

- . Titre: **BigData en NeuroScience**
- . Thématique: **bases de données**
- . Laboratoire, institution et université: **Inria**
- . Ville et pays: **Sophia-Antipolis, France**
- . Équipe ou projet dans le labo: **équipe mnemosyne**
- . Nom et adresse électronique du directeur de stage: Thierry.Vieville@inria.fr
- . Nom et adresse électronique du directeur du laboratoire: Frederic.Alexandre@inria.fr

. Présentation générale du domaine: *Les modèles informatiques étudiés en neurosciences computationnelles sont nourris de faits obtenus à partir de la littérature. Cette connaissance est mise en mots et en chiffres, conduisant à des descriptions dites "phénoménologiques", puis une partie de cette information est intégrée dans des modèles, donc mise en équations et formalisée sous forme d'algorithmes distribués pour vérifier certains aspects quantitatifs et qualitatifs de la compréhension globale présumée. Curieusement, ce travail est réalisé de manière artisanale et on est un peu en retard par rapport à d'autres domaines de connaissance qui ont, eux, pris en compte les développements récents et puissant en terme de représentations sémantiques. De très vastes et hétérogènes corpus comme ceux de Wikipedia ont été reformulés en terme de base de données sémantiques, avec des connaissances formalisées à partir d'ontologies. Ce stage va aider à changer de paradigme.*

- . Objectifs du stage: Le but du travail est d'adresser cette question en trois étapes:
 - 1 / Considérant un sujet bien défini (soit le traitement visuel précoce de la rétine au thalamus, soit la sélection de l'action au niveau des ganglions de la base et des structures neuronales liées, à choisir) on construira un formalisme informatique décrivant les connaissances actuelles du sous-système du cerveau.
 - 2 / En ce basant sur cet exemple, on proposera un langage formel intermédiaire (disons, ``turtle``) pour que les neuroscientistes annotent facilement leurs publications afin d'en construire une ontologie qui puissent être traitée ensuite par des outils sémantiques existants.
 - 3 / Considérant les outils et initiatives existants pour une plate-forme collective de travail sur ces données, on analysera la complexité (d'usage et algorithmique) d'un tel projet et on en dressera les possibilités et les limites.

. Références:

- [0] Plus d'information sur ce projet: <http://www-sop.inria.fr/members/Thierry.Vieville/projects/neurobigdata>
- [1] Site de l'équipe : <https://team.inria.fr/mnemosyne>
- [2] Le projet DBpedia : <http://fr.dbpedia.org>
- [3] Présentation grand public d'une ontologie : <http://interstices.info/ontologie>
- [4] Le langage Turtle : http://fr.wikipedia.org/wiki/Turtle_%28syntaxe%29
- [6] Un outil autoantique d3js (<http://d3js.org>) dont l'aspect dynamique pourraient aider à fouiller l'ontologie.
- [7] Quelques références complémentaires:
 - Using Semantic [for .] an Ontology of Brain-Cortex Anatomy Olivier Dameron et al Stanford, Inria Rennes, 200x
 - Foundations for an ontology of brain areas, circuits and functions Bénédicte M Batrancourt, CRIM Paris ; CNS'2013
 - Cognitive Ontologies Mapping structure and function of the brain from a systemic view Jaime Gomez, U. Polytechnica Madrid, 2008

- . Compétences espérées: **Stage de M1 d'initiation à la recherche, c'est la motivation qui est le levier.**
 - **Intérêt vis à vis des neurosciences et de la compréhension du cerveau.**
 - **Bases sur la représentation et manipulation de données, sensibilisation aux Web sémantique.**
 - **Un peu de programmation Web HTML5/JS.**

Ce stage pourra se poursuivre en thèse.