AMMICO: recommandation sociale pour la visite de musée

Raphaël Fournier*, Emmanuel Viennet*

*L2TI – Institut Galilée, Univ. Paris-Nord, Paris-Cité prenom.nom@univ-paris13.fr

Résumé. Nous présentons les techniques de recommandations mises en œuvre pour la plateforme AMMICO qui permet une visite de musées enrichie. Nous proposons diverses techniques de recommandations reposant sur l'analyse de graphes sociaux afin de répondre aux problématiques spécifiques de la visite de musées.

1 Introduction

Les musées disposent de collections de très grandes taille et leur mission consiste à les mettre à disposition du public. Les visiteurs ne peuvent donc pas accéder à l'intégralité des œuvres disponibles et il est naturel d'envisager de compléter le travail de sélection des conservateurs avec des systèmes de recommandations informatisés.

Lors de la visite d'un musée, ou dans la planification d'activités touristiques en général, un visiteur est en effet confronté à un certain nombre de problèmes. Tout d'abord, il dispose généralement d'un temps limité à passer dans le musée. Il doit donc choisir, plus ou moins consciemment, les œuvres qu'il va voir (Tsai et Chung, 2012). Le visiteur qui vient pour la première fois ne sait pas forcément ce qu'il devrait voir ou ce qu'il va aimer. Le parcours qu'il effectue n'est donc généralement pas très réfléchi (Kuflik et al., 2011). En conséquence, il se peut que les œuvres qu'il ait vues ne l'aient pas particulièrement intéressé. Inversement il se peut aussi qu'il ait raté des œuvres qui l'auraient beaucoup intéressé.

Une approche maîtrisée de l'espace du musée a aussi une importance significative sur le déroulement de la visite et la satisfaction du visiteur (Tsai et Chung, 2012). En particulier, l'ordre dans lequel sont découvertes les œuvres est très important. Si cet ordre est mal choisi, un visiteur peut se retrouver à faire un nombre important d'allers-retours, ce qui va non seulement le fatiguer rapidement mais aussi lui faire perdre du temps pour découvrir de nouvelles œuvres. Dans certains musées et dans la plupart des expositions temporaires à forte affluence, l'ordre de visite est imposé à tous les visiteurs.

Un musée peut proposer un ensemble de parcours destinés à différents types de visiteurs. Cependant chacun de ces parcours est figé, et donc pas parfaitement adapté à un visiteur spécifique. Par ailleurs, si d'autres visiteurs suivent le même parcours, il est possible que le visiteur soit constamment confronté à une affluence de personnes pour chacune des œuvres contenues dans le parcours. Un objectif consiste donc à personnaliser les propositions de parcours, afin de mieux correspondre aux affinités de chacun.

Ces dernières années, des téléphones disposant d'une connexion Internet et de fonctionnalités avancées (dits *smartphone*) sont apparus. Ils permettent à leurs utilisateurs d'obtenir rapidement des réponses à leurs questions, et d'utiliser divers types de services à tout moment, quel que soit l'endroit où ils se trouvent. Ces services ont aussi la possibilité de s'adapter à la position de l'utilisateur grâce à la fonctionnalité de géolocalisation dont bénéficient les *smartphones*. Ces appareils ont permis à de nombreux visiteurs de musées de transformer leur expérience de visite : au lieu d'utiliser les audioguides traditionnels, un certain nombre de personnes, en particulier les jeunes, utilisent leurs propres *smartphones* pour obtenir des informations supplémentaires sur les œuvres qui les intéressent (Analytics et Frankly, 2012). Il devient donc pertinent de proposer des services dédiés à la visite de musées pour enrichir l'expérience des visiteurs. Au-delà de la fonctionnalité traditionnelle de l'audioguide (renseigner le visiteur sur une œuvre en particulier), ces dispositifs peuvent aussi être utilisés pour fournir des services supplémentaires lors de la visite (carte du musée avec œuvres localisées, calcul d'itinéraires, recherche d'œuvres dans le musée, etc.), éventuellement personnalisés en fonction du profil, des préférences et du contexte dans lequel se trouvent les utilisateurs.

Un visiteur peut par exemple se faire recommander : des œuvres adaptées au contexte présent (humeur, activité sociale, raison de la visite, ...), un parcours construit à partir d'une liste d'œuvres (recommandées ou non), un parcours adapté à ses contraintes (en particulier les contraintes de temps ou la distance de marche) (Baltrunas et al., 2011; Tsai et Chung, 2012).

Nous présentons dans cet article les techniques de recommandations que nous développons dans le cadre du projet AMMICO, de fournir aux visiteurs des recommandations personnalisées aux visiteurs. Dans la section 2, nous détaillons quelles sont les problématiques de recherche induites par le contexte de la visite de musée. Dans la section 3, nous proposons un aperçu des techniques proposées dans la littérature, avant de présenter notre approche du problème dans la section 5.

2 Problématiques scientifiques

Proposer des recommandations dans les musées soulève un certain nombre de problématiques, certaines étant génériques (Tintarev et al., 2010), d'autres plus spécifiques (Ruotsalo et al., 2009).

Cold start Le problème du démarrage à froid, ou du *cold start*, est double, affectant à la fois les utilisateurs et les objets. Il désigne le manque d'information généralement disponible sur un objet qui vient d'être ajouté au catalogue ou sur un utilisateur qui vient d'arriver dans le musée. On ne sait pas quels sont les utilisateurs intéressés par l'objet en question, ou quels objets intéresse l'utilisateur donné, ce qui rend difficile une prédiction efficace et donc des recommandations pertinentes. Ce problème est bien connu et étudié, il affecte avant tout les systèmes de filtrage collaboratif, qui reposent sur l'historique du comportement des utilisateurs pour calculer des recommandations pertinentes. Dans le contexte des musées, où les visiteurs viennent pour la plupart une fois seulement dans une exposition, la situation du *cold start* est très fréquente pour le système de recommandation.

Pénurie de notes Le problème de la pénurie de notes tient à la taille des ensembles des éléments sur lesquels opèrent le système de recommandation : il est courant que des millions d'objets soient disponibles et la probabilité que deux utilisateurs tirés au hasard aient noté le

même objet en est d'autant plus faible. Pour prédire si un objet plaira à cet utilisateur, le système doit alors choisir parmi un ensemble restreint d'utilisateurs similaires et risque de rater des utilisateurs plus pertinents. De fait, de nombreuses notes deviennent très dures à prédire. Dans les musées, ce qui peut amener à des difficultés pour trouver des utilisateurs similaires est moins lié à la taille des ensembles {objets,utilisateurs} qu'au nombre très limité d'interactions explicites dont le système dispose pour proposer des recommandations : contrairement à un système en ligne où on peut mettre en avant certaines parties d'une page en vue de collecter des retours explicites (clics sur "favoris" ou "like"), un dispositif de visite n'est pas en permanence sous les yeux du visiteur et celui-ci va sans doute peu indiquer ces préférences via cette interface.

Sur-spécialisation, absence de diversité Comme l'on souhaite proposer des recommandations personnalisées, on a pour objectif de recommander à un utilisateur les objets qui s'approchent le plus de ses préférences ou des objets *similaires* à ceux qu'il a aimés. Cependant, ce type d'approche a plusieurs inconvénients. Tout d'abord, il n'est pas nécessairement bon de recommander à un utilisateur un objet *trop similaire* à un objet qu'il a aimé, de la même façon qu'acheter la copie d'un tableau que l'on possède déjà a un intérêt discutable. On risque de recommander toujours le même type d'objets, ressemblant beaucoup à ceux que l'utilisateur connaît déjà. Cela aurait pour effet d'empêcher l'utilisateur de découvrir de nouveaux objets, ce qui aurait donné plus d'attrait à sa visite.

Longue traîne et popularité Pour un catalogue donné d'objets, il existe généralement une minorité d'objets très populaires que la majorité des gens aime, quels que soient leurs profils. Les autres objets ne plaisent qu'à des profils particuliers de personnes. Les objets très populaires ont, en raison même de cette popularité, de grandes chances d'être recommandés à un visiteur. Dans le pire des cas, les recommandations calculées ne sont même plus personnalisées : elles correspondent uniquement à la liste figée des objets les plus populaires. Recommander avec succès des objets peu populaires présente pourtant des avantages certains : pour les musées, cela permet de mieux valoriser l'ensemble du catalogue, ainsi que de limiter l'accès aux objets populaires. Pour l'utilisateur, cela permet de découvrir de nouveaux objets qui lui plaisent, potentiellement plus accessibles. Néanmoins, même si la recommandation est parfaitement en accord avec le profil de l'utilisateur, si celui-ci n'a jamais entendu parler de l'objet recommandé, il est possible qu'il ignore la recommandation. L'intégration d'objets très peu populaires dans la liste de recommandation est donc délicate.

Recommandations sous contraintes Dans le cas d'un musée, l'environnement spécifique apporte quelques problématiques supplémentaires. Il est notamment souhaitable de pouvoir proposer un parcours à un visiteur, c'est-à dire une liste d'œuvres à visiter dans un ordre défini. La recommandation ne porte plus alors seulement sur une œuvre en particulier, mais sur des ensembles d'œuvres. Une problématique est alors de savoir comment pondérer les scores de chacune des œuvres d'un parcours donné pour donner un score au parcours (et permettre la comparaison de plusieurs parcours possibles). Une autre problématique réside dans la prise en compte de la position des œuvres dans l'espace du musée, afin que non seulement les œuvres soient en rapport avec les goûts du visiteur mais aussi que ce dernier ne soit pas contraint à d'incessants aller-retour pour voir successivement les œuvres. De manière similaire, dans

certains musées, l'affluence devant les œuvres est une contrainte supplémentaire à prendre en compte, *en temps réel* ou non.

3 Recommandations et visite de musées

3.1 Formalisation du problème

Considérant l'ensemble U d'utilisateurs et l'ensemble I des objets (items que l'on peut proposer, effectuer des recommandations, c'est se donner une fonction d'utilité $R:U\times I\to\mathbb{R}$ où R(u,i)) est l'utilité – ou score – associée à (u, i), typiquement un nombre réel, pris dans un intervalle généralement fixé. La fonction n'est pas connue pour l'ensemble des éléments de $U\times I$, mais pour un sous-ensemble seulement. Pour chaque utilisateur u de U, on veut être capable d'estimer R(u,i)) pour les objets i de i pour lesquels i est encore inconnu, mais aussi trouver un ensemble d'objets qui maximisent i i est encore inconnu,

Les données que l'on peut utiliser pour effectuer les recommandations peuvent provenir du contenu, des utilisateurs, ou des interactions entre utilisateurs et objets. Pour modéliser le contenu, des mots-clefs sont généralement extraits de la description pour caractériser au mieux chacun des objets. Ceux-ci peuvent être complétés par des méta-données telles que des ontologies, qui structure et classifie des relations entre objets. Pour les utilisateurs, des attributs démographiques tels que le sexe et l'âge peuvent être pris en compte, ainsi que des centres d'intérêts. Enfin, différentes interactions entre utilisateurs et objets peuvent être collectées, qu'elles soient exprimées de façon explicites (œuvres marquées comme favorites) ou implicites (temps passé devant une œuvre, par exemple).

3.2 Approches dédiées

Générer des recommandations personnalisées dans le contexte de visite de musées a déjà été traité dans un nombre importants de projets, dont nous ne présenterons pas le détail dans cet article. On pourra se référer avec profit à l'étude comparative des différents systèmes existants présentée dans (Ardissono et al., 2012).

Historiquement centrée autour de la visite virtuelle sur des ordinateurs de bureau à la fin des années 90, cette problématique s'est progressivement déplacée vers des systèmes mobiles, à mesure que des dispositifs incorporaient des connexions Internet (du WAP jusqu'à la 4G actuelle). Dès 1999, les auteurs de (Bianchi et Zancanaro, 1999) ont classé des visiteurs selon leurs trajectoires afin de proposer des présentations personnalisées (projet HIPS). Leur projet a été prolongé par le projet HIPPIE qui utilisait une approche de modélisation dynamique du profil de l'utilisateur (Oppermann et Specht, 2000).

Diverses techniques classiques de recommandations ont été utilisées, qu'il s'agisse d'approches reposant purement sur le contenu ("content-based") (citons par exemple (Cramer et al., 2008)), sur le filtrage collaboratif ¹ (Bohnert et al., 2008) voire hybrides (Stock et al., 2007; Schmidt-Belz et al., 2003). En particulier, des techniques sociales, proches de celles que nous détaillons dans la section suivante ont été proposées dans iCity (Carmagnola et al., 2008) et

^{1.} Celui-ci pouvant être "user-based" ou "item-based".

Chat (De Gemmis et al., 2008). Mais celles-ci reposent avant tout sur des *tags* et des commentaires déposés par les utilisateurs, alors que nous proposons une méthodologie sociale plus générale.

4 Approche de recommandation élémentaire

Une approche simple pour proposer des recommandation d'œuvres aux visiteurs est d'utiliser des règles d'associations. Comme dans le cas d'un magasin où l'on peut analyser quels produits sont achetés ensemble, en vue de proposer l'un si le premier est acheté, on cherche à déterminer quelles sont la ou les œuvres qui sont les plus fréquemment vues avec une œuvre donnée.

La forme générale d'une règle d'association r est $r:m_i\to m_j$ où m_i est la prémisse de la règle et m_j est sa conclusion. Une règle d'association est classiquement évaluée par deux mesures, son support et sa confiance. Le support est défini, par extension de la notion de support d'un motif ensembliste, de la façon suivante :

$$supp(m_i \to m_j) = supp(m_i \cap m_j)$$

La confiance étant définie comme :

$$conf(m_i \to m_j) = \frac{supp(m_i \cap m_j)}{supp(m_i)}$$

La manière la plus classique de construire un ensemble de règles d'association consiste à produire dans un premier temps un ensemble d'itemsets fréquents (des œuvres fréquemment vues ensembles). À partir d'un itemset fréquent $\{A_1,A_2,\cdots,A_n\}$, on peut construire puis évaluer toutes les règles $m_i \to m_j$ avec $m_i,m_j \in \{A_1,A_2,\cdots,A_n\}$. Afin de contenir la complexité du processus d'extraction de règles, on s'en tiendra à des itemsets fréquents de taille 2. Les recommandations seront ensuite produites en ordonnant les règles d'associations par confiance décroissante.

5 Approches de recommandation sociale

L'approche de la recommandation que l'on souhaite développer et mettre en avant dans le projet AMMICO est une approche sociale : elle repose sur l'utilisation des informations du réseau social d'un individu afin d'améliorer la pertinence des recommandations qui lui sont faites, pour compléter ou supplanter les recommandations issues du filtrage collaboratif classique. On peut ainsi recommander des objets à un utilisateur en privilégiant les objets qui ont plu à ses amis, les objets qui ont plu à d'autres utilisateurs qui partagent les mêmes centres d'intérêt que lui, ou encore les objets similaires à ceux qui lui plaisent déjà. Pour cela, on va utiliser le réseau social pour constituer des segments que l'on va appeler communautés. Une communauté regroupe des entités (dans notre cas utilisateurs ou objets) que l'on va considérer comme similaires ou interdépendants. Une fois ces communautés obtenues, il devient ainsi aisé d'identifier un petit groupe d'objets qui ont de fortes chances de plaire à un utilisateur.

Réseaux utilisés Cette approche implique que l'on utilise des réseaux sociaux pour identifier les recommandations à faire à un utilisateur. Dans notre contexte, un réseau social sera mis en place autour de chaque établissement et permettra à terme une connexion avec des réseaux sociaux tiers (pour éviter de recréer ex-nihilo un réseau d'amitié). Cependant, nous utiliserons aussi des graphes issus du comportement des visiteurs, qui ne seront donc pas explicitement sociaux.

Nous proposons d'utiliser à la fois des graphes où les entités sont des visiteurs mais aussi des œuvres, lesquels pourront être issus de la projection d'un graphe biparti visiteurs-œuvres. La Figure 1 illustre cela : les visiteurs sont liés aux œuvres qu'ils ont vu dans l'exposition; on obtient le graphe "visiteurs" en projetant ce graphe biparti, c'est-à-dire en ne prenant en compte que les entités "visiteurs" et en les reliant s'ils ont vu au moins 3 œuvres communes (paramètre Ko). De même, on peut construire un graphe d'œuvres en reliant deux œuvres si elles ont été vues par au moins 3 visiteurs communs (paramètre Kv).

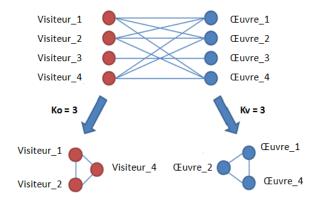


FIG. 1: Graphe biparti Visiteurs-Œuvres, et projections en graphe "visiteurs" et graphe "œuvres"

Détection de communautés Une communauté dans un réseau social se définit comme "un ensemble de nœuds plus densément connectés entre eux qu'avec le reste du graphe". Nous considérons, selon le graphe étudié, différents types de communautés : des communautés explicites d'amis, des communautés de centres d'intérêt communs, voire des communautés homogènes pour un certain type d'attributs (voir (Viennet et al., 2012)). Il existe de nombreuses techniques pour partitionner un graphe en communautés (voir (Fortunato, 2009)). Nous utilisons la méthode dite "de Louvain" présentée dans (Blondel et al., 2008), qui permet d'optimiser la modularité. Dans certains cas (lorsque l'on ne dispose que d'un graphe limité pour un visiteur donné par exemple), nous étudions les communautés locales auquel appartient cet utilisateur, à l'aide des techniques présentées dans (Ngonmang et al., 2012).

Ces méthodes nous permettent d'obtenir des visiteurs *proches* en terme d'amitié, de goûts ou d'attributs démographiques, à partir desquels on proposera des recommandations.

Procédure de recommandation En fonction des graphes analysés, on obtient différents types de communautés, desquels vont découler le type de recommandations que l'on va faire à un utilisateur : soit une recommandation communautaire *item-based* ou soit *user-based*. Le schéma illustrant la procédure générale est présenté sur la Figure 2.

Dans le premier cas, on cherche des similarités entre objets sur un graphe qui les relie, par exemple à partir des liens prédéfinis par les conservateurs du musées si l'exposition comporte un nombre limité d'œuvres. On utilise un algorithme de détection de communautés pour segmenter les objets du graphe. Pour chacun des objets d'une communauté, celle-ci représente l'ensemble des recommandations potentielles. Les recommandations à un utilisateur sont faites à partir des communautés auxquels appartiennent les objets qui ont plu à celui-ci. On peut recommander les objets de la communauté, ou bien affiner les performances en utilisant de façon additionnelle d'autres indices de similarité, ou encore combiner les résultats obtenus avec d'autres graphes.

De façon similaire, on propose une approche de recommandation communautaire *user-based*, où ce sont les liens entre visiteurs qui sont utilisés pour former le graphe, à partir duquel on partitionne l'ensemble des visiteurs en communautés. On utilise ensuite les préférences des utilisateurs de la communauté à laquelle appartient le visiteur pour lui proposer des recommandations.

L'utilisation de graphes sociaux pour faire de la recommandation comporte plusieurs avantages. Comme les relations sociales se nouent en général entre personnes qui se ressemblent (homophilie), ces relations sont supposées aider à fournir de bonnes recommandations. De plus, comme pour un visiteur donné, ces relations ne sont pas uniquement avec des visiteurs aux goûts *similaires*, on peut espérer davantage de diversité que dans le cas d'un filtrage collaboratif classique.

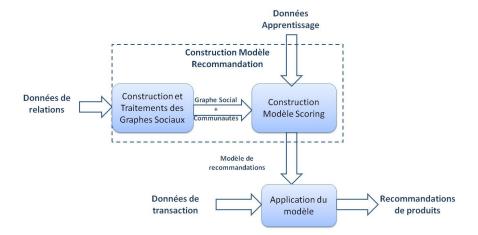


FIG. 2: Illustration de la procédure de recommandation

6 Conclusion et perspectives

Nous avons présenté dans cet article le contexte de la visite des musées et les possibilités d'enrichissement de celle-ci à l'aide de techniques de recommandations. Nous proposons une approche reposant sur l'analyse de graphes sociaux pour proposer des recommandations aux visiteurs, en espérant notamment s'affranchir des limitations traditionnelles telles que le *cold-start*.

Les perspectives de poursuites de ce travail sont nombreuses. Un logiciel prototype fonctionnant sur des appareils Android et embarquant un nombre de fonctionnalités limitées a été mis en place pour l'exposition "Great Black Music" à la Cité de la Musique de Paris, avec pour objectif principal de collecter à grande échelle des *données réelles* de visite de musée. L'analyse de ces données devrait permettre d'incorporer les fonctionnalités de création à la volée de parcours en tenant compte de contraintes telles que l'affluence. Des fonctionnalités de géolocalisation en intérieur sont développées par des partenaires du projet et, une fois intégrées, elles devraient également permettre d'améliorer la précision disponible quant à la position des visiteurs, et la pertinence des recommandations.

Références

- Analytics, F. R. . et G. . W. Frankly (2012). Understanding the mobile v&a visitor. English
- Ardissono, L., T. Kuflik, et D. Petrelli (2012). Personalization in cultural heritage: the road travelled and the one ahead. *User Modeling and User-Adapted Interaction* 22(1-2), 73–99.
- Baltrunas, L., B. Ludwig, S. Peer, et F. Ricci (2011). Context-aware places of interest recommendations for mobile users. In A. Marcus (Ed.), *Design, User Experience, and Usability. Theory, Methods, Tools and Practice*, Volume 6769 of *Lecture Notes in Computer Science*, pp. 531–540. Springer Berlin Heidelberg.
- Bianchi, A. et M. Zancanaro (1999). Tracking users' movements in an artistic physical space. In *Proc. i3 Annual Conference*, pp. 103–106.
- Blondel, V. D., J.-L. Guillaume, R. Lambiotte, et E. Lefebvre (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics : Theory and Experiment 2008*(10), P10008.
- Bohnert, F., I. Zukerman, S. Berkovsky, T. Baldwin, et L. Sonenberg (2008). Using interest and transition models to predict visitor locations in museums. *AI Communications* 21(2), 195–202.
- Carmagnola, F., F. Cena, L. Console, O. Cortassa, C. Gena, A. Goy, I. Torre, A. Toso, et F. Vernero (2008). Tag-based user modeling for social multi-device adaptive guides. *User Modeling and User-Adapted Interaction* 18(5), 497–538.
- Cramer, H., V. Evers, S. Ramlal, M. Someren, L. Rutledge, N. Stash, L. Aroyo, et B. Wielinga (2008). The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction* 18(5), 455–496.

- De Gemmis, M., P. Lops, G. Semeraro, et P. Basile (2008). Integrating tags in a semantic content-based recommender. In *Proceedings of the 2008 ACM conference on Recommender systems*, pp. 163–170. ACM.
- Fortunato, S. (2009). Community detection in graphs. *Physics Reports*.
- Kuflik, T., O. Stock, M. Zancanaro, A. Gorfinkel, S. Jbara, S. Kats, J. Sheidin, et N. Kashtan (2011). A visitor's guide in an active museum: Presentations, communications, and reflection. *J. Comput. Cult. Herit.* 3(3), 11:1–11:25.
- Ngonmang, B., M. Tchuente, et E. Viennet (2012). Local community identification in social networks. *Parallel Processing Letters* 22(01).
- Oppermann, R. et M. Specht (2000). A context-sensitive nomadic exhibition guide. In *Hand-held and Ubiquitous Computing*, pp. 127–142. Springer.
- Ruotsalo, T., E. Mäkelä, T. Kauppinen, E. Hyvönen, K. Haav, V. Rantala, M. Frosterus, N. Do-koohaki, et M. Matskin (2009). Smartmuseum personalized context-aware access to digital cultural heritage.
- Schmidt-Belz, B., H. Laamanen, S. Poslad, et A. Zipf (2003). Location-based mobile tourist services-first user experiences. *Information and communication technologies in tourism* 2003, 115–123.
- Stock, O., M. Zancanaro, P. Busetta, C. Callaway, A. Krüger, M. Kruppa, T. Kuflik, E. Not, et C. Rocchi (2007). Adaptive, intelligent presentation of information for the museum visitor in peach. *User Modeling and User-Adapted Interaction* 17(3), 257–304.
- Tintarev, N., A. Flores, et X. Amatriain (2010). Off the beaten track: A mobile field study exploring the long tail of tourist recommendations. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, New York, NY, USA, pp. 209–218. ACM.
- Tsai, C.-Y. et S.-H. Chung (2012). A personalized route recommendation service for theme parks using rfid information and tourist behavior. *Decis. Support Syst.* 52(2), 514–527.
- Viennet, E. et al. (2012). Community detection based on structural and attribute similarities. In *ICDS 2012*, *The Sixth International Conference on Digital Society*, pp. 7–12.

Summary

Nous présentons les techniques de recommandations mises en œuvre pour la plateforme AMMICO qui permet une visite de musées enrichie. Nous proposons diverses techniques de recommandations reposant sur l'analyse de graphes sociaux afin de répondre aux problématiques spécifiques de la visite de musées.