

Unveiling Linguistic Harmony: Asserting Interlingual Synchronicity in Documents

Balisage 2023, August 3

Srikanth Venkata Subramanian (Cognizone)

Geert Bormans (C-Moria)

Introduction



Srikanth Venkata Subramanian

knowsAbout

Angular, Frontend technologies;

memberOf

[Cognizone](#).



Geert Bormans

knowsAbout

Markup, Semantic Web technologies;

memberOf

[C-Moria](#), [MarkupUK](#), [WCG Xproc Next](#).

Setting the scene

Allgemeine Informationen

Dieser Text ist in Kraft

Abkürzung	ZGB
Beschluss	10. Dezember 1907
Inkrafttreten	1. Januar 1912
Quelle	AS 24 233
Beschluss auf.	19. Dezember 2008
Quelle auf.	AS 2010 1739
Sprache(n) der Veröffentlichung	DE FR IT RM EN
Chronologie	Chronologie
Änderungen	Änderungen
Zitate	Zitate

210

[Alles einblenden](#) | [Artikelübersicht](#) | [Alles ausblenden](#)

Schweizerisches Zivilgesetzbuch

vom 10. Dezember 1907 (Stand am 23. Januar 2023)

Die Bundesversammlung der Schweizerischen Eidgenossenschaft, gestützt auf Artikel 64 der Bundesverfassung^{1,2} nach Einsicht in eine Botschaft des Bundesrates vom 28. Mai 1904³, beschliesst:

¹ [BS 1 3]. Dieser Bestimmung entspricht Artikel 122 der Bundesverfassung vom 18. April 1999 (SR 101).
² Fassung gemäss Anhang Ziff. 2 des Gerichtsstandsgesetzes vom 24. März 2000, in Kraft seit 1. Jan. 2001 (AS 2000 2355; BBl 1999 2829).
³ BBl 1904 IV 1, 1907 VI 367

- [Einleitung](#)
- [A. Anwendung des Rechts](#)
- [Art. 1](#)

Note:
Fedlex, publication platform for legislation of the Swiss Federal administration
 - up to 5 languages
 - metadata driven from a knowledge graph
 - documents available in PDF, HTML, XML Akoma Ntoso and MS Word

Some project figures and facts

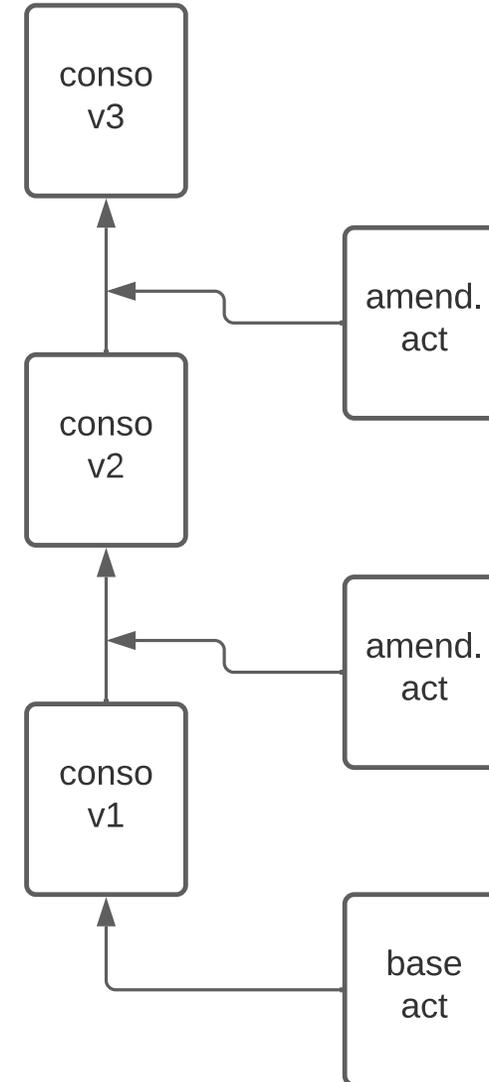
2019	Project start			
Jan 2021	Consolidations (.docx) => HTML (via Information Model)	5000 documents (dynamic) (up to 5 languages)	From well-crafted docx to XML to HTML XProc 3.0 pipeline (50 steps, XSLT,...)	[Balisage 2021]
Jun 2022	<ul style="list-style-type: none"> • Consolidations • Official collection (law) • Gazette (all .docx) => HTML	<ul style="list-style-type: none"> • 5000 documents (dynamic) • 1000 docs in '22 • 3000+ docs in '22 (per language)	From (well-crafted) docx to XML to HTML XProc 3.0 pipeline(s) (70 steps, XSLT,...)	
Jun 2022	=> Akoma Ntoso XML		AKN XSD + fedlex schematron restrict. (<i>publ. Dec 2023</i>)	[...] [Balisage 2022 DR] [...]
Aug 2023	Now all the above works reasonably well... it is time we start making life easier for the analysts (authors)			[Balisage 2023 LB] Work in progress

A word about « consolidations »

Small step aside

(a) definition

A “consolidation” is a unified and updated version of a law, created by integrating various amendments, modifications, and repeals that have been made over time.

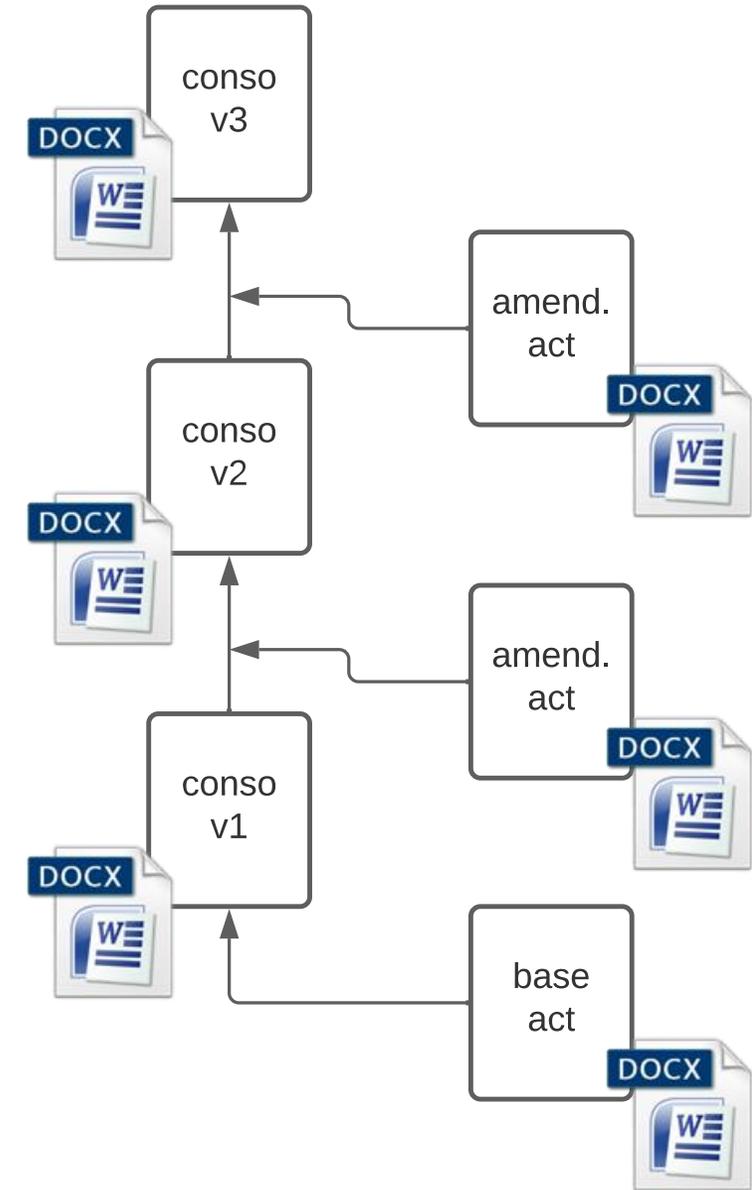


Process (CH)

A manual process through legal analysis by jurists.

NOT copy / paste !

When not sent to Parliament =>
convenient but unofficial



Tutte le versioni	
● 01.01.2024	HTML XML PDF DOC
● 01.09.2023	HTML XML PDF DOC
● <u>23.01.2023</u>	HTML XML PDF DOC
● 01.01.2023	HTML XML PDF DOC
● 01.07.2022	HTML XML PDF DOC
● 01.01.2022	HTML XML PDF DOC
● 01.01.2021	HTML PDF DOC
● 01.07.2020	PDF DOC
● 01.01.2020	PDF DOC
● 01.01.2019	PDF DOC

« | 1 2 ... 7 | »

³ Egli si attiene alla dottrina ed alla giurisprudenza più autorevoli.

- **B. Limiti dei rapporti giuridici**

- **I. Osservanza della buona fede**

- **Art. 2**

¹ Ognuno è tenuto ad agire secondo la buona fede così nell'esercizio dei propri diritti come nell'adempimento dei propri obblighi.

² Il manifesto abuso del proprio diritto non è protetto dalla legge.

- **II. Effetti della buona fede**

- **Art. 3**

¹ Quando la legge fa dipendere un effetto giuridico dalla buona fede di una persona, la buona fede si presume.

² Nessuno può invocare la propria buona fede quando questa sia incompatibile con l'attenzione che le circostanze permettevano di esigere da lui.

- **III. Apprezzamento del giudice**

- **Art. 4**

¹ Il giudice è tenuto a decidere secondo il diritto e l'equità quando la legge si rimette al suo prudente criterio o fa dipendere la decisione dall'apprezzamento delle circostanze, o da motivi gravi.

- **C. Rapporti col diritto cantonale**

*Note:
different versions of a
consolidation. The green
bullet indicates the version
currently "applicable"*

Code civil suisse (RS 210)
10 Dec. 1907 (state 23 Jan. 2023)

Code civil suisse (RS 210)
10 Dec. 1907 (state 1 Sept. 2023)

5. les mesures techniques et organisationnelles
assurer la protection et la sécurité des données

6. l'archivage des données.

⁶ Il peut prévoir que les coûts engendrés par des prestations en faveur de tiers à des fins qui ne relèvent pas du domaine de l'état civil sont facturés aux bénéficiaires.

Art. 46

5.⁶⁹ les mesures organisationnelles et techniques nécessaires pour
assurer la protection et la sécurité des données ainsi que la surveillance du respect des dispositions de protection des données;

6. l'archivage des données.

⁶ Il peut prévoir que les coûts engendrés par des prestations en faveur de tiers à des fins qui ne relèvent pas du domaine de l'état civil sont facturés aux bénéficiaires.

Art. 46

¹ Quiconque subit un dommage illicite causé, dans l'exercice de leur

⁶⁹ 1^{er} janv. 2019 (RO 2018 4017; FF 2014 3395).
Nouvelle teneur selon l'annexe 1 ch. II 16 de la LF du 25 sept. 2020 sur la protection des données, en vigueur depuis le 1^{er} sept. 2023 (RO 2022 491; FF 2017 6565).

17 / 370

Two versions of the same consolidated act ([RS 210](#))

Expiration du délai référendaire et entrée en vigueur

¹ Le délai référendaire s'appliquant à la présente loi a expiré le 14 janvier 2021 sans avoir été utilisé.¹⁹

² La présente loi entre en vigueur le 1^{er} septembre 2023.

In force date (pdf page 30/92)

Abrogation et modification d'autres actes

Annexe 1
(art. 68)

Annex 1 (pdf page 31/92)

16. Code civil⁶⁷

Art. 45a, al. 5, ch. 5

Impact on civil code (pdf page 45/92)

⁵ Le Conseil fédéral règle en collaboration avec les cantons:

5. les mesures organisationnelles et techniques nécessaires pour assurer la protection et la sécurité des données ainsi que la surveillance du respect des dispositions de protection des données;

*Note:
Fragments of the law that triggered the change...indicates the complexity of the "legal analysis". This is not "just copy and paste"*

And how we got there (Protection des données. LF [RO 2022 491](#))

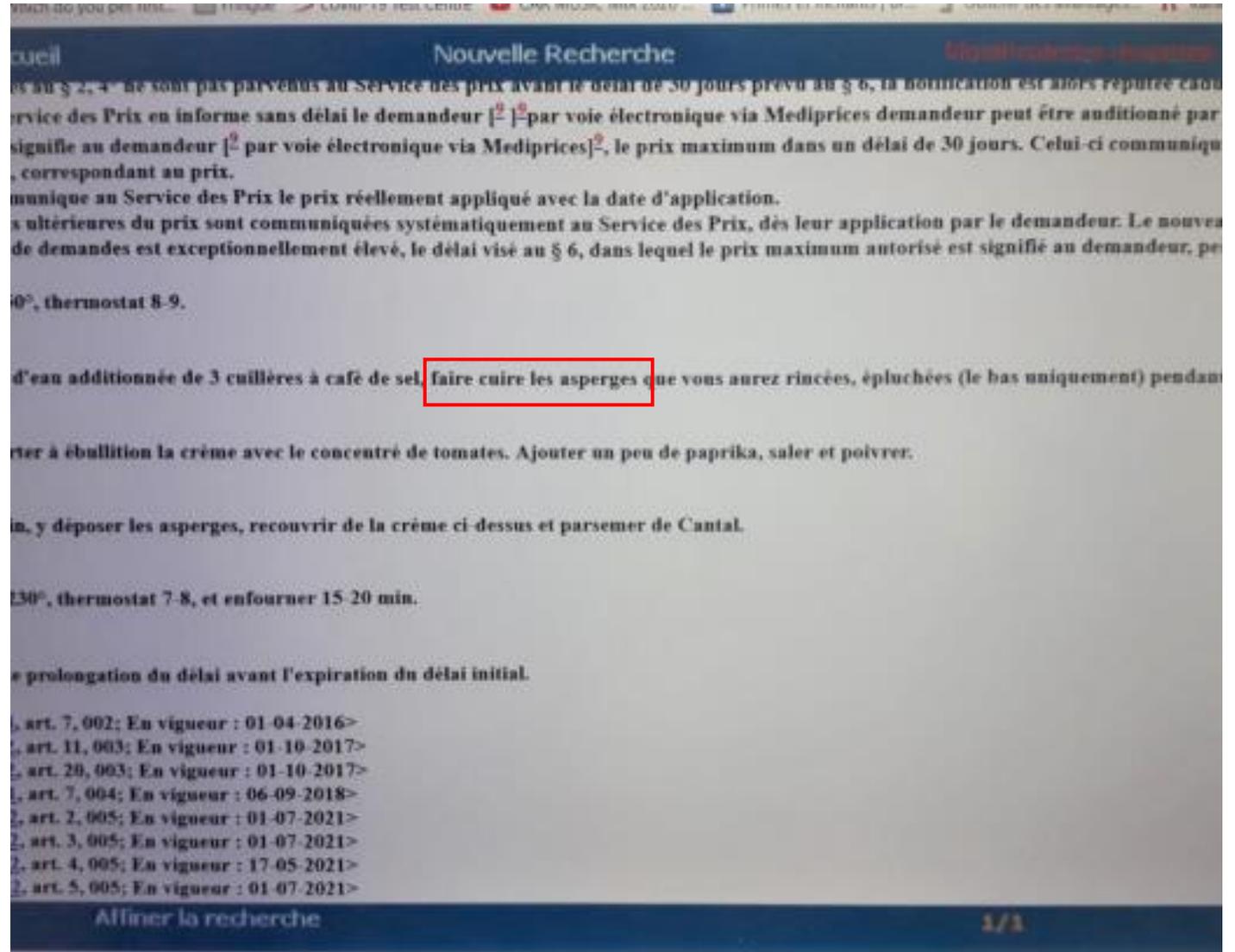
Quality Assurance...

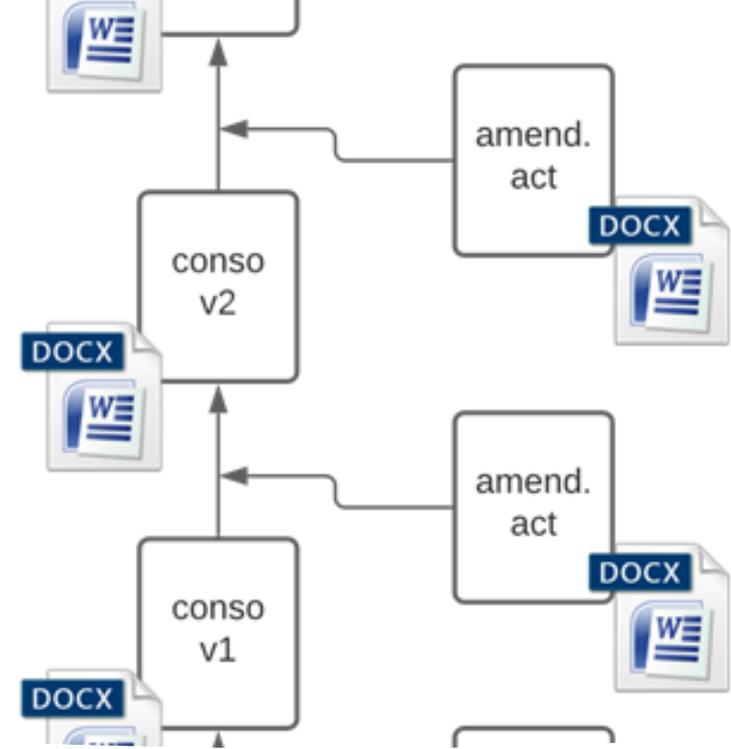
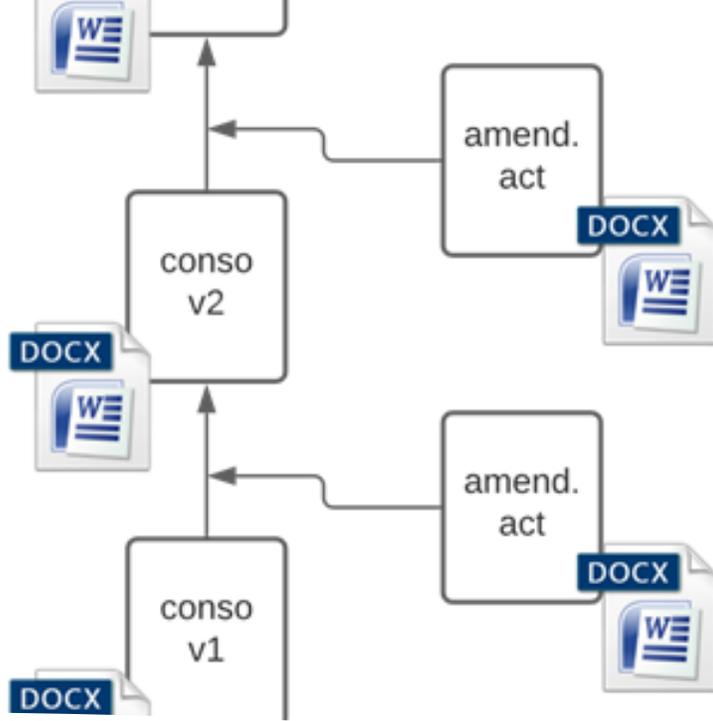
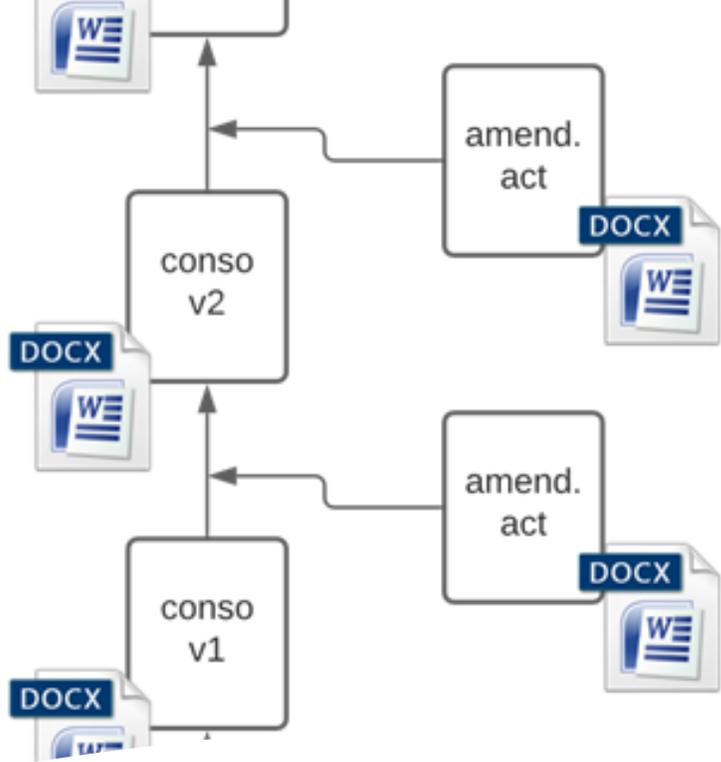
How can we help the legal analysts to spend less time on QA?

Note:

Importance of Quality Assurance: An asparagus recipe appeared in the publication of a consolidation on the website that publishes the official Belgian legislation

Yes, the director of [Moniteur Belge](#) was forced to talk to the press. “Technically,” he said, the online version “is not the official publication of the law; it is a consolidated database, with no legal basis.” The official Moniteur, he assures us, is proofread and approved.
(The Bulletin, May 31, 2021)





What are we facing?

- FR, DE, IT are equivalent, not translations (RM, EN are translations)
- 3 equivalent consolidation efforts happen in parallel
- That is a challenge...
- BUT also an opportunity...

Some of the QA check points

- Make sure that each chapter has the correct amount of articles
- Make sure that a newly added blacklist item has the correct label
- Make sure that a specific footnote references the correct law
- Make sure that a certain number was updated correctly (eg. speeding ticket fine)
- Make sure that an image is present and correct
- Make sure that dates in the document are correct
- ...

⁴ Abrogé par l'annexe 1 ch. II 3 du CPC du 19 déc. 2008, avec effet au 1^{er} janv. 2011 (RO 2010 1739; FF 2006 6841).

The opportunity: “two out of three ain’t bad”

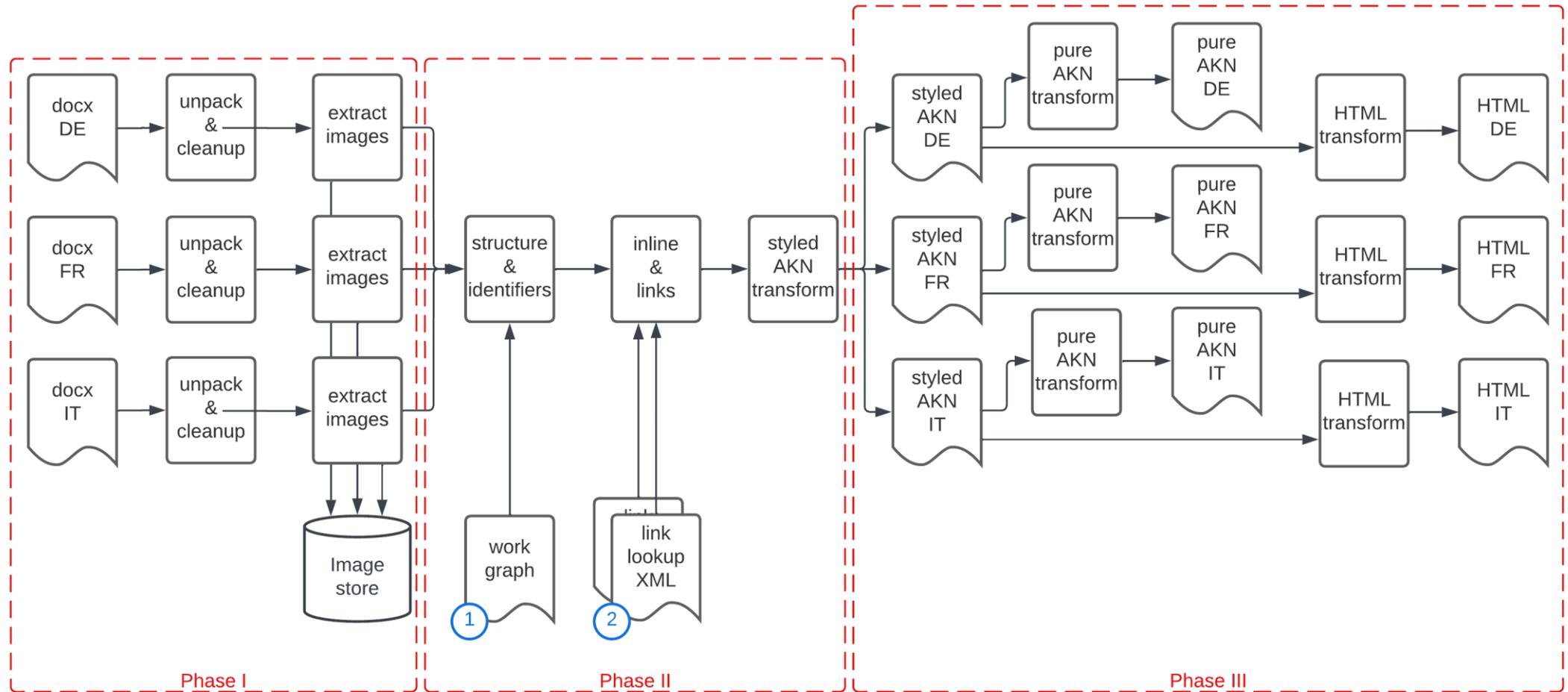
Note:

Stated this way, the “rules” can be expressed as a structural XML compare. So, the complexity becomes a convenience

- Make sure that a certain chapter in each language has the same amount of articles
- Make sure that the newly added blacklist item in a certain article has the same label in each language
- Make sure that a specific footnote references the same law in every language
- Now the fines for speeding have changed, make sure that the fines mentioned in a certain paragraph are aligned (e.g.)
- Make sure that the images are present and the same in each language
- Make sure that dates in the document are the same in each language
- ...

Complexity => convenience

Opportunity: use the existing transformer



.docx QA by XML compare

Requirements again?

- We don't want a long report
- We want to see where the errors are
- We want to have control over the amount of messages

Options to consider:

- Inject reporting messages as comments back in Word document
=> rejected by customer
- Inject reporting messages in HTML
=> the next best thing for side-by-side evaluation

- We want to add rules over time
(now: 30 and counting...)

Required:

- Rule independence

What if we would go “All-in”?

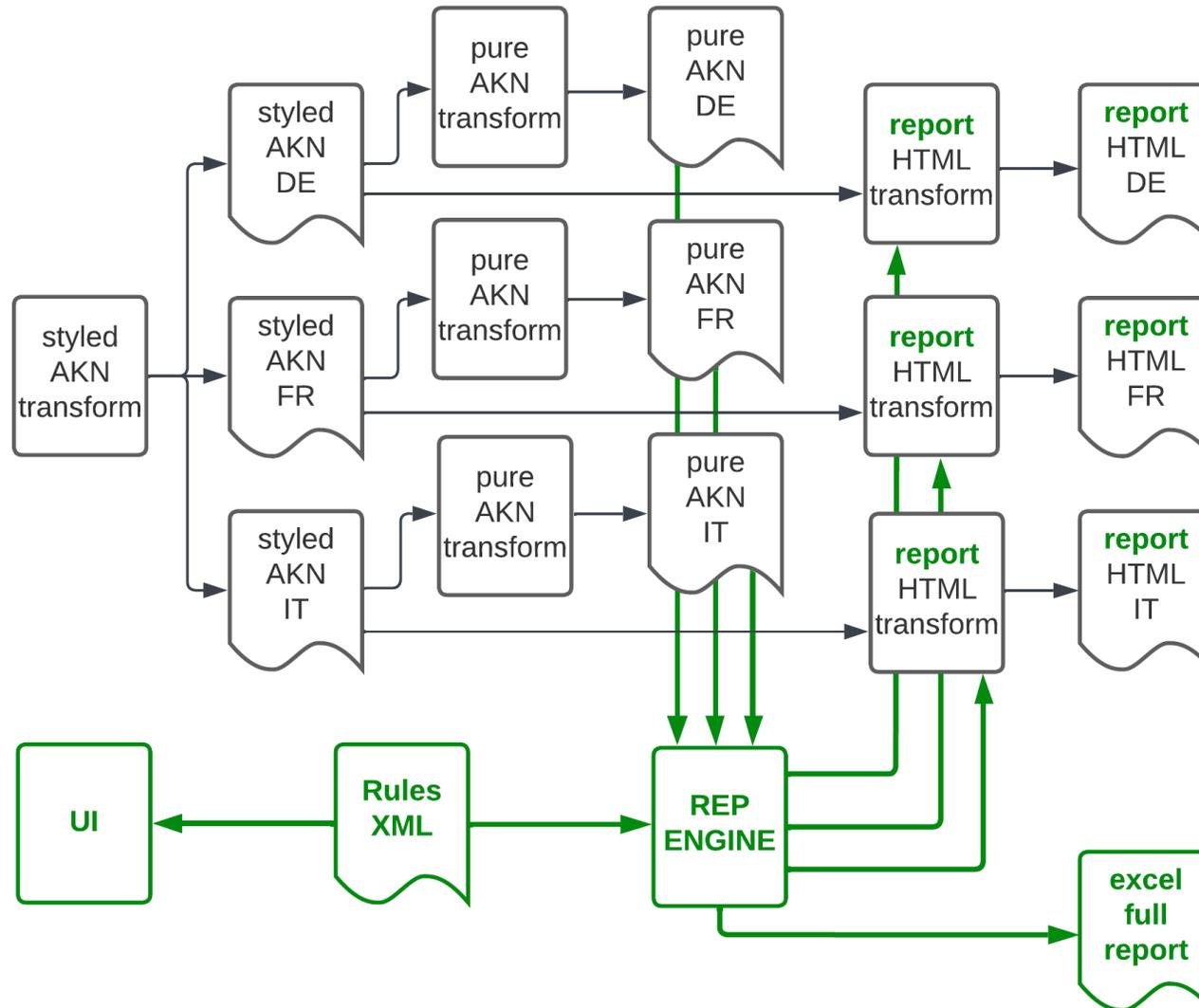
Let's build a full-blown reporting engine

- Transformer / Reporting Tool on a server
- Interface to upload documents
 - And select the rules / group of rules to check for (rules pre-filter)
- Review the transformed document
 - Side by side review with source .docx
 - Concentrate on specific issues (rules post-filter in result)
 - Navigate between reporting messages
 - Future: assign issues, mark as done, mark priority, ...
- Download cross language full report
 - MS Excel (*and sneak transformer error messages in there as well*)

From “we need beter error messages” to this...

Solution overview: engine perspective

Plug in a reporting engine



Note:

The existing transformer is based on an [Xproc 3.0](#) pipeline and as such has the flexibility for plugging in the reporter engine easily.

The “styled” Akoma Ntoso XML representation does exist. However, it is not published (but it still has the original MS Word template style information).

Moving the reporter to an earlier phase for style related rules, would be straightforward.

Splitting the reporter, depending on the rule type, would be straightforward too.

Technology options considered

<p>XML compare the different languages (eg. using Delta XML)</p>	<ul style="list-style-type: none">• Rules logic in the post filtering pipeline• Only useful really for the structure in the project's context (less so for text nodes in different languages)• Would require rule dependent preprocessing
<p>Schematron</p>	<ul style="list-style-type: none">• Rule independence ?• Some rules are extremely complex to develop (let alone maintain)• Performance
<p>Role our own</p>	<ul style="list-style-type: none">• That is what we ended up doing• Performance to be seen...

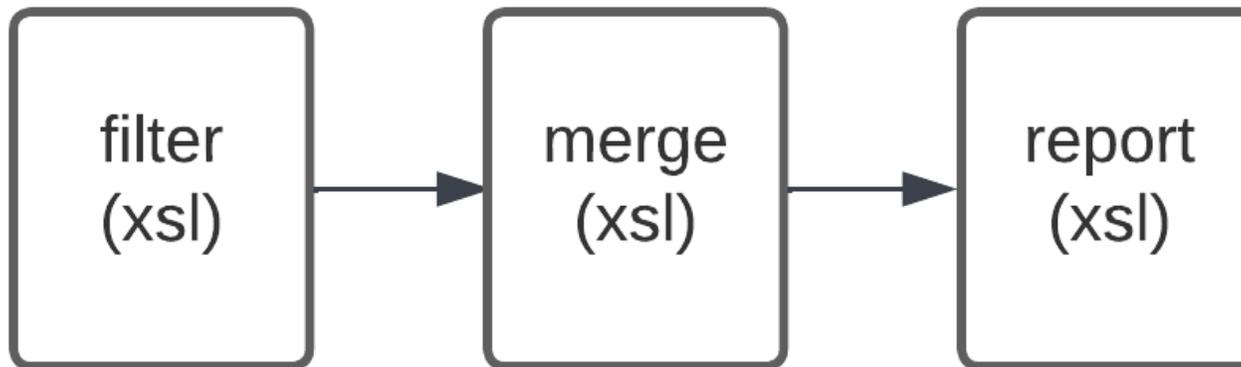
one rule to rule them all

- We saw a pattern for rule processing
- Do that for each rule
- Wrap the messages
- And inject them in the HTML

Note:

The pattern:

- **Filter** relevant nodes
- **Merge** the languages (per context)
- **Report** a node with an inconsistency

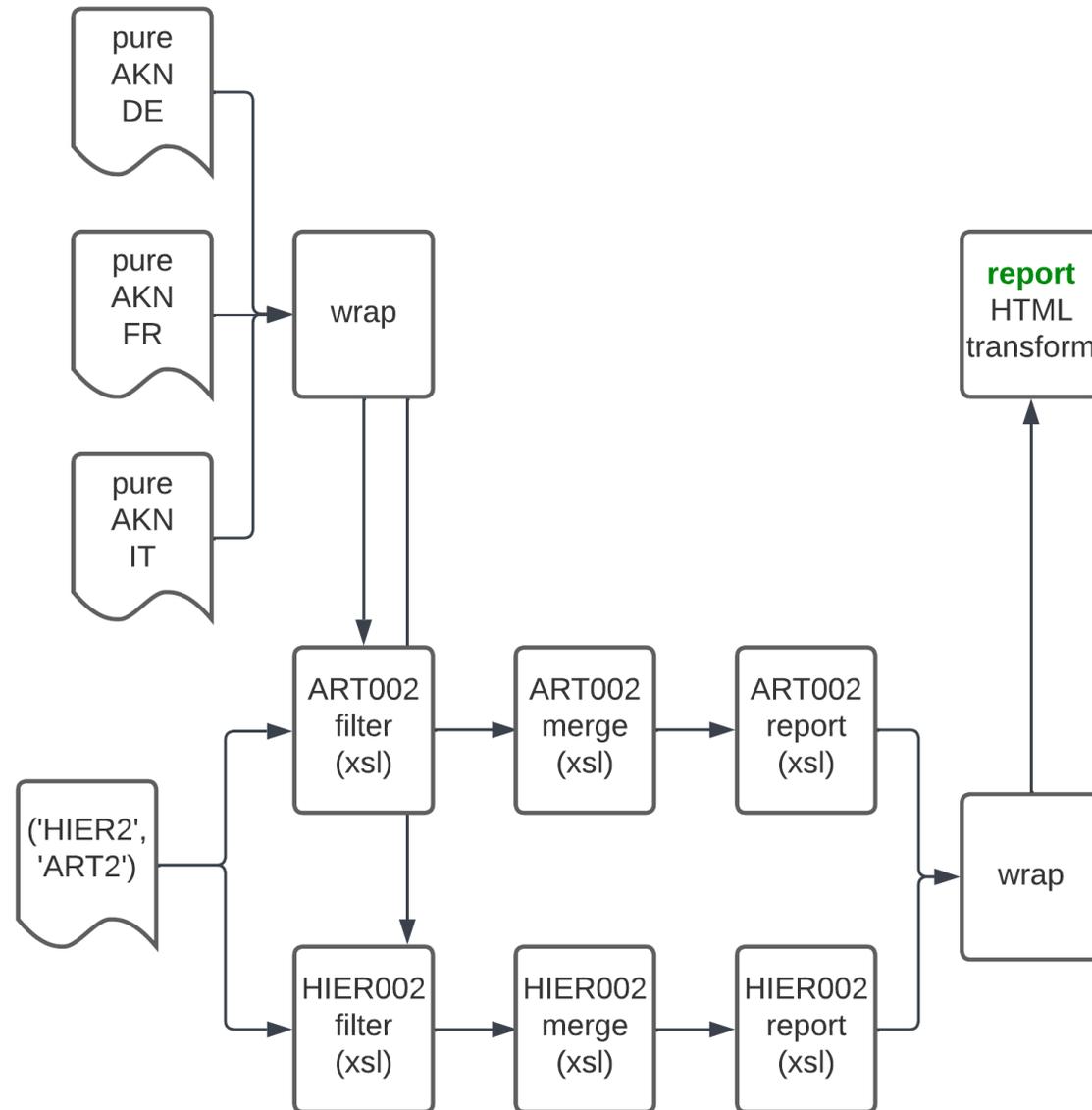


Actual engine

Note:

A p:for-each triggers a pipeline of bespoke XSLT (as per pattern), per activated (UI) rule. The source of the sub-pipeline is a wrapped document containing the different language variants of the AKN XML.

Each sub-pipeline returns a sequence of messages. All messages get wrapped and then injected in the HTML transformer for visualisation in the HTML representation... allowing the author to review side by side with the source document



Xproc 3: execute for every active rule

Note:
XProc 3.0 snippet for
illustration only

```
<p:for-each>
  <p:with-input select="$reporter-rules/rules/@enabled/tokenize(., '\s')">
    <doc/>
  </p:with-input>
  <p:variable name="rule-token" select="."/>
  <p:identity>
    <p:with-input pipe="@preproc-wrapped-akn"/>
  </p:identity>
  <cma:reporter-rule>
    <p:with-option name="rule-token" select="$rule-token"></p:with-option>
    <p:with-option name="store-path" select="$store-path"/>
    <p:with-option name="reporter-rules" select="$reporter-rules"/>
  </cma:reporter-rule>
</p:for-each>

<p:wrap-sequence wrapper="messages"/>
```

E.g. ('HIER002', 'ART002')

Context switch to wrapped AKN document

Wrap all the messages from all the sub-pipeline executions

Xproc 3: the pattern per rule

*Note:
XProc 3.0 snippet for
illustration only*

```
<p:xslt version="3.0">  
  <p:with-input port="stylesheet"  
    href="../../90-xslt/cma-w80-reporter-engine/{$rule-token}-filter.xsl"/>  
</p:xslt>
```

Rule dependent XSLT file URI

```
<p:xslt version="3.0">  
  <p:with-input port="stylesheet"  
    href="../../90-xslt/cma-w80-reporter-engine/{$rule-token}-merge.xsl"/>  
</p:xslt>
```

```
<p:xslt version="3.0">  
  <p:with-input port="stylesheet"  
    href="../../90-xslt/cma-w80-reporter-engine/{$rule-token}-report.xsl"/>  
  <p:with-option name="parameters"  
    select="map{'rules' : $reporter-rules}"/>  
</p:xslt>
```

Observations

- No conclusions yet
 - Demo in May was very well perceived (LB)
 - End-user testing in August
- Some warnings
 - Make sure messages in HTML are not obfuscated by a missing context
 - We have yet to experience the effect of “a lot” of rules (performance)
- Future
 - Combine rules for optimisation
 - Formal expression language for rules (could still be Schematron)
 - Autogenerate the XSLTs from that (we know, lot of XSLTs)

Solution overview: user perspective

Importer des documents pour générer un rapport



Méthode d'importation

↳ Insérer URI

↳ Importer document

Type de publication

Recueil systématique

Versions linguistiques

Allemand

Français

Italien

Romanche

Anglais

Français

FR fedlex-data-admin-ch-eli-cc-2015-705-20211103-fr-doc (1).docx

Allemand

DE fedlex-data-admin-ch-eli-cc-2015-705-20211103-de-doc (2).docx

Italien

IT fedlex-data-admin-ch-eli-cc-2015-705-20211103-it-doc (2).docx

Annuler

→ Suivant

Générer un rapport

12:44 Mettre à jour

✓ Réussite

✓ Réussite

✓ Réussite

✓ Réussite

✓ Réussite

✓ Réussite

Note:

- Upload a document or pass the URI of a document uploaded before
 - Indicate the type of the document
 - Indicate the language variant
- (all language variants of a document are processed as a set)

Importer des documents pour générer un rapport



Générer un rapport

Méthode d'importation

-
-

Paramètres de vérification

Sélectionner tous les paramètres

- HIER001 - HIER001
- HIER002 - HIER002



Structure des sous-articles

- ART001 - Correspondance des paragraphes
- ART002 - Éléments (liste) correspondants
- BLCK002 - BLCK002
- BLCK004 - BLCK004
- FN001 - FN001

← Précédent

Annuler

▶ Générer le rapport



Note:

- User can select all the rules that should be evaluated
- And then generate the report

Dernier rapports générés

8945 rapports

Dernière mise à jour des status: 03.08.23 à 12:49 [Mettre à jour](#)

	7e55095d-d8d2-4ee6-86ad-fc718129c59c	srikanth.v@cogni.zone	03.08.23	12:49	En cours
	f3173ecd-0c01-4d01-96e1-d5b37ddc9a85	srikanth.v@cogni.zone	03.08.23	12:42	Réussite
	https://fedlex.data.admin.ch/eli/cc/1998/3033_3033_3033/20230728 ecc3ace4-e368-40ca-9dac-12c72026ee00	user	28.07.23	16:22	Réussite
	https://fedlex.data.admin.ch/eli/cc/1998/3033_3033_3033/20250426 0a4d715a-6c84-4512-bad8-1c508e390dae	user	26.07.23	13:36	Réussite
	https://fedlex.data.admin.ch/eli/cc/2020/954/20230725 51327171-4512-bad8-1c508e390dae	user	26.07.23	00:51	Réussite



Note:
List of reports that are available for viewing (top one still running in the background)

0.672.925.81

AS 2015 4205; BBI 2014 9125

Übersetzung

Abkommen zwischen der Schweizerischen Eidgenossenschaft und der Republik Zypern zur Vermeidung der Doppelbesteuerung auf dem Gebiet der Steuern vom Einkommen und vom Vermögen

Abgeschlossen am 25. Juli 2014

Von der Bundesversammlung genehmigt am 19. Juni 2015¹

In Kraft getreten durch Notenaustausch am 15. Oktober 2015

(Stand am 3. November 2021)

¹ AS 2015 4203

Der Schweizerische Bundesrat
und

die Regierung der Republik Zypern,

vom Wunsch geleitet, ein Abkommen zur Vermeidung der Doppelbesteuerung auf dem Gebiet der Steuern vom Einkommen und vom Vermögen abzuschliessen,

vom Wunsch geleitet, ihre wirtschaftlichen Beziehungen weiterzuentwickeln und die Zusammenarbeit in steuerlichen Angelegenheiten zu vertiefen,

in der Absicht, in Bezug auf die Steuern vom Einkommen und vom Vermögen eine Doppelbesteuerung zu beseitigen, ohne Möglichkeiten zur Nichtbesteuerung oder reduzierten Besteuerung durch Steuerhinterziehung oder -umgehung (unter anderem durch missbräuchliche Gestaltungen mit dem Ziel des Erhalts von in diesem Abkommen vorgesehenen Erleichterungen zum mittelbaren Nutzen von in Drittstaaten ansässigen Personen) zu schaffen.

*Note:
It seems the
German variant
has no issues
that need
reporting*



HIER002 Component avec identifiant "art_VI" pas présent

Art. VI, par. 2 et 3, du Prot. du 20 juil. 2020⁹

⁹ RO 2021 810, 809; FF 2020 9091

2. Le présent protocole entrera en vigueur à la date de réception de la seconde de ces notifications et les dispositions du présent protocole de modification seront applicables:

- a) s'agissant des impôts retenus à la source sur les revenus mis en paiement ou attribués le 1^{er} janvier de l'année civile suivant l'entrée en vigueur du présent protocole, ou après cette date;
- b) s'agissant des autres impôts, aux années fiscales qui débutent le 1^{er} janvier de l'année civile suivant l'entrée en vigueur du présent protocole, ou après cette date.

3. Nonobstant les dispositions du par. 2, les modifications prévues aux art. II, III et IV du présent protocole prennent effet dès l'entrée en vigueur du présent protocole, sans égard à la période fiscale à laquelle le cas se rapporte.

Protocole

à la Convention entre la Confédération suisse et la République de Chypre en vue d'éviter les doubles impositions en matière d'impôts sur le revenu et sur la fortune

ART002 Élément (liste)
"lvl_u2/list_u1/lbl_Es_besteht_Einvernehmen_dar_ber_dass_das_Kriterium_der_Mindesthaltedauer_nach_Artikel_10_Absatz_3_Buchstabe_a_erf_ilt_ist_auch_wenn_die_Haltedauer_nach_dem_Datum_der_Zahlung_der_Dividende_erf_ilt_wird" pas trouvé

ART002 Élément (liste) "lvl_u2/list_u1/lbl_a" pas trouvé

ART002 Élément (liste) "lvl_u2/list_u1/lbl_b" pas trouvé

ART002 Élément (liste) "lvl_u2/list_u3/listintro" pas trouvé

ART002 Élément (liste) "lvl_u2/list_u4/lbl_c" pas trouvé

Note:
*First message is in focus (better visibility)
The messages are added to the HTML in context (so the error is reported where it happens)
The messages are available in multiple languages.
Configuration in this example is set to visualizing the French language message*

préavis de dénonciation a été donné;

b) aux autres impôts pour toute année fiscale commençant le 1^{er} janvier de l'année civile suivant celle au cours de laquelle le préavis de dénonciation a été donné, ou après cette date.

Signatures

En foi de quoi, les soussignés, dûment autorisés, ont signé la présente Convention.

Fait en deux exemplaires à Nicosie, le 25 juillet 2014, en langues française, grecque et anglaise, chaque texte faisant également foi. En cas d'interprétation divergente entre le texte français et le texte grec, le texte anglais fera foi.

Pour le Conseil fédéral suisse: Gabriela Nützi Sulpizio	Pour le Gouvernement de la République de Chypre: Harris Georgiades
---	--

HIER002 Component avec identifiant "art_VI" pas présent

Art. VI, par. 2 et 3, du Prot. du 20 juil. 2020⁹

⁹ RO 2021 810, 809; FF 2020 9091

2. Le présent protocole entrera en vigueur à la date de réception de la seconde de ces notifications et les dispositions du présent protocole de modification seront applicables:

a) s'agissant des impôts retenus à la source sur les revenus mis en paiement ou attribués le 1^{er} janvier de l'année civile suivant l'entrée en vigueur du présent protocole, ou après cette date;

b) s'agissant des autres impôts, aux années fiscales qui débutent le 1^{er} janvier de l'année civile suivant l'entrée en vigueur du présent protocole, ou après cette date.

3. Nonobstant les dispositions du par. 2, les modifications prévues aux art. II, III et IV du présent protocole prennent effet à l'entrée en vigueur du présent protocole, sans égard à la période fiscale à laquelle le cas se rapporte.

*Note:
There seems to be only one message on rule "HIER002" as selected on the bottom right*

- a)
aux impôts retenus à la source sur les montants mis en paiement ou attribués à partir du 1^{er} janvier de l'année civile suivant celle au cours de laquelle le préavis de dénonciation a été donné;
- b)
aux autres impôts pour toute année fiscale commençant le 1^{er} janvier de l'année civile suivant celle au cours de laquelle le préavis de dénonciation a été donné, ou après cette date.

Signatures

En foi de quoi, les soussignés, dûment autorisés, ont signé la présente Convention.

Fait en deux exemplaires à Nicosie, le 25 juillet 2014, en langues française, grecque et anglaise, chaque texte faisant également foi. En cas d'interprétation divergente entre le texte français et le texte grec, le texte anglais fera foi.

Pour le Conseil fédéral suisse:	Pour le Gouvernement de la République de Chypre:
Gabriela Nützi Sulpizio	Harris Georgiades

HIER002 Component avec identifiant "art_VI" pas présent

Art. VI, par. 2 et 3, du Prot. du 20 juil. 2020⁹

⁹ RO 2021 810, 809; FF 2020 9091

2. Le présent protocole entrera en vigueur à la date de réception de la seconde de ces notifications et les dispositions du présent protocole de modification seront applicables:

- a)
s'agissant des impôts retenus à la source sur les revenus mis en paiement ou attribués le 1^{er} janvier de l'année civile suivant l'entrée en vigueur du présent protocole, ou après cette date;
- b)
s'agissant des autres impôts, aux années fiscales qui débutent le 1^{er} janvier de l'année civile suivant l'entrée en vigueur du présent protocole, ou après cette date.

Note:
Navigation at the bottom:

- *Left: navigate between messages in the document*
- *Middle: select document language*
- *Right: select which messages to view (rule based)*

Top: download an MS Excel report collecting all messages from all languages + the transformation warnings report

Thank you

Some references	
Srikanth	srikanth.v@cogni.zone
Geert	geert@c-moria.com
Cognizone BV	https://www.cogni.zone/
C-Moria BV	https://c-moria.com/
Fedlex Publication platform (consolidations)	https://www.fedlex.admin.ch/de/cc
Xproc 3.0 specification	https://xproc.org/specifications.html
Morgana Xproc III (Xproc 3.0 Processor)	https://www.xml-project.com/morganaxproc-iiise.html
Markup UK Xproc 3.0 tutorial by Erik Siegel	https://mu-2023-xproc.xatapult.com/
Sample projects (in the making...)	https://github.com/Gertone/XProcLS