

Nous proposons au travers d'un outil comme *Github* un catalogue de scénarios d'analyse qui permette aux différents acteurs réalisant des analyses de traces d'activité (learning analytics) de partager au plus grand nombre les différentes procédures mises en œuvre dans l'analyse d'un dispositif particulier. Nous nous proposons dans les lignes qui suivent de revenir sur les fonctions de l'outil, les publics correspondant, les différents éléments que nous nous proposons de partager avec les internautes.

Le catalogue sur *Github* que nous nous proposons de développer dans le cadre de l'ANR Hubble a plusieurs fonctions et s'adresse à plusieurs publics de manière simultanée. Il s'adresse d'une part aux individus susceptibles de réaliser eux-mêmes des analyses sur la base des traces d'activité de dispositifs analogues, les analystes, et d'autre part à des décideurs de toutes sortes, enseignants, administrateurs des sites hébergeant les dispositifs, susceptibles de solliciter des analystes. L'outil Github est justifié par le bon référencement du site par les moteurs de recherche, ce qui contribue à la meilleure diffusion des scénarios d'analyse.

L'adresse du catalogue en question est

<https://github.com/hubble-learning-analytics/learning-analytics-catalog>.

On retrouve sur la page principale les hyperliens menant vers les différents scénarios d'analyse que comporte le catalogue, chacun de ces liens amenant sur la page d'un scénario d'analyse comme suit :

<https://github.com/hubble-learning-analytics/learning-analytics-catalog/tree/master/echantillonnage-au-sein-d-une-sequence-pedagogique> Dans le site actuel, trois exemples de scénarios ont été décrits, l'échantillonnage d'une séquence pédagogique, une analyse de séquence de modes d'appropriation d'un dispositif, et la détection d'un pattern prédéfini dans une séquence d'actions.

On trouve sur cette page, une page de présentation du scénario d'analyse, un fichier contenant les données réelles et/ou virtuelles sur lequel se base le scénario, et un fichier comportant les codes (en Python, en R, etc.) à mettre en œuvre pour réaliser le scénario d'analyse.

La page de présentation d'un scénario d'analyse comporte les éléments suivants :

- courte description du scénario
- tags (ex : analyse de séquence)
- cas d'utilisation, orientés vers les chercheurs ou les concepteurs de dispositif
- Inputs (variables nécessaires à la réalisation du scénario)
- Outputs (variables issues de la réalisation du scénario)
- Des indicateurs recommandés
- Des liens vers les données réelles (issues notamment d'un cas Hubble) ou factices permettant de mettre en œuvre le scénario
- Un exemple éventuel d'un graphique issu d'une mise en œuvre du scénario
- Des éléments de légitimation du scénario d'analyse : références dans la littérature scientifique, extraits d'entretiens avec des utilisateurs du dispositif

Nous nous proposons maintenant de revenir sur certains de ces éléments. Les cas d'utilisation ont vocation à fournir à l'analyste ou au décideur le ou les contextes dans lequel le scénario a vocation à être mis en œuvre. Un concepteur de MOOC pourra par exemple chercher à identifier les ressources pédagogiques du cours qui sont peu ou pas utilisées par les participants, tandis qu'un chercheur pourra davantage chercher à typologiser les participants selon les actions qu'ils réalisent au sein du dispositif.

Les inputs servent avant tout à informer l'analyste des éléments qu'il doit rassembler. La structure des données correspondantes peut être fournie sous forme de tableaux ; il en va de même pour les *outputs*. Ces outputs peuvent être nombreux, correspondre à des valeurs uniques, des matrices, des vecteurs. À partir de ces derniers, on pourra préciser quelques indicateurs pertinents dans le cadre du scénario, qu'on exemplifiera sur des données *Hubble* dès que l'opportunité se présente. Des données factices peuvent supplanter ou compléter de telles données lorsque la sensibilité, la complexité, ou le temps de préparation de données réelles justifie une telle démarche. Elles peuvent alors avoir une valeur didactique, dans la mesure où la construction même de ces données par l'utilisateur facilite la compréhension des données comme des scénarios. Dans le cas de données réelles, l'utilisateur doit disposer de tous les éléments pour mettre en œuvre de bout en bout le scénario d'intérêt.

Enfin, bibliographie comme entretiens avec des utilisateurs du dispositif analysé peuvent servir à légitimer un scénario d'analyse donné. La bibliographie peut se rapporter à des analyses identiques similaires à celles rapportées dans le scénario du catalogue réalisées sur des dispositifs analysés ou non dans le cadre de Hubble. Enfin, des entretiens éventuels peuvent illustrer de manière qualitative une analyse réalisée de manière quantitative via les traces d'activité.

Le projet initial consistait à réaliser un catalogue de scénarios basé quasi-exclusivement sur des entretiens avec des utilisateurs des dispositifs, mais la démarche ne semblait appropriée que pour quelques cas Hubble, dans la mesure où cette démarche de réalisation d'entretiens ne concernait qu'un nombre réduit de cas. Le projet de catalogue actuel, exemplifié sur Github – sachant que l'outil utilisé au final est susceptible de changer - dépasse en ambition le projet initial, dans la mesure où il a vocation à fournir tous les éléments nécessaires au partage et à la mise en œuvre de ces scénarios d'analyse, l'accent étant mis sur leur diffusion auprès d'utilisateurs potentiels.