dernières spécifications sont publiées en 2014 sous le nom d'Experience API ou *xAPI*.

xAPI a permis d'inclure des nombreuses fonctionnalités dans la spécification qui n'étaient pas présentes dans *SCORM* comme la séparation entre les ressources et la navigation et l'utilisation des téléphones portables avec la possibilité de combiner une activité à la fois en mobilité et sur un ordinateur. Mais l'apport

² <http://www.adlnet.gov/adl-research/scorm>

³ <https://standards.ieee.org/findstds/standard/1484.12.1-2002.html>

⁴ <https://experienceapi.com/>

⁵ <https://rusticissoftware.com/>

qui nous intéresse le plus concerne la spécification particulièrement avancée des traces d'usages, pour permettre un meilleur suivi du travail en équipe, des plannings et des interactions des élèves avec des ressources pédagogiques, des jeux et des simulateurs.

Ce standard se focalise sur la collecte d'informations des expériences d'apprentissage des apprenants, que ce soit quand ils interagissent avec des contenus éducatifs ou avec d'autres apprenants. Les informations sont envoyées pour stockage dans un *LRS* (Learning Record Store) (voir Figure 3) sous la forme de triplets (sujet, verbe, objet), comme par exemple, (Pierre, a complété, exercice1).

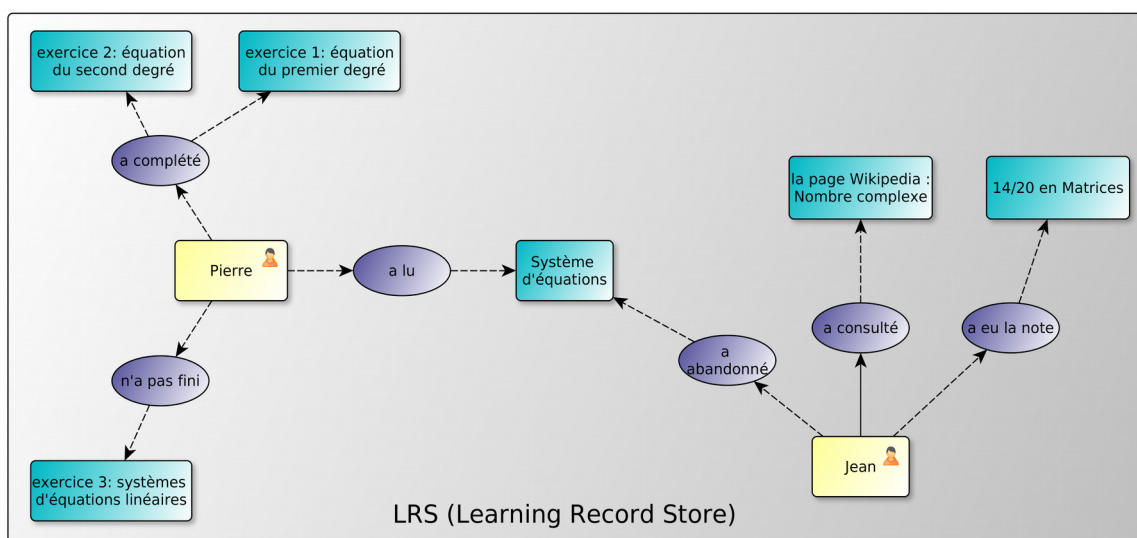


Figure 3: Un exemple de triplets xAPI

À la différence de son prédécesseur *SCORM*, le vocabulaire pour les verbes dans les triplets n'est pas restreint à quelques mots, permettant ainsi une adaptabilité à toute sorte d'Environnement Numérique de Travail et une grande flexibilité pour représenter les actions des utilisateurs.

xAPI est d'ores et déjà adopté par de grands acteurs dans le domaine de l'éducation (*Moodle, Adobe, iSpring...*).

Dans le cadre du projet METAL, nous sommes amenés à nous appuyer à la fois sur les activités en ligne, bien prises en compte par le standard *xAPI*, mais aussi sur les activités classiques qui sont aussi susceptibles d'engendrer des

informations numériques comme par exemple la planification d'un devoir ou la saisie d'une note.

Il est donc nécessaire d'intégrer l'ensemble des activités dans un modèle de données plus étendu.

OneRoster

OneRoster⁶ est un standard, publié par l'association collaborative à but non lucratif *IMS Global Learning Consortium*, comprenant une infrastructure de données et de développement (framework), dont le but est d'offrir une solution pour l'importation/exportation d'informations de vie scolaire, de ressources pédagogiques et de notes de manière sécurisée, standard et fiable. Sa version la plus récente date de l'année 2013. Elle se distingue par une double possibilité de format d'échange entre deux systèmes d'information, l'une basée sur la création de fichiers au format standard *CSV* (Comma-separated values) et l'autre sur le recours à des services web (basés sur les technologies *REST* et *JSON*) complètement décrites dans les spécifications du standard.

De manière implicite, *OneRoster* s'appuie sur un modèle de données qui intègre à la fois l'organisation d'un établissement scolaire, la vie scolaire et les activités pédagogiques.

Ce modèle est particulièrement orienté vers une problématique d'établissements américains, que ce soit en terme d'organisation structurelle ou juridique. Par conséquent, il nécessite une adaptation au système national et aux contraintes du projet METAL, notamment en ce qui concerne le respect de la vie privée.

Le modèle conceptuel

Sur le plan méthodologique, notre démarche consiste donc à adapter le modèle standard *OneRoster* aux exigences du projet METAL.

Une première partie du modèle des données dérivant du modèle *OneRoster* est présentée à la Figure 4 au format UML où chaque classe représente une entité *OneRoster*.

Remarques :

⁶ <https://www.imsglobal.org/activity/onerosterlis>

- Toutes les entités *OneRoster* n'apparaissent pas sur ce schéma, notamment celle qui décrit les résultats d'évaluation autrement dit les notes. Nous verrons plus loin, une manière plus générique de l'intégrer, en étendant l'entité au concept plus général d'activité.
- La classe *User*, représentant à l'origine tous les utilisateurs d'un établissement, a été modifiée afin d'exclure les personnels non-enseignants, non pris en compte dans le cadre du projet METAL. Cette classe se décompose, par héritage, en 4 sous-classes représentant les élèves (*Student*), les enseignants (*Teacher*), les parents (*Parent*) et les tuteurs légaux (*Guardian*).
- La classe *PersonalData* est rendue indépendante de la classe *User* car elle ne devrait pas intégrer l'entrepôt METAL pour répondre à la contrainte de pseudonymisation des données.
- La classe *Demographics* a été remaniée, d'une part, par le retrait d'attributs illégaux en France (comme par exemple l'origine ethnique, la religion...) et, d'autre part, par l'ajout d'attributs (schématisés en rouge) souhaités par les responsables des autres tâches du projet METAL.
- La classe *Class* associe un cours (*Course*) à une période (*AcademicSession*).
- Et enfin un établissement (*Org*) est constitué de *Class*.

Une deuxième partie du modèle est présentée à la *Figure 5*. Elle résulte d'une adaptation plus profonde du modèle original *OneRoster* plutôt axé sur la notion de notation, tandis que notre souhait est de l'étendre aux concepts plus riches apportés par le standard *xAPI*.

Remarques :

- Les classes symbolisées en vert sont directement issues du standard *xAPI*
- La classe *Activity* décrit les activités sur lesquelles portent des actions pédagogiques pouvant éventuellement déboucher sur une évaluation notée. La classe *Statement* regroupe les instances horodatées des actions effectuées par un utilisateur (*User*) sur une activité. Les actions sont représentées par la classe *Verb*.

- La classe *AssignableUnit* est composée d'activités ; dans la version présentée, nous nous limitons à ce stade, à un seul type d'activités, à savoir celles qui intéressent un utilisateur et un objet pédagogique au sens large. Cette manière de modéliser facilite une extension future du modèle de données en introduisant de nouvelles classes représentant d'autres types d'activités comme par exemple des activités collaboratives qui touchent deux ou plusieurs utilisateurs. Dans la version proposée, nous mettons en suspens cette possibilité dans la mesure où les informations provenant des applications existantes n'ont pas encore mis en évidence ce type d'activités. La classe composite *AssignableUnit* est associée à la classe *Class* décrite plus haut.
- Le modèle de référence *OneRoster* ne prend pas en compte la représentation des informations relatives au concept d'emploi du temps pourtant très importante au sein des applications observées. Une option serait d'étendre le modèle avec des classes spécifiques à cette notion. Une autre option serait de traduire la présence ou l'absence en cours d'un élève par la création d'une activité adéquate. A l'heure actuelle, le choix n'est pas définitivement déterminé et la question nécessite encore de mieux définir les besoins des concepteurs des outils.

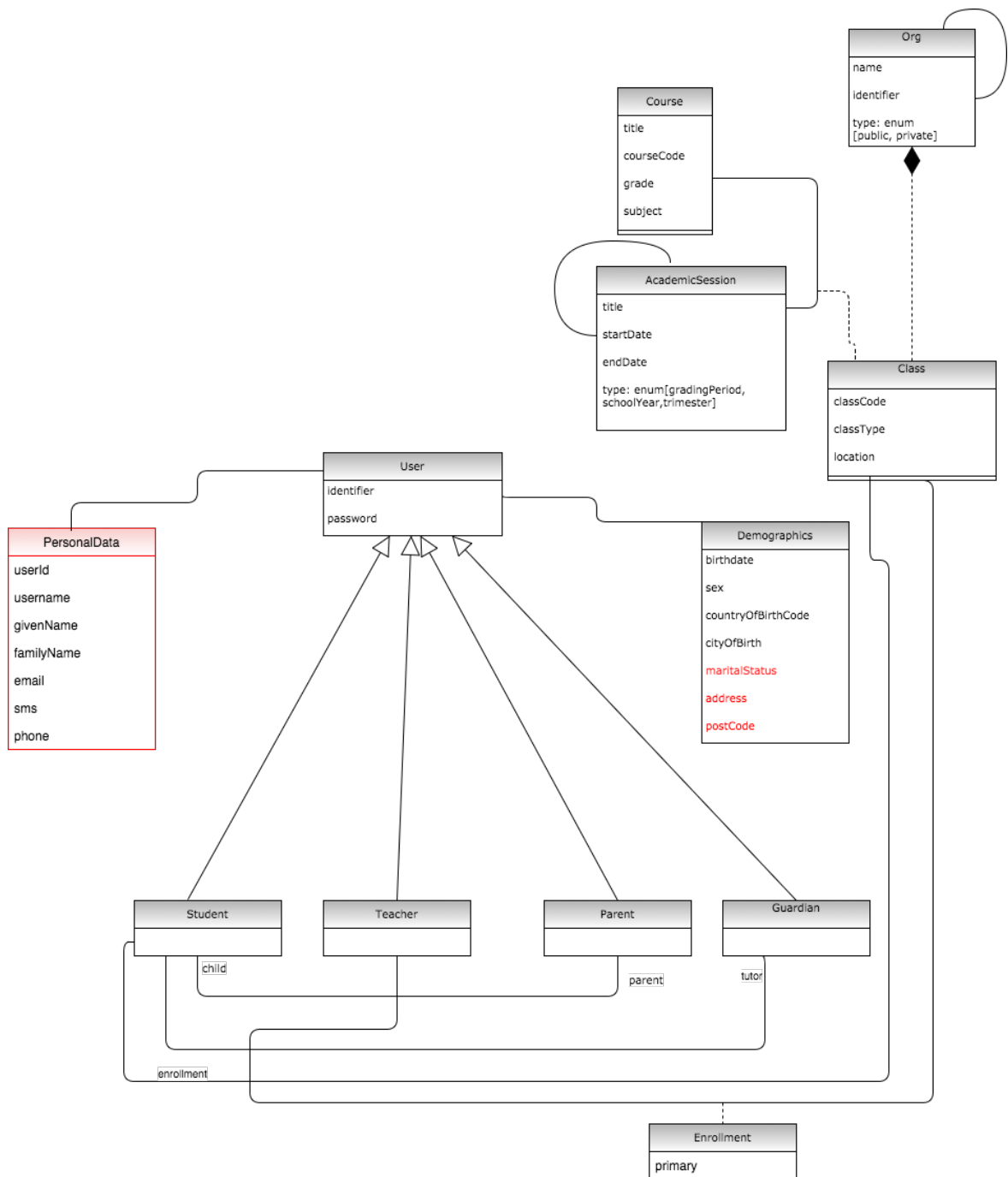


Figure 4: modèle de données (partie 1)

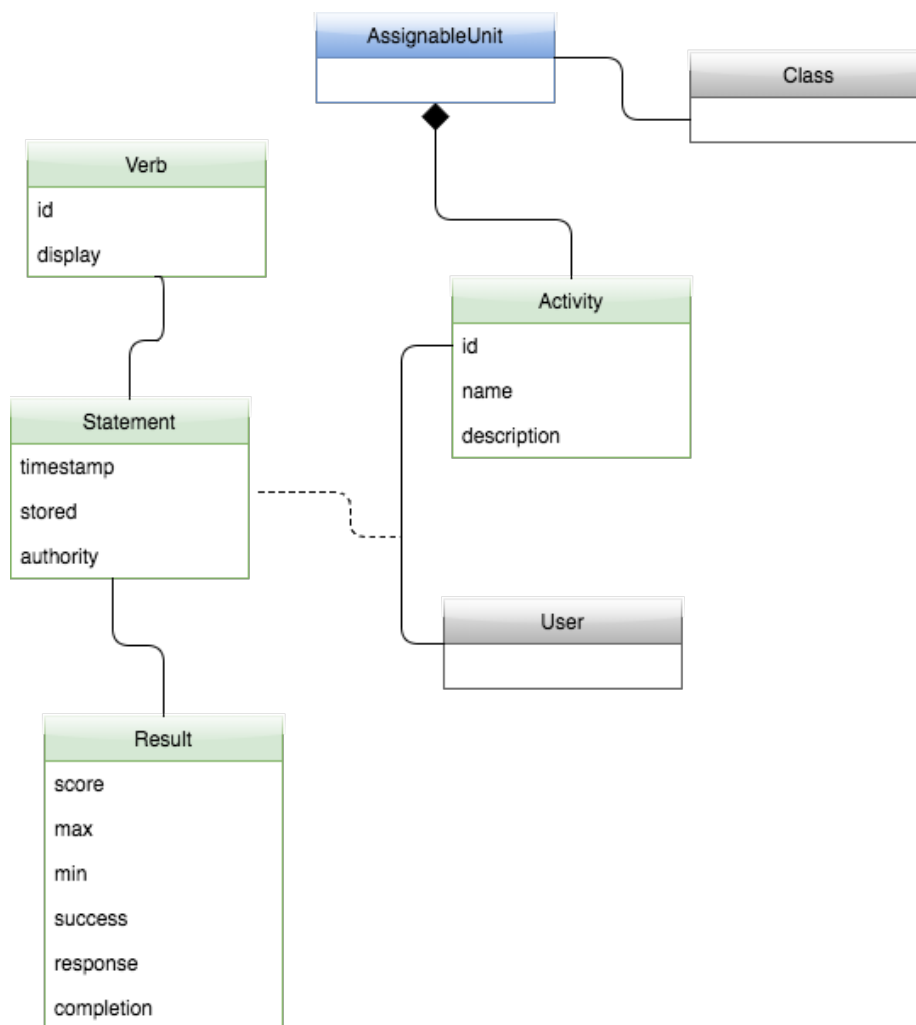


Figure 5: modèle de données (partie 2)

Conclusion

Dans la suite du projet, le modèle ainsi défini va permettre d'implémenter un entrepôt de données regroupant l'ensemble des informations sur lesquelles vont pouvoir s'appuyer les outils de *Learning Analytics* envisagés. L'alimentation de cet entrepôt se fera par l'intermédiaire des applications disponibles qui devront être, dans une prochaine étape, sélectionnées en

collaboration avec les organismes responsables (juridiquement et techniquement) de ses applications.

Ces prochaines étapes doivent tenir compte d'un certain nombre de questionnement que nous résumons ci-dessous.

Précision sur l'approche

L'approche dominante pour la conception de notre modèle est résolument descendante dans la mesure où nous partons des standards pour déterminer les données. Outre le fait que le rapprochement avec les standards vise à faciliter les processus de collecte, notre approche se justifie aussi par le fait que le contexte conjoncturel de l'évolution à court terme de l'espace numérique de travail ne permet pas de constituer un corpus de données observables avec un niveau de qualité suffisant pour une approche ascendante. Cependant, nous avons effectué systématiquement, dans la mesure du possible, une confrontation entre nos propositions et les échantillons de données fournis par les partenaires du projet.

Interdépendance des actions

Notre modèle devrait être validée par la collecte de données réelles. La période actuelle d'évolution de l'environnement numérique d'une part, et le délai nécessaire à l'implication de classes expérimentales d'autre part, ne nous permet pas d'envisager de définir une réelle collecte de données qui permettrait à la fois de valider le modèle, mais aussi d'avoir une idée des répartitions volumétriques, primordiales pour le choix des techniques expérimentales à envisager.

Culture du standard et de l'analytics

L'effort de développement des processus automatique de collecte devrait dépendre de l'aptitude de ces applications mises à la disposition des enseignants et des élèves à héberger une composante d'*analytics* qui consiste à tracer les usages des acteurs. Cet effort devrait être d'autant moindre que cette composante s'appuie sur des standards reconnus.

Une autre qualité à observer concerne la capacité des applications à exporter les informations qu'elles utilisent.

Compromis entre collecte “raisonnée” et recherche

Sur le plan de l'éthique et du respect de la vie privée, une collecte “raisonnée”, c'est-à-dire limitée aux seules informations directement exploitées par les traitements, est fortement préconisée. Dans le contexte du projet METAL, nous nous plaçons dans une démarche de recherche reposant sur une approche guidée par les données, ce qui rend complexe la justification *a priori* des données collectées.

Enjeux juridiques et éthiques

Le plan de collecte des données sur le plan technique s'accompagne de considérations juridiques et éthiques. Ainsi il faut dans un premier temps répondre à la question préalable suivante : quel est le cadre juridique permettant la mise en application des *Learning Analytics* ?

L'essentiel des problématiques juridiques spécifiques aux *Learning Analytics* sont liées au traitement des données.

Les enjeux juridiques des Learning Analytics liés à la nature des données

Les données identifiées précédemment pour la mise en œuvre des procédés de *Learning Analytics* peuvent être considérées comme des « informations relatives à une personne physique identifiée ou identifiable », elles sont donc définies comme des données à caractère personnel au sens de l'article 4 du règlement (UE) 2016/679 du Parlement européen et du Conseil. Elles sont donc concernées par cette réglementation et par tous les textes traitant de la question des données personnelles.

Les textes applicables sont donc dans un premier temps le Règlement général européen relatif à la protection des données à caractère personnel. Le règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 ou RGPD constitue l'essentiel du régime juridique applicable aux données personnelles. Les règles qu'il pose ne s'appliquent pas aux données anonymisées.

S'appliquent également notamment en France la Loi pour une République Numérique (Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique) ainsi que la Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

Les enjeux juridiques et éthiques concernant les *Learning Analytics* sont intimement liés à la nouveauté de ces règles juridiques, qu'il est nécessaire d'appréhender de manière rapide et efficace, aussi est-il nécessaire de considérer les divers avis rendus par les autorités nationales et internationales -à minima européennes- sur la question du traitement des données personnelles et de l'éthique.

La responsabilisation des responsables de traitement

Le nouveau régime juridique du traitement des données à caractère personnel n'implique plus de déclaration préalable. Cependant, cet allègement des obligations s'accompagne d'un certain nombre d'obligations nouvelles à la charge des responsables de traitement.

Le responsable de traitement est défini comme "la personne, l'autorité publique, le service ou l'organisme qui détermine les finalités et les moyens" du traitement. En cela, l'unité de recherche du projet METAL peut être considérée comme responsable de traitement dans la mesure où elle détermine seule les finalités et les moyens du traitement nécessaire au projet.

Consentement

En premier lieu, le traitement des données à caractère personnel nécessite un consentement prévu dans une forme particulière.

Le projet doit donc veiller à ce que le consentement au traitement des données personnel d'une personne soit librement donné, qu'il soit spécifique, informé et non ambiguë selon l'article 4§11 de la directive. Par ailleurs, l'article 7 précise que «la demande de consentement doit être présentée d'une manière qui se distingue clairement des autres, sous une forme intelligible et facilement accessible, dans un langage clair et simple».

Puisque les données concernent ici des enfants, des enjeux éthiques et juridiques spéciaux doivent être considérés. C'est notamment le cas du consentement, sur lequel la réglementation européenne considère qu'en deçà

de 16 ans, le consentement doit être donné par le titulaire de la responsabilité parentale.

Des consultations nationales dont les résultats seront publiés prochainement doivent sous peu apporter d'autres éclaircissements sur le traitement de données impliquant des personnes mineures.

Privacy by design

Il est donc primordial d'intégrer dès la conception du plan de collecte de données le respect des nouvelles obligations, auxquelles s'ajoutent celles de privacy by design (ou respect de la vie privée dès la conception) selon l'article 25 du règlement. Cette protection des données intégrée dans le traitement nécessite de mettre en œuvre des mesures techniques, telles que la pseudonymisation et la minimisation des données collectées. Cependant, comme il a été précisé dans les remarques conclusives précédentes, l'approche du projet METAL ne permet pas de justification à priori des données utilisées, et donc de minimisation des données.

Accountability et transparence algorithmique

Le projet nécessite également de prendre en compte les nouveaux enjeux d'accountability (« *l'obligation de mettre en œuvre des mécanismes et des procédures internes permettant de démontrer le respect des règles relatives à la protection des données* ») et d'explicabilité (les actions effectuées par les algorithmes doivent pouvoir être comprises par les utilisateurs, à qui une documentation compréhensible doit être mise à disposition). Ces notions doivent donc prendre en compte les enjeux éthiques de transparence algorithmique ainsi que de neutralité algorithmique.

L'empowerment des utilisateurs

La responsabilisation des responsables de traitement s'accompagne d'un certain nombre de droits accordés aux utilisateurs. Le responsable de traitement doit faciliter l'exercice de ces droits par les utilisateurs.

Droit à la portabilité (Art 20 RGPD)

La personne concernée doit pouvoir recevoir les données à caractère personnel les concernant qu'elles ont fournies à un responsable du traitement, dans un

format structuré, couramment utilisé et lisible par machine, et ont le droit de transmettre ces données à un autre responsable du traitement sans que le responsable du traitement auquel les données à caractère personnel ont été communiquées y fasse obstacle.

Il s'agit des données que la personne concernée a transmises mais également des données générées et collectées à partir de ses activités et en particulier de par l'utilisation du service fourni par le responsable du traitement.

Droit d'accès (Art 15 RGPD) (Art 40 LRN)

La personne concernée a le droit d'obtenir du responsable du traitement la confirmation que des données à caractère personnel la concernant sont ou ne sont pas traitées et, lorsqu'elles le sont, l'accès auxdites données à caractère personnel ainsi que diverses informations supplémentaires.

Droit de rectification (Art 16 RGPD) (Art 40 LRN)

La personne concernée a le droit d'obtenir du responsable du traitement, sans retard injustifié, la rectification des données personnelles inexactes la concernant.

Droit à l'effacement ou droit à l'oubli (Art 17 RGPD) (Art 40 LRN)

La loi française pour une République numérique (LOI n ° 2016-1321 du 7 octobre 2016) a modifié l'article 40 de la loi sur la protection des données. Il comprend désormais un «droit à l'oubli» spécifique aux enfants et une procédure accélérée pour l'exercice de ce droit. Lorsque la personne concernée était un enfant au moment de la collecte des données, la personne concernée peut obtenir des plates-formes l'effacement des données problématiques "le plus tôt possible".

Droit à la limitation de traitements (Art 18 RGPD)(Art 40 LRN)

Ce droit permet de stopper temporairement le traitement tout en conservant les données pour diverses raisons(rectifications des informations, organisation droit de la défense, etc..)

Le cas de la recherche scientifique

Le projet METAL pouvant être considéré comme effectué dans un cadre de recherche scientifique, il est nécessaire d'étudier les exceptions juridiques découlant de cette situation particulière.

Ainsi, lorsque des données à caractère personnel sont traitées à des fins scientifiques, la législation de l'Union ou des États membres peut prévoir des dérogations aux droits visés aux articles 15 (droit d'accès), 16 (rectification) 18 (restriction) et 21 (droit d'opposition).

Cela n'est possible que si ces droits sont susceptibles de rendre impossible ou d'entraver sérieusement la réalisation des objectifs spécifiques.

Enjeux éthiques particuliers

Enfin, des enjeux éthiques se posent sur divers aspects du traitement. Ainsi, le risque de ré-identification doit être une problématique connue par le responsable de traitement et évitée autant que l'état de l'art le permet.

Le profilage est nécessaire dans la plupart des cas de *Learning Analytics* en ce qu'il se définit comme « le traitement automatisé de données à caractère personnel qui consiste à utiliser ces données pour évaluer certains aspects de la personne concernée, et analyser ou prédire ses intérêts, son comportement et d'autres attributs ». L'article 22§4 du règlement les autorise, mais conditionne le profilage à une autorisation préalable, consentement fondé sur des informations significatives sur la logique impliquées.

Ceci induit un risque de biais dans la pratique des utilisateurs et donc des résultats obtenus qui doit être pris en compte.

Il doit également être noté que la génération de données sensibles via le profilage ne sont pour le moment pas considérés par le règlement. La société civile préconise cependant d'éviter une telle génération de données sensible suite à l'utilisation du profilage, notamment par les *Learning Analytics*.