
Un système de recommandation sensible au contexte pour la visite de musée

Idir Benouaret ¹

*Sorbonne universités, Université de technologie de Compiègne, CNRS, Heudiasyc
UMR 7253, CS 60 319, 60 203 Compiègne cedex.*

idir.benouaret@hds.utc.fr

RÉSUMÉ. Notre travail concerne les systèmes d'aide à la visite de musée et l'accès au patrimoine culturel. Notre but est de concevoir un système de recommandation, implémenté sur dispositifs mobiles adaptable au profil de l'utilisateur et sensible à son contexte, afin d'améliorer l'expérience de l'utilisateur et de l'aider à construire ses parcours en musée en fonction de ses préférences et ses contraintes.

ABSTRACT. Our work is in the field of the support systems to museum visits and access to cultural heritage. Our goal is to design a recommendation system implemented on mobile devices adaptable to the user's profile and sensitive to its context, to improve the user experience and help build his visits to the museum based on his preferences and constraints.

MOTS-CLÉS : systèmes de recommandation, web sémantique, contexte, musées.

KEYWORDS: recommender systems, semantic web, context, museums.

1. Encadré par : Dominique Lenne

1. Introduction

La recherche sur les systèmes mobiles d'assistance à la visite de musée a été particulièrement importante ces dernières années. En effet, les musées disposent en général de collections de très grande taille mises à disposition du grand public. Cependant, les visiteurs ne peuvent pas accéder à l'intégralité des œuvres présentées dans le musée. Le visiteur d'un musée est confronté à un certain nombre de problèmes. Tout d'abord, il dispose en général d'un temps limité à passer dans le musée et 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 le visiteur perde du temps en regardant des œuvres qui ne l'intéressent pas beaucoup. Inversement, il se peut qu'il ait manqué des œuvres qui auraient pu l'intéresser.

Pour répondre à la diversité des préférences des visiteurs, les conservateurs de musées proposent des visites sur différents thèmes. Cependant, ces thèmes sont généralement choisis en fonction des œuvres les plus connues de la collection et les visites comprennent une séquence fixe et prédéfinie d'œuvres d'art à voir, qui est la même pour tous les visiteurs. Or, imposer un parcours est une contrainte forte pour les visiteurs, qui aiment à faire usage de leur liberté de visiter (Gob et Drouguet, 2006).

Ces dernières années, des dispositifs mobiles disposant de fonctionnalités avancées (smartphones) sont apparus. Ils permettent à leurs utilisateurs d'accéder à des services à tout moment, quel que soit l'endroit où ils se trouvent. Ces services peuvent aussi s'adapter à la position de l'utilisateur grâce à la fonctionnalité de géolocalisation. Pour faire face aux différents problèmes cités précédemment et aller au-delà des fonctionnalités traditionnelles qu'offre l'audio guide, nous proposons de mettre en place un système de recommandation appliqué à la visite de musées. Le but principal est de construire un système sur dispositif mobile capable d'améliorer l'expérience de l'utilisateur lors de la visite de musées en lui recommandant des œuvres qui correspondent à ses goûts et en l'aidant à construire un parcours de visite adapté à son contexte et à ses préférences. Notre proposition consiste à concilier deux approches (méthode hybride) : une approche basée sur la représentation sémantique du domaine muséal ainsi que sur les préférences de l'utilisateur et une approche collaborative qui utilise l'avis d'utilisateurs similaires. Nous cherchons également à prendre en compte le contexte (localisation, temps, environnement physique...) pour la génération de parcours de visite dynamiques et personnalisés. Nous considérons comme domaine d'application le musée du palais impérial de Compiègne situé en Picardie.

L'article est organisé comme suit. Nous présentons dans la section 2 un bref état de l'art sur les systèmes de recommandation et sur les systèmes d'aide à la visite de musées. Dans la section 3, nous présentons l'architecture de notre système tout en détaillant les composants de cette architecture ainsi que les approches proposées. Nous présentons, en particulier, le modèle de connaissance proposé pour représenter le domaine muséal, le modèle du contexte, le système de recommandation et le générateur de parcours. La section 4 souligne les conclusions de ce travail et la motivation pour les travaux futurs.

2. Etat de l'art

Nous présentons ci-dessous un bref état de l'art sur les systèmes de recommandation ainsi que quelques systèmes d'aide à la visite de musées.

2.1. Systèmes de recommandation

Les systèmes de recommandation peuvent être classés de différentes manières. La classification la plus fréquente est une classification selon deux approches : les recommandations basées sur le contenu et le filtrage collaboratif (Balabanović et Shoham, 1997).

Les approches basées contenu n'utilisent que l'avis de l'utilisateur pour générer des recommandations. Elles sont focalisées sur l'historique de l'utilisateur en utilisant les informations qu'il a vues ou notées auparavant. On cherche ainsi à déterminer des caractéristiques communes d'intérêt pour un utilisateur (Pazzani et Billsus, 2007). Le filtrage collaboratif génère quant à lui des recommandations personnalisées en se basant sur les opinions des autres utilisateurs ayant des intérêts proches. Le filtrage collaboratif a été utilisé pour la première fois par (Resnick *et al.*, 1994) afin de générer des recommandations d'actualités sur la plate-forme GroupLens. Cette plate-forme collectait des notes sur des articles et générait une prédiction de la note d'un utilisateur en moyennant les notes données par les utilisateurs les plus similaires à cet individu. On parle alors de méthode de recherche des plus proches voisins.

Pour proposer un système de recommandation dans le cadre de la visite de musées, il nous faut répondre à de nombreux problèmes. Certains sont génériques et concernent n'importe quel système de recommandation (Adomavicius et Tuzhilin, 2005), d'autres sont plus spécifiques au cas des musées.

– Démarrage à froid (Cold start) : le problème du démarrage à froid est très fréquent dans les systèmes de recommandation. Ce problème est double, il affecte à la fois les utilisateurs mais aussi les objets. Il désigne le manque d'informations sur un utilisateur ou sur un objet qui vient d'être ajouté à la collection.

– Sur-spécialisation (Over specialization), absence de diversité : lorsque le système ne peut recommander à un utilisateur que des éléments qui sont en relation avec son profil, l'utilisateur est limité à des recommandations de ressources trop similaires à celles qu'il a déjà aimées. En revanche, la diversité des recommandations est souvent une caractéristique souhaitable pour les systèmes de recommandation.

– Parcimonie (Sparsity) : Dans tout système de recommandation, le nombre de notes déjà obtenues est généralement très faible par rapport au nombre de notes qui doivent être prédites. Les objets très populaires ont plus de chances d'être recommandées par rapport aux autres. Cependant, si un utilisateur a des préférences insolites par rapport au reste de la communauté, il sera difficile de faire des recommandations pertinentes pour ce genre d'utilisateurs.

Dans le cas des musées, l'environnement physique apporte quelques problématiques supplémentaires. Il est important de pouvoir proposer un parcours de visite, c'est-à-dire une liste d'œuvres à regarder successivement dans un ordre défini. La recommandation peut alors porter sur un ensemble d'œuvres et pas seulement sur une œuvre en particulier. Une problématique est alors de savoir comment faire des comparaisons entre plusieurs parcours. Il faut aussi prendre en compte le contexte de l'utilisateur. Par exemple, le contexte de localisation ainsi que la position des œuvres dans le musée, afin que le visiteur puisse se déplacer de manière efficace et ne soit pas contraint à d'incessants aller-retour, ce qui réduirait sa satisfaction.

2.2. Systèmes d'aide à la visite de musées

Parmi les travaux qui se sont intéressés à la recommandation lors de la visite de musée, nous décrivons ci-dessous brièvement les projets Hippie et CHIP, ainsi que les limites que nous avons pu noter dans ces projets.

Le projet Hippie (Oppermann et Specht, 1999) a été l'un des premiers systèmes de recommandation basé contenu dans le cadre de la visite de musée. Hippie utilise une caractérisation des travaux basés sur la taxonomie ICONCLASS, une classification exhaustive des différents thèmes de l'art occidental. Les visiteurs sont caractérisés par des scores d'intérêt pour les différents thèmes de la taxonomie. Quand un visiteur se déplace dans le musée, le système détecte sa position. Par conséquent, Hippie est en mesure d'informer le visiteur sur les œuvres qui peuvent l'intéresser et qui se situent autour de lui. Cependant, le modèle des œuvres est basé uniquement sur ICONCLASS, ce qui en limite l'intérêt. On peut noter également que le système ne propose pas de parcours au visiteur.

Le projet CHIP (Wang *et al.*, 2009) est l'un des projets les plus aboutis en ce qui concerne les approches basées contenu pour l'aide à la visite de musées. Le système recommande des œuvres et des thèmes qui correspondent au profil de l'utilisateur. Le profil est basé sur les notes d'intérêt données par l'utilisateur aux différentes œuvres ou à leurs caractéristiques (artiste, style,...). À partir des notes positives données, le système peut alors proposer à l'utilisateur des œuvres ou des informations en lien avec celles qu'il a aimées ou qui l'ont intéressé. Le modèle de l'œuvre est plus riche dans le projet CHIP comparé au modèle proposé dans HIPPIE. Cependant, il ne permet la comparaison qu'entre les œuvres d'art (et non pas entre artistes par exemple). De plus, le nombre de recommandations est trop élevé (toutes les recommandations ne sont pas pertinentes pour l'utilisateur). Le contexte ainsi que l'avis des autres utilisateurs n'est pas pris en compte.

3. Un système de recommandation pour la visite de musée

3.1. Architecture générale

Comme il a été dit précédemment, nous avons choisi de nous orienter vers la conception d'un système de recommandation en musée. La scénarisation des parcours est de ce fait réalisée dynamiquement en fonction des recommandations issues de ce système. L'architecture générale du prototype que nous concevons est présentée en figure 1.

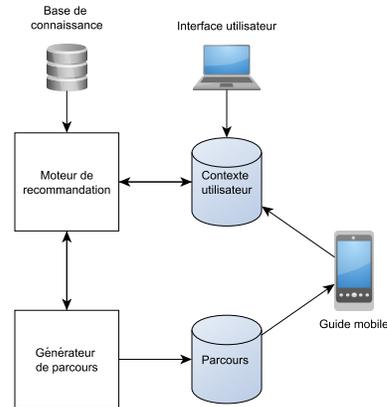


Figure 1 – Architecture générale

3.2. Modèle de connaissance

Le modèle de connaissances que nous proposons (figure 2) repose essentiellement sur le modèle défini par (Gicquel *et al.*, 2013) qui utilise l'ontologie CIDOC-CRM (Center for Intercultural Documentation-Conceptual Reference Model) (Doerr *et al.*, 2007) qui est l'ontologie de référence pour la description sémantique du patrimoine culturel, la taxonomie ICONCLASS (CLASSification iconography) qui permet de caractériser les thèmes des œuvres ainsi que le thesaurus AAT (Art & Architecture Thesaurus) pour la description des modes de production des œuvres. Dans le but de proposer un modèle de connaissances plus riche, nous proposons d'ajouter les deux thesaurus ULAN (Union List of Artist Names) et TGN (Thesaurus of Geographic Names).

3.3. Modèle du contexte

Le modèle du contexte que nous proposons s'inspire des travaux de (Zimmermann *et al.*, 2007). En l'adaptant à nos besoins spécifiques dans le cadre de la visite de musée, il se compose de cinq catégories :

Individualité : regroupe toutes les informations personnelles sur un utilisateur (âge, niveau d'expertise, etc) utilisées lors du filtrage collaboratif. Plus les informations entre deux utilisateurs sont proches, plus ils auront tendance à avoir les même

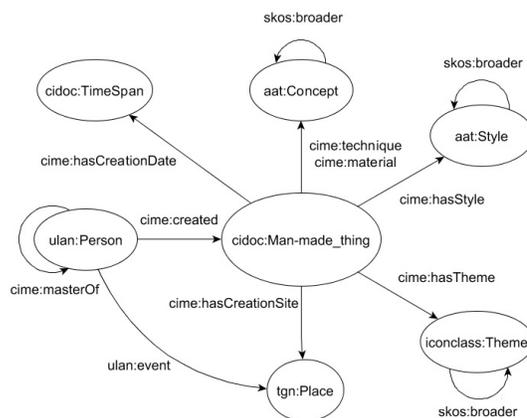


Figure 2 – Modèle de l'œuvre

préférences.

Localisation : localisation de l'utilisateur dans le musée ainsi que le positionnement des œuvres. Ces informations interviennent lors de la construction du parcours en musée pour détecter la position du visiteur et lui permettre d'effectuer un parcours qui minimise son temps de déplacement.

Temporalité : informations concernant le temps que souhaite passer le visiteur dans le musée, le temps qu'il passe devant chaque œuvre, etc. Afin de pouvoir calculer le nombre d'œuvres du parcours proposé à l'utilisateur.

Activité : œuvres qu'a consulté un utilisateur, notes d'appréciation qu'il a donné pour celles-ci ainsi que leurs différentes caractéristiques, etc. Ces informations sont principalement utilisées dans l'approche basée contenu, pour pouvoir recommander à l'utilisateur des œuvres en lien avec celles qu'il a déjà appréciées.

Relations : liste des utilisateurs qui sont similaires à l'utilisateur considéré. Cette liste est utilisée pour recommander à l'utilisateur des d'œuvres que des utilisateurs proches ont regardées

3.4. Moteur de recommandation

3.4.1. Approche basée contenu

Cette approche recommande à l'utilisateur des œuvres en lien avec celles qu'il a déjà aimées. Par exemple, si un aime « La Joconde », son créateur « Léonard de Vinci » ainsi que d'autres œuvres du même style ou du même artiste peuvent éventuelle-

ment lui être recommandés. Cette approche utilise les informations présentes dans le contexte de l'utilisateur considéré (notamment les œuvres notées positivement) ainsi que le modèle sémantique des œuvres présenté précédemment, pour suggérer à l'utilisateur des œuvres similaires à celles présentes dans son profil. Il faut alors définir une mesure permettant de calculer la similarité entre deux œuvres de la base de connaissance muséale. Cependant, pour les systèmes de recommandation, l'utilisation des relations sémantiques peut poser un problème. En effet dans le projet CHIP (Wang *et al.*, 2009), le nombre de recommandations faites à l'utilisateur est très grand, et toutes les recommandations ne sont pas pertinentes. Nous proposons donc une mesure de similarité entre les œuvres de la base de connaissance qui utilise les propriétés sémantiques des œuvres pondérées en fonction des goûts de l'utilisateur :

$$SIM(o_i, o_j) = \sum_k W_k * SIM_k(o_i, o_j) \quad [1]$$

où $SIM_k(o_i, o_j)$ représente la similarité entre l'œuvre o_i et l'œuvre o_j par rapport à la propriété k . si le domaine de la propriété k consiste en des instances organisées hiérarchiquement, le similarité de Wu et Palmer est utilisée. Sinon, on utilise l'indice de Jaccard. W_k est un poids représentant le facteur d'importance de la propriété k , avec $\sum_k W_k = 1$. Les poids attribués à chaque relation peuvent varier entre différents utilisateurs. En effet, un visiteur peut préférer des recommandations du même artiste tandis qu'un autre des recommandations du même thème.

Nous pouvons alors estimer si une nouvelle œuvre o_j non encore consultée peut être intéressante pour un utilisateur u . On calcule alors la similarité entre cette nouvelle œuvre et celles présentes dans son profil et on détermine la prédiction de la manière suivante :

$$profil(u) = \{o_i / u \text{ aime } o_i\} \quad [2]$$

$$pred(u, o_j) = \frac{\sum_{o_i \in profil(u)} SIM(o_i, o_j)}{|profil(u)|} \quad [3]$$

$|profil(u)|$ étant le nombre d'œuvres présentes dans le profil de l'utilisateur. La formule adoptée pour calculer $pred(u, o_j)$ prend en compte la similarité entre les œuvres présentes dans le profil de l'utilisateur et l'œuvre o_j candidate à la recommandation. si la valeur de $pred(u, o_j)$ est supérieure à un certain seuil, alors l'œuvre o_j est recommandée à l'utilisateur.

3.4.2. Approche collaborative

Cette approche recommande à l'utilisateur des œuvres que les utilisateurs les plus similaires ont aimées. Une approche basique pour calculer la similarité entre deux utilisateurs consiste simplement à compter la proportion d'œuvres communes dans

leurs historiques. La méthode PCC (Pearson's Correlation Coefficient) (Herlocker *et al.*, 1999) a particulièrement retenu notre attention dans la littérature :

$$sim_1(a, b) = \frac{\sum_i (r_{ai} - \bar{r}_a)(r_{bi} - \bar{r}_b)}{\sqrt{\sum_i (r_{ai} - \bar{r}_a)^2 (r_{bi} - \bar{r}_b)^2}} \quad [4]$$

Avec \bar{r}_a et \bar{r}_b la moyenne des notes données par deux utilisateurs a et b respectivement, et r_{ai} et r_{bi} les notes données par les deux utilisateurs à l'œuvre i

Cette mesure est celle qui est utilisée dans les méthodes de filtrage collaboratif traditionnel. Cependant, son utilisation pose un certain nombre de problèmes. La similarité entre deux utilisateurs ne peut être significative que si les deux utilisateurs ont un nombre important d'œuvres en commun dans leur historiques. Il est aussi connu que cette méthode souffre du problème de parcimonie. Dans le but de faire face à ces problèmes, nous proposons d'intégrer l'information sémantique entre les œuvres lors du calcul de la similarité entre utilisateurs. Il s'agit de dire que même si deux utilisateurs n'ont pas beaucoup d'œuvres en commun, si les œuvres consultées sont proches sémantiquement, la similarité entre les deux utilisateurs est assez importante :

$$sim_2(a, b) = \max(sim(a * b), sim(b * a)) \quad [5]$$

avec :

$$sim(a * b) = \frac{\sum_{k_1} \max_{k_2} (SIM(o_{k_1}, o'_{k_2}))}{k_1} \quad [6]$$

$$sim(b * a) = \frac{\sum_{k_2} \max_{k_1} (SIM(o'_{k_2}, o_{k_1}))}{k_2} \quad [7]$$

Sachant que l'utilisateur a a noté les œuvres o_1 jusqu'à o_{k_1} , l'utilisateur b a noté les œuvres o'_1 jusqu'à o'_{k_2} , SIM représente la mesure de similarité entre les œuvres.

Finalement, la similarité entre deux utilisateurs est définie comme la moyenne pondérée des deux mesures de similarité présentées précédemment :

$$Sim(a, b) = \sum_{k=1}^2 \alpha_k Sim_k(a, b) \quad [8]$$

Nous pouvons maintenant évaluer si une nouvelle œuvre o_j non encore consultée peut être intéressante pour un utilisateur u . On calcule alors la prédiction de la note que pourrait donner l'utilisateur à cette œuvre en moyennant les notes que ses K utilisateurs les plus similaires ont donné, il s'agit d'une méthode des K plus proches voisins bien connue dans le domaine de la classification.

$$pred(u, o_j) = \frac{\sum_{i=1}^K r(u_i, o_j) * Sim(u_i, u)}{\sum_{i=1}^K Sim(u_i, u)} \quad [9]$$

Avec $r(u_i, o_j)$ la note donnée par l'utilisateur u_i à l'œuvre o_j et $Sim(u_i, u)$ la similarité entre les deux utilisateurs u_i et u . Si la valeur de $pred(u, o_j)$ est supérieure à un certain seuil, alors l'œuvre o_j est recommandée à l'utilisateur.

3.5. Génération dynamique de parcours de visite

Le module "générateur de parcours" propose un parcours aux visiteurs en fonction des résultats du moteur de recommandation, en utilisant un post filtrage contextuel et les informations sur l'environnement du musée. Nous proposons de définir le parcours en musée comme un graphe orienté, où les nœuds du graphe représentent les œuvres et les arcs le temps de déplacement entre les œuvres. Ainsi, le parcours en musée est similaire au problème classique du voyageur de commerce. Généralement, le musée est organisé en salles. Dans chaque salle sont exposées un ensemble d'œuvres, et une salle peut être adjacente à une ou plusieurs salles. En supposant que l'utilisateur ne voie qu'une seule fois la même œuvre, une approche de type "greedy" peut être utilisée. Cela signifie sélectionner toujours l'œuvre qui est la plus proche de l'utilisateur. Nous supposons que l'utilisateur n'est pas en soi intéressé à voir les œuvres qu'il aime le plus, mais plutôt à voir l'ensemble des œuvres qui recouvrent le mieux ses intérêts. Par exemple, un utilisateur qui aime les "portraits" et les "paysages" ne serait pas satisfait par un parcours contenant seulement des "paysages".

Nous construisons un parcours de visite contenant N œuvres à voir dans un ordre bien défini, à partir des œuvres présentes dans le profil de l'utilisateur, les œuvres qui lui sont recommandées ainsi que les œuvres majeures du musée. La valeur de N dépend du temps que le visiteur souhaite passer dans le musée. Ainsi, l'estimation de la valeur de N dépend principalement des contraintes temporelles : le temps passé devant chaque œuvre et le temps d'aller d'une œuvre à une autre.

4. Conclusions et perspectives

Nous avons présenté dans cet article le contexte de la visite de musée et les possibilités d'enrichissement de celle-ci à l'aide de techniques de recommandation. Nous proposons une approche hybride reposant à la fois sur la modélisation sémantique du domaine muséal et sur la modélisation du contexte de l'utilisateur en essayant de faire face aux limites traditionnelles des systèmes de recommandation ainsi qu'à celles des systèmes d'aide à la visite de musées. Nous prévoyons de réaliser une expérimentation de notre système au musée du palais impérial de Compiègne. L'objectif principal est de collecter à grande échelle des données réelles de visites de musée. Ces données devront permettre de réaliser des premiers tests sur le prototype et avoir un premier retour sur la satisfaction des visiteurs ainsi que l'avis des experts des musées quant à la qualité des recommandations et parcours proposés. Des fonctionnalités de localisation en intérieur sont développées par des partenaires du projet qui, une fois intégrées, devraient permettre d'améliorer la précision actuelle quant à la position des visiteurs dans le musée. Notons que le parcours doit s'adapter dynamiquement à l'utilisateur.

En effet, l'utilisateur a le choix à tout moment de suivre à la lettre le parcours recommandé ou pas. Par exemple, s'il remarque au cours de la visite une œuvre qui l'attire, il peut aller la contempler et donner son avis sur celle-ci. Le système doit alors mettre à jour le profil de l'utilisateur et en conséquence mettre à jour le parcours recommandé.

5. Bibliographie

- Adomavicius G., Tuzhilin A., « Toward the Next Generation of Recommender Systems : A Survey of the State-of-the-Art and Possible Extensions », *IEEE Trans. on Knowl. and Data Eng.*, p. 734-749, 2005.
- Balabanović M., Shoham Y., « Fab : Content-based, Collaborative Recommendation », *Commun. ACM*, vol. 40, n° 3, p. 66-72, 1997.
- Doerr M., Ore C.-E., Stead S., « The CIDOC conceptual reference model : a new standard for knowledge sharing », *Conference on Conceptual modeling-Volume 83*, Australian Computer Society, Inc., p. 51-56, 2007.
- Gicquel P., Lenne D., Moulin C., « Using Semantic Proximities to Control Contextualized Activities during Museum Visits », *Artificial Intelligence in Education*, 2013.
- Gob A., Drouguet N., *La muséologie : histoire, développements, enjeux actuels*, 2006.
- Herlocker J. L., Konstan J. A., Borchers A., Riedl J., « An algorithmic framework for performing collaborative filtering », *the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, ACM, p. 230-237, 1999.
- Kuflik T., Stock O., Zancanaro M., Gorfinkel A., Jbara S., Kats S., Sheidin J., Kashtan N., « A Visitor's Guide in an Active Museum : Presentations, Communications, and Reflection », *J. Comput. Cult. Herit.*, p. 11 :1-11 :25, 2011.
- Oppermann R., Specht M., « A nomadic information system for adaptive exhibition guidance », *Archives and museum informatics*, vol. 13, n° 2, p. 127-138, 1999.
- Pazzani M., Billsus D., « Content-Based Recommendation Systems », *The Adaptive Web*, vol. 4321, p. 325-341, 2007.
- Resnick P., Iacovou N., Suchak M., Bergstrom P., Riedl J., « GroupLens : an open architecture for collaborative filtering of netnews », *The 1994 ACM conference on Computer supported cooperative work*, p. 175-186, 1994.
- Wang Y., Stash N., Aroyo L., Hollink L., Schreiber G., « Using semantic relations for content-based recommender systems in cultural heritage », *The Workshop on Ontology Patterns (WOP) at ISWC*, p. 16-28, 2009.
- Zimmermann A., Lorenz A., Oppermann R., « An operational definition of context », *Modeling and using context*, Springer, p. 558-571, 2007.