



d'observation mis en place à l'occasion du séminaire annuel de recherche de ce groupement.

L'utilisation effective des œuvres

Ce premier poste de développement prend l'œuvre elle-même pour objet, pour caractériser les contenus en ligne. L'objectif est d'obtenir des résultats précis et factuels sur la réalité des contenus culturels dématérialisés, puis de les traiter à la lumière d'études sur les pratiques telles que déclarées par les utilisateurs. Les premiers des travaux prévus sur ce poste sont les qualifications et quantifications des contenus présents sur les plateformes de streaming. Les premiers cas traités à ce titre sont les plateformes YouTube et Dailymotion.

Les supports ont été sélectionnés sur la base de leur popularité. Les travaux de recherche en informatique et mathématiques appliquées présentés ici visent à apporter des éléments de compréhension précis et factuels des contenus publics présents et consommés sur les principales plateformes de streaming en France, YouTube et Dailymotion. Compte tenu de la quantité des contenus présents sur ces plateformes et de la diversité des usages qui en sont

faits, il est en effet difficile d'en avoir une vision précise. Pourtant, il s'agit de supports très populaires, dont la place dans les usages est importante.

Le protocole consiste à réaliser un échantillonnage par *random walk* des vidéos publiques des plateformes. Cette méthode est utilisée pour déterminer les caractéristiques d'un graphe lorsqu'il est impossible de réaliser un échantillonnage aléatoire sur ce graphe (*i.e.* de sélectionner de façon aléatoire un nœud du graphe parmi l'ensemble des nœuds). Pour schématiser, cela correspond à une navigation classique d'un utilisateur qui circulerait un grand nombre de fois au hasard d'une vidéo à une autre parmi celles qui sont suggérées sur la page qu'il consulte.

L'échantillon (environ 3000 vidéos) est analysé pour induire les caractéristiques des vidéos hébergées. Les vidéos sont classées selon une première série de catégories *ad hoc* permettant d'en préciser le contenu (clips musicaux, tutoriels, par exemple). Les catégories dont la diversité le permettait ont fait l'objet d'une sous-catégorisation plus détaillée : clips musicaux, films, séries et média. Quelques précautions sont nécessaires à la bonne lecture de ces travaux afin de tenir compte de la surreprésen-

sentation potentielle de certaines vidéos « suggérées », de la limitation aux vidéos publiques, de la non-localisation des contenus et de son influence sur leur popularité, et enfin de leur catégorisation manuelle.

Sur la plateforme YouTube, les travaux de recherche mettent en exergue une présence notable de contenus culturels, dont la popularité est particulièrement importante.

On relève notamment que les clips musicaux représentent 13 % de l'ensemble des contenus. Un quart d'entre eux (25,47 %) sont « officiels » c'est-à-dire mis à disposition par les comptes officiels des artistes, des sociétés, des maisons de production, des organismes de diffusion, etc. Les trois quarts restants sont notamment des versions originales, ou des versions dont les bandes son ou l'image ont été modifiées, n'ayant pas été postées par des comptes officiels. Les clips musicaux dans l'ensemble sont de loin les contenus les plus populaires de la plateforme. Ils totalisent en moyenne 40 000 vues par jour. Parmi ces clips, ce sont les versions « officielles » qui totalisent le plus d'audience (plus de 110 000 vues en moyenne par jour). Le reste de l'audience se partage principalement entre les versions non officielles, dont celles dont la bande-son ou la vidéo support ont été modifiées. Parmi les contenus non officiels, on relève que 36,68 % des pages présentent des liens vers une plateforme permettant d'acheter l'œuvre relative à la vidéo.

- Les contenus relatifs aux films ou aux séries sous forme de parties organisées pour permettre le visionnage complet ou en intégralité constituent environ 9,65 % des vidéos présentes. Les extraits représentent 2,66 % des vidéos (les bandes annonces sont intégrées dans la catégorie « publicités »). La catégorie regroupant les parties de films organisées pour permettre le visionnage complet est la deuxième catégorie la plus populaire. Ses vidéos font l'objet d'environ 12 000 vues en moyenne chaque jour.
- Enfin les « médias », extraits ou intégralité de programmes audiovisuels initialement