



**15<sup>e</sup> conférence francophone sur l'Extraction et la Gestion des Connaissances  
avec Session Industrielle : “Big Data Analytics”**

*La Session Industrielle aura lieu entièrement en anglais*

**EGC 2015**  
27-30 janvier 2015, Luxembourg

<http://egc2015.lippmann.lu>

---

**Président du comité de programme :**

Jérôme Darmont, ERIC – Université de Lyon, France

**Responsables du comité d'organisation :**

Benoît Otjacques, Centre Gabriel Lippmann, Luxembourg

Thomas Tamisier, Centre Gabriel Lippmann, Luxembourg

**Responsable de la Session Industrielle :**

Patrik Hitzelberger, Centre Gabriel Lippmann, Luxembourg

**Responsable de la Session Démonstrations :**

Mohammad Ghoniem, Centre Gabriel Lippmann, Luxembourg

L'objectif d'EGC est de rassembler des chercheurs de disciplines connexes à l'extraction et la gestion de connaissances (apprentissage automatique, ingénierie des connaissances, statistique et analyse de données, systèmes d'information, bases de données) et des spécialistes du monde industriel qui déploient des méthodes d'extraction et de gestion de connaissances sur des données réelles.

Comme chaque année, la conférence EGC est ouverte à la présentation de travaux de recherche récents et de développements industriels dans le domaine de l'extraction et de la gestion des connaissances. Elle permet ainsi d'exposer des résultats de travaux théoriques et appliqués et d'échanger autour de ces derniers. Merci de se référer à la liste des thèmes ci-dessous.

Plusieurs types de communications sont possibles : travaux de recherche originaux (publiés sous forme d'articles longs de 12 pages, d'articles courts de 6 pages ou de posters accompagnés de résumés de 2 pages, selon leur maturité) ; travaux de recherche déjà publiés dans de bonnes conférences internationales mais inédits en français (publiés sous forme de résumé de 6 pages) ; démonstrations de logiciels (présentées dans des articles de 6 pages). Les articles peuvent être soumis en français ou en anglais (pour les non-francophones). En outre, EGC 2015 accueillera plusieurs ateliers thématiques, avec des appels à contribution séparés. Les informations sur les ateliers sont disponibles sur le site de la conférence.

Deux prix du meilleur article académique et appliqué seront décernés, ainsi que le prix de thèse (appel séparé à venir). Les règlements des prix sont disponibles sur le site de l'association. Les actes d'EGC 2015, comprenant les articles de recherche et les démonstrations, paraîtront dans un numéro de la revue RNTI. Les meilleurs articles seront sélectionnés pour être publiés dans des post-actes internationaux édités par Springer Verlag.

## **Session Industrielle : Big Data Analytics**

EGC 2015 comportera une Session Industrielle entièrement en anglais sur le thème : Big Data Analytics. Cette session permettra la rencontre et les échanges entre participants académiques et industriels.

La Session Industrielle se spécialisera sur la mise en pratique et le déploiement opérationnel d'applications et de projets en relation avec les thèmes d'EGC (voir ci-dessous), avec un intérêt particulier pour les expériences dans le domaine des données massives. Les scénarios de réussites et les retours d'expérience concernant la mise en production de technologies analytiques dans l'industrie sont attendus. La Session Industrielle s'intéressera également aux pistes de recherche ouvertes, aux perspectives industrielles de l'analyse des données massives, ainsi qu'à l'utilisation pratique, en général, des outils et méthodes d'extraction et de gestion de connaissances.

Les contributions doivent avant tout porter sur l'intérêt pratique des travaux, qu'elles viennent de l'industrie ou du monde académique, ou en particulier présentent des collaborations entre les deux. Les articles courts (maximum 6 pages) ou longs (maximum 12 pages) peuvent être soumis, uniquement en anglais, en se conformant au format et règles de la conférence : voir plus bas et sur le site de la conférence.

## **Démonstrations**

Une session spéciale de démonstrations de logiciels et de prototypes de recherche sera organisée dans le cadre de la conférence EGC 2015.

Au même titre que pour les articles de recherche, ces démonstrations devront relever des thèmes de la conférence, à savoir : l'extraction et la gestion de connaissances (fondements, méthodologies, domaines émergents, applications...). Merci de se référer à la liste des thèmes ci-dessous.

Les papiers (maximum 6 pages) présentant ces logiciels seront inclus dans les actes de la conférence.

Chaque proposition acceptée fera l'objet d'une présentation rapide en session plénière, ainsi que d'un temps de démonstration approfondie lors d'une rencontre avec les participants de la conférence

---

## **Thématiques de la conférence**

### **Fondements de l'extraction et de la gestion de connaissances**

#### Tâches

- Apprentissage supervisé : machines à noyaux, machines à vecteurs supports, réseaux de neurones formels, modèles probabilistes, apprentissage de règles, méthodes d'ensembles, régression...
- Classification non supervisée : regroupement conceptuel, méthodes de partitionnement et de recouvrement, méthodes neuronales, méthodes hiérarchiques, *clustering* multivue, multistratégie, incrémental, collaboratif
- Sélection de modèles
- Extraction de motifs : *itemset*, séquence, arbre, graphe...
- Détection d'*outliers*, d'inattendus et d'anomalies

- Système de recommandations
- Autres paradigmes : apprentissage semi-supervisé, apprentissage actif, apprentissage multistratégie...

### Méthodes

- Méthodes statistiques en fouille de données
- Apprentissage symbolique, programmation logique inductive
- Apprentissage topologique, variétés mathématiques
- Fouille visuelle de données
- Fouille de données et contraintes
- Méthodes incrémentales de fouille de données
- Algorithmes de fouille de données robustes au passage à l'échelle
- Systèmes distribués/parallèles pour la fouille de données
- Analyse de données symboliques

### Données

- Données structurées, semi-structurées, textuelles...
- Données hétérogènes sémantiquement / multimédia, spatiales, images, vidéo, audio... / peuplant des modèles de connaissances, données relationnelles, données en réseau, données de graphes
- Données géolocalisées, temporelles, spatiales
- Données annotées à l'aide d'ontologies, exploitées dans le cadre du Web sémantique
- Données volumineuses, complexes, dynamiques...
- Données ouvertes
- Données sociales

### **Méthodologies d'extraction et de gestion de connaissances**

- Acquisition, recueil, prétraitement des données, filtrage, réduction de données, sélection et modification des caractéristiques
- Critères et évaluation de la qualité des données et des connaissances
- Intégration de données (ETL, entrepôts de données, médiation...)
- Intégration de connaissances dans le processus d'extraction
- Gestion des connaissances et d'ontologies (acquisition, stockage, mise à jour, interopérabilité, interconnexion, évolution)
- Visualisation analytique, OLAP, interaction homme-machine en fouille de données
- Traçabilité, sécurité et intégrité de l'information et des données
- Plateformes et systèmes pour l'ECD
- Études comparatives par banc d'essais
- Protocoles d'évaluation et validation de modèles à partir d'échantillons d'utilisateurs
- Études expérimentales sur des données volumineuses

### **Extraction et gestion de connaissances dans des domaines émergents**

- Social : analyse de liens, communautés en ligne, réseaux sociaux et autres
- Mobilité, géolocalisation, ambient, ubiquitaire
- Modélisation de comportements
- Commerce électronique, publicité en ligne
- Fouille de données d'opinions, de dépêches, de *microblogging*
- *Open data*
- *Linked Data*
- *Crowdsourcing*

### **Applications de l'extraction et de la gestion de connaissances**

- Mémoires d'entreprise, veille technologique
- Détection d'intrusion, prévention de fraude
- Modélisation des épidémies, recherche clinique, surveillance de points de santé
- Gestion de la relation client, des réseaux et des systèmes de gestion
- Développement durable, transports intelligents

- Autres applications dans des domaines comme la médecine, la biologie, la chimie, la finance, les assurances...

---

## Soumission

Les soumissions doivent être anonymes et se faire par voie électronique exclusivement à l'adresse ci-dessous :

- Pour les articles longs, courts, et les posters :  
<https://www.easychair.org/conferences/?conf=egc2015>
- Pour les démonstrations :  
<https://easychair.org/conferences/?conf=egc2015demos>

Elles doivent être envoyées au format PDF exclusivement et doivent impérativement utiliser le format LaTeX RNTI accessible à l'adresse suivante.

<http://www.editions-rnti.fr/files/RNTI-X-Y2.1.zip>

*Les soumissions qui dépasseront 12 pages ou ne respecteront pas le format LaTeX RNTI ne seront pas évaluées.*

---

## Dates importantes

*Toutes les dates sont fermes et définitives. Il ne sera accordé aucun délai ou report.*

### **Articles de recherche** (long, short, posters, resp. 12, 6, 2 pages)

- Résumés des articles : 10 octobre 2014
- Textes complets des articles : 17 octobre 2014
- Interactions auteurs/membres du comité de programme : 12-14 novembre 2014
- Notification aux auteurs : 22 novembre 2014
- Version finale des articles : 10 décembre 2014

### **Démonstrations de logiciels** (6 pages)

- Date limite de soumission : 17 octobre 2014
- Notification aux auteurs : 22 novembre 2014
- Version finale des articles : 10 décembre 2014