

**Titre long** Extraction de données textuelles et métadonnées depuis des ressources informatives en santé en Suisse Romande pour la constitution d'une base de connaissances et sa représentation (carto-)graphique interactive, pour une santé intégrative pour le citoyen

**Titre court Intégramap** : Un explorateur dynamique et graphique de la Santé Intégrative en Romandie

**Laboratoire d'accueil.** Département SEN (Santé Energie Environnement)

**Suiveur.** Adrien UGON adrien.ugon@esiee.fr (chercheur associé au LIP6, Sorbonne Université)

**Équipe de recherche.** Adrien UGON (LIP6 – Sorbonne Université/CNRS), Jean-Philippe GOLDMAN (HUG/UNIGE), Christian LOVIS (HUG/UNIGE)

**Filières visées.** Biotechnologies et e-santé / DSIA / Informatique

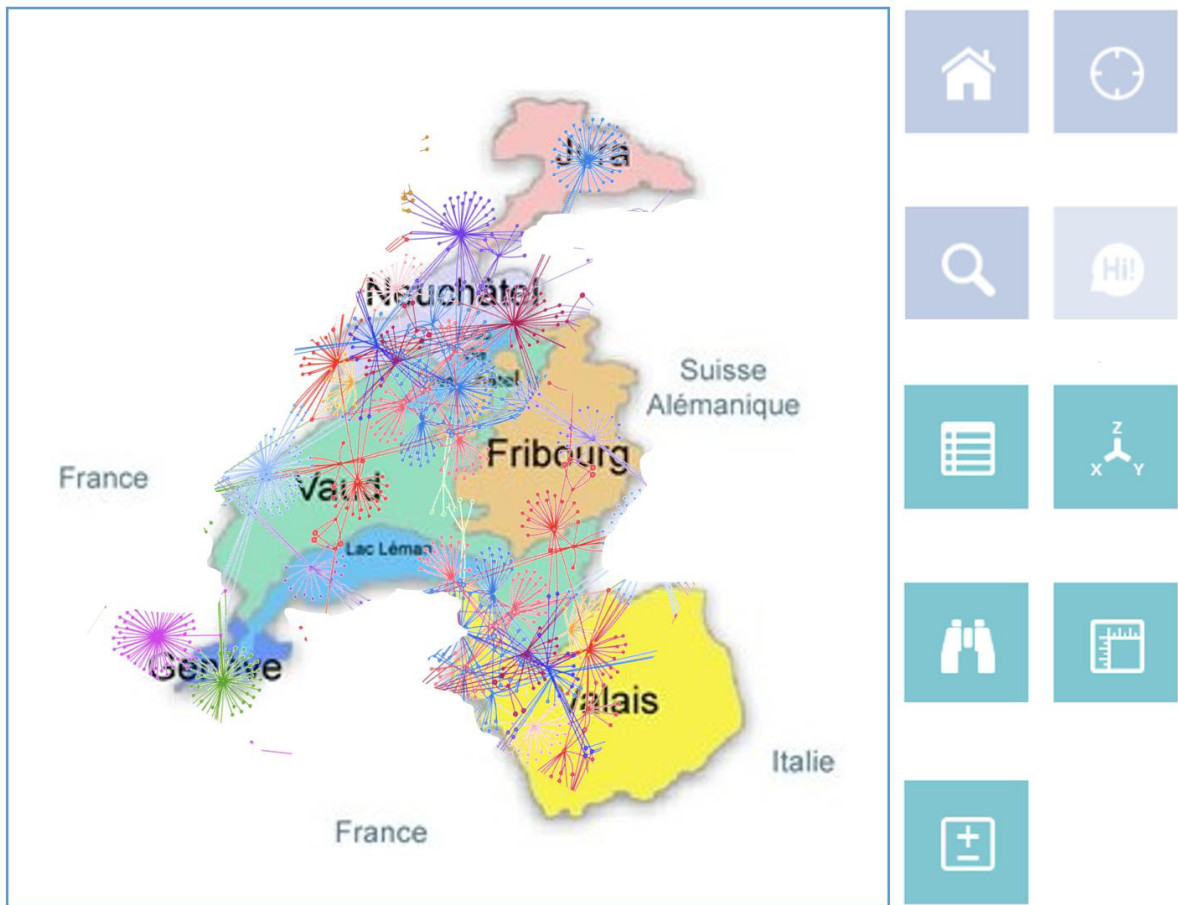
**Equipe d'accueil.** L'étudiant sélectionné sera intégré au département SEN, sous l'encadrement d'Adrien UGON, chercheur associé au LIP6, unité mixte de recherche de Sorbonne Université et CNRS. Le travail se fera en collaboration avec le service des sciences de l'information médicale des Hôpitaux Universitaires de Genève, dirigé par le professeur Christian LOVIS où il sera accueilli de mai 2022 à juillet 2022

### **Présentation générale du sujet**

La suisse romande possède une source d'information pour la santé intégrative hors norme avec des éditions comme Médecine et Hygiène qui investissent pour rendre numériques les contenus de leurs revues. Les silos informationnels et la multitude de revues spécialisées, sont des éléments qui au lieu de servir la société en proposant un éventail d'approches sur la santé plus complet, complexifie pour le citoyen la tâche d'accéder à la globalité des informations utiles pour augmenter son adhésion et sa participation dans les choix thérapeutiques qui sont à sa disposition.

Comment rendre accessibles les informations existantes sur la santé intégrative pour le citoyen ?

Le projet **IntégraMap** consiste à créer un explorateur dynamique et graphique de la Santé Intégrative en Romandie.



## Objectif du projet

L'interopérabilité des données constitue un défi dans les soins de santé. Plus précisément, la mise en relation du texte libre reste un problème non résolu. La nomenclature systématisée des termes cliniques de la médecine (SNOMED CT) est devenue le langage universel des soins de santé et présente les caractéristiques d'un langage naturel. Afin de faire « parler la même langue » les différents domaines de santé intégrative tout en garantissant à chacun la liberté d'exprimer des concepts spécifiques et potentiellement de manière très différente il est possible de traduire de manière invisible pour les utilisateurs ni les fournisseurs de contenus, les concepts principaux en des concepts SNOMED CT. Ce mapping a déjà été fait pour quelques centaines d'articles de la Revue médicale Suisse et permet de mettre en lien des articles dont le vocabulaire utilisé ne permettrait pas autrement de recouper l'information commune.

Une base de données en graphe basée sur les concepts SNOMED CT ainsi que sur les données structurées disponibles permettra de naviguer à travers les silos informationnels des différents domaines de la santé. Une interface graphique « user

friendly » permettra à chaque citoyen de s'informer rapidement sur l'ensemble des contenus disponible dans cette base de données sans se soucier de la source ni du phrasé exact à un domaine spécialisé. La plateforme veut rendre possible aux professionnels de la santé ainsi qu'au reste des citoyens d'accéder à l'ensemble de l'offre de santé intégrative disponible en Suisse romande. Ce projet veut démocratiser l'accès à l'information afin d'améliorer la visibilité de toutes les approches de santé grâce à un explorateur intelligent, dynamique et facile d'utilisation pour in fine améliorer la prise en charge globale de la population romande.

L'objectif principal du projet est de développer un navigateur intelligent, graphique et dynamique de la Romandie qui permette d'explorer l'information disponible en santé de manière générale et de façon inclusive.

Concrètement, l'étudiant Tremplin Recherche participera au projet de recherche et développement **Intégramap** qui a lieu au sein du SIMED, l'équipe Systèmes d'Information Médicale, du Prof. Lovis, aux Hôpitaux Universitaires de Genève d'abord à Paris (nov 21 - avril 22), puis à Genève (mai 22 - juillet 22). Il se familiarisera d'abord avec le langage Python et les ontologies (bases informatiques de connaissances). Il accèdera ensuite au corpus de la Revue Médicale Suisse (consultable en ligne <https://www.revmed.ch/>) et s'intégrera dans l'équipe DataTeam du Simed, afin d'extraire les données textuelles ainsi que les données d'auteurs (auteurs, instituts d'affiliation, spécialités, période temporelle). La ressource textuelle servira au développement d'un encodage automatique en concept SNOMED CT. Les métadonnées serviront à construire un graphe de connaissances sur le milieu professionnel de la santé en Suisse Romande. Dans les deux cas, l'exploitation de fichiers XML et l'utilisation d'ontologies seront au cœur du travail. La période de stage aux HUG correspondra également au développement de prototypes graphiques pour une application web interactive en collaboration avec l'Evalab, la seconde équipe du Simed, qui est un laboratoire de conception centrée utilisateur, interactions homme-machine et ingénierie des facteurs humains.

## Bibliographie

- Robert A, Damachi F, Bjelogrljic M, Goldman J-P, Lovis C. *Evaluation of Document Retrieval Systems on a Medical Corpus in French: Indexation vs. Feature Learning*. Stud Health Technol Inform 2020 Jun 16;270:208–212. PMID:32570376
- Ehrler F, Lovis C, Rochat J, Schneider F, Gervaix A, Galetto-Lacour A, Siebert JN. *[InfoKids: changing the patients' journey paradigm in an Emergency Department]*. Rev Med Suisse 2018 Sep 5;14(617):1538–1542. PMID:30226668

- Gaudet-Blavignac C, Foufi V, Bjelogrljic M, Lovis C. *Use of the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) for Processing Free Text in Health Care: Systematic Scoping Review*. *Journal of Medical Internet Research* 2021 Jan 26;23(1):e24594. [doi: 10.2196/24594]
- Lamy Jean-Baptiste *Ontologies with Python: Programming OWL 2.0 Ontologies with Python and Owlready2* , Released December 2020, Publisher(s): Apress, ISBN: 9781484265529