

Emmanuel ROUSSEAU

Address

Geneva School of Economics and Management
& Institute of Demography and Socioeconomics
University of Geneva
Uni-Mail, Office 3252
CH-1211 Geneva 4
Switzerland

Age 31
Nationality French, C permit
Website <http://emmanuel.rousseau.me>
E-mail emmanuel.rousseau@gmail.com
E-mail (pro) emmanuel.rousseau@unige.ch
Phone (pro) +41 22 379 82 31

Education

- 2017** PhD in Information Systems, University of Geneva.
(*expected*) Co-advised by Gilbert RITSCHARD and Giovanna DI MARZO SERUGENDO.
- 2010** MSc. (research) in Data Mining and Knowledge Management, Nantes-Lyon.
Upper second class Honours, top of the year (1/17).
- 2010** MSc. (applied) in Computer Science, with a specialization in Business Intelligence.
From the Graduate School of Engineering of the University of Nantes, Polytech’Nantes.
Upper second class Honours, top of the year (1/72).
- 2007** MSc. in Fundamental and Applied Mathematics (*Maîtrise*), with distinction.
- 2006** BSc. in Mathematics (*Licence*), with distinction.

Work experience

- 2015-Now** **Head of IT, Geneva School of Economics and Management**
(*2 years, 60%*) Our team of 10 collaborators provide Professors, researchers, lecturers and students of GSEM with an high level IT support. We also continuously design, create and maintain IT systems and services of GSEM.
- 2010-2015** **Teaching and Research Assistant, University of Geneva**
(*5 years*)
Research: Methodological approaches for mining critical events in life course data.
Teaching: please refer to the Teaching Section.
Administrative responsibilities: IT Manager of the Department of Economics, then IT Manager ad interim of GSEM. Representative of the Intermediary Staff at the SES Faculty board (2011-2013), then representative of the Intermediary Staff at the University board (2013-2015).
- 2009-2010** **Internship at Department of Econometrics, University of Geneva**
(*6 months*)
Topic of the intership: “*Decision trees on unbalanced data sets*”. Literature review, solution proposal, implementation, assessment on real data sets, draft of a paper.
- 2008-2009** **Internship at Center for Clinical Research, UHC of Angers**
(*3 months*)
Study and development of a clinical decision support for suspected venous thrombo-embolism to help diagnosis work in emergency departments. Specifications, software design with an integrated client/server data synchronisation and private data encryption system.
- 2007-2009** **Mathematics and Computer Science teacher**
(*2 years*)
Teaching to 2 post-baccalaureat classes at Lycée Clémenceau, Nantes. Private tutor for *Academia* and *Complétude* companies, Nantes.

PhD thesis project

Methodological strategies to support the discovery of underlying factors of vulnerability in life courses

Conducted within the NCCR LIVES, this thesis aims at providing some methodological strategies for the study of the vulnerability across the life course. Ideally, we aim to be able to explain how some individuals experience critical life situations, while some others succeed in avoiding them. Reasons of experiencing critical life situations are most often complex and generally result from the interaction of several factors. This makes the task of modeling the vulnerabilization process difficult. In addition, critical life situations are usually (and hopefully) rare. The consequence is that we have less occasions to observe them. Figuring out what combination of factors may lead to experience a critical life situation is for the mind even more difficult.

In addition, population studies strongly rely on survey data. Studying individual from a life course perspective requires to set up complex databases, as for instance by collecting longitudinal data or network data. The increasing volumes of structured data complicate the task of both documenting and manipulating data. However, in social sciences, the description of data is crucial to allow the social scientist to draw conclusions. Furthermore, as social analysis are conducted on a sample of the population studied, it is mandatory to take the sampling design of the survey into account, and also how respondents answer to the survey. Especially, it is important to take non-response weights into account, and to document the different reasons an individual didn't answer to a specific question.

In this thesis I focus on (1) the issue of handling life course survey data in statistical software, (2) the strategies allowing to highlight variable interactions, (3) how to handle the potential rarity of vulnerable situations, and, on a more general aspect, (4) how to support the researcher in social sciences in her daily work to help her figuring out what underlying factors could explain the social phenomenon studied. The originality of our approach is to formulate the issue of finding underlying factors in life course data as both an information system and an data analysis issue. We aim at introducing innovative strategies in these two respective fields to propose an ad hoc methodological framework to place the practitioner in a favorable mind context to facilitate her task of figuring out what are the factors underlying the social phenomenon studied.

Research projects

- **NCCR LIVES, IP214: Measuring vulnerability** (*position: member*)
Led by Gilbert RITSCHARD is designed to provide the participants in the NCCR LIVES with high methodological competences in data collection and data management and with advanced statistical tools for the analysis of vulnerability and resilience processes in the life course.
- **Rsocialdata: Handling survey data in R** (*position: co-leader*)
Pre-processing survey data before running into analysis is not an easy task. It is even harder when working on longitudinal data or network data. The Rsocialdata project aims at providing researchers in social sciences high-level tools for storing, sharing, and pre-processing survey data in a secure and efficient way. The project introduces, in particular, a variable and value labels management system, a specific framework for handling weights, and a data consistency and representativity to the initial population check system. The software is developed as a series of R packages:
 - **Rsocialdata** for cross-sectional and longitudinal survey data.
 - **Rsocialdata.network** for network survey data.
- **The package trim: TRees for IMbalanced data** (*position: co-leader*)
In many dataset coming from real-life situations, class distributions are imbalanced. This especially occurs when studying group of vulnerable people, as for instance people experiencing long-term unemployment, or people experiencing a critical health problem (cancer, etc.). When we want to extract profiles of people falling in these critical situations, we often perform a classification task. But this imbalance impacts the classification quality. Although we succeed in getting a good overall classification rate, most of the error is done on the rare class. However, our interest is more on the rare class (vulnerable people) than on the majority class (well-being people). In the past decade, several methods has been designed to reduce/overcome the class imbalance problem, especially for decision tree methods. This package aims at providing these decision tree methods especially designed for classification of imbalanced categorical data.
- **The package debate: Measuring content and relational in debates** (*position: co-leader*)
Several dimensions are in play during a debate. This is especially the case for political debates. This

package provides tools for measuring both the content and the relational dynamics of debates. For measuring content we suggest to use a thematic coding based on a dictionary. From this dictionary, the list of the actors of the debate, and the memorials in PDF files, the package can help you to count occurrences of dictionary items in each actor speeches. The relational is measured by count approval, disapproval and neutral references between actors. Using the `spnet` package as a back-end for the visualization, the user can display both the content and the relational aspects of the debate on the assembly map.

- **The package `spnet`: Plotting (social) networks on maps** (*position: co-leader*)

Social networks analysis has received special attention over the past decade, and a lot of tools for manipulating and rendering social networks have emerged. In several situations a social network is associated with a spatial dimension, and behaviors observed within the network cannot be interpreted without taking into account the location of each of its nodes regarding to the other nodes. This is the case, for example, when studying inflows/outflows between cities or companies, or when studying people debating in a room. Based on the `sp` package, which provides efficient classes for storing spatial data and methods for handling and rendering them, the `spnet` package aims at facilitating the rendering of (social) networks on maps. Furthermore, fixing network node positions allows to more easily monitor time-varying networks and observe how connections and flows evolve over time.

Articles

- Deville, Marion and Emmanuel Rousseaux (submitted). “Dynamique entre parlementaires à l’Assemblée constituante de Genève : Méthode de visualisation en réseaux pour un récit interactionniste.” In: *Revue suisse de sociologie*.
- Cullati, Stéphane, Emmanuel Rousseaux, Alexis Gabadinho, Delphine S Courvoisier, and Claudine Burton-Jeangros (2014). “Factors of change and cumulative factors in self-rated health trajectories: a systematic review.” In: *Advances in Life Course Research, Elsevier*. 19, pp. 14–27.
- Bürgin, Reto, Gilbert Ritschard, and Emmanuel Rousseaux (2012a). “Exploration graphique de données séquentielles.” In: *Atelier Fouille visuelle de données : Méthodologie et évaluation. Conférence EGC2012*, pp. 39–50.
- (2012b). “Visualisation de séquences d’évènements.” In: *Extraction et Gestion des Connaissances. Revue des Nouvelles Technologies de l’Information E.23*. Ed. by Yves Lechevallier, Guy Melançon, and Bruno Pinaud, pp. 559–560.

Book chapters

- Guarin, Andres and Emmanuel Rousseaux (accepted, to be published in 2017). “Labor market insertion risk factors for second-generation immigrants in Switzerland.” In: ed. by C. Bolzman, L. Bernardi, and J-M. Le Goff. *Life course studies and social policies*. Springer. Chap. 3.

Software documentations and user manuals

- Rousseaux, Emmanuel and Marion Deville (2016a). *Overview of the `spnet` R package*. Vignette of the R package ‘`spnet`’, version 0.9.1.

Conferences, workshops and invited talks

- Deville, Marion, Pierre-Alexandre Fonta, Roy Gava, Yannick Rochat, and Emmanuel Rousseaux (2017). “Le secret bancaire au Parlement suisse: Une analyse des débats aux chambres fédérales 1995-2015.” In: *XIVème Congrès de l’Association Française de Science Politique, Montpellier, France*.
- Rousseaux, Emmanuel and Marion Deville (2016a). *Overview of the `spnet` R package*. Vignette of the R package ‘`spnet`’, version 0.9.1.
- (2016b). *La suite R debate: Outils méthodologiques pour l’analyse du débat politique*. Research Seminars of the Institute for Demographics and Socioeconomics Studies, Geneva School of Social Sciences,
- Rousseaux, Emmanuel and Gilbert Ritschard (2015a). *The `Rsocialdata.network` extension: handling and analysing egocentric network survey data in R*. Research Seminars in Statistics for Social Sciences, Institute for Demographics and Socioeconomics Studies,
- (2015b). “The `Rsocialdata.network` extension: Handling and analysing egocentric network survey data in R - beta release.” In: *4èmes Rencontres R, Grenoble, France*.

- Rousseaux, Emmanuel and Gilbert Ritschard (2015c). “Decision tree methods for discovering (cross-national) interactions in longitudinal data with imbalanced data.” In: *Society for Longitudinal and Life Course Studies Annual Conference, Dublin, Ireland*.
- Deville, Marion and Emmanuel Rousseaux (2014a). “A Spatial Network Approach for Measuring the Differentiation between Content and Relational Dynamics in the Political Debate.” In: *7th Political Networks Conference, McGill University, Montreal, Canada*.
- Rousseaux, Emmanuel and Marion Deville (2014). *The R spnet package, plotting social networks on maps*. Research Seminars in Statistics for Social Sciences, Institut d’études démographiques et du parcours de vie, Université de Genève,
- Deville, Marion and Emmanuel Rousseaux (2014b). “A Spatial Network Approach for Measuring the Differentiation between Content and Relational Dynamics in the Political Debate.” In: *4th Annual General Conference Of The European Political Science Association, Edinburgh, Scotland*.
- Rousseaux, Emmanuel, Marion Deville, and Gilbert Ritschard (2014). “The SPNET package: Plotting social networks on maps with R.” In: *3èmes Rencontres R, Montpellier, France*.
- Rousseaux, Emmanuel and Gilbert Ritschard (2014a). “An association rule miner for unbalanced data based on artificial bee colony optimization.” In: *21st International Conference on Computational Statistics, Geneva, Switzerland*.
- (2014b). “The Rsocialdata project: handling and sharing longitudinal and network survey data with R.” In: *Society for Longitudinal and Life Course Studies Annual Conference, Geneva, Switzerland*.
- Rousseaux, Emmanuel, Danilo Bolano, and Gilbert Ritschard (2013a). “Dataset: An efficient and secure software framework for handling survey data in R.” In: *Population Days 2013, 10th edition, Brixen, Italy*, p. 44.
- Rousseaux, Emmanuel and Gilbert Ritschard (2013). “The Dataset project: Handling survey data in R.” In: *7th International Conference of Panel Data Users in Switzerland, Lausanne*, pp. 37–38.
- Rousseaux, Emmanuel (2013). *Handling non-response in survey data: on the usage of weights*. Research Seminars in Statistics for Social Sciences, Institute for Demographic and Life Course Studies, University of Geneva,
- Guarin, Andres and Emmanuel Rousseaux (2013). “Access to the labour market among second-generation immigrants in Switzerland.” In: *9e Congrès de la Société Suisse de Sociologie*. Bern.
- Rousseaux, Emmanuel, Danilo Bolano, and Gilbert Ritschard (2013b). “The Rsocialdata package: Handling survey data in R.” In: *XXVII IUSSP International Population Conference, Busan, Republic of Korea*.
- Rousseaux, Emmanuel and Andres Guarin (2013). *Employment attainment of 2nd generation immigrants in Switzerland*. Research Seminars in Statistics for Social Sciences, Institute for Demographic and Life Course Studies, University of Geneva, Switzerland.
- Rousseaux, Emmanuel (2012a). *Handling survey data in R: the Dataset Project, first steps*. Research Seminars in Statistics for Social Sciences, Institute for Demographic and Life Course Studies, University of Geneva,
- Rousseaux, Emmanuel and Gilbert Ritschard (2012). “The Dataset project: Handling survey data in R.” In: *LaCoSA 2012’s Conference, Lausanne*.
- Rousseaux, Emmanuel (2012b). *First steps towards a software framework for handling life course survey data in R*. Rencontres Recherche et Méthodes, University of Lausanne, Institute for Demographic and Life Course Studies, University of Geneva,
- (2012c). *Advances in the Dataset R package*. Research Seminars in Statistics for Social Sciences, Institute for Demographic and Life Course Studies, University of Geneva,
- (2011). *Arbres de décision pour la découverte de profils vulnérables*. Research Seminars in Statistics for Social Sciences, Institut d’études démographiques et du parcours de vie, Université de Genève,

Research seminars

- Rousseaux, Emmanuel and Marion Deville (2016b). *La suite R debate: Outils méthodologiques pour l’analyse du débat politique*. Research Seminars of the Institute for Demographics and Socioeconomics Studies, Geneva School of Social Sciences,
- Rousseaux, Emmanuel and Gilbert Ritschard (2015a). *The Rsocialdata.network extension: handling and analysing egocentric network survey data in R*. Research Seminars in Statistics for Social Sciences, Institute for Demographics and Socioeconomics Studies,
- Rousseaux, Emmanuel and Marion Deville (2014). *The R spnet package, plotting social networks on maps*. Research Seminars in Statistics for Social Sciences, Institut d’études démographiques et du parcours de vie, Université de Genève,
- Rousseaux, Emmanuel (2013). *Handling non-response in survey data: on the usage of weights*. Research Seminars in Statistics for Social Sciences, Institute for Demographic and Life Course Studies, University of Geneva,

- Rousseaux, Emmanuel and Andres Guarin (2013). *Employment attainment of 2nd generation immigrants in Switzerland*. Research Seminars in Statistics for Social Sciences, Institute for Demographic and Life Course Studies, University of Geneva, Switzerland.
- Rousseaux, Emmanuel (2012a). *Handling survey data in R: the Dataset Project, first steps*. Research Seminars in Statistics for Social Sciences, Institute for Demographic and Life Course Studies, University of Geneva,
- (2012b). *First steps towards a software framework for handling life course survey data in R*. Rencontres Recherche et Méthodes, University of Lausanne, Institute for Demographic and Life Course Studies, University of Geneva,
- (2012c). *Advances in the Dataset R package*. Research Seminars in Statistics for Social Sciences, Institute for Demographic and Life Course Studies, University of Geneva,
- (2011). *Arbres de décision pour la découverte de profils vulnérables*. Research Seminars in Statistics for Social Sciences, Institut d'études démographiques et du parcours de vie, Université de Genève,

Affiliations

- **NCCR LIVES, IP 214**
NCCR LIVES: "Overcoming Vulnerability: Life Course Perspectives"
IP 214: "Measuring Vulnerability", led by Gilbert Ritschard
 University of Geneva, Switzerland
- **Institute for Demographic and Life Course Studies**, University of Geneva, Switzerland
- **Institute of Information Systems and Service Science**, University of Geneva, Switzerland
- **IEEE**, Institute of Electrical and Electronics Engineers
 IEEE Computer Society, IEEE Computer Society Technical Committee on Pattern Analysis and Machine Intelligence, IEEE Computational Intelligence Society, IEEE Systems, Man, and Cybernetics Society.

Grands and Awards

- 2014** **Travel fellowship** for presenting to the 2014 Political Networks Conference, by the Fellowship Committee for the 2014 Political Networks Conference and the National Science Foundation, 1'000 USD.
- 2013** **Travel fellowship** for presenting to the XXVII International Population Conference, by the Société Académique de Genève, 1'500 CHF.
- 2010** **Study fellowship** for a Master internship, by the Région des Pays de la Loire, France, 1'000 EUR.

Teaching

Course title	Main topics	Level	Lang.	Years
<i>Mathématique II</i>	Nombres complexes, dérivée des fonctions de plusieurs variables, théorème des fonctions implicites, approximation linéaire, extremums libres, extremums sous contraintes, théorème de l'enveloppe, valeurs propres d'une matrice, formes quadratiques, convexité et concavité, qualification des extremums.	Ba	FR	2010
<i>Analyse statistique de données catégorielles</i>	Mise en place d'une problématique en socio-économique, introduction à R, gestion des données d'enquête sous R, recodages, analyses univariées, analyses bivariées (symétriques, directionnelles, ordinales), arbres de décisions, régression logistique, rédaction d'un article scientifique.	Ma	FR	2010-2014

<i>Statistiques pour sciences sociales</i>	Statistiques descriptives, probabilités, probabilités conditionnelles, variables aléatoires discrètes, variables aléatoires continues, loi normale, estimation ponctuelle, intervalles de confiance, théorie des tests d'hypothèses, test de moyenne avec variance connue, test d'indépendance du Chi-carré.	Ba	FR	2011
<i>Outils informatique pour l'économie</i>	Calcul formel sous Maple (polynômes, dérivation, intégration, variables aléatoires, limites de suites). Introduction à L ^A T _E X. Calcul numérique sous Matlab (calcul matriciel, boucles, conditions, graphiques, variables aléatoires, quadrature, méthode de Newton, algorithme du point fixe, limites de suites).	Ba	FR	2011
<i>Statistical Inference</i>	Random variables, moments, sampling methods, Central-limit theorem, method of moments estimators, maximum likelihood estimators, least square estimators, bias, efficiency, convergence, Cramer-Rao bound, linear regression, hypothesis testing, Newman-Pearson lemma, mean test (variance known/unknown), variance test, ANOVA, non-parametric statistics (Mann Whitney, Wilcoxon, Kruskal Wallis).	Ma	EN	2012-2014

Skills

Datamining: Decision Trees, Association Rules, Induction Graphs, Neuronal Networks and Machine Learning techniques, Knowledge Validation, Knowledge Visualisation. *Tools : R, SAS Enterprise Miner, Weka, Tanagra, KXEN.*

Knowledge-based systems: Bayesians Networks, basics Artificial Intelligence methods, Data Quality and Knowledge Validation, Information Visualisation, Semantic Web and Ontologies.

Mathematics for the decision: Optimisation and Meta-Heuristics, Multicriteria Decision-Aide, Game Theory.

Languages: PL-SQL, R, C++/C#, JavaSE, JavaME, Maple/Matlab, PHP, Lisp, OCaml, Prolog, Ruby, L^AT_EX.

Development and office tools: Visual Studio, Eclipse, MS-Office (Word, Excel, ...), MS-Project, VMware.

Corporate skills: Team management, negotiation, quality assurance, innovation, accounting, and good skills in Project Management.

Languages:

- French native speaker.
- German (A2)
- English (C1)
- Spanish (A2)

Social skills: Dynamic, enthusiastic, organised, I like working in a team. I enjoy challenges and thrive in problem solving situations. Hard worker.

Driving license (A,B).