

# Application BASTRI

## Fiches Equipes

### MAGNET (SR0550PR)

Machine Learning in Information Networks  
MOSTRARE (SR0039GR) □ MAGNET □ MAGNET (SR0739PR)

**Statut:** Terminée

**Responsable :** Marc Tommasi

**Mots-clés de "A - Thèmes de recherche en Sciences du numérique - 2024" :** *Aucun mot-clé.*

**Mots-clés de "B - Autres sciences et domaines d'application - 2024" :** *Aucun mot-clé.*

**Domaine :** Perception, Cognition, Interaction

**Thème :** Représentation et traitement des données et des connaissances

**Période :** 01/01/2013 -> 01/05/2016

**Dates d'évaluation :**

**Etablissement(s) de rattachement :** <sans>

**Laboratoire(s) partenaire(s) :** <sans UMR>

**CRI :** Centre Inria de l'Université de Lille

**Localisation :** Centre Inria de l'Université de Lille

**Code structure Inria :** 101044-0

**Numéro RNSR :** 201321079K

**N° de structure Inria:** SR0550PR

### Présentation

Le projet de l'équipe Magnet consiste en la définition de méthodes et modèles d'apprentissage automatique au sein de réseaux d'informations. Ces réseaux sont des collections d'informations hétérogènes organisées sous la forme de grands graphes dont les liens sont explicitement fournis (réseaux d'interactions, hyperliens, réseaux sociaux) ou sous-jacents induits par une similarité. On s'intéresse de façon privilégiée aux réseaux d'information présents dans Internet qui se caractérisent par de grands volumes de données hétérogènes où la dimension textuelle est omniprésente. L'étude de ces réseaux d'informations textuelles pose des défis importants aux méthodes traditionnelles d'apprentissage. D'une part, la prise en compte des relations entre les objets remet en cause l'hypothèse d'indépendance et d'identique distribution habituellement postulée. D'autre part, l'analyse de données textuelles pose des problèmes de grande dimensionalité, d'annotations coûteuses, et lorsqu'elle dépasse l'approche par mots-clés, fait intervenir l'aspect structuré des entrées et des sorties. Les objectifs envisagés sont les systèmes de veille, de recherche et d'extraction d'informations, et de recommandation. Dans ce but, on entend étudier les tâches de détection d'entités et de co-références, d'extraction de structures temporelles et rhétoriques, et de fouille d'opinions dans un cadre multi-documents.

### Axes de recherche

Nous proposons des approches basées sur l'apprentissage statistique, en présence de données représentées sous forme de graphes, dans un cadre semi-supervisé et non supervisé. Les recherches peuvent se décliner selon les axes suivants:

- Découverte de structures : prédiction structurée pour les textes, prédiction de liens, clustering, estimation de densités dans les graphes.
- Combinaisons de représentations et de méthodes d'apprentissage pour la manipulation de données hétérogènes.
- Exploitation de la structuration sous forme de graphes des données pour la classification, la diffusion et la recommandation.

### Relations industrielles et internationales

- SAP
- Music Story
- Clic and Walk
- Projet Hermès (Pôle de compétitivité Picom)

### Contact

- **Responsable :** Marc Tommasi
- **Tél :** + 33. 3. 5.9 .35. 8.7 .23
- **Secrétariat Tél :** + 33. 3. 5.9 .35. 8.6 .16

### En savoir plus

- Site de l'équipe
- Site sur [inria.fr](http://inria.fr)
- Site du [responsable](#)
- Derniers Rapports d'Activité : [2016](#) , [2017](#) , [2018](#) , [2019](#) , [2020](#) , [2021](#) , [2022](#) , [2023](#)

### Documents sur la structure

- [Intranet](#)
- [Privés](#)

### Décisions

- **9085** (03/01/2013) : création
- **10149** (01/07/2014) : prolongation
- **11052** (30/06/2015) : prolongation
- **11365** (16/12/2015) : prolongation
- **11592** (26/04/2016) : fermeture

### Localisation

- **Adresse postale :** Centre Inria de l'Université de Lille Parc Scientifique de la Haute Borne 40, avenue Halley Bât.A, Park Plaza 59650 Villeneuve d'Ascq France
- **Coordonnées GPS :** 50.606, 3.149

