

Mars 2024  
RAPPORT N°20.44



Institut des Études  
et de la Recherche  
sur le Droit et la Justice

# Outil d'analyse de conformité des stipulations (OACS)

Sous la direction de

**Malo DEPINCÉ, Agnès ROBIN,  
Laurent FAURÉ, Mathieu LAFOURCADE  
et Anne LAURENT**



LICeM



LIRMM



UNIVERSITÉ DE  
MONTPELLIER



**Sr**

**SYNTHÈSE DE  
RECHERCHE**



## Sous la coordination :

du LICeM Laboratoire Innovation Communication et Marché UR\_UM213, Université de Montpellier

## Membres de l'équipe

**Malo DEPINCÉ,**

Professeur à l'université de Montpellier (spécialisé en droit de la consommation et droit des contrats)

**Gwenaëlle DONADIEU,**

Docteure de l'université de Montpellier (spécialisée en droit de la consommation)

**Laurent FAURÉ,**

Maître de conférences à l'université de Montpellier 3 (spécialisé en linguistique)

**Mathieu LAFOURCADE,**

Maître de conférences HDR à l'université de Montpellier (spécialisé en informatique, intelligence artificielle et traitement du langage naturel)

**Anne LAURENT,**

Professeure à l'université de Montpellier (spécialisée en informatique, intelligence artificielle et données<sup>1</sup>)

**Agnès ROBIN,**

Maîtresse de conférences HDR à l'université de Montpellier (spécialisée en droit de la propriété intellectuelle et du numérique).

**Philippe HUGOT,**

Technicien et responsable de la communication du LICeM a rejoint l'équipe pour aider à la coordination de la saisine des références de la base de données.

---

<sup>1</sup> Le Professeur Anne Laurent et Madame Donadieu ont quitté le projet en mai 2023 manquant de temps pour y rester pleinement investies.



# Note de synthèse du projet OACS

*Outil d'Analyse du Contenu des Stipulations*, Université de Montpellier

Pour l'IERDJ 21 novembre 2023

## **Ambition**

Le projet interdisciplinaire OACS entendait développer un système informatique capable d'analyser des clauses contractuelles pour identifier toute non-conformité avec les réglementations en vigueur. Financée par l'IERDJ, la première phase se concentre sur la création d'un outil d'analyse de conformité des clauses contractuelles à la réglementation applicable aux « clauses abusives » dans les contrats entre professionnels et consommateurs. Les chercheurs à l'initiative du projet avaient constaté que les consommateurs éprouvaient des difficultés légitimes à assimiler les informations dans les conditions générales de vente, et que les autorités de contrôle pouvaient parfois manquer de moyens pour les analyser systématiquement.

L'outil informatique élaboré vise par conséquent à rendre les informations juridiques accessibles, en évaluant la conformité des offres de produits ou services aux réglementations en vigueur. À terme, le système doit filtrer l'ensemble des pratiques contractuelles en ligne. La première étape, objet de la présente note de synthèse, impliquait la création d'un outil d'analyse des clauses considérées individuellement avant d'appliquer un contrôle systématique des sites et des conditions générales d'utilisation. Plus précisément il s'agissait de permettre à un interrogateur d'entrer le texte d'une clause d'un contrat sur un site internet pour que le système l'analyse au regard d'une base de données complète et lui réponde en retour sur le caractère potentiellement abusif de la clause renseignée en lui donnant toutes les informations utiles à la compréhension de la situation.

Le projet rassemble plusieurs laboratoires de recherche, le laboratoire Innovation Communication et Marché, le Laboratoire d'Informatique, de Robotique et de Micro-électronique de Montpellier, et le laboratoire LHUMAIN, avec des enseignants-chercheurs de l'université de Montpellier et de l'université de Paul Valéry Montpellier 3 (pour plusieurs disciplines : droit de la consommation, droit des contrats, linguistique, informatique, droit de la propriété intellectuelle et numérique), ainsi que des étudiants stagiaires. L'objectif est de fournir des solutions aussi pertinentes que possible, compréhensibles aussi bien par des spécialistes que par les consommateurs.

## **Moyens**

L'équipe a pu se constituer, se réunir, et travailler grâce au soutien de l'IERDJ. Elle a pu également acquérir le matériel nécessaire à la pérennisation de son travail.

**Moyens humains** - Six enseignants-chercheurs, un technicien de recherches et une quinzaine d'étudiants ont participé à cette première phase du projet dans l'objectif d'une initiation à la recherche.

**Moyens techniques** – Le projet a bénéficié de l’expérience du laboratoire LIRMM en matière d’outils d’analyse de textes et de l’outil jeuxdemots.org particulièrement. Il a été possible sur cette base de construire un algorithme complexe d’analyse des clauses soumises au regard de celles renseignées dans la base de données spécialisée.

Pour pérenniser le système (au moins sur les cinq prochaines années), l’université de Montpellier a pu acquérir grâce au financement de l’IERDJ un serveur dédié à OACS pour accueillir les données et les outils de leur traitement. Des postes de travail vont en outre être mis en place pour le traitement de la base de données par les étudiants (hors financement de l’IERDJ).

### **Difficultés**

Comment construire un système en mesure de penser de la manière la plus proche possible de celle d’un être humain (un juge) ? Un effort pédagogique a d’abord été entrepris, nous avons ainsi refusé de construire un système brut qui n’aurait pas été spécialement adapté à notre sujet. À ce propos, il a fallu dans un premier temps transposer en une présentation la plus segmentée possible l’ensemble du dispositif de qualification de clauses abusives, pour éviter qu’un système informatique ait à manier des concepts trop ambigus ou imprécis. L’analyse des stipulations contractuelles pour en apprécier la conformité à la loi est en outre un exercice difficile, qui suppose de connaître autant la réglementation que la jurisprudence et le contexte de la relation des parties. Cela est particulièrement vrai s’agissant de l’analyse du caractère potentiellement abusif d’une stipulation contractuelle. L’essentiel du dispositif législatif et réglementaire a longtemps porté sur la notion de « *déséquilibre significatif* » qui pouvait se manifester dans une seule stipulation (un « article » du contrat) ou au contraire résulter de l’étude conjointe de stipulations inscrites dans des articles distincts du même contrat ou encore dans des stipulations inscrites dans des documents différents mais intégrant le même champ contractuel (par exemple pour un lien à opérer entre des conditions générales et des conditions particulières). Pour autant, l’analyse de la jurisprudence montre que l’essentiel des clauses stigmatisées relevaient bien d’un seul et même article contractuel. Des formules sémantiques récurrentes (des clauses types) permettaient de relever des clauses à risques. L’étude par les spécialistes du langage de la lecture des contrats par plusieurs juristes (cette étude se poursuit encore) confirme que certaines expressions pouvaient être considérées par des lecteurs spécialisés et porteuses d’un risque de clause abusive.

Puis en 2008, le pouvoir réglementaire a expressément (et pour la première fois de manière complète en 30 années d’existence d’une réglementation sur les clauses abusives) dressé une liste de clauses présumées (de manière simple ou irréfragable) abusives. Le raisonnement sur la qualification de clause abusive a ainsi pu être segmenté en plusieurs autres, sur la base de ces considérations plus spécifiques. Cela nous a amenés à proposer des sous-catégories de clauses abusives. Si l’analyse des clauses par le système se fait selon un corpus de données (clauses types) général, la présentation des résultats se fait lui en fonction de ces sous-catégories.

Dans un premier temps, une classification des clauses en 25 catégories a été élaborée. Cette mise en ordre s'est avérée complexe, la réglementation en la matière ayant procédé à plusieurs renvois vers d'autres textes et exonérations.

Il a fallu ensuite construire une base de données en récupérant les informations les plus pertinentes des avis et décisions relevant de la qualification des clauses abusives. Certaines informations étaient toutefois déjà interprétées dans les sources auxquelles les chercheurs ont eu accès, de sorte qu'il a été souvent impossible d'inscrire dans la base de données la clause elle-même : l'interprétation donnée de la clause par un tiers (un juge par exemple) devait par conséquent souvent être retraduite en une formule relevant du discours contractuel.

L'analyse du langage juridique a, en outre, nécessité des présentations interdisciplinaires pour tenter de construire les outils de détermination des nœuds (ces formules langagières que le système extrait de la clause présentée pour en faire un index de sa recherche) les plus efficaces. Le travail d'analyse linguistique a porté sur la forme des contrats et leur lecture pour mieux comprendre la construction des stipulations contractuelles. Il a consisté à interroger directement la pratique juridique, en complément et en opposition à l'analyse automatisée des clauses abusives. Une telle approche consiste à examiner les méthodes utilisées par les juristes, appelées « ethnométhodes », et à confronter celles-ci à la corrélation probabiliste du système informatique.

Le focus initial porte sur la lecture experte d'un contrat par un juriste, qui implique une maîtrise technique du discours juridique et une capacité à distinguer entre langage spécialisé et langage ordinaire. L'expertise juridique repose sur des hypothèses inférentielles inductives et abductives, utilisant à la fois le raisonnement et l'intuition.

La méthodologie de l'étude a été qualitative, reposant sur la saturation ethnographique des données, avec des enregistrements d'entretiens et d'analyse de contrats. La première phase de l'étude, portant sur six heures d'enregistrements transcrits, a permis de stabiliser la méthode de collecte de la parole experte, mais des enquêtes supplémentaires seraient nécessaires pour varier les perspectives et les styles de lecture analytique, et ainsi adapter encore l'outil informatique. Ce travail est toujours en cours.

## **Le système**

*L'analyse* – OACS est un système basé sur des techniques de traitement du langage naturel et d'intelligence artificielle qui utilise une approche abductive pour qualifier des propositions. Le système se concentre sur l'extraction de termes pertinents à partir des clauses, qualifiés de ("termes d'entrée"), et utilise une représentation de graphes basée sur un recueil de connaissance.

Pendant l'apprentissage, les termes d'entrée sont reliés à des « neurones » ou « nœuds » créés automatiquement par le système, avec des réponses attendues associées. Lors de l'exploitation, les neurones correspondant à une partie significative des termes extraits sont activés, et la réponse obtenue est une combinaison des réponses associées à chaque neurone activé.

Le réseau neuronal ainsi construit n'est pas spécifique à l'identification des clauses abusives, mais plutôt générique pour diverses tâches. Les réponses attendues sont annotées avec un terme indiquant la tâche liée. Une notice explicative est recommandée pour contextualiser les réponses, compte tenu de l'incapacité du système à produire une analyse contextuelle similaire à celle du cerveau humain.

Cette notice explicative est constituée des fiches pédagogiques qui accompagnent le résultat présenté.

***La présentation du résultat*** – le système ne peut être compris que comme un outil d'aide à la qualification et à la compréhension de celle-ci. Les résultats sont donc présentés sous une forme brute qui se veut simple : « *potentiellement abusive* » ; « *a priori non abusive* » ou « *En l'état des informations fournies, le système est incapable de donner une réponse* » mais qui est accompagnée de fiches pédagogiques pour expliquer la solution à l'interrogateur afin qu'il puisse lui-même juger si la proposition de qualification qu'il reçoit est pertinente et le satisfait. Il y a autant de fiches que de classes de clauses abusives reconnues par le système (25).

### **Résultats**

À l'heure actuelle, la base de données n'est pas encore complète sur les clauses abusives (il nous faudra doubler ou tripler le contenu de cette dernière). Pour autant le système est opérationnel et lorsqu'un consommateur interroge le système il obtient une réponse qui est toujours accompagnée d'une notice explicative. Si aujourd'hui la notice a une place importante dans la réponse donnée, nous considérons que plus la base de données sera complète, plus le système aura de points de comparaison et plus la fiche perdra en utilité.











LICeM



LIRMM



UNIVERSITÉ DE  
MONTPELLIER



Institut des Études  
et de la Recherche  
sur le Droit et la Justice