Un modèle multi points de vue pour représenter les trajectoires de vie

Cahier des charges

BENKAOUR Salwa

FLEURY Pierre

JORDAN Célia

Table des matières

[1 Introduction](#_heading=h.gjdgxs) **4**

[2 Guide de lecture](#_heading=h.30j0zll) **4**

[2.1 Maitrise d’œuvre](#_heading=h.1fob9te) **4**

[2.1.1 Responsable](#_heading=h.3znysh7) 4

[2.1.2 Personnel administratif](#_heading=h.2et92p0) 4

[2.1.3 Personnel technique](#_heading=h.tyjcwt) 4

[2.2 Maîtrise d’ouvrage](#_heading=h.3dy6vkm) **4**

[2.2.1 Responsable](#_heading=h.1t3h5sf) 4

[2.2.2 Personnel administratif](#_heading=h.4d34og8) 4

[2.2.3 Personnel technique](#_heading=h.2s8eyo1) 4

[3 Concept de base](#_heading=h.17dp8vu) **4**

[4 Contexte](#_heading=h.3rdcrjn) **6**

[6 Description de la demande](#_heading=h.lnxbz9) **6**

[6.1 Les objectifs](#_heading=h.35nkun2) 6

[6.2 Produit du projet](#_heading=h.1ksv4uv) 6

[6.3 Les fonctions du projet](#_heading=h.44sinio) 6

[6.4 Critères d’acceptabilité et de réception](#_heading=h.2jxsxqh) 6

[7 Contraintes](#_heading=h.z337ya) **6**

[7.1 Contraintes de coûts](#_heading=h.3j2qqm3) 6

[7.2 Contraintes de délais](#_heading=h.1y810tw) 6

[7.3 Contraintes matérielles](#_heading=h.4i7ojhp) 6

[7.4 Autres contraintes](#_heading=h.2xcytpi) 6

[8 Déroulement du projet](#_heading=h.1ci93xb) **7**

[8.1 Planification](#_heading=h.3whwml4) 7

[8.2 Ressources](#_heading=h.2bn6wsx) 7

[9 Annexes](#_heading=h.qsh70q) **7**

[10 Glossaire](#_heading=h.3as4poj) **7**

[11 Références](#_heading=h.1pxezwc) **7**

[12 Index](#_heading=h.49x2ik5) **7**

# 1 Introduction

Le projet est réalisé par trois étudiants du master MIASHS (Mathématique et Informatique Appliqués aux Sciences Humaines et Sociales) parcours IC (Informatique et Cognition), il a lieu dans le cadre d’études.

Ce document a pour objectif de mettre au clair les attentes de notre projet en termes de réalisation et les moyens que nous mettons en œuvre pour y parvenir. On cherche un modèle de trajectoire de vie.

Des modèles existent déjà mais on vise plus de généricité ici, avec la prise en compte d’un modèle multidimensionnel et de facteurs explicatifs autres que ceux déjà rapportés dans des travaux. (Thériault, par exemple, cherche à analyser la dimension résidentielle des trajectoires de vie.)

Pour réaliser le modèle, un fichier csv nous est fourni, il contient des données, récoltées par l’INED, qui sont à analyser. Le document qui est fourni a été réalisé grâce à un questionnaire, celui-ci traite 3 grandes dimensions, la dimension familiale, résidentielle et professionnelle.

Ces données ont déjà été exploitées par M. David Noel qui a fait une étude sur le cas de la trajectoire résidentielle.

Il est important pour ce projet de comprendre les notions qui font être utilisées, en effet, la notion de trajectoire de vie est complexe. C’est un objet d’études qui sert de support à l’analyse des motivations qui déterminent les choix d’un individu dans son parcours biographique. C’est-à-dire en mettant en perspective dans le temps les informations disponibles sur les individus pour mieux comprendre et expliquer ce qui déterminent leurs choix et, par conséquent, leurs parcours.

Le but de ce projet étant d’étudier les facteurs explicatifs, c’est-à-dire l’ensemble des motivations qui guident, consciemment ou non, le comportement d’un individu.

# 2 Guide de lecture

# 2.1 Maîtrise d’œuvre

## 2.1.1 Responsable

Jérôme Gensel dirige nos travaux.

## 2.1.2 Personnel administratif

Notre contact pour les détails administratifs du stage du côté du Laboratoire Informatique de Grenoble (LIG) est Laurence Schimicci.

## 2.1.3 Personnel technique

L’équipe autour du projet est composée de Marlène Villanova-Oliver, Camille Bernard,Pierre Le Queau et Jérôme Gensel. Un doctorant rejoint cette équipe en novembre.

# 2.2 Maîtrise d’ouvrage

## 2.2.1 Responsable

Damien Pellier est le responsable administratif du groupe pour le projet. Nous le voyons toutes les semaines pour rendre compte de l’avancement du projet.

## 2.2.2 Personnel administratif

Notre contact pour les détails administratifs du stage du côté de l’Université Grenoble Alpes (UGA) est Fatima Belounis.

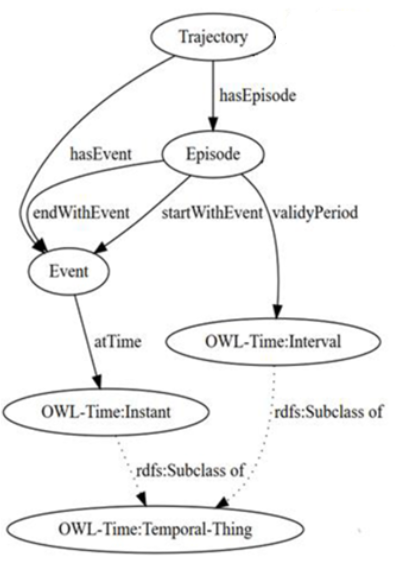
## 2.2.3 Personnel technique

Le groupe est composé de Salwa Benkaour, Pierre Fleury, Célia Jordan. Nous sommes trois étudiants du Master Mathématiques et Informatique Appliqués aux Sciences Humaines et Sociales mention Informatique & Cognition (MIASHS - IC).

# 3 Concept de base

**Trajectoires de vie :** Les trajectoires de vie sont au centre du sujet. On les désigne comme des objets qui permettent d’analyser la vie d’un individu. Grâce à ces dernières, on cherche à comprendre pourquoi l’individu à effectuer un changement à un moment donné de sa vie. Le temps est donc une notion centrale dans ce concept. Hélardot (2006) définit les trajectoires de vie comme « un entrecroisement de multiples lignes biographiques plus ou moins autonomes ou dépendantes les unes des autres ». Les lignes de vie sont alors des domaines de l’existence. Dans ce sujet les domaines auxquels nous nous intéressons sont les domaines familiaux, professionnels, résidentiels et de voyage. Une trajectoire peut être considérée comme composée de plusieurs sous-trajectoires. Toutes ont en commun de posséder des *épisodes* et des *événements*. Le premier représente une période stable et le second un changement entre deux de ces périodes.

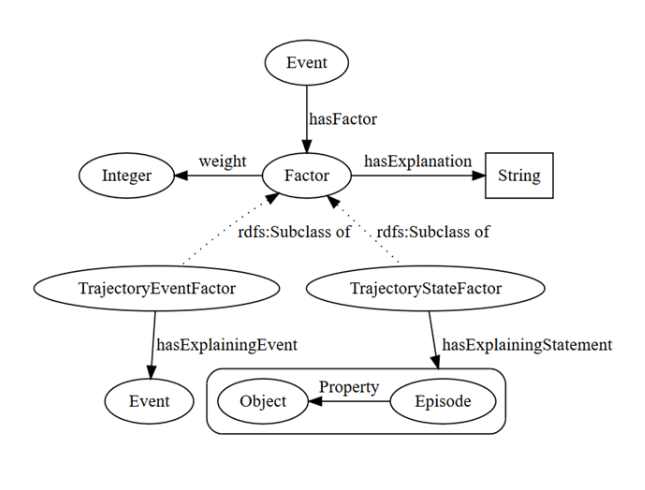
Ce modèle multipoint de vue a été développé par David Noël en 2019 et se nomme LTOP (voir figure 1), *Life Trajectory Ontlogy Pattern*, et sert à modéliser des trajectoires de vies composées de sous-trajectoires.



*Figure 1 Patron de conception LTOP extrait du sujet*

La notion de facteur explicatif cherche simplement à donner des explications cohérentes aux événements ou aux épisodes. Ces explications sont internes si l’explication appartient à la même trajectoire et externe sinon. Toutes les trajectoires ont en commun le temps, ce qui permet de croiser les données et trouver des facteurs explicatifs. Il est indiqué que l’espace peut aussi être commun à plusieurs trajectoires, mais pas à toutes. Les facteurs explicatifs sont donnés par l’individu.

Enfin, le modèle LTOP prend en compte différents niveaux de granularité. Ces derniers déterminent la précision de l’information à disposition. L’explication peut donc varier selon ces niveaux.



*Figure 2 Modèle de facteur explicatif extrait du sujet*

# 4 Contexte

Notre projet se base sur une étude qui a été réalisée en 1981 par l’institut national d’études démographiques (INED). C’est un questionnaire nommé « questionnaire 3B » pour 3 biographies. En effet, il y a la biographie familiale, professionnelle et migratoire. Ce questionnaire à été réalisé sur une population âgée de 45 à 69 ans. Nous n’avons pas toutes les données qui ont été demandées à l'enquêté, toutefois, le fichier qui nous est fourni contient 10 colonnes et 493856 lignes.

Un travail a été effectué par rapport à cette base de données. En effet, Monsieur David Noël a produit une thèse concernant les trajectoires de vie et plus particulièrement le cas de l’étude de la trajectoire résidentielle. Pour cela il a utilisé un patron d'une ontologie de trajectoire de vie nommée LTOP (Life Trajectory Ontology Pattern), ce qui lui permet de représenter les différentes dimensions des lignes de vie.

Un framework basé sur les technologies du web a été créé. Il permet de modéliser l’ontologie d’une trajectoire de vie, c'est à dire qu’il donne une représentation de la trajectoire selon la définition du modèle présenté plus haut. Ensuite, on récolte des données pour documenter l’ontologie pour les stocker dans un *Triple Store*. Enfin on analyse les données pour mettre en évidence certains facteurs explicatifs plutôt que d'autres, s’il y en a. **(Je sais pas si on met plus de détails comme les langages)**

# 6 Description de la demande

## 6.1 Les objectifs

Nous devons créer un modèle UML des lignes de vies, et à partir de celui-ci, produire une base de données relationnelle.

L’obtention de cette base de données peut se faire de deux façons différentes.

**Soit** en utilisant un parseur et ainsi on crée la base de données à partir du fichier csv.

**Soit** en utilisant une autre base de données qui regroupe toutes les informations du fichier csv que l’on interroge afin d’obtenir une base de données en accord avec notre modèle.

On décrit ci-dessous la nature des données qui nous on été fournies dans grâce à l’enquête 3B.

|  |  |
| --- | --- |
| Donnée | Signification |
| v002 | identification de la personne interrogée |
| Annee | année de l'événement |
| DepCom | lieu de l'événement |
| nom\_commune | commune d'habitat au moment de l'événement |
| longitude | longitude de la commune |
| latitude | latitude de la commune |
| Pers | personne concernée par l'événement |
| Type.Lieu | type de lieu de l'événement |
| Rang | nombre de fois que l’événement en question s’est produit |
| TU | - |

A partir de maintenant, on cherche à créer le dictionnaire des données, qui nous permet ensuite de créer le modèle conceptuel. Les cases en vert du tableau ci-dessus représentent les informations que nous avons jugées utiles et que nous retenons. Celles en rouge ne sont pas gardées.

Ainsi, à l’aide du tableau ci-dessus et du modèle UML (cf. ressources) on retient le dictionnaire suivant :

|  |  |  |
| --- | --- | --- |
| Propriété | Type | Description |
| id | entier | id de la personne |
| annee | entier | année de l'événement |
| dep | texte | département de l'événement |
| commune | texte | commune d'habitat au moment de l'événement |
| longitude | double | longitude de la commune |
| latitude | double | latitude de la commune |
| personne | texte | personne concernée par l'événement |
| type\_lieu | texte | type de lieu de l'événement |
| rang | texte | nombre de fois que l’événement en question s’est produit |
| evenement | texte | description de l'événement |
| episode | texte | description de l'épisode |
| localisation | texte | description de la localisation de l'événement |
| rang | texte | nombre de fois que l’événement en question s’est produit |
| mere | texte | Mère de la personne interrogée |
| pere | texte | Père de la personne interrogée |
| enfant | texte | Enfant de la personne interrogée |
| conjoint | texte | Conjoint de la personne interrogée |
| f\_trajectory | abstrait | trajectoire familliale |
| f\_episode | texte | description de l'épisode |
| f\_event | texte | description de la localisation de l'événement |
| p\_trajectory | abstrait | trajectoire profesionnelle |
| p\_episode | texte | description de l'épisode |
| p\_event | texte | description de la localisation de l'événement |
| l\_trajectory | abstrait | trajectoire de loisirs |
| l\_episode | texte | description de l'épisode |
| l\_event | texte | description de la localisation de l'événement |
| r\_trajectory | abstrait | trajectoire résidentielle |
| r\_episode | texte | description de l'épisode |
| r\_event | texte | description de la localisation de l'événement |
| f\_event | texte | description de la localisation de l'événement |

On propose maintenant le modèle conceptuel qui suit.

Dans un premier temps, on veut pourvoir répondre à des questions comme :

- Combien de fois l’individu X a changé d’emploi ?

- Combien d’enfants à l’individu Y ? etc…

La construction de requêtes nous permettra de répondre à ces questions.

## 6.2 Produit du projet

Le produit est une base de données relationnelle. Des requêtes pourront ensuite lui être posées afin de pouvoir étudier certains comportements. Idéalement, le produit doit pouvoir nous donner une interface qui permettra de mieux visualiser les informations qu’il contient. On doit pouvoir exploiter les données obtenues afin de les étudier.

## 6.3 Les fonctions du projet

Le produit doit pouvoir fonctionner peu importe le fichier csv que l’on lui fournit. Il doit pouvoir ajouter de nouveaux éléments mais aussi en supprimer. On doit aussi pouvoir modifier ce que l’on a déjà, si par exemple on a de nouvelles informations.

## 6.4 Critères d’acceptabilité et de réception

Le produit doit correspondre au modèle UML choisi, c’est-à-dire au modèle 3B tout en se basant sur le modèle LTOP. Si celui-ci ne correspond pas à l’un des deux, il n’est pas acceptable.

# 7 Contraintes

## 7.1 Contraintes de coûts

A priori aucun coût pour nous.

## 7.2 Contraintes de délais

Le stage en plein temps a lieu de mi-mai 2022 jusqu’au 30 juin 2022. C’est à la fin de cette période que le produit doit être fourni. On entend par produit la création de la base de données et les requêtes qui permettent d’exploiter les tables.

## 7.3 Contraintes matérielles

Pour faire fonctionner le produit l’utilisateur doit bien sur avoir un ordinateur fonctionnel.

## 7.4 Autres contraintes

Contraintes juridiques sur les données ? Les données 3B sont fournies par l’INED à l’équipe et nous sont prêtés lors du stage. Nous n’avons pas le droit de les utiliser en dehors du cadre du stage.

Contraintes logicielles ? Nous utiliserons Postgresql comme SGBD car il contient l’extension PostGIS qui permet d’avoir une représentation spatiale de ses données.

# 8 Déroulement du projet

## 8.1 Planification

Le projet se découpe en deux grandes parties. La première a lieu d’octobre 2021 à janvier 2022 et consiste en la planification de ce qui sera fait durant la seconde période. Cette dernière consiste au développement du produit.

La première phase débouche sur une soutenance intermédiaire qui a lieu le 17 ou 18 janvier 2022. La seconde phase se termine par une soutenance le 23 juin 2022.

## 8.2 Ressources

- thèse de David Noel

- modèle 3B.jpg

- algo 3B

- diapo présentation

- plan de sondage

- grille recueil lieux 3B

- résumé du projet

- infos sur l’enquête IE0161C

- données 3B

- 3 questionnaires

- dépliant ined insee

- instruction générale sur le déroulement de l’enquête

- quelques remarques sur la présentation du questionnaire

- présentation de l’enquête 3B

- bibliographie de l’enquête

# 9 Annexes

# 10 Glossaire

# 11 Références

# 12 Index