



# TextMine

## une communauté entre recherche et industrie

---

**Pascal Cuxac**    **INIST Nancy**  
pascal.cuxac@inist.fr



# La communauté EGC

---

**EGC : Extraction et Gestion des Connaissances**

<https://www.egc.asso.fr/presentation>

*L'association EGC se fixe pour objectif de promouvoir les échanges multidisciplinaires , en connexion avec les spécialistes d'entreprises qui déploient les méthodes et les outils adaptés à leurs besoins, afin de contribuer à la formation d'une communauté scientifique dans le monde francophone autour de cette double thématique de l'extraction et de la gestion de connaissances.*

# Atelier TextMine - EGC

---

Chaque année depuis **2017**

L'**objectif** de cet atelier est de favoriser les discussions et les échanges entre participants sur la thématique de la **fouille de textes (FdT)** s.l.

Les **contributions** sont publiées :

- sous forme d'**actes en ligne** : <https://textmine.sciencesconf.org/>
- dans une collection **HAL** : <https://cnrs.hal.science/TEXTMINE>

# Groupe de Travail EGC TextMine

---

Créé en 2022

Pascal Cuxac, INIST

Vincent Lemaire, Orange Labs

Cédric Lopez, Emvista



# Groupe de Travail EGC TextMine

---

Le but de ce groupe de travail est de réunir des **acteurs** (chercheurs, ingénieurs, doctorants ...) sur la thématique large de la **fouille de textes**.

Ce groupe vise à offrir une occasion de **rencontres** pour les **universitaires** et les **industriels**, appartenant aux différentes communautés de l'intelligence artificielle, l'apprentissage automatique, le traitement automatique des langues, pour discuter des méthodes de fouille de textes au sens large et de leurs applications.

# Activités du GT TextMine

---

- **Fédérer** la collectivité scientifique **francophone** autour de la fouille de textes au sens large (TAL, Machine Learning, Web sémantique, Traduction automatique...),
- Promouvoir la **Science Ouverte** (publications en OA, ouverture des codes)
- **Proposer des journées d'échanges** autour d'invités (monde académique et industriel) avec des conférenciers invités et des présentations **d'étudiants**,
- Organiser des **challenges**,
- Animer l'**atelier TextMine** en conjonction avec la conférence EGC,
- Être une pépinière pour la préparation de soumissions à **appels à projets**,
- **Communiquer** autour de la fouille de textes et de nos activités.

# Une synergie monde académique - industrie

## Entreprise :

ACS / Akio / Berger-Levrault / BIMBAMJOB / CEGEDIM / Dassault Systèmes / EDF / Groupe Crédit-Mutuel / Kernix / Nextino / Orange / ReportLinker / SNCF / Thales / Worldline

## EPST :

CEA / Cergy Paris Université / CIRAD / CSTB / INA / INRAE / INRIA / INSA Rennes / IRIT / ISC-PIF / Léonard De Vinci Pole Universitaire / LIG / LIST-Luxembourg / Loria / Mine Alès / MIRACL FSEGS / Nantes Université / Sorbonne Université / Universität Tübingen / Université Aix-Marseille / Université d'Avignon / Université de Caen / Université de Franche-Comté / Université de Genève / Université de La Réunion / Université de Strasbourg / Université du Mans / Université Lille / Université Lorraine / Université Lyon / Université Montpellier / Université Paris 1 / Université Paris 8 / Université Paris-Nanterre / University of Tunis

# Défi TextMine

## Les défis

2022/2023 : <i>Reconnaissance d'entités d'intérêts dans les signatures d'e-mails</i>	2023/2024 : <i>Reconnaissance d'entités géographiques dans un corpus des Instructions Nautiques</i>
Kevin Cousaut Cédric Lopez	Helen Mair Rawsthorne Adrien Guille Nathalie Abadie

## Favoriser la Science Ouverte :

- dépôts des actes dans la **collection HAL-TextMine** (<https://cnrs.hal.science/TEXTMINE>) ,
- “avantage” accordé aux participations au défi avec **mise à disposition du code.**



# TextMine

---

## TextMine :

- un lieu d'échanges
  - entre étudiants, chercheurs académiques, chercheurs en entreprises, usagers de la FdT
  - afin de partager des expériences, avancées, données, problématiques
  - apporter des réponses, des solutions
- un lien entre d'autres associations/groupes de travail (GDR TAL, Madics, AFIA...)
- un espace pour la communauté FdT francophone