

Projet TCHEK-OFF

Traitement automatique de la langue, Culture, Heuristiques, Exploration, Kaléidoscope, Optimisation, Facilités, Festivals

Metteur en scène

Didier Josselin^{2,(1)}

Acteurs à l'affiche

Juan-Manuel Torres-Moreno^{1,4}, Frédéric Gimello³, Fen Zhou¹

Avec la généreuse participation

des doctorants :

Mayeul Mathias¹, Amine Ait Ouahmed¹, Eloi Flesch³

d'étudiants de Master 1 de Géographie : Dephine Picolot² et Alexis Marcon²

Compagnies

¹ Laboratoire Informatique d'Avignon (LIA/UAPV)

² UMR ESPACE, 7300 – CNRS/UAPV

³ CNE, équipe Culture et Communication, UMR 8562 – CNRS/UAPV

⁴ École Polytechnique de Montréal (Canada)

Évènement : festival *LeOff* d'Avignon

Marion Foliasson, attachée de direction : marion@avignonleoff.com

Logistique : TCRA Avignon

Stéphane Deplace, directeur marketing, stephane.deplace@transdev.com

Projet Agorantic 2015

Éléments de présentation, nom du porteur, laboratoire(s) associé(s)

Le projet Tchek-off est porté par Didier Josselin de l'UMR ESPACE. Il associe l'Université d'Avignon, dont 3 laboratoires (cf casting ci-dessus liés par une nouvelle convention-cadre d'échange et de collaboration scientifique avec le OFF), en géographie, informatique et sciences de l'information et de la communication, avec deux partenaires opérationnels :

- LeOff pour le festival Off d'Avignon

- l'entreprise TCRA pour les transports publics et le vélo-partage (Vélopop) d'Avignon.

Descriptif du projet

L'équipe interdisciplinaire du projet TCHEK-OFF l'Université d'Avignon est composée d'informaticiens du traitement automatique de la langue (Juan-Manuel Torres, Mayeul Mathias), de l'optimisation des réseaux (Fen Zhou, Amine Ait Ouahmed), de sociologues des publics de la culture (Frédéric Gimello et Eloi Flesch) et de géographes (Didier Josselin, Delphine Picolot et Alexis Marcon). Cette équipe, ayant déjà travaillé sur la personnalisation des visites de musées et également sur les festivals de films, propose de commencer une collaboration autour des outils de préconisation de parcours dans le festival d'Avignon, avec l'association Avignon-le-off, en charge de l'organisation et de la communication du programme du festival Off d'Avignon. Ce projet se situe à l'interface entre les questions de transport et de « routage » dans les réseaux (d'où le partenariat avec TCRA), d'extraction de l'information textuelle des événements et des sites culturels, de connaissance des publics et des réseaux sociaux des festivals et de l'accès (social et physique) à la culture. Le terrain d'analyse spécifique est le festival de spectacles vivants Off d'Avignon.

Un « parcours » est une séquence riche et variée de sites ou d'évènements culturels, entrecoupés d'autres activités (repos, restauration, rendez-vous professionnels ou d'agrément, etc.) et insérés dans une chaîne de déplacement pouvant faire appel à plusieurs modes de transport successifs, dont le bus, les transports à la demande et le vélopartage (par exemple). Les lieux de résidence ou d'hébergement (hôtels, locations) sont également pris en compte dans l'établissement des Origines-Destinations du parcours.

Le parcours s'approche des définitions des pérégrinations et de la Time-geography. Les parcours stipulent une forme de rationalité des visiteurs visant à maximiser leur temps festivalier, même si les critères à prendre en compte peuvent être issus de minimisation de distances, temps ou de coûts combinés à des critères plus qualitatifs qu'il faut intégrer dans le calcul de l'optimum (avec une notion intéressante de chemins alternatifs).

L'objectif de ce projet est :

- de fournir des propositions de parcours culturels statiques et dynamiques, incluant les profils et les goûts des festivaliers par enquête ;
- tout en connectant ces parcours à des modes de transport permettant d'accéder au festival dans les meilleures conditions; on inclut dans ces modes les bus, mais également les transports à la demande, le vélopartage;
- d'analyser en retour, via les échanges dans l'internet et par les réseaux sociaux (notamment de type twitter, facebook, blogs, etc), et via les sites dédiés à la culture, la perception et l'opinion des spectateurs.

De même, la dimension culturelle, motivation première des festivaliers, est considérée à plusieurs niveaux, de l'œuvre dans la salle de spectacle ou dans le musée, jusqu'au monument historique à l'échelle de la ville. C'est la combinaison de ces facilités qui constitue in fine l'attrait du parcours proposé. On pourra ici manipuler des graphes multi-niveaux.

L'approche prônée par l'équipe doit rester sensible à trois aspects fondamentaux qui font l'essence du festival d'Avignon Off :

l'équité, à savoir le fait qu'aucun spectacle n'est a priori mis en valeur au détriment d'autres, l'effet de découverte de « pépite » devant rester intact ;

le libre choix du festivalier, qu'il soit touriste ou professionnel, en recherche de spectacles à programmer (on ne doit pas lui imposer des parcours, mais plutôt lui suggérer en fonction de ses attentes) ;

l'importance du bouche à oreilles, qui reste un des moteurs importants de la découverte culturelle, impliquant par exemple des propositions de parcours souples et/ou composées d'alternatives (ce processus est-il modélisable via les réseaux sociaux ?).

Deux cibles et pistes de travail sont pressenties :

La première piste concerne les publics et spectacles du festival 2015 ; il s'agit ici de tester et proposer des parcours aux (groupes de) festivaliers et d'évaluer leur faisabilité/intérêt, avec le recul sur la mouture précédente ;

La seconde piste touche aux acheteurs / programmeurs du festival 2016, qui viennent faire leur « shopping » culturel ; l'objectif consiste alors à leur proposer une série de spectacles en fonction de leurs préférences (genre, langues, pays, type de public, tarif) et contraintes (horaire, budget, capacité de salle, etc.), compte tenu du temps relativement court dont ils disposent et du nombre important de spectacles auxquels ils doivent assister pour faire leur choix.

Dans les deux cas, nous proposons de tester des parcours à plusieurs niveaux, normatifs vs ouverts :

Un premier niveau généraliste et normative qui fixera une échelle de valeur dans les spectacles catégorisés ; cette approche est utile et nécessaire pour un objectif de recherche, mais est relativement en décalage par rapport à la philosophie du festival décrite ci-dessus ;

Une approche plus souple de recommandation, considérant les avis et les choix de festivaliers de profil similaire (e.g. festivalière de plus de 50 ans) dans les réseaux sociaux ; ici on tient compte du fait qu'il peut exister des profils cohérents de festivaliers (personnes ou équivalents) qui pratiquent la culture de façon assez similaire et ont des goûts communs ;

Une approche totalement guidée par les souhaits exprimés des festivaliers, qui s'affranchit de tout type ; la difficulté ici est que le festivalier ne voudra/saura pas forcément exprimer ses souhaits complètement : il peut s'agir de texte libre ou de mots clés, par exemple.

Dans le paragraphe qui suit, nous proposons une approche en phase avec la philosophie du festival. Ces approches peuvent être mises en œuvre de deux façons :

statique : on prédéfinit des préconisations en fonction des contraintes et des objectifs, préconisations qui peuvent être remodelées en fonction de l'évolution des publics d'une année à l'autre par exemple ;

dynamique : on fait évoluer par simulation ou observation ces parcours en fonction du parcours réel pratiqué par le festivalier, éventuellement selon des informations que celui/celle-ci nous fournit (bouche à oreille, échange dans des réseaux sociaux, validation de check-point sur le réseau via le smartphone, choix de flashs codes, par exemple).

Résultats

D'une part, une convention cadre est actuellement en cours de signature entre l'UAPV et LeOff pour ce projet. Contre toute attente, elle est très difficile à obtenir, compte tenu des nombreux changements à la direction de l'association, qui ont bloqué le processus successivement depuis début 2015. La rencontre avec le tout nouveau président en décembre 2016 devrait débloquer définitivement la situation (*inchallah*:).

Pour pouvoir travailler concrètement, l'équipe aurait dû disposer des bases de données numériques du festival Off 2013 -2014 -2015, mises à disposition dans le cadre de la convention de recherche encadrant les termes de la mise à disposition de données de fréquentation :

- carte des théâtres et capacité des salles
- programmation complète par théâtre (horaires, tarif, accès handicapé, entre autres)
- typologies des spectacles (littéralement la version numérique du livre du Off)
- description complète des œuvres (résumés et documents textuelles disponibles, le cas échéant)

Mais face au blocage de la signature de convention, nous avons dû « pomper » les données en ligne et la « parser » avec des méthodes usitées basées sur des expressions régulières afin de reconstruire la base de données.

Dans le cadre de la seconde année de thèse de Mayeul Mathias, cela a constitué un premier (long et fastidieux) travail dans les domaines suivants :

- fouille et extraction de données textuelles et géographiques en ligne
- parsing d'informations issues des notices résumées des spectacles
- reconstruction des catégories de spectacles en fonction de la localisation des théâtres
- structuration d'une base de données dédiée (sur la base d'un diagramme UML adapté à la problématique).

Les premiers traitements d'analyse des textes des notices de spectacles montrent une très grande disparité des contenus et une certaine difficulté à identifier des cooccurrences de termes, rendant très délicate la catégorisation automatique. L'usage des données sources telles qu'elles sont stockées au sein de la BD du Off semble ainsi nécessaire. Le traitement a été effectué sur deux années de festival : 2014 et 2015.

Les algorithmes de calcul de « chemins culturels optimaux » ont été dessinés (ils sont en cours de tests et d'implémentation) à partir des informations extraites de la BD parsée et d'une fonction d'évaluation de score construite à cet effet et comparant la proposition aux profils des utilisateurs (cf ci-dessous). Ce travail constituera la troisième partie de la thèse de Mayeul Mathias.

Une autre partie importante du travail a été réalisée dans le cadre du mémoire de master 1 et du stage de géographie de Delphine Picolot. Tout d'abord, il a été constitué une base de données géographique géocodée très précise qui a corrigé les coordonnées des plans récupérés en ligne, parfois entachés d'erreurs au niveau des localisations des théâtres par exemple. Nous disposons aujourd'hui d'un SIG à jour et complet des lieux clés du festival d'Avignon, ainsi que du réseau d'Avignon de TCRA. L'ensemble des informations est en effet librement utilisable auprès de la TCRA (décision actée lors d'une réunion des partenaires en octobre 2016).

Les travaux de Delphine Picolot ont également apporté des éléments très intéressants sur plusieurs points complémentaires. Une typologie des personas (profils types des festivaliers) y compris les professionnels a été établie suite à un questionnaire réalisé pendant le festival (plus de 250 personnes interrogées). Les questions portaient sur la façon dont les festivaliers ont consommé le festival (provenance, fréquence, préférences, organisation du loisir, budget temps, temps de repos, réseaux sociaux, etc.). À partir de ces informations, il a été défini une série de séquences spatio-temporelles (exemple : *point de départ, période de mobilité, spectacle 1, mobilité, restauration, mobilité, spectacle 2, mobilité, spectacle 3, retour point transport...*). types en fonction de chaque profil (personne seule, famille, groupe, professionnel...) intégrant les points de départ et d'arrivée sur le réseau. Celles-ci ont été cartographiées en exemple et

constituent des « patterns » qui permettent d'évaluer les propositions faites par l'algorithme de chemin culturel développé. Les premiers résultats sont prometteurs et des publications sont prévues pour les mettre en valeur.

Lors de la dernière réunion avec les partenaires, il a été évoqué la question de l'exploitation de l'outil prototypé, preuve que l'association du festival leOff est clairement intéressée. Un autre fait marquant est qu'une ANR, baptisée Tchekoff, a été déposée avec l'assentiment et la participation de l'ensemble des partenaires de ce projet qui ont fourni leurs lettres d'intention.

Apport au travail interdisciplinaire

La dimension interdisciplinaire apparaît dans les différents volets réalisés du projet.

Un premier volet (traitement du langage) consiste à extraire des textes et des catégories décrivant les spectacles du Off, la substantifique moelle du spectacle en question, afin de le catégoriser, voire de lui accorder un crédit théorique par rapport à son type, selon les pistes explorées. Ce volet est réalisé en lien avec la thèse de Mayeul Mathias, co-encadré entre informatique et géographie.

Un second volet (recherche opérationnelle) concerne l'optimisation des parcours (routage multi-critères) qui combine des informations sur les préférences, les facilités et les réseaux (de transport et de communication). Des objectifs doivent être maximisés et des contraintes respectées, afin de fournir des solutions optimales ou excellentes, selon les impératifs de temps de calcul (par exemple, on dispose d'un temps très court pour un web service de recommandation de parcours, tandis que l'on peut avoir suffisamment de temps pour optimiser des parcours proposés). Ce volet est réalisé en lien avec la thèse d'Amine Ait Ouahmed, co-encadré entre informatique et géographie.

Un troisième volet plus socio-spatial concerne les personas et les parcours culturels préconisés. Il associe les trois partenaires académiques (informatique, géographie, sociologie). Il constitue également le lien le plus fort avec l'opérationnalité du projet (partenariat avec leOff et TCRA).

Intérêt par rapport à l'axe identitaire "Culture, Patrimoines, Sociétés numériques"

Le projet s'insère effectivement à l'interface des axes :

- culture et numérique
 - mobilités et transports intelligents
 - réseaux sociaux, structures, contenus et usages
- et, par ricochet :
- accès au savoir, éthique et méthodologies

Budget prévisionnel et financements envisagés

Le budget total demandé à Agorantic était de 8440 € et a été réduit à 6425 €. Voici la répartition des dépenses effectives :

Déplacements (conférences, dont 3 prévues en 2017) : 3890 €

Stages (2 mois) : 1100 €

Matériel divers (1 ordinateur portable, un écran, une station d'accueil, des ouvrages, un accès pro à Overleaf, etc.) : 1430 €

Reliquat : 5 € ;-)

Publications

- (soutenu), Delphine Picolot, Recherche et création de parcours culturels au sein du festival OFF d'Avignon, mémoire de master 1 de Géographie, UAPV, 75 pages
- (en préparation) Path optimisation methods regarding festival user profiles – a geographical issue in the « Festival d'Avignon »