



# LOLA, une plateforme de valorisation de la donnée éducative

**Auteur :** Anne Boyer

**Nom de l'organisation :** UL/Loria/CNRS

**Date de création :** 21-02-2022

**Date de modification :** 21-02-2022

**Mots-clés :** Learning Analytics, données éducatives, mutualisation et partage



## Table des matières

<b>1</b>	<b>Contexte</b>	<b>3</b>
<b>2</b>	<b>Organisation du projet LOLA</b>	<b>4</b>
<b>3</b>	<b>Le site web LOLA</b>	<b>4</b>
<b>4</b>	<b>La plateforme LOLA</b>	<b>5</b>
<b>5</b>	<b>La collection LOLA</b>	<b>6</b>



## 1 Contexte

La généralisation de l'usage du numérique dans l'enseignement conjointement à l'apparition de technologies Big Data a permis la constitution de corpus massifs de données sur les profils et les comportements numériques des apprenants. Les travaux de recherche permettant l'exploitation conjointe ou non de ces sources d'informations sur l'apprentissage se sont multipliés, dans le but d'améliorer la qualité des dispositifs, de permettre aux enseignants de mieux évaluer les processus d'apprentissage et aux apprenants d'être plus efficaces dans leur démarche d'apprentissage. Les Learning Analytics (ou analyse des données de l'apprentissage) œuvrent à rendre lisible et interprétable cette énorme quantité de données d'apprentissage collectées. En effet, les Learning Analytics s'intéressent à la collecte, l'analyse, la modélisation, la synthèse et la communication des données relatives aux apprenants et à leurs environnements, afin de mieux comprendre et améliorer l'apprentissage. L'exploitation des traces d'interaction en situation d'apprentissage à l'aide notamment de techniques d'intelligence artificielle et de fouilles de données permet différents types de retours d'expérience, vers l'apprenant, le tuteur ou le formateur, les pairs, le concepteur de la ressource ou l'institution, le plus souvent sous la forme synthétique d'indicateurs, de prédicteurs, d'outils de visualisation (notamment de tableaux de bords) ou de systèmes d'intervention comme les systèmes de recommandations personnalisés.

Les défis scientifiques majeurs des Learning Analytics concernent notamment les problématiques suivantes :

- Big et Small Learning Analytics : l'abondance des données éducatives conduit à exploiter les méthodes du Big Data pour fouiller des données complexes issues de multiples sources souvent bruitées et hétérogènes, éventuellement temporelles. Toutefois, sur chaque apprenant ne sont le plus souvent disponibles que des données parcimonieuses desquelles il faut tirer des modèles de comportement individuel néanmoins performants.
- Modélisation du contexte : le contexte d'apprentissage est une source d'informations essentielle, sa modélisation et son exploitation sans être intrusif deviennent donc des champs d'investigation d'intérêt majeur, comme en FTLV où l'apprenant doit combiner vie professionnelle, vie privée et apprentissage.
- Personnalisation de la formation : les Learning Analytics sont un apport évident à l'industrialisation massive de l'individualisation de la formation, qui sans automatisation est impensable. De plus, comme les objectifs et les acteurs des Learning Analytics sont multiples et divers, une personnalisation en fonction des besoins, des intérêts ou du profil de l'utilisateur est nécessaire.
- Learning Analytics éthiques : préserver la vie privée est notamment un défi fondamental des Learning Analytics qui impacte significativement les algorithmes d'IA et de fouille de données conçus.
- Des Learning Analytics transparents et explicables : justifier les informations, indicateurs, recommandations et autres résultats fournis est un point essentiel. L'impact des Learning Analytics sera d'autant plus grand que les approches et les outils offerts auront la capacité à expliquer leur raisonnement, à tracer les données utilisées, à justifier leurs résultats et par la même à les rendre plus exploitables.
- Learning Analytics et visualisation : une attention toute particulière doit être portée à l'échange d'informations avec les acteurs concernés : les visualisations doivent être claires, lisibles par tout un chacun, compréhensibles et exploitables aisément, par tout un chacun, quels que soient son âge, son degré d'expertise et sa spécificité.

Les Learning Analytics sont ainsi un thème en plein essor, à la fois comme une thématique de recherche pluridisciplinaire (intelligence artificielle, informatique, sciences de l'éducation, droit, psychologie et ergonomie cognitive, ...), comme un service à valeur ajoutée pour les entreprises de l'Edtech et comme une demande des institutions éducatives confrontées à des défis aussi divers que



la massification, la flexibilité, la personnalisation ou l'individualisation de la formation. Les Learning Analytics touchent tous les secteurs de la formation, de la maternelle à l'enseignement supérieur, en formation initiale, en alternance ou continue, en formation professionnelle ou Formation Tout au Long de la Vie (FTLV), en présence ou à distance.

**L'objectif du projet LOLA (Laboratoire Ouvert en Learning Analytics) est la création et la mise en place d'un environnement ouvert et partagé de recherche, développement et accompagnement au déploiement des Learning Analytics.**

## 2 Organisation du projet LOLA

Le projet LOLA constitue un des éléments de la valorisation des données éducatives. Il participe, à terme, au développement de technologies pour l'éducation qui s'appuient sur l'exploitation des données numériques. Initié par le laboratoire LORIA, et soutenu par le MENJS dans le cadre d'une convention de recherche, LOLA a vocation à accueillir des contributions de multiples acteurs des Learning Analytics. L'objectif est d'avoir à terme une plateforme qui soit un lieu d'échange et de partage pour une communauté française de recherche qui se structure et une Edtech qui se développe, un lieu d'appropriation et d'accompagnement pour les institutions éducatives qui s'interrogent, un lieu de sensibilisation et d'information pour les enseignants qui expérimentent.

En effet, il apparaît nécessaire que 1) les équipes de recherche et les Edtech puissent disposer de corpus permettant d'apprendre des modèles, de les tester et de les comparer sur des données partagées, à l'aide d'indicateurs communs ; 2) les institutions éducatives disposent d'un lieu où trouver des informations notamment autour des aspects méthodologiques, déontologiques et éthiques, techniques, standardisation, ..., puissent partager des retours d'expérience (hub, d'expériences) sur les expérimentations menées ; 3) tous puissent accéder à une boîte à outils libres.

C'est pourquoi le projet LOLA est organisé de la manière suivante :

- un site Web ([lola.loria.fr](http://lola.loria.fr)) permettant d'obtenir les informations générales sur le projet LOLA et d'accéder à la plateforme LOLA.
- la plateforme LOLA qui offre un accès identifié aux différents services offerts par la plateforme (partage et mutualisation de corpus de données en environnement hautement sécurisé, définition ou utilisation de scénarios d'exploitation des données éducatives, rapport d'évaluation des scénarios permettant de comparer des approches de Learning Analytics entre elles à l'aide d'indicateurs sélectionnés dans une bibliothèque, outils spécifiques comme LolaMeter, ...).
- la collection LOLA, une documentation disponible sur le site web, décrivant les aspects essentiels du projet LOLA. Certains chapitres concernent les aspects scientifiques, d'autres les aspects organisationnels ou des retours d'expériences.

## 3 Le site web LOLA

A l'adresse <http://lola.loria.fr> se trouve le site web LOLA dont la page d'accueil se trouve ci-dessus. Sur ce site web, vous pouvez trouver l'information sur le projet LOLA, sur les Learning Analytics, sur les outils, données et scénarios disponibles, ainsi que les ressources mises à disposition des utilisateurs. En particulier, est accessible la collection LOLA qui propose des chapitres d'informations sur l'organisation, les fonctionnalités, services, outils du projet LOLA.



FIGURE 1 – Site LOLA

## 4 La plateforme LOLA

La plateforme LOLA se propose de mettre à disposition des chercheurs académiques, des décideurs éducatifs et des partenaires de l'EdTech, des observations d'expérimentation sous la forme de jeu de données. L'intérêt de la plateforme est double :

- Mettre à disposition des données réalistes, nettoyées et conformes à la réglementation (pour information, de nombreuses études rapportent que le temps consacré à collecter et à nettoyer les données représentent entre 40% et 70% du temps d'un projet de fouille de données).
- Fournir un environnement permettant l'accès non seulement à des données respectant des standards internationaux, mais aussi à des méthodes (basées sur des formats d'échange et des mesures de qualité des indicateurs) et des outils (notamment de reporting) permettant de comparer et d'objectiver les améliorations permanentes des techniques utilisées en éducation.

Il s'agit en particulier de permettre l'évaluation et la comparaison d'approches en Learning Analytics, en fournissant des indicateurs de qualité clairs et interprétables par des experts pédagogiques ou des chercheurs du domaine.

La Figure 2 montre schématiquement le workflow de la plateforme LOLA, tel qu'il a été imaginé.

D'un côté, il y a les outils qui produisent des données d'observations sous forme de traces numériques d'une population (ici essentiellement les apprenants). Ces données sont transmises (1) de manière normalisée et sécurisée à la plateforme. Ces données constituent (2) des jeux de données. Ces jeux de données sont exploitables (3) par des chercheurs habilités afin qu'ils effectuent (4) leur traitement analytique. Il est alors possible de soumettre (5) le résultat des traitements à la plateforme afin de produire (6) un rapport contenant les mesures de qualité communes à l'intention des chercheurs (7).

La plateforme LOLA entretient donc un ensemble de jeux de données. Un jeu de données regroupe les observations concernant un ensemble d'individus (appelés ici aussi communautés) comme par exemple les apprenants, leur environnement et leurs activités. Un jeu de données peut être utilisé dans un (ou plusieurs) scénarios.

Un scénario est formalisé à partir d'une problématique (par exemple, monitorer le parcours des

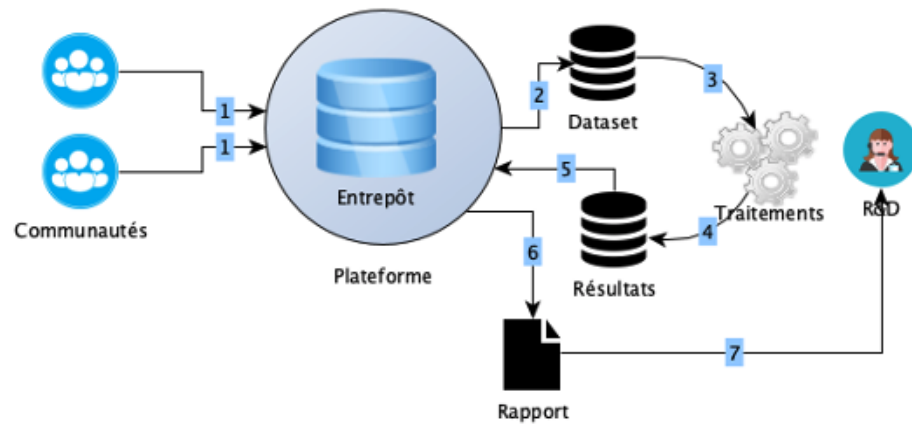


FIGURE 2 – WORKFLOW DE LOLA

élèves). Cette problématique s’appuie sur des indicateurs (par exemple, l’estimation des chances, à un instant de la formation, pour parvenir à l’objectif). Il est alors nécessaire d’utiliser des mesures de qualité de ces indicateurs.

Ainsi un scénario peut se définir comme l’addition de plusieurs fonctionnalités :

- un ou plusieurs jeux de données, parmi ceux disponibles sur la plateforme,
- des primitives d’accès aux données pertinentes,
- un ou plusieurs indicateurs, parmi ceux proposés sur la plateforme,
- des mesures de qualité pour évaluer et comparer ces indicateurs, sélectionnés dans une bibliothèque,
- un format d’échange des résultats,
- et un outil de reporting reproductible faisant le bilan d’une expérimentation.

Le scénario est exécuté sur la plateforme LOLA, à aucun moment les données ne quittent la plateforme LOLA et elles ne sont jamais accessibles aux utilisateurs. La prise en main de la plateforme LOLA par un utilisateur est facilitée par la mise à disposition de tutoriaux, expliquant pas à pas les concepts et les actions à entreprendre.

## 5 La collection LOLA

La collection LOLA regroupe la documentation sur le projet LOLA. Elle est accessible à l’adresse [lola.loria.fr/ressources](http://lola.loria.fr/ressources)