

Vers une consolidation automatique du droit français

GEORGES-ANDRÉ SILBER, CRC, Mines Paris, PSL University, France

CCS Concepts: • **Information systems** → **Information extraction**; • **Software and its engineering** → **Domain specific languages**; • **Applied computing** → **Law**.

1 INTRODUCTION

Les textes composant le droit français sont modifiés par des textes modificateurs publiés dans le journal officiel de la République française (JORF). Le cycle de vie d'un texte juridique législatif ou réglementaire démarre par la publication de sa version intégrale dans le JORF et se poursuit par la publication éventuelle de textes le modifiant. Le texte intégral modifié, appelé sa version *consolidée*, n'est jamais publié dans le JORF et n'a pas de valeur juridique: seule la version initiale et la suite des modifications ordonnées du texte font foi [8]¹.

Le site Légifrance [4] présente depuis 2008 la majeure partie des textes juridiques français dans leurs versions d'origines ainsi que dans leurs versions successives, conséquences des modifications apportées à ces textes dans le temps. L'opérateur français du site Légifrance, la Direction de l'Information Légale et Administrative (DILA), reporte manuellement les modifications décrites en langage naturel dans les textes afin de d'obtenir, à chaque date de modification, la version consolidée complète du texte.

Cette commodité d'accès aux textes dans une version plus simple à lire et à utiliser a de facto changé le statut de ces versions consolidées: elles sont vues par la plupart des utilisateurs, y compris les professionnels du droit, comme le reflet du droit applicable [7]. De plus, les rédacteurs de nouveaux textes, au parlement français ou dans les ministères, partent de cette version consolidée pour concevoir les textes modificateurs. Il est ainsi de toute première importance que ce travail de consolidation soit exempt d'erreurs et disponible le plus rapidement possible.

Nous présentons ici des travaux préliminaires intégrés dans l'outil Legistix que nous développons, dont l'objectif est de créer un système de consolidation des textes de loi français automatisé et fiable. Ces travaux se fondent à la fois sur des expressions régulières utilisées dans plusieurs grammaires composées, similaires aux passes successives d'un compilateur, et sur un nouveau langage spécialisé de type fonctionnel, permettant de décrire les changements appliqués aux textes.

Pour chaque texte modificateur, notre outil génère de manière complètement automatique un programme informatique dans ce nouveau langage qui, lorsqu'il est exécuté, permet d'effectuer les changements induits par le texte modificateur sur les textes cibles. Dans les travaux antérieurs sur ce sujet, présentés par exemple dans [6], seul le problème de classification des types de modification est abordé. À notre connaissance, nos travaux sont les premiers à présenter une approche complète permettant d'identifier les textes cibles et de transformer les instructions en langage naturel du texte modificateur en un programme informatique formalisant les règles effectives de transformation.

Contrairement à d'autres sujets d'études liés aux textes juridiques, comme extraire automatiquement des raisonnements, des règles de droit ou tenter de comprendre la sémantique du texte, nous nous plaçons dans une perspective purement *légistique* [8], c'est-à-dire que nous ne cherchons pas à comprendre le sens du texte, mais uniquement sa structure et ses relations à d'autres textes. Notre

¹Extrait du site Légifrance: "*La consolidation consiste à intégrer dans un acte unique, sans valeur officielle, les modifications et les corrections successives apportées à un texte ; son objectif étant de faciliter la connaissance de leurs droits et obligations par les citoyens.*" <https://www.legifrance.gouv.fr/contenu/en-tete/informations-de-mises-a-jour>

problème étant clairement défini, il se prête bien à une automatisation, car la consolidation doit veiller à ne pas interpréter les modifications mais à uniquement les appliquer à la lettre [7].

2 EXEMPLE DE CONSOLIDATION

Afin d'illustrer la complexité d'un outil comme Legistix, nous allons utiliser dans ce document un exemple de consolidation partant de la loi 2022-1348 du 24 octobre 2022 publiée au JORF du 25 octobre 2022 et présentée fig. 1. Comme tous les textes publiés au JORF, sauf indication contraire, cette loi est entrée en vigueur le lendemain de sa publication [8], soit le 26 octobre 2022. Nous pouvons également noter dans le paragraphe III de la loi une *entrée en vigueur différée* le 1er janvier 2023 de certains changements induits. L'unique texte cible des changements est l'article L723-4 du code de commerce, présenté fig. 2.²

Article 1 de la loi 2022-1348 du 24 octobre 2022 (partiel)

I. - L'article L723-4 du code de commerce est ainsi modifié:

- 1° Au début du premier alinéa, est ajoutée la mention: « I. - » ;
- 2° Au 1°, la deuxième occurrence du mot : « *et* » est remplacée par le mot : « *ou* » ;
- 3° Aux 3° et 4°, après le mot : « *procédure* », sont insérés les mots : « *de sauvegarde*, »;
- 4° Au 4° bis, la première occurrence du mot: « *fait* » est supprimée;
- 5° Au 5°, après le mot: « *qualités* », sont insérés les mots : « *et fonctions* »;
- 6° Le dernier alinéa est remplacé par un II ainsi rédigé:
«II. - Sont également éligibles, [...]»
«2° [...] des établissements inscrits au répertoire des métiers ou dans le ressort des tribunaux limitrophes. [...]»

II. - A la première phrase du 2° du II de l'article L723-4 du code de commerce, dans sa rédaction résultant du I du présent article, les mots : « *répertoire des métiers* » sont remplacés par les mots : « *registre national des entreprises en tant qu'entreprise ou établissement du secteur des métiers et de l'artisanat* ».

III. - Le II du présent article est applicable à compter du 1er janvier 2023. La présente loi sera exécutée comme loi de l'Etat.

Fig. 1. Reproduction partielle de l'article 1 de la loi 2022-1348 du 24 octobre 2022 publiée dans le journal officiel du 25 octobre 2022 [14] (fr/loi/2022-1348/1/20221025).

Article L723-4 du code de commerce, en vigueur depuis le 13 octobre 2021

Sont éligibles aux fonctions de juge d'un tribunal de commerce les personnes âgées de trente ans au moins:

- 1° Inscrites sur les listes électorales des chambres de commerce et d'industrie et des chambres de métiers et de l'artisanat [...]
- 3° A l'égard desquelles une procédure de redressement judiciaire ou de liquidation judiciaire [...]
- 4° [...] une procédure de redressement judiciaire ou de liquidation judiciaire est en cours au jour du scrutin;
- 4° bis Qui n'ont fait pas fait l'objet des sanctions [...]
- 5° Et qui justifient [...] des qualités énumérées [...]

Sont également éligibles les membres en exercice des tribunaux de commerce, [...].

Fig. 2. Reproduction partielle de l'article L723-4 du code de commerce français [13], en vigueur depuis le 13 octobre 2021 (fr/code/commerce/L723-4/20211013).

²Ces changements peuvent être visualisés en ligne sur le site Légifrance, avec les changements appliqués manuellement par la DILA: https://www.legifrance.gouv.fr/codes/section_lc/LEGITEXT000005634379/LEGISCTA000006161381/2022-10-26/#LEGISCTA000006161381

L'article de loi, le texte modificateur qui modifie l'article de code cible, contient des instructions destinées à des humains pour modifier le texte original et obtenir le nouveau texte. Ces instructions sont des insertions, suppressions, remplacements, etc. On peut noter également au 4^e bis de l'article une erreur du législateur d'une loi modificative antérieure votée au parlement [11]: "*Qui n'ont fait pas fait...*". L'objectif de notre outil est de transformer ces instructions destinées à des humains en un programme informatique permettant de les appliquer automatiquement.

3 DONNÉES, IDENTIFICATEURS ET VERSIONS

Les données intégrées dans Legistix sont constituées de l'intégralité des bases JORF [1] et LEGI [2] fournies en open data par la DILA, et mises à jour quotidiennement. Ces bases contiennent l'intégralité des textes publiés au JORF depuis 1990 ainsi que tous les textes qui ont été consolidés par la DILA. Les données sont intégrées au fil de l'eau dans la base de données de Legistix, grâce à un processus d'intégration améliorant les données, reconstruisant notamment les paragraphes et les tables des matières complètes versionnées, non fournies dans les données originales. Par exemple, le code de commerce a plus de 600 versions depuis l'an 2000.

Chaque document reçoit un identificateur unique sous la forme d'un URI. La forme de cet URI est une poursuite des travaux menés avec la DILA dans le cadre du projet Légimobile [9]. Ces identificateurs ont la particularité d'être pérennes et de pouvoir être dérivés systématiquement à partir des informations du document lui-même. Ainsi, par exemple, l'article 1 de la loi 2022-1348 du 24 octobre 2022 présenté fig. 1 a comme identificateur unique l'URI `fr/loi/2022-1348/1`, représentant l'ensemble des versions de cet article au cours du temps.

Pour préciser une version particulière, les identificateurs permettent également de prendre en charge les nombreuses dates régissant le cycle de vie d'un texte de loi: date de signature, date de publication dans le JORF, dates de modification induisant des versions, date d'abrogation. Ainsi, la version publiée au JORF du 25 octobre 2022 de l'article précédent a comme URI `fr/loi/2022-1348/1/20221025`. Suivant la même logique, l'article L723-4 du code de commerce dans sa version en vigueur depuis le 13 octobre 2021 présenté fig. 2 a comme identifiant `fr/code/commerce/L723-4/20211013`.

4 DÉTECTION DES TEXTES MODIFIÉS ET CRÉATION DE NOUVELLES VERSIONS

Dans l'exemple de la loi 2022-1348, Legistix détecte la référence "*article L723-4 du code de commerce*" présente deux fois. Grâce au mécanisme d'identificateurs décrit plus haut, elle va être résolue en l'identificateur `fr/code/commerce/L723-4`, sans nécessité d'interroger un mécanisme de résolution externe: la chaîne de caractères de la référence suffit.

Nous pouvons ici commencer à décrire le programme obtenu automatiquement par Legistix: présenté ici pour plus de clarté dans un langage fonctionnel avec une syntaxe à la *Python*, il va nous permettre d'illustrer certains des mécanismes que nous mettons en œuvre. Tout d'abord, le texte source `s` peut être défini comme étant l'article 1 de la loi 2022-1348 dans sa version au 25 octobre 2022 telle que publiée au JORF, avec `db` un objet représentant la base de données des textes et `get_version` une méthode renvoyant une référence vers une version existante:

```
s = db.get_version("fr/loi/2022-1348/1", Date(2022, 10, 25))
# s is fr/loi/2022-1348/1/20221025
```

Legistix doit maintenant trouver les différentes versions du texte cible dont il va être question dans les modifications décrites. La première version dont nous avons besoin est celle dont on part, c'est-à-dire la version sur laquelle les changements doivent s'appliquer. Pour la première référence se trouvant au paragraphe I, la date est celle de la publication au JORF du texte source, soit le 25 octobre 2022. La référence:

```
t = db.get_version("fr/code/commerce/L723-4", Date(2022, 10, 25))
```

```
# t is fr/code/commerce/L723-4/20211013 (!)
# that is, the version in force on October 10, 2022
```

est positionnée à l'article L723-4 dans sa version au 13 octobre 2021. En effet, la version en vigueur de ce texte au 25 octobre 2022 est bien celle du 13 octobre 2021 [13], telle que modifiée par l'article 1 de la loi 2021-1317 entrée en vigueur le même jour [12].

Un texte au JORF entrant en vigueur par défaut au lendemain de sa publication, soit ici le 26 octobre 2022, la nouvelle version de l'article L723-4 décrite en I de la loi 2022-1348 doit être créée à cette date avec la fonction `new_version`, à partir de la version représentée par `t`:

```
v1 = t.new_version(Date(2022, 10, 26))
# v1 is fr/code/commerce/L723-4/20221026, created from target t
```

En II, des modifications sont décrites en faisant référence à l'article L723-4 du code de commerce dans sa rédaction résultant du I du présent article, soit *après* les changements appliqués. En combinant ces informations avec celle en III qui indique que les changements en II doivent s'appliquer au 1er janvier 2023, il faut créer une nouvelle version à cette date:

```
v2 = v1.new_version(Date(2023, 1, 1))
# v2 is fr/code/commerce/L723-4/20230101, created from version v1
```

Les noms `s`, `t`, `v1` et `v2` représentent des références abstraites à des documents. Comme on le verra dans la section suivante, ces références peuvent être utilisées pour référencer de manière abstraite des portions de texte. Par exemple, les méthodes `par` et `sen` vont permettre de générer des fonctions qui vont respectivement renvoyer des références vers un paragraphe ou une phrase lorsqu'elles seront évaluées. Ainsi, `s.par("I")` génère une fonction qui lorsqu'elle sera évaluée renverra une référence vers le paragraphe I de l'article 1 de la loi 2022-1348. Ces générateurs peuvent être combinés comme dans:

```
v1.par("II").par("2°").sen(1)
```

qui génère une fonction permettant d'obtenir la première phrase du 2° du II de l'article L723-4 du code de commerce dans sa version au 26 octobre 2022.

5 DÉTECTION DES CHANGEMENTS DANS LES TEXTES MODIFICATEURS

Après avoir identifié les versions des textes cibles mentionnés dans les textes modificateurs, Legistix détecte les changements et les transforme en une suite de fonctions permettant de les appliquer. À partir du texte de la fig. 1, Legistix génère le programme complet présenté fig 3.

Il crée deux nouvelles versions de l'article L723-4 du code de commerce, `v1` entrant en vigueur au 26 octobre 2022, avec les changements décrits dans le paragraphe I de l'article 1 de la loi 2022-1348 (`s.par("I")`), s'appliquant à la version du code de commerce au 13 octobre 2021, et `v2` entrant en vigueur au 1er janvier 2023 avec les changements décrits dans le paragraphe II (`s.par("II")`). On peut noter que la deuxième version est créée à partir de la première. Les méthodes `prepend`, `replace`, `insert`, `suppress`, `replace_par` sont des générateurs de fonctions qui ne sont pas appliquées directement par le programme, tant que la fonction `apply_changes()` n'est pas appelée.

Chaque méthode utilisée dans le programme présenté ci-dessus traduit une opération depuis le langage naturel. Chaque fonction décrivant un changement est de la forme `action(source, target, what...)`, où `source` et `target` indiquent respectivement le fragment de texte à l'origine du changement et la cible. Les paramètres `what...` décrivent les changements, avec par exemple `("et", 2)` qui représente la deuxième occurrence du mot "et" et "de sauvegarde" un mot de remplacement.

```

s = db.get_version("fr/loi/2022-1348/1", Date(2022, 10, 25))
t = db.get_version("fr/code/commerce/L723-4", Date(2022, 10, 25))

v1 = t.new_version(Date(2022, 10, 26))
pI = s.par("I")
v1.schedule_changes(s, [
    v1.prepend(pI.par("1"), v1.par(1), "I. -"),
    v1.replace(pI.par("2"), v1.par("1°"), ("et", 2), "ou"),
    v1.insert(pI.par("3"), v1.par("3°"), "procédure", "de sauvegarde,"),
    v1.insert(pI.par("3"), v1.par("4°"), "procédure", "de sauvegarde,"),
    v1.suppress(pI.par("4"), v1.par("4° bis"), ("fait", 1)),
    v1.insert(pI.par("5"), v1.par("5°"), "qualités", "et fonctions"),
    v1.replace_par(pI.par("6"), v1.lastpar(), "II. - Sont également[...]")
])
v1 = db.add_version(v1.apply_changes())

v2 = v1.new_version(Date(2023, 1, 1))
pII = s.par("II")
v2.schedule_changes(s, [
    v2.replace(pII, v2.par("II").par("2°").sen(1), "répertoire des métiers",
        "registre national des entreprises ...")
])
v2 = db.add_version(v2.apply_changes())

```

Fig. 3. Programme Legistix dérivé de l'article 1 de la loi 2022-1348 du 24 octobre 2022 publiée dans le journal officiel du 25 octobre 2022 [14].

6 CONCLUSION

Nous sommes partis d'une base de données de référence de plusieurs dizaines d'années où la consolidation a déjà été effectuée manuellement. Cette consolidation historique nous sert de référence pour mesurer la fiabilité de notre approche. L'outil Legistix est capable d'automatiser 93% des opérations de consolidations qui ont été effectuées manuellement. Nos efforts vont se poursuivre pour arriver à un taux de 100%, en analysant les cas non détectés, qui sont pour certains des cas ambigus même pour un humain, contredisant ainsi l'aspect mécanisable souhaité par le législateur [7].

L'étape suivante sera d'étendre nos travaux aux données de la base EUR-Lex [3] contenant le droit de l'Union européenne et où les règlements et directives sont publiés sur un modèle similaire au droit français.

Troisièmement, nous allons proposer un système expert [5] pour assister le législateur lors de la rédaction des textes modificateurs où les règles de modification formelles (le programme) pourraient être générées directement lors de la rédaction du texte. Cela mettrait en place un cercle vertueux où les impacts des modifications pourraient être immédiatement visualisés, améliorant et fiabilisant ainsi la production du droit.

Quand le taux de 100% aura été atteint, il n'y aura plus de délai entre la publication d'un texte modificateur et la disponibilité du texte modifié consolidé. Il ne sera plus qu'un sous produit de l'application d'une série de programmes partant du texte original. Ces programmes pourraient

être votés en même temps que le texte modificateur, rendant juridiquement réel l'adage "*code is law*" [10].

REMERCIEMENTS

Merci à Salomé Ouaknine, étudiante à Mines Paris, pour son aide pendant ses deux mois de stage de recherche sur le sujet de la consolidation automatique, notamment sur le défrichage de la future approche hybride mêlant expressions régulières et apprentissage automatique.

REFERENCES

- [1] Direction of Legal and Administrative Information (DILA) (Eds.). 2022. *Data in XML format containing all laws and regulations published in the Official Journal of the French Republic (JORF) since 1990*. <https://www.data.gouv.fr/fr/datasets/jorf-les-donnees-de-l-edition-lois-et-decrets-du-journal-officiel/>
- [2] Direction of Legal and Administrative Information (DILA) (Eds.). 2022. *Data in XML format containing French law consolidated by the Direction of Legal and Administrative Information (DILA)*. <https://www.data.gouv.fr/fr/datasets/legi-codes-lois-et-reglements-consolides/>
- [3] Publications Office of the European Union (Ed.). 2022. *EUR-Lex, access to European Union Law*. <https://eur-lex.europa.eu>
- [4] Direction of Legal and Administrative Information (DILA) (Eds.). 2022. *Légifrance*. <https://www.legifrance.gouv.fr>
- [5] Alain Delaët, Denis Merigoux, and Aymeric Fromherz. 2022. Turning Catala into a Proof Platform for the Law. In *POPL 2022 - Programming Languages and the Law*. Philadelphia, United States. <https://hal.inria.fr/hal-03447072>
- [6] Samuel Fabrizi, Maria Iacono, Tesei Andrea, and Lorenzo De Mattei. 2022. A First Step Towards Automatic Consolidation of Legal Acts: Reliable Classification of Textual Modifications. In *Proceedings of the Eighth Italian Conference on Computational Linguistics*. Milan, Italy. <http://ceur-ws.org/Vol-3033/paper26.pdf>
- [7] Thierry-Xavier Girardot. 2014. Accéder au droit: importance et défis de la consolidation. *Documentaliste – Sciences de l'Information* 51, 4 (2014), 30–32. <https://www.cairn.info/revue-documentaliste-sciences-de-l-information-2014-4-page-30.htm>
- [8] Secrétariat général du gouvernement and Conseil d'État. 2017. *Guide de légistique*. La documentation française.
- [9] Pierre Larrède and Georges-André Silber. 2011. Un plan de classement des données juridiques françaises: l'expérience Légimobile. In *iExpo2011*. <https://www.gfii.fr/uploads/docs/un-plan-de-classement-des-donnees-juridiques-francaises-l-experience-legimobile-laureat-au-prix-de-la-meilleure-contribution-scientifique-i-expo-2011.pdf>
- [10] Lawrence Lessig. 2000. Code Is Law. *Harvard Magazine* (2000). <http://harvardmagazine.com/2000/01/code-is-law.html>
- [11] French Republic. 2019. *Article 69 of law 2019-486 of may 22, 2019, published in the french official journal of may 23, 2022*. https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000038496214
- [12] French Republic. 2021. *Article 1 of law 2021-1317 of october 11, 2021, published in the french official journal of october 12, 2021*. https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000044183798
- [13] French Republic. 2021. *Article L723-4 of the French Commercial Code, in force on October 13, 2021, following the modifications made by art. 1 of the Law 2021-1317 of October 11, 2021*. https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000044191424/2021-10-13/
- [14] French Republic. 2022. *Article 1 of law 2022-1348 of october 24, 2022, published in the french official journal of october 25, 2022*. https://www.legifrance.gouv.fr/jorf/article_jo/JORFARTI000046480974