Table des matières

Qu’est-ce que le text and data mining (TDM) ? 2

Quelles sont les données concernées ?..... 2

Qui peut utiliser la technique du TDM ? 2

Qu’est-il permis de faire ? 2

- Accès licite : 3
- Intégrité et sécurité : 3

Quand le TDM ne peut-il pas être utilisé ? 3

Crédits 3

Qu’est-ce que le text and data mining (TDM) ?

T : Text, D : Data, M : Mining

Le text and data mining ou fouille de textes et de données est selon l’ordonnance n°2021-1518 du 24 novembre 2021 : *« On entend par fouille de textes et de données […] la mise en œuvre d’une technique d’analyse automatisée de textes et données sous forme numérique afin d’en dégager des informations, notamment des constantes, des tendances et des corrélations ».*

Quelles sont les données concernées ?

Les opérations de TDM peuvent porter sur des textes et sur tous les types de contenus numériques : données, images fixes ou animées, sons, musiques, logiciels, etc.

Qui peut utiliser la technique du TDM ?

Les organismes de recherche, les bibliothèques ouvertes au public, les musées, les services d’archives et les institutions dépositaires du patrimoine cinématographique, audiovisuel ou sonore.
Aucune compensation n’est à verser aux titulaires de droits.

Qu’est-il permis de faire ?

Il est permis de « **reproduire des contenus protégés** par des droits de propriété intellectuelle dans le but de conduire des activités de fouille à des fins de recherche scientifique, **sans avoir à recueillir d’autorisation préalable** des « titulaires de droits » (les producteurs des bases de données, les propriétaires des textes et/ou des données ciblés par le TDM :

entreprises, éditeurs, …) ou à obtenir des licences de leur part ».

Quelles sont les conditions à respecter ?

- **Accès licite :** il est possible d'utiliser « des contenus protégés disponibles uniquement par le biais d'un accès licite ». Exemples : ressources acquises ou mises à disposition par le biais d'un abonnement.
- **Intégrité et sécurité :** les titulaires de droits sont autorisés à mettre en œuvre « des mesures proportionnées et nécessaires afin d'assurer la sécurité et l'intégrité des réseaux et des bases de données dans lesquels les œuvres sont hébergées ». Exemples : système d'authentification.

Quand le TDM ne peut-il pas être utilisé ?

- Dans le cadre d'un **partenariat à but lucratif entre acteurs publics et privés** : « lorsqu'une entreprise est actionnaire ou structurellement associée à un organisme de recherche et bénéficie par ce biais d'un accès privilégié aux données, alors cet organisme ne peut mobiliser l'exception à des fins de recherche scientifique ».
- Lorsque les contenus à fouiller contiennent des **données protégées**. Exemple : données personnelles.

Crédits

Maurel L., Rennes S. La fouille de textes et de données à des fins de recherche : une pratique confirmée et désormais opérationnelle en droit français. 2021. [En ligne]. Consulté le 17 février 2023.