

Acquisition de connaissances par l'exploration visuelle et interactive de traces

Mots clés :

- Learning Analytics,
- Ingénierie des connaissances,
- Découverte interactive / acquisition de connaissances à partir de traces.

Laboratoire/Entreprise :

[LIRIS](#) , UMR 5205 CNRS, Lyon, équipe [TWEAK](#) et/ou IAE – Université Lyon 3
Le stage est supervisé par [Béatrice Fuchs](#)

Durée : 5/6 mois, dès que possible à partir de 2021

Contact : [Béatrice Fuchs](#)

Profil du candidat : 2^{ème} année de master / dernière année d'école d'ingénieur, le candidat doit être motivé pour le développement, avoir des compétences en programmation web, Javascript et Python et des notions en fouille de données.

Contexte :

Les traces numériques d'interactions sont des informations laissées par les utilisateurs d'un système informatique lors de leurs activités. Elles témoignent de l'expérience des utilisateurs dans un environnement numérique et constituent des sources de connaissances intéressantes qui peuvent être exploitées à des fins variées en accord avec l'utilisateur. Une fois collectées et modélisées, les traces sont des conteneurs de connaissances riches en informations contextuelles et utilisables à la fois à des fins d'analyse pour inférer des connaissances pertinentes sur l'activité menée par les utilisateurs et également à des fins d'assistance à l'utilisateur. L'équipe [TWEAK](#) s'intéresse à des systèmes à base de traces où les traces sont associées à un modèle explicite avec une sémantique expressive sous la forme de *traces modélisées*. L'équipe a développé des approches visuelles et interactives pour analyser les traces [1].

Sujet du stage :

Dans ce contexte, le stage a pour objectif de contribuer au développement de prototypes d'exploration de traces. Plusieurs sujets sont possibles :

- 1°) Visualisation multidimensionnelle de traces : il s'agit d'étudier et concevoir une visualisation interactive de traces en fonction de dimensions paramétrables, et de proposer une première implémentation.

- 2°) Contribution au développement d'un *laboratoire d'analyse de traces* à partir de la plateforme TRANSMUTE d'exploration visuelle et interactive de traces. TRANSMUTE s'appuie sur un environnement visuel et interactif basé sur html/CSS/JS pour l'exploration visuelle et interactive de traces et exploite un framework de visualisation de traces, un algorithme d'extraction d'épisodes séquentiels (DMT4SP) encapsulé dans un module Python appelé DISKIT qui assure les pré- et post- traitements. Une ou plusieurs directions peuvent être étudiées, parmi lesquelles : le choix des mesures d'intérêt (actuellement prédéfinies et limitées) afin d'ordonner des résultats de la fouille, l'enrichissement des possibilités de pré-traitement, de post-traitement et de paramétrage (découpage des traces, patterns négatifs, combinaison d'attributs en pré-traitement, *etc.*), l'assistance à la construction d'un *modèle* à partir des épisodes séquentiels retenus à l'issue de la fouille et mémorisation dans un système de gestion de base de traces.
- 3°) Assistance à l'acquisition de connaissances à partir de traces : KATIE [1] est une approche d'acquisition de connaissances qui vise à assister, en interaction avec l'utilisateur, le processus de modélisation et d'intégration des traces dans un système à base de traces, en détectant et corrigeant les erreurs résiduelles dans les données (données bruitées ou manquantes). Cette approche utilise l'analyse de concepts formels pour générer un modèle des données analysées. Le processus d'acquisition de connaissances est interactif et itératif et permet à l'utilisateur d'introduire des contraintes afin d'identifier des erreurs dans les données et les corriger.

Références

- [1] Pierre-Antoine Champin, Béatrice Fuchs, Nathalie Guin, and Alain Mille. “Explicabilité : vers des dispositifs numériques interagissant en intelligence avec l'utilisateur”. In: *Workshop HIA 2020 – Humains et IA, travailler en intelligence, EGC 2020*. Ed. by David Gross-Amblard. Bruxelles, 2020, pp. 4–15. URL: <https://hal.inria.fr/hal-02462059>.
- [2] Béatrice Fuchs and Amélie Cordier. “Interactive Interpretation of Serial Episodes: experiments in Musical Analysis.” In: *Knowledge Engineering and Knowledge Management, 21st International Conference - EKAW-2018*. Ed. by C. Faron-Zucker and C. Ghidini. LNAI 11 313. Nancy, France: Springer, 2018, pp. 131–146.
- [3] Rémi Casado, Nathalie Guin, Pierre-Antoine Champin, and Marie Lefevre. “kTBS4LA : une plateforme d'analyse de traces fondée sur une modélisation sémantique des traces”. In: *Méthodologies et outils pour le recueil, l'analyse et la visualisation des traces d'interaction - ORPHEE-RDV*. Font-Romeu, France, Jan. 2017. URL: <https://hal.archives-ouvertes.fr/hal-01637548>.
- [4] Pierre-Loup Barazzutti, Amélie Cordier, and Béatrice Fuchs. “Transmute : un outil interactif pour assister l'extraction de connaissances à partir de traces”. In: *Extraction et Gestion des Connaissances - EGC 2016*. Ed. by Bruno Crémilleux and Cyril de Runz. Vol. RNTI-E-30. Reims, France, 2016, pp. 463–468. URL: <https://hal.archives-ouvertes.fr/view/index/docid/2481079>.
- [5] Mirco Nanni and Christophe Rigotti. “Extracting Trees of Quantitative Serial Episodes”. In: *Knowledge Discovery in Inductive Databases: 5th International Workshop, KDID 2006 Berlin, Germany, September 18, 2006 Revised Selected and Invited Papers*. Ed. by Sašo Džeroski and Jan Struyf. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 170–188.
- [6] Heikki Mannila, Hannu Toivonen, and A. Inkeri Verkamo. “Discovery of Frequent Episodes in Event Sequences”. In: *Data Mining and Knowledge Discovery 1 (1997)*, pp. 259–289.