

Towards an Automatic Consolidation of the French Law

The Legistix Project

Georges-André Silber, CRC, Mines Paris, PSL University, France

ProLaLa 2023 – January 15, 2023



Is it possible to **build** an **automatic consolidation system** of the French and the European Law with a **100% reliability** and to **maintain** this reliability **over time**?

Context and Motivation



- **20+ years of technical work** on french and european law data
- Academic/technical partnership with French Official Journals (DILA) [2002-2009]
- **Légimobile**, created with DILA, a mobile version of **Légifrance** [2010-2017]
- Created a company on this subject : **Regmind** (previously **Luxia**) [2009-2019]
- Member of the SPDDI committee, advisor for the Prime Minister [2007-2019] : **Public Service for the Broadcast of Law on the Internet**
- Advisor of the French Official Journals to understand the manual process of **consolidation** and propose some automatization ideas (confidential report) [2017]



- A french law enters **in force** only if it has been **published** in the official journal *Journal Officiel de la République Française (JORF)*
- A published law enters in force **the day after** its **publication** or **later**
- Successive **modifications** of the law are published in the official journal under the form of **patches**
- The consolidation process **applies** those patches to the initial published version of the law, and creates **versions of the law** over time, with only one version in force at a given time
- Because the consolidated versions **are not published** in the official journal, they have **no legal values**
- This process is currently essentially **manual**
- **BUT** lawyers and law makers use **only** those **consolidated versions** to work
- The process must be **as reliable as possible** (should be 100% reliable)



Article 1 (published oct. 25, 2022, in force oct. 26, 2022)

I. - Article L723-4 of the Commercial Code is amended as follows :

- 1° At the beginning of the first paragraph, the following is added : *"I. -"*;
- 2° In 1°, the second occurrence of the word : *"and"* is replaced by the word : *"or"*;
- 3° In 3° and 4°, after the word : *"judicial"*, if inserted the word : *"rescue,"*;
- 4° In 4° bis, the second occurrence of the word : *"were"* is suppressed;
- 5° In 5°, after the word : *"qualities"*, are inserted the words : *"and duties"*;
- 6° The last paragraph is replaced by a II redacted as follows :
"II. - Also eligible are, [...] 2° [...] establishments [...]"

II. - In the first sentence of 2° of II of article L723-4 of the French Commercial Code, as amended by I of this article, the words : *"trade directory"* are replaced by the words : *"national register of company or establishment in the trades and crafts sector"*.

III. - II of this article shall apply as of January 1, 2023. This act shall be executed as a law of the of the state.



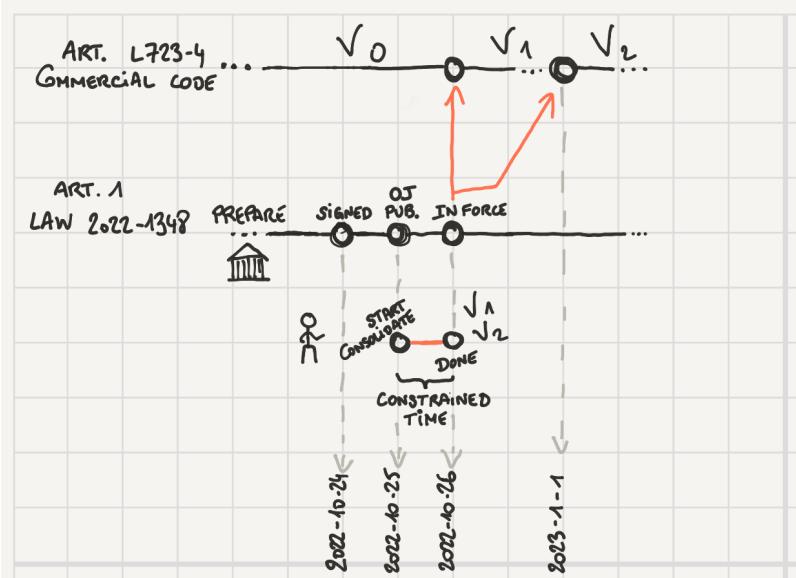
Art. L23-4 of Commerce Code (in force since oct. 13, 2021)

Persons who are at least thirty years of age shall be eligible for election as a judge of a commercial court :

- 1° Registered on the electoral lists of the chambers of commerce and industry and the chambers of trade and crafts [...]
- 3° In respect of which a judicial recovery or liquidation procedure [...]
- 4° [...] a judicial recovery or liquidation procedure is in progress on the day of the vote ;
- 4° bis Who were not were subject to the sanctions [...]
- 5° And that justify [...] of the qualities listed [...]

The following are also eligible for election as members of the commercial courts, [...].

Consolidation Timeline for Art. L723-4





Art. L23-4 of Commerce Code (in force since oct. 26, 2022)

I. - Persons who are at least thirty years of age shall be eligible for election as a judge of a commercial court :

- 1° Registered on the electoral lists of the chambers of commerce and industry **or** the chambers of trade and crafts [...]
- 3° In respect of which a judicial **rescue**, recovery or liquidation procedure [...]
- 4° [...] a judicial **rescue**, recovery or liquidation procedure is in progress on the day of the vote;
- 4° bis Who were not **were** subject to the sanctions [...]
- 5° And that justify [...] of the qualities **and duties** listed [...]

II. - Also eligible are, [...] trade directory [...]



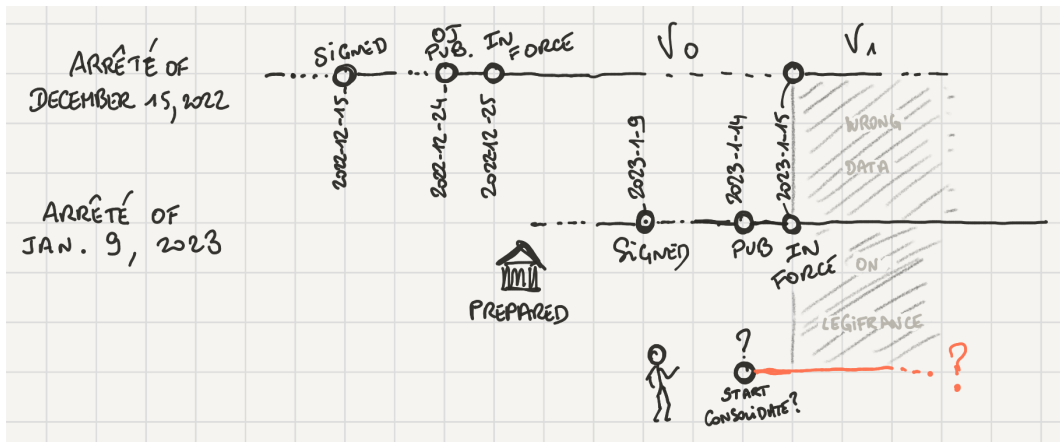
Art. L23-4 of Commerce Code (in force since jan. 1, 2023)

- I. - Persons who are at least thirty years of age shall be eligible for election as a judge of a commercial court :
- 1° Registered on the electoral lists of the chambers of commerce and industry or the chambers of trade and crafts [...]
- [...]
- II. - Also eligible are, [...] **trade directory**
national register of company or establishment in the trades and crafts sector [...]



1. The **constrained time** to manually consolidate texts **cannot always be respected** (e.g. saturday and sundays, huge texts) → **misleading information** on Légifrance.
2. The manual consolidation process **does not capture the exact actions** that could be traced back to the original modifying text → **verification is difficult**.
3. There can be **consolidation errors**, corrected afterwards when (if) detected.
4. **Redaction** of new texts or changes are also **manual** : there are **errors** there too (like the *were not were* in the preceeding example).
5. The changes to texts should be **coded** when **changes are redacted**, not using spreadsheets and word processors.

Example of Misleading Information on Légifrance



Accessed on January 15, 2023, [Arrêté of December 15, 2022](#) has not the correct content on Légifrance (it has been changed by [Arrêté of Jan. 9, 2023](#)).

Example of Misleading Information on Légifrance (snapshot)



← → ↻ <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000046791990> ⓘ ☆

Informations de mises à jour | Gestion des cookies | Nous contacter

RÉPUBLIQUE FRANÇAISE **Légifrance**
Le service public de la diffusion du droit

Liberté
Égalité
Fraternité

BULLETINS OFFICIELS | **BULLETINS OFFICIELS DES CONVENTIONS COLLECTIVES** | **JOURNAL OFFICIEL** | **DÉBATS PARLEMENTAIRES** | **QUESTIONS ÉCRITES PARLEMENTAIRES** | **DOCUMENTS ADMINISTRATIFS**

DROIT NATIONAL EN VIGUEUR | **PUBLICATIONS OFFICIELLES** | **AUTOUR DE LA LOI** | Droit et jurisprudence de l'Union européenne | Droit international

Arrêté du 15 décembre 2022 fixant un contingent exprimé en puissance et en jauge pour la réservation de capacités aux fins de délivrance de permis de mise en exploitation de navires de pêche au mois de décembre 2022

NOR : PRMM2233719A
ELI : https://www.legifrance.gouv.fr/eli/arrete/2022/12/15/PRMM2233719A/jo/texte/JORF_n°0298_du_24_décembre_2022
Texte n° 7

↓ Extrait du Journal officiel électronique authentifié
PDF - 222,1 Ko

Consolidation actions between 2009 and 2016



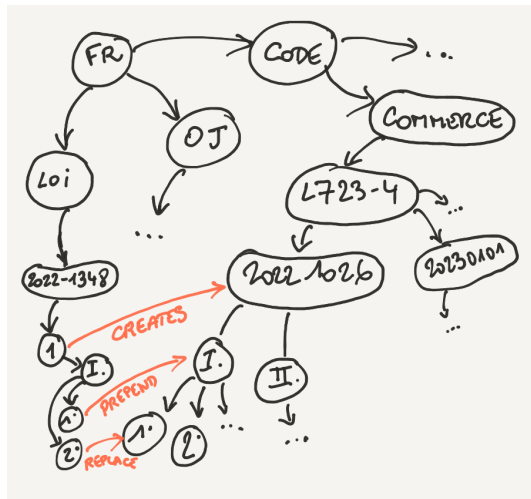
Year ≈ 300 oj	OJ texts	OJ art.	Src texts	Src art.	Consol. actions	Target texts	Target sect.	Target art.
2009	28 394	71 824	2 643	8 311	43 164	4 441	3 916	34 776
2010	28 429	75 414	2 827	9 384	44 320	4 334	3 614	36 911
2011	29 159	82 571	3 259	9 601	45 452	3 810	4 598	34 967
2012	27 282	77 323	2 717	8 066	43 814	3 532	5 007	35 454
2013	26 451	72 301	2 570	7 375	41 079	3 330	4 649	32 874
2014	25 971	78 943	2 888	9 237	46 197	3 421	4 586	34 756
2015	26 140	76 748	3 083	9 317	50 400	3 553	5 169	38 504
2016	30 722	88 660	3 105	11 170	65 638	3 717	7 916	50 094
M.≈	28 000	78 000	3 000	9 000	48 000	3 800	5 000	38 000
/ 300	94	260	10	30	160	13	17	127

Automatic Consolidation with Legistix



- Research project started in 2021 at Mines Paris
- Question : is it possible to **build and maintain automatically** a **graph** representing the "full" **structure** of the French and European **Law** ?
- **Vertices** are legal **documents** or **sets** of documents : versions of texts, sections, articles, paragraphs, ...
- **Edges** are **relations** between **documents** or **sets**
- Kind of relations : **creates, contains, cite, change, abrogate, ...**
- Each vertex is identified by a unique hierarchical **Legistix Id** (an URI)
- **Not** focused on the **meaning** of the law, but on its **legistic** structure
- Capture the **evolution** of the relations between **documents** and **sets** over **time**
- **Automatic Consolidation** is one of the sub-problems of the graph construction

Very Partial View of the Legistix Graph





- Legistix uses two set of **open data** provided by **Légifrance**
- JORF : all documents published in the official journal
 - Approx. : 1 000 000 texts, 140 000 sections, 2 300 000 millions articles
- LEGI : all manually consolidated documents (including 78 legal Codes)
 - Approx. : 130 000 texts, 1 600 000 articles versions, 220 000 sections versions
- Data are **imported daily** and analyzed, in a versioned database.



- Basically, in the previous example, needs to **recognize the reference to L723-4** article :

I. - Article L723-4 of the Commercial Code is amended as follow
and **retrieve the ID** in the system.

- This phase basically use **regular expressions** and simple **grammars** with simple ad-hoc **backtracking**.
- "Predictible" Legistix IDs allow to **infer the ID from text** (almost) without searching in some global search engine :

Article L723-4 of the Commercial Code → fr/code/commerce/L723-4

- **Determine the version(s)** of the referenced texts : here we need to find the version of the code article one day before the modifying text enters in force (October 25, 2022).
- The III. of the modifying text implies another version at January 1, 2023.



- The **creation of the modification program** can start.
- Once we have the target versions, we need to **schedule the changes**
- The **source of modifications and the target(s)** can be defined :

```
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))
```
- And the two versions :

```
v1 = t.new_version(Date(2022, 10, 26))  
# ... changes will be added here  
v2 = v1.new_version(Date(2023, 1, 1))  
# ... changes will be added here
```



```
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°"), ("and", 2), "or"),
  v1.insert(pI.par("3"), v1.par("3°"), "judicial", "rescue,"),
  v1.insert(pI.par("3"), v1.par("4°"), "judicial", "rescue,"),
  v1.suppress(pI.par("4"), v1.par("4° bis"), ("were", 2)),
  v1.insert(pI.par("5"), v1.par("5°"), "qualities", "and duties"),
  v1.replace_par(pI.par("6"), v1.lastpar(),
    "II. - Also eligible[...]")
])
```



```
# Remember: v1 = t.new_version(Date(2022, 10, 26))
#           Changes on v1 scheduled here

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), "trade directory",
              "national register of company[...]")
])

# Actually applying the changes will be performed by the Legistix
# engine (there can be several texts changing the same
# version of L723-4)
```



- Legistix uses the following process, **for each** published official **journal** :
 - **generate** all the **programs induced** by modifying texts
 - **apply the changes** scheduled for each new version
- The consolidation system has been **validated gradually**
- Starting with the initial version of a text we can **compare the successive manual and automatic versions**
- If they differ, there are two choices :
 - The **manual consolidation is wrong** (validate the program)
 - The **automatic rules are false** : change the rules
- The full process can be restarted to check for **regression issues** when changing rules
- Did that for **all** texts

Conclusion and Next steps



- **Fully automatic** creation of programs that modify texts and create versions
- New **functional API** to express **lazy modification programs**
- Engine to **apply** those transformation programs
- Implemented in Python (with some C code under the hood for speed)
- Framework to **compare manually and automatically consolidated texts**



- Continue to **refine rules** to achieve 100% (right now at 93%)
- Use an **hybrid** rules / machine learning model trained with already consolidated data
 - Help find new **future form** of changes
 - **Detect errors** that occurred at the same time in manual and automatic consolidation
 - **Reinforce** the formal system of **rules**
- Apply the techniques to **European data** : adapt the identifiers, changes mining phases
 - API and programs to express the changes **stays the same**