

## Proposition d'exposé - GDR I3 Journée fouille de données

- **Thématique principale** : extraction de motifs, applications pour la fouille de données
- **Titre** : “Le sport électronique : des données et des challenges”
- **Orateur** : Mehdi Kaytoue
- **Affiliation** : LIRIS - INSA de Lyon
- **Durée estimée** : 30 minutes

Dans cet exposé, nous présenterons le sport électronique comme une application d'intérêt à la fouille de données et à l'extraction de connaissances. Il s'agit d'un phénomène nouveau en plein essor qui pourrait d'ici 10 à 20 ans concurrencer les sports traditionnels. Il s'agit avant tout d'un sport (professionnels, amateurs, équipes, championnats, commentateurs, sponsors, etc.) qui se différencie par l'utilisation d'un support électronique (jeu vidéo sur un ordinateur, une console) au cœur d'une industrie florissante ciblant les hommes de 15 à 35 ans.

Ancré dans le Web et les technologies sociales, le sport électronique génère une quantité importante de données de diverses origines et natures (complexes, structurées, dynamiques, spatio-temporelles, bruitées et inter-corrélées). Nous avons proposé une première étude de cette communauté afin de mesurer l'ampleur de ce phénomène via les TVs sociales, un service Web où des milliers de spectateurs assistent à des matchs de sport électronique et interagissent ensemble en direct [2].

Nous montrerons alors que les données qui gravitent autour de cette communauté attirent déjà divers domaines de recherche, comme l'étude des interactions hommes-machines [1] et les sciences cognitives [4, 3], mais aussi de nombreux industriels. En sciences cognitives par exemple, il est possible d'observer et comparer des comportements humains sur des données de jeux vidéos (valides, gratuites, facile d'accès), et d'utiliser les résultats dans un autre contexte faisant intervenir les mêmes sens de l'Homme.

En informatique, ces données sont sources de problèmes intéressants en intelligence artificielle [5], Web sémantique, traitement automatique des langues, “serious-games” et “gamification”, et systèmes de recommandation personnalisée. Ces problèmes sont suffisamment généraux pour que leurs solutions s'adaptent à d'autres problèmes (e.g. suivi de cible, analyse de trajectoires, aide à la ré-éducation, etc.).

Enfin, et surtout, on fait face dans tous les cas à d'important problèmes de fouilles de données, ces dernières pouvant se formaliser de diverses manières comme des grands graphes attribués et dynamiques, des séquences d'événements bruitées longues, etc. pour lesquels l'extraction de motifs et l'analyse de données en général est une piste de recherche d'actualité.

## Références

- [1] Gifford Cheung and Jeff Huang. Starcraft from the stands : understanding the game spectator. In Desney S. Tan, Saleema Amershi, Bo Begole, Wendy A. Kellogg, and Manas Tungare, editors, *CHI*, pages 763–772. ACM, 2011.
- [2] Mehdi Kaytoue, Arlei Silva, Loïc Cerf, Wagner Meira Jr., and Chedy Raïssi. Watch me playing, i am a professional : a first study on video game live streaming. In Alain Mille, Fabien L. Gandon, Jacques Misselis, Michael Rabinovich, and Steffen Staab, editors, *WWW (Companion Volume)*, pages 1181–1188. ACM, 2012.
- [3] Benjamin Law, M. Stella Atkins, Arthur E. Kirkpatrick, and Alan J. Lomax. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In Andrew T. Duchowski and Roel Vertegaal, editors, *ETRA*, pages 41–48. ACM, 2004.
- [4] Simon Fraser University. <http://skillcraft.ca/about-the-project/>.
- [5] Ben George Weber, Michael Mateas, and Arnav Jhala. A particle model for state estimation in real-time strategy games. In Vadim Bulitko and Mark O. Riedl, editors, *AIIDE*. The AAAI Press, 2011.