



E-DOC LABS

ASSISTANT D'ANONYMISATION

*L'assistant d'anonymisation repère
les données à anonymiser, vous
permet de les valider, et génère un
fichier anonymisé prêt à la
publication !*

L'assistant d'anonymisation e-doc Labs vous permet en trois phases très simples de rendre anonymes les données personnelles que contient votre base documentaire.

Identifiez automatiquement les informations à anonymiser¹

La technologie e-doc Labs vous permet de repérer automatiquement les informations à anonymiser : noms de personnes mais aussi, si besoin, adresses ou autres données personnelles. Ces informations sont mises en évidence par une coloration du texte et une synthèse est proposée.

Elements	Number
Persons	7
Organisations	6
Locations	2
Time	2
Measure	0
Relations	0

CONSEIL DE LA CONCURRENCE

(...)

CONCLUSION

Sur la saisine effectuée le **10 janvier 1971** par la société **Gemard-Lepine**, par son pdg Michel **Varnou**, qui se plaint de pratiques de concurrence déloyale de la part de la société **Walter Dessoult** à travers son établissement secondaire situé à Rivesaltes dans deux points de vente, situés **rue Saint-Georges**.

L'examen des pratiques mises en oeuvre par la société **Walter Dessoult** de **Londres** sur le marché des sacs plastique dans la région parisienne, fait apparaître que la société **Walter Dessoult**, représenté par son directeur Gérard **Blaine**, a exploité d'une façon abusive sa position dominante en organisant une situation qui se caractérise par :

(...)

et qui a eu pour objet et a pu avoir pour effet de restreindre le développement de la concurrence sur ce marché, ce qui constitue une pratique prohibée par l'article 17 de l'ordonnance du **23 décembre 1957**.

D'autre part, l'intervention de la directrice générale de **Keinju SA**, Géraldine **Gérard**, concurrent direct des deux sociétés sus-nommées, ne fait apparaître aucune concurrence déloyale, eu égard aux pratiques de cette société :

Délibéré, sur le rapport de Mme **Marie-Laure Hervit**,

par M. Georges **Valley** président,

(...)

Le rapporteur général, Micheline **Frait**

Le président, Georges **Valley**

Chaque onglet synthétise les informations précises extraites par le logiciel. Ici, les **noms Michel Varnou, Gérard Blaine, Géraldine Gérard, Marie-Laure Hervit, George Valley et Micheline Frait** sont identifiés, ainsi que leur prénom. Chaque type d'information comporte ainsi une liste d'attributs. Pour les personnes, le logiciel repère le prénom, la fonction, le genre et la nationalité.. Ces attributs peuvent être adaptés aux besoins des utilisateurs.

¹ Le fichier exemple est fictif et toute ressemblance avec un fichier et/ou des personnes existantes serait fortuite.



Pour plus de sécurité, une interface intuitive vous permet de valider les données repérées

e-doc Labs effectue pour vous le repérage des données personnelles mais une interface simple et intuitive permet à un opérateur humain de valider définitivement les données « candidates » à l'anonymisation.

The screenshot shows a web browser window with the URL `http://localhost/files/edoclabs/docValidate/parseXML.php`. The page title is "e-doc Labs DocumentStore : Pick up a file, validate the data and store it into a DB". Below the title, there is a "Choose a file:" section with a text input and "Parcourir..." and "View" buttons. The main content area is divided into four tabs: "Persons" (selected), "Organisations", "Locations", and "Terminology". The "Persons" tab displays a table with the following data:

Person	Count	Name	Role
VARNOU	1	Michel	pdg
BLAINE	1	Gérard	directeur
GERARD	1	Géraldine	
VALLEY	2	Georges	président
FRAIT	1	Micheline	

Below the table are buttons for "Add Row", "Delete Row", a dropdown menu showing "0", and "Anonymize Data". Below the table, there is a text document with the following content:

(...)

CONCLUSION

Sur la saisine effectuée le 10 janvier 1971 par la société **Gemard-Lepine**, par son pdg Michel **Varnou**, qui se plaint de pratiques de concurrence déloyale de la part de la société **Walter Dessoult** à travers son établissement secondaire situé à Rivesaltes dans deux points de vente, situés **rue Saint-Georges**.

L'examen des pratiques mises en oeuvre par la société **Walter Dessoult de Londres** sur le marché des sacs plastique dans la région parisienne, fait apparaître que la société **Walter Dessoult**, représenté par son directeur Gérard **Blaine**, a exploité d'une façon abusive sa position dominante en organisant une situation qui se caractérise par :

(...)

et qui a eu pour objet et a pu avoir pour effet de restreindre le développement de la concurrence sur ce marché, ce qui constitue une pratique prohibée par l'article 17 de l'ordonnance du 23 décembre 1957.

D'autre part, l'intervention de la directrice générale de **Keiniu SA**, Géraldine **Gérard**, concurrent direct des deux sociétés sus-nommées, ne fait apparaître aucune concurrence déloyale, eu égard aux pratiques de cette société.

(...)

At the bottom of the browser window, there is a "Terminé" status and an "Intranet local" icon.

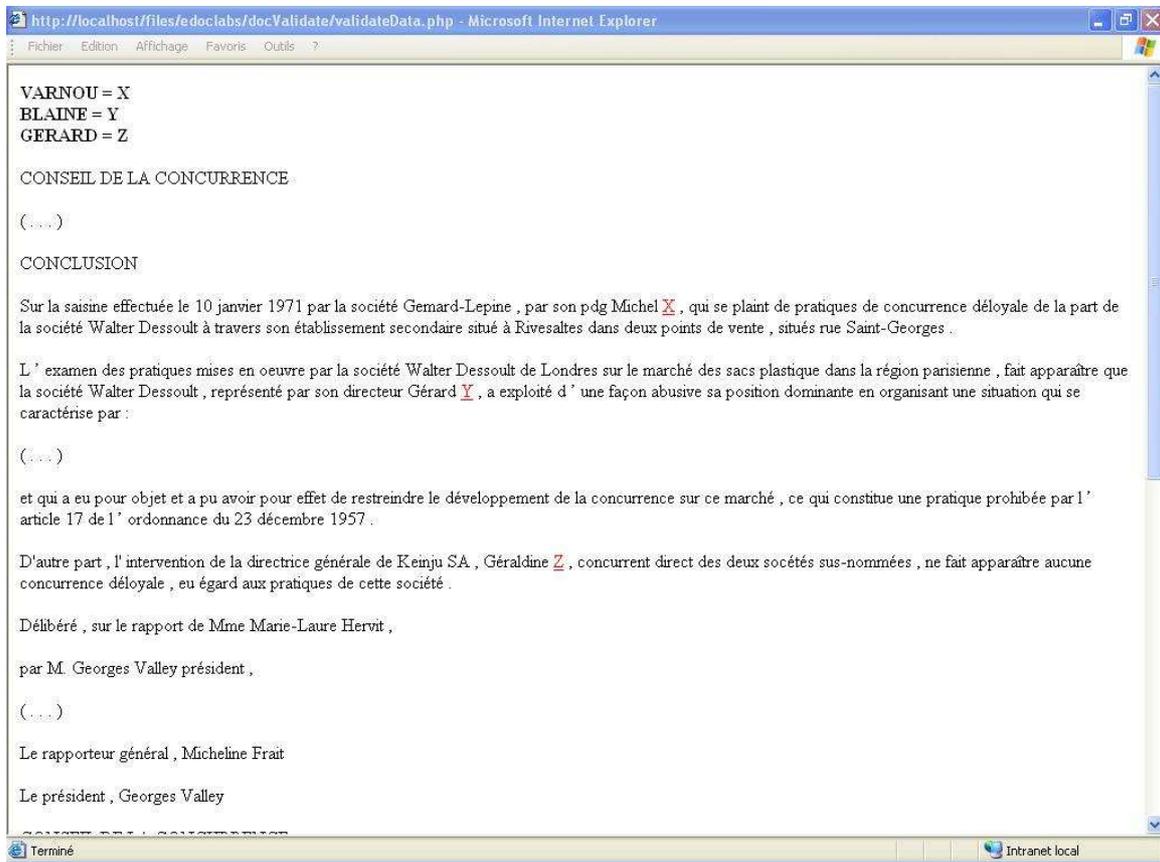
Zone
d'édition des
informations
repérées

Commandes
de la
validation des
données



Remplacement automatique des noms repérés et validés par des données anonymes :

Finalement, une fois validées les données à anonymiser, le programme génère le fichier de départ avec les données anonymisées, selon les recommandations de la CNIL.



AVANTAGES DE LA SOLUTION

- ✓ *Traitement de tous les fichiers bureautiques classiques (office, pdf, html...)*
- ✓ *Précision du repérage par l'utilisation de techniques éprouvées en Traitement Automatique des Langues Naturelles*
- ✓ *Interface de validation*
- ✓ *Possibilité de publication du document anonymisé en différent format (pdf, html, etc.)*
- ✓ *Possibilité de stockage des documents dans une base de données XML ou relationnelle pour exploitation ultérieure (moteur de recherche etc.)*

CARACTERISTIQUES TECHNIQUES

Langues traitées	Français, Anglais, Espagnol
Formats d'entrée	TXT, WORD, PDF,HTML,XML
Plate-formes	Windows, Linux
Technologies	Java, XML,Unicode, Technologies linguistiques
Ressources linguistiques	Externes et facilement adaptables
Configuration minimale	Pentium 3, 10 Mo , RAM 256 Mo

SOCIETE

e-doc Labs est une société innovante qui développe et commercialise des solutions multilingues d'analyse du langage et de recherche documentaire.

Ses fondateurs sont issus des entreprises leader du domaine du Traitement Automatique des Langues (TALN) et de la recherche universitaire. Ils possèdent une grande expérience des projets d'entreprise et ont mené à bien de nombreuses missions de conseils pour des grands comptes en France et à l'international. Ils ont également participé à des projets de recherche européens et internationaux d'envergure.

e-doc Labs est actuellement une société en incubation à la Technopole de l'Aube en Champagne. (<http://www.technopole-aube.fr>).

CONTACT :

e-doc Labs

Technopole de l'Aube en Champagne

BP 601

10901 TROYES Cedex 9

emmanuel.cartier@e-doclabs.com

