

## Inter - Text Mining ou Analyse de données appliquée à des données textuelles

Les 19 et 20 mars 2020

<p><b>Objectifs</b></p>	<ul style="list-style-type: none"> <li>• Importer une base de données textuelles sous R</li> <li>• Représenter les données textuelles sous un format approprié</li> <li>• Préparer les données en effectuant les principaux pré-traitements spécifiques aux données textuelles</li> <li>• Visualiser les données textuelles</li> <li>• Effectuer une recherche d'informations</li> <li>• Effectuer une classification de documents (apprentissage supervisé)</li> <li>• Extraire des thématiques d'une base de données textuelles (apprentissage non supervisé)</li> </ul>		
<p><b>Public</b></p> <p><b>Pré-requis</b></p>	<p>Toute personne (ITA, Chercheurs, doctorants) souhaitant découvrir le text mining et sa mise en application sous R.</p> <p>Avoir des notions de bases en statistique (statistique descriptive, estimation), en data mining (apprentissage supervisé, non supervisé, ACP) ainsi que des notions de R. Joindre le questionnaire complémentaire à votre demande d'inscription</p>		
<p><b>Programme</b></p>	<p><b>Introduction au text mining</b> : Contexte ; Exemple de données ; Application du text mining</p> <p><b>Particularité des données textuelles</b> : Multiplicité des représentations ; Fléau de la dimension ; Subjectivité des données</p> <p><b>Application du text mining</b> : Recherche d'information ; Classification de documents ; Analyse de sentiments</p> <p><b>Représentations des données textuelles</b> : Sac de mots ; Vectorisation</p> <p><b>Pré-traitements</b> : Tokenization ; Nettoyage (majuscule, nombre, ponctuations...) ; Stopwords ; Stemming</p> <p><b>Visualisation</b> : Histogramme ; Nuage de mots</p> <p><b>Recherche d'informations</b> : Comparaison de textes ; Pondération des termes (TF-IDF...) ; Text ranking</p> <p><b>Application du text mining</b> : Classification de documents ; Extraction de thématiques avec le modèle LDA</p>		
<p><b>Dates et lieu</b></p>	<p>Date : <b>les 19 et 20 mars 2020</b> Inscription avant le : <b>20 janvier 2020</b> sur <a href="https://www.sirene.inserm.fr/">https://www.sirene.inserm.fr/</a></p> <p>Lieu : <b>Paris 13<sup>e</sup></b></p>		
<p><b>Contact</b></p>	<table border="0"> <tr> <td data-bbox="432 1939 794 2051"> <p><b>Nathalie Suzanne</b> Chargée de développement RH <a href="mailto:nathalie.suzanne@inserm.fr">nathalie.suzanne@inserm.fr</a> Tel 01 85 55 38 18/44</p> </td> <td data-bbox="922 1955 1150 2038"> <p>INSERM DR Paris 11 48-50 rue Albert 75013 PARIS</p> </td> </tr> </table>	<p><b>Nathalie Suzanne</b> Chargée de développement RH <a href="mailto:nathalie.suzanne@inserm.fr">nathalie.suzanne@inserm.fr</a> Tel 01 85 55 38 18/44</p>	<p>INSERM DR Paris 11 48-50 rue Albert 75013 PARIS</p>
<p><b>Nathalie Suzanne</b> Chargée de développement RH <a href="mailto:nathalie.suzanne@inserm.fr">nathalie.suzanne@inserm.fr</a> Tel 01 85 55 38 18/44</p>	<p>INSERM DR Paris 11 48-50 rue Albert 75013 PARIS</p>		