REFERENCES
Main activities
In the framework of the project described above, the postdoc can work on a combination of the three following aspects.

I. Design of online learning algorithms.
We plan to evaluate three different frameworks for learning online how to populate the local datastore.
1. Adapt existing caching policies like LRU, e.g., by inserting the content on the basis of its marginal utility (i.e., its contribution to inference quality). Ad-hoc policies in this spirit have been proposed to support image similarity search in [5] and in [8]. This framework leads usually to a combinatorial analysis with a focus on expected performance under a stochastic request process.
2. Study the problem as a discrete-space metrical task system (MTS) [2], where the state of the system is the set of instances in the datastore. Each state has a corresponding service cost (the loss of inference quality due to running a simpler model at the edge) and updating the datastore generates so-called movement costs. Competitive analysis is the common approach to study this setting.
3. When the set of possible instances is very large and roughly homogeneously distributed, at least over a low-dimension manifold, it is possible to consider the state space to be continuous. This setting is closer to online machine learning with regret as its main performance metric.

At the methodological level, we will explore gradient-based approaches. They are common in online machine learning, but, more recently, they have also been effectively employed to study combinatorial problems in the other two settings [1, 3, 14].

II. Characterization of datasets' topological properties.
Which framework, among the three described above, is the most appropriate? The answer depends to a large extent on the topological properties of the space where instances lie. Whereas we are looking for collaborations with other research teams studying the topological and geometric structure of data, we will push a practical approach, starting from real traces. Many traces are available for recommender systems based on ML predictors. This application is particularly interesting for MAMMALS, as recommendations need to be customized to the user (a particular example of domain adaptation) and constantly updated to follow dynamic popularities of media contents or products.

III. Prototype implementation.
We plan to provide practical evidence of the potential improvements from MAMMALS new algorithms in a simpler context. In many ML and information retrieval applications it is required to retrieve fast the k nearest neighbours (k-NN) of a given point in a dataset. When the number of dimensions exceeds 10, exact k-NN computation essentially requires to scan the whole dataset [17], so specialized approximate indexing structures have been proposed and are currently implemented in libraries like Facebook Faiss [7]. Now, these systems can also benefit from a fast memory that stores a small subset of the whole repository. Managing this memory dynamically presents many of the challenges described above with the advantage of 1) avoiding the additional complexity of the interaction with the model, and 2) having a clear evaluation framework with well established benchmarks and performance metrics.

Skills
We are looking for one of the following profiles:
1) a candidate with solid analytical skills to design algorithms with strong performance guarantees,
2) a candidate expert on high-dimensional data analysis,
3) a candidate with hands-on experience on machine learning, able to reproduce state-of-the-art results like those in [12] and in [19].

Benefits package
- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Remuneration
Gross Salary: 2653 € per month

General Information
Theme/Domain: Networks and Telecommunications
System & Networks (BAP E)
Inria Center: CRI Sophia Antipolis - Méditerranée
Starting date: 2021-09-01
Duration of contract: 1 year, 6 months
Deadline to apply: 2021-07-31

Contacts
Inria Team: NEO
Recruiter: Giovanni Neglia / Giovanni.Neglia@inria.fr

About Inria
Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions: 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

Instruction to apply
Defence Security:
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy:
As part of its diversity policy, all Inria positions are accessible to people with disabilities.

Warning: you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.