We are opening **two doctoral fellowships (3 years) in computer science and/or computational social science** under the umbrella of the ANR-funded project RECORDS (2020-23) focused on the understanding of practices surrounding music streaming platforms.

One fellowship will be based at **Géographie-cités** in **Paris**, under the supervision of Thomas Louail, and will focus on the **spatial dynamics underlying content consumption** on streaming platforms, with music streaming as a primary case study. The other fellowship will be based at **Centre Marc Bloch** in **Berlin**, under the supervision of Camille Roth in his **Computational Social Science team**, and will focus on the large-scale and longitudinal **study of algorithmic guidance** in the context of music streaming platforms.

Specific details on each position may be found on p. 2 for Géographie-cités and on p. 3 for Centre Marc Bloch.

Context. RECORDS is a collaborative research project funded through an ANR grant (2020-23). It involves several CNRS units, along with a partnership with Orange and with one of the major music on-demand platforms, Deezer. The goal is to improve our understanding of (i) the diversity of users practices and consumptions on streaming platforms (ii) the effects of manual and algorithmic content recommendation (iii) the space-time diffusion of music. This interdisciplinary research articulates quantitative and qualitative methods. It relies on an original data collection protocol that combines anonymized individual listening history data on the platform; detailed self-declared information collected through a large-scale web survey; in-depth interviews with volunteer participants.

Project members come from different academic backgrounds, some from the social sciences (mainly sociology and geography), some from computer science and signal processing. The project is led by Thomas Louail (CNRS, Géographie Cités), Philippe Coulangeon (CNRS, Observatoire Sociologique du Changement), Camille Roth (CNRS, Centre Marc Bloch), Jean-Samuel Beuscart (Orange Labs SENSE) and Manuel Moussallam (Deezer R&D). More information (in french) is available on the project website, records.huma-num.fr

Conditions. Both fellowships would start in September 2020, for a duration of three years. Conditions are aligned with French doctoral contracts at CNRS, i.e. roughly 1700 euros net/month.

Requirements. We expect a strong profile either in quantitative social science, or in computer science, applied mathematics or statistical physics. The work will more specifically involve modeling and data analysis skills (including spatial analysis/GIS at Géographie-Cités). A strong interest for interdisciplinary collaborations is mandatory.

Application process. Candidates should send a cover letter and a resume by **May 31** to **records@parisgeo.cnrs.fr** for Fellowship A, or to **records@cmb.hu-berlin.de** for Fellowship B. The letter must contain a brief research proposal in the framework of RECORDS that demonstrates some prior understanding of the issues pertaining to the study of spatial dynamics (especially for Fellowship A) and/or online behavior (especially for Fellowship B).

To check and discuss the adequacy between their profile and these topics, candidates should feel free to contact beforehand, respectively, for Fellowship A, Thomas Louail ⟨thomas.louail@cnrs.fr⟩, CC: records@parisgeo.cnrs.fr and for Fellowship B, Camille Roth ⟨roth@cmb.hu-berlin.de⟩, CC: records@cmb.hu-berlin.de

Fellowship A: spatial dynamics at Géographie-Cités in Paris

General objectives

The fellowship focuses on the spatial dynamics underlying content consumption on streaming platforms, with music listening as a primary case study. The largest music-on-demand streaming services are accessible in almost every country, through a standard interface, interoperable technologies, and almost identical subscription plans. A large part of platforms catalogs is common to all countries. As such they constitute a great benchmark to investigate the spatial dimension of online cultural consumption. While online information consumption has been heavily studied, its spatial dimensions are much less understood quantitatively. Individuals stream music on a daily basis, and like any other ordinary activity it is embedded in space through social contexts and habits. Applicants' research proposals may focus on different topics related to the project. Possible directions include (but are not restricted to):

- *the role of context in individual consumption.* The listening context (where you are, what you are doing, the people you are with) is assumed to determine the choice of content you listen to. However there are few results that corroborate this hypothesis thanks to detailed, georeferenced individual activity data. Additionally, recommender systems hardly integrate this information so far. Log files include low-level information (timestamp, ip, device, platform features) that are useful to infer higher-level spatial contexts thanks to statistical learning (e.g. 'at home', 'at work', 'on the move', 'partying'). Such higher-level information may be used to compare intra-individual and inter-individual variability, and suggest models useful to relate context, musical content and listeners satisfaction.

- *the geographical diversity of tastes and practices on global platforms.* If we spatially aggregate users streaming data and calculate the dominant music genres and artists in a given area, clusters will appear on the map. Along with the long established connections between social status/origins and taste preferences —the geographical space being socially organized, we can expect to uncover spatial consumption patterns—, other factors may explain the observed differences between areas. They include local music scenes or the presence of cultural infrastructures. What are the geographical scales at which location explains music consumption? How strong remain local genres and traditions, including country-specific, on global platforms?

- *the effects of external shocks/perturbations.* How do users change their habits during extraordinary events? How useful are the activity data collected by platforms to retrieve higher-level information in such situations, e.g. human mobility?

- *the spatial diffusion of songs, artists and genres.* Many models have been proposed to capture the spatial diffusion of people, information, innovations, epidemics, etc. at different scales. How useful are these models to understand how songs and artists popularity grow through internet platforms? For example, a common strategy/constraint among artists to gain popularity is to move to a larger city ('if i can make it there i'll make it anywhere'). We could, for a limited number of cases, combine artists biographical data with georeferenced listening data spanning several years. How musicians trajectories relate to the number and location of their listeners on the internet? What are the respective contributions of social networks and space in shaping one's digital trajectory?

Scientific environment

The PhD student will join Géographie-cités, a research unit located on Campus Condorcet, a social science research centre that opened in September 2019. It is located in Aubervilliers, in the inner northern suburbs of Paris, 20 minutes away from the city center by public transportation. The campus hosts eleven research institutions and universities, more than sixty research units and thousands of people. The PhD student could also work part-time in Deezer R&D offices in Paris, 20 minutes away from the campus. She/he will be provided a laptop and secure access to all the sensitive data of the project that were collected either by the company or through the RECORDS websurvey. She/he will benefit from the project dynamics, which include possible collaborations with all the project participants (see the list on the project website). The thesis will be co-advised by Thomas Louail (PhD in computer science) and by another co-advisor, chosen according to the candidate's research interests.

Fellowship B: algorithmic guidance at Centre Marc Bloch in Berlin

General objectives

The fellowship at Centre Marc Bloch focuses on the large-scale and longitudinal study of algorithmic guidance in the context of music streaming platforms. Deezer, Spotify and other global music-on-demand platforms are accessible in hundreds of countries. They host music libraries that are so large that no human could fully explore them in an entire life (40M tracks represent several hundred years of recorded sound). However, only a very small portion of these catalogs are explored by users, in a rather heterogeneous fashion (some songs and artists are very popular, while others — the large majority — are being listened to a much more smaller number of times). In this context, the role of algorithmic recommendation remains to be appraised, especially in terms of the consumption of popular vs. niche content, of well-known vs. surprising content.

Broadly, the effects of recommendation algorithms on the access to information and cultural goods is at the center of a growing debate which aims at assessing whether they rather contribute to enlarge or to restrain the horizon of users with respect to their "organic" behavior, i.e. absent algorithms. A growing literature aims at empirically comparing what happens when users do rely, at least in part, on the output of some recommendation algorithm *vs.* when they do not. Contrarily, perhaps, to intuitions related to the popularization of so-called "filter bubbles", several recent studies on a variety of online platforms appear to show that algorithmic recommendation does not necessarily contribute to restrict the horizon of users. Nevertheless, user reactions to algorithmic curation may not be deemed to be homogeneous: users may variously seek diversity, be variously responsive to recommendation, use it for various purposes or have various expectations about it — in these respects, the "average user" does not really exist.

This kind of scientific endeavor need not venture into knowing or reverse-engineering which principles drive these algorithms, even though it may benefit from such insider information. In general, such studies nonetheless require sophisticated experimental protocols or privileged access to private company data. Within RECORDS, we will precisely have a prime access to usage data on a major music streaming platform, leaving much freedom to explore the complex interplay between user preferences and algorithmic guidance.

The applicants research proposals may focus on different topics related to the project. Possible starting questions include (but are not restricted to):

- are there various types of attitudes toward algorithmic recommendation ?

- what distinguishes organic vs. algorithmically mediated behaviors ?

- how is organic serendipity distributed among users and content types ? (organic: natural or absent algorithms)

- what is the contribution of algorithmic guidance in the formation of taste over a certain period of time ?

Scientific environment

The PhD student will benefit from the dynamics surrounding this project and the various perspectives developed in the consortium to shed light on music listening practices. These include frequent meetings with the participants and possible collaborations with researchers belonging to the different research institutions participating to the project (CNRS, Sciences Po), including the R&D departments of both companies (Deezer and Orange).

The fellowship will be based at Centre Marc Bloch in Berlin, an international and interdisciplinary CNRS research center — technically a dual unit of CNRS and the French Foreign Affairs Ministry, further affiliated with the German Federal Ministry of Research (BMBF), as well as an An-Institut of the Humboldt Universität. Centre Marc Bloch is located in a very central and accessible location in Berlin, right in front of the metro station Stadtmitte.

The selected candidate will join the larger Computational Social Science team (*cmb.huma-num.fr*) at Centre Marc Bloch under the supervision of Camille Roth, which hosts an interdisciplinary group of doctoral and post-doctoral researchers generally coming from sociology or computer science, as well as an ERC grant called Socsemics.

A quick presentation of the team may be found at: youtube.com/watch?v=idmreTQzeg0