



EAP XML Legislator

Tworzenie aktów prawnych w postaci tekstów strukturalnych w XML przy użyciu Edytora XML Legislator

ABC PRO Sp. z o.o.

NAJCZĘŚCIEJ POPEŁNIANE BŁĘDY – JAK ICH UNIKAĆ

Data: 2017-06-13

Autor:

Piotr Jegorow



Spis treści

WPROWADZENIE	3
Formatowanie dokumentu przed wklejeniem do Legislatora	4
Przeklekanie tekstu do NOTATNIKA – jako „wyjście awaryjne”	4
Przygotowanie dokumentu do „wklejenia” do Edytora XML:	5
Informacje podstawowe	5
Używanie „twardych spacji”	7
Konsekwencje używania ENTER w obrębie jednostki redakcyjnej	8
Podsumowanie	9



WPROWADZENIE

Edytor XML Legislador umożliwia tworzenie aktów prawnych w postaci tekstów strukturalnych w formacie XML:

- **poprzez napisanie aktu prawnego bezpośrednio w Edytorze XML** (pisze się równie łatwo i wygodnie jak w popularnych edytorach biurowych);
- **poprzez „przeklejenie” do Edytora XML treści aktu z innego edytora biurowego** (np. MS Word) – CTR C / CTR V.
 - wytworzony w ten sposób akt prawny jest zgodny z obowiązującymi wymogami prawnymi w tym zakresie – po jego podpisaniu (podpisem elektronicznym może zostać przekazany do wojewody celem ogłoszenia w wojewódzkim dzienniku urzędowym (lub w każdym innym dowolnym dzienniku).

W tej części „poradnika praktycznego” zajmiemy się właśnie „przeklejeniem” treści aktów do Edytora XML Legislador i uzyskiwanie aktów prawnych w postaci tekstów strukturalnych przy użyciu tej właśnie metody.

Sposób realizacji jest bardzo wygodny i prosty. Wystarczy skopiować całą treść aktu z edytora biurowego (CTR-V) i wkleić go do Edytora XML Legislador (CTR-V) – poza tekstem, kopiują się w ten sposób również tabele, rysunki.. Z jednym wszak zastrzeżeniem – kopiowany tekst aktu prawnego powinien być napisany poprawnie, w szczególności chodzi o zachowanie wymogów wynikających z „Zasad techniki prawodawczej”. Jeżeli tak nie jest, wówczas tekst może zostać rozpoznany w sposób inny od oczekiwanego – w konsekwencji zostanie nieprawidłowo „przetłumaczony” na tekst strukturalny w XML.

Bardzo ważne:

Tekst „wklejony” do Edytora Legislador, jest odpowiednio analizowany i zapisywany w postaci tekstu strukturalnego XML. Formatowanie dokumentu odbywa się w oparciu o strukturę XML, z uwzględnieniem przyjętych zasad (domyślnie jest to formatowanie zgodne z zasadami techniki prawodawczej, ale można wprowadzić także własne ustawienia).

W efekcie, wszystkie przekonwertowane do XML akty prawne mogą mieć identyczne formatowanie, i co najważniejsze – automatycznie są formatowane w sposób zgodny z zasadami techniki prawodawczej – bez względu na to w jaki sposób zostały napisane w edytorze biurowym. Analogiczny efekt uzyskujemy oczywiście w sytuacji, gdy akt prawny jest od razu pisany w Edytorze XML – zawsze będzie prawidłowo sformatowany.

Tak więc, stosując powszechnie (np. w odniesieniu do wszystkich tworzonych aktów) Edytor XML Legislador, możemy w łatwy sposób zapewnić standaryzację tworzonych aktów, z zachowaniem wymogów wynikających z zasad techniki prawodawczej.

Poniżej zawarto kilka wskazówek praktycznych, które mają na celu wsparcie użytkowników w zakresie prawidłowego przygotowywania tekstów napisanych w popularnych edytorach biurowych do automatycznego przetworzenia na XMLe.

Prosimy o zapoznanie się z tymi wskazówkami.

Formatowanie dokumentu przed wklejeniem do Legislatora

Przeklejanie tekstu do NOTATNIKA – jako „wyjście awaryjne”

Standardowym działaniem jest bezpośrednie „przeklejenie” tekstu z Edytora biurowego do Edytora XML Legislators. Jeżeli tekst jest przygotowany prawidłowo, bez użycia różnego rodzaju szczególnych formatowań, stylów, podziałów sekcji, etc., wówczas efekt powinien być zgodny z oczekiwanym – tekst powinien zostać poprawnie przekonwertowany na tekst strukturalny w formacie XML.

Co jednak zrobić, jeżeli widzimy, że nasz „źródłowy” dokument, np. napisany w MS WORD, nie konwertuje się poprawnie? W takim przypadku, skutecznym rozwiązaniem może okazać się przekopiowanie go najpierw do „notatnika” (aplikacja systemu Windows – dostępna na każdym komputerze) i dopiero z „notatnika” przekopiowanie do Edytora XML Legislators. Dlaczego właśnie tak? Chodzi o to, że po skopiowaniu tekstu z Edytora biurowego do „notatnika”, jest on zapisany jako „czysty tekst” – bez formatowań, stylów, podziałów sekcji, etc., czyli wszystkiego co w konsekwencji może utrudniać prawidłową konwersję na XML. Oczywiście, można próbować poprawiać tekst w samym edytorze biurowym, ale z doświadczenia wynika, że znacznie prościej i szybciej jest po prostu przekleić tekst do notatnika i następnie z „notatnika” do Edytora XML.

Uwaga:

Takie działanie możemy zastosować jedynie w odniesieniu do tekstu, ale nie do tabel czy obrazków. Gdy do „notatnika” przekleimy tabelę, to się „rozsypie” – pozostanie jedynie tekst, ale bez komórek. Co zatem robić w takiej sytuacji, gdy mamy „trudny” tekst zawierający także tabele i obrazy? Do notatnika należy skopiować jedynie tekst i następnie „wkleić” go do Edytora XML. Natomiast tabele i obrazy należy skopiować i wkleić z edytora biurowego bezpośrednio we właściwe miejsce do Edytora XML Legislators.

Przygotowanie dokumentu do „wklejenia” do Edytora XML:

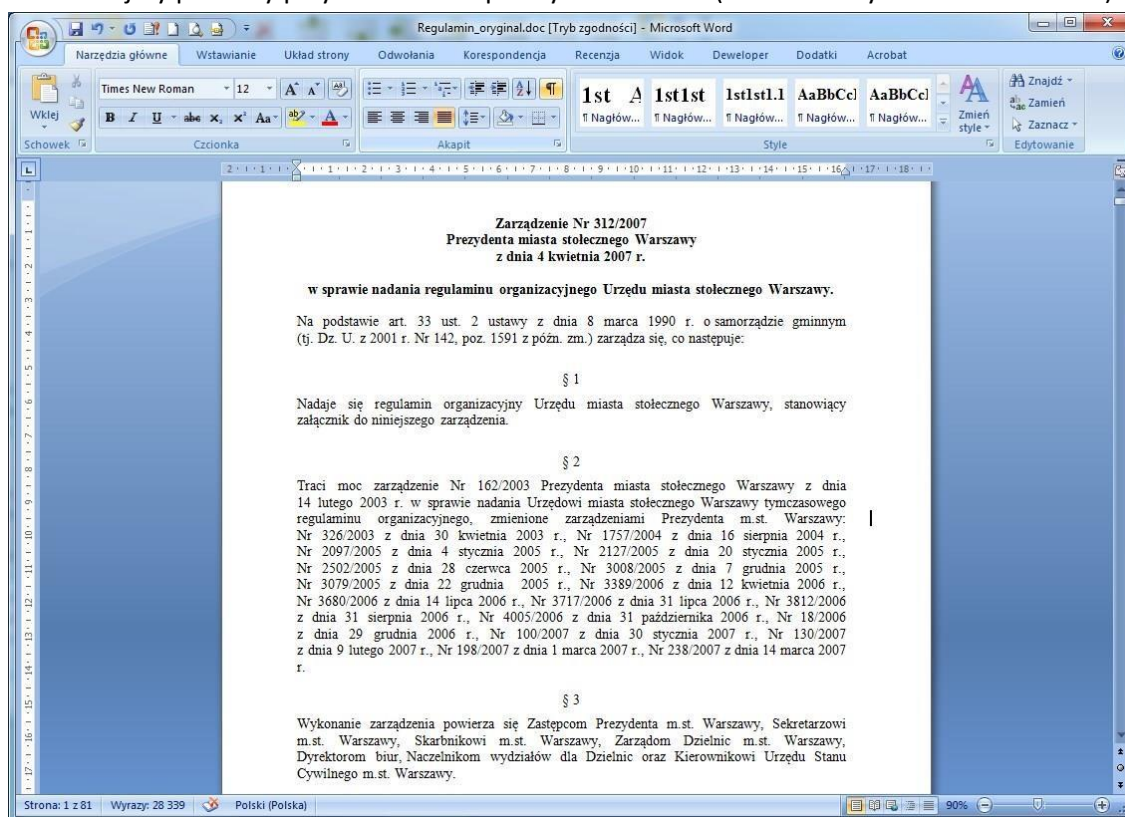
Informacje podstawowe

Tworząc dokument przy pomocy programu WORD / Writer, zazwyczaj formatujemy go:

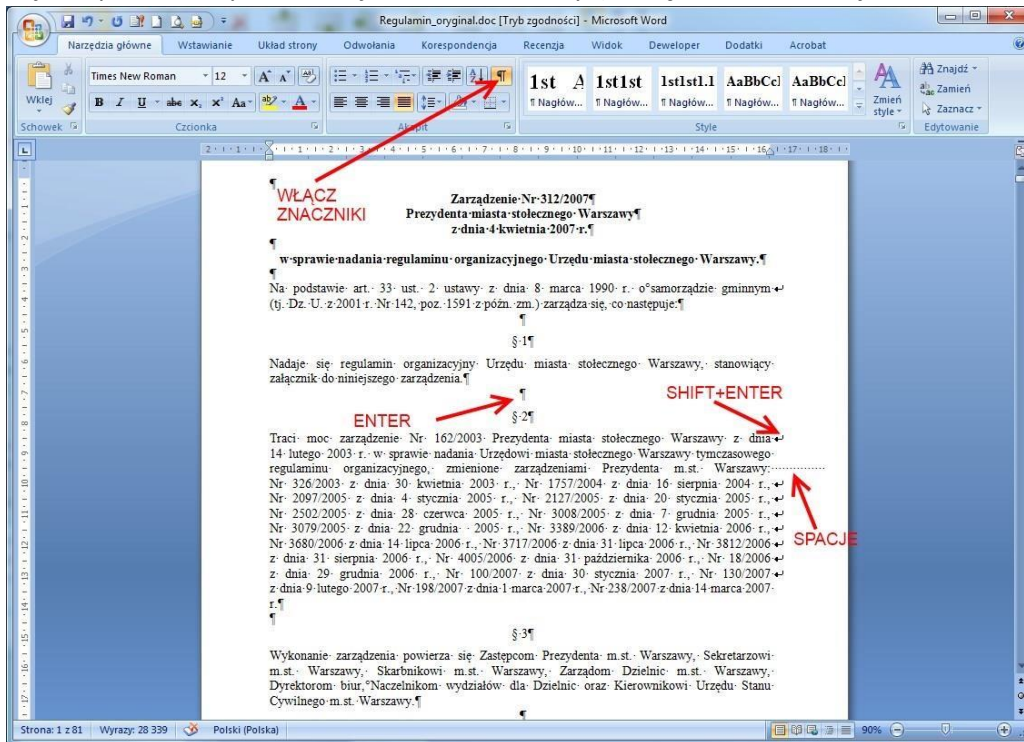
- ręcznie przenosząc zaimki do nowej linii poprzez skrót klawiszowy **Shift+Enter**,
- ręcznie przenosząc zaimki do nowej linii poprzez wstawienie kilku / kilkunastu spacji aż żądany wyraz „przeskoczy” do nowej linii,
- używając **Enterów** zrobienia odstępów pomiędzy wybranymi jednostkami redakcyjnymi,

Wszystkie te sposoby powodują, że wprowadzony tekst wygląda bardzo dobrze, równo, i możemy go wydrukować. Jednakże każda zmiana czcionki czy dodanie wyrazu może spowodować konieczność ponownego formatowania tekstu.

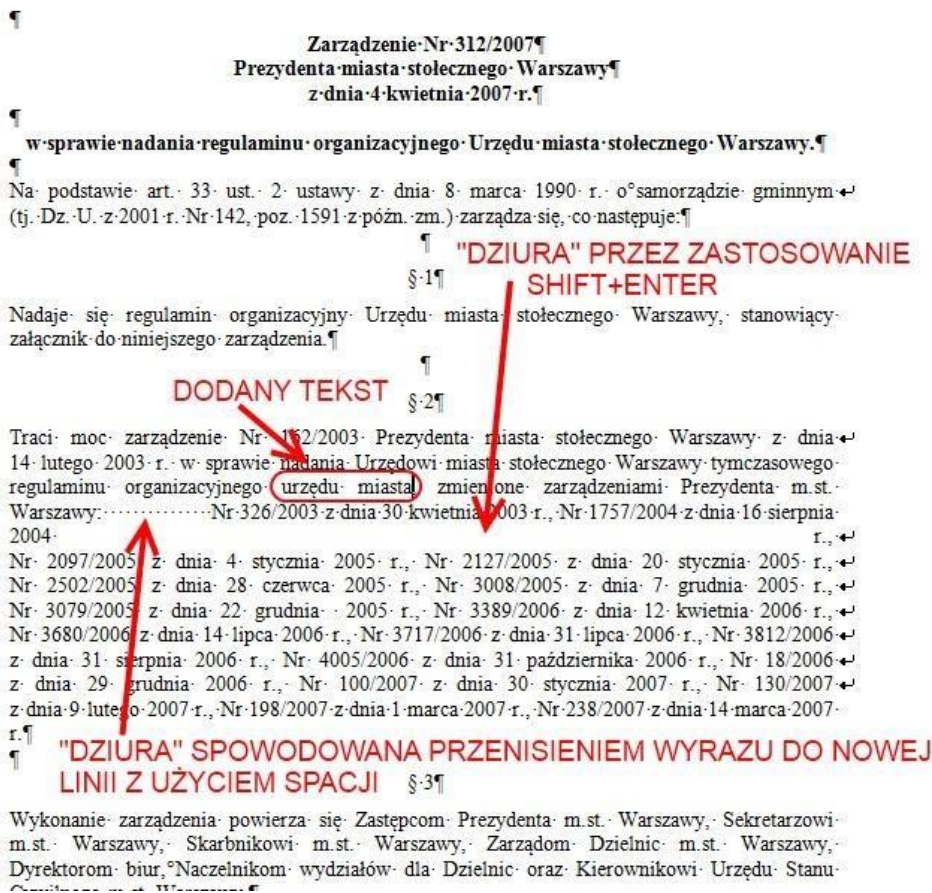
Przeanalizujemy poniższy przykład – akt napisany w MS Word (standardowy widok w MS Word):



Na pierwszy rzut oka wszystko wydaje się w porządku. Włączmy jednak wyświetlanie znaczników, aby zobaczyć w jaki sposób faktycznie akt jest sformatowany (co będzie miało swoje konsekwencje przy konwersji na XML):



Spróbujmy teraz dodać jakiś wyraz w trzeciej linijce paragrafu drugiego:



Jak widzimy, dodaliśmy tekst „urzędu miasta”. Ale w efekcie „rozsyłał” się nam cały układ tekstu w par.2:

- **wyrażenie „m.st. Warszawy” zostało rozdzielone** – „m.st.” pozostało w trzeciej linii, natomiast „Warszawy” zostało przeniesione do czwartej linii (zapewne wyrażenie „m.st.Warszawy” chcielibyśmy mieć zawsze zapisane w jednej linii);
- **w czwartej linii powstała nam „dziura”** – jest to spowodowane przeniesieniem wyrazu „Warszawy” do nowej linii, przy czym po tym wyrazie były wstawione „spacje” (użyte celem „poprawienia” wyglądu „oryginalnego” dokumentu);
- **w piątej linii powstała nam kolejna „dziura”** - w wyniku zastosowania określonego formatowania (efekt polecenia SHIFT+ENTER).

Tak więc, musimy sobie zdawać sprawę z tego, że używanie różnego rodzaju formatowań, zawsze będzie rodziło określone konsekwencje. Dokument „oryginalny” będzie co prawda dzięki temu wyglądał dobrze, ale jeżeli zechcemy go zmodyfikować (dodać/usunąć wyrazy, zastosować inną wielkość lub krój czcionki, inne rozmiary marginesów lub wcięć akapitów, etc.), to dokument najprawdopodobniej się nam „rozsyple” – zarówno w „Wordzie” jak i tym bardziej w XML-u.

Jak sobie zatem radzić w takich sytuacjach? – poniżej kilka przykładów.

Używanie „twardych spacji”

Poniżej ten sam tekst, ale utworzony przy użyciu „twardych spacji”¹

Zarządzenie Nr 312/2007
Prezydenta miasta stołecznego Warszawy
z dnia 4 kwietnia 2007 r.

w sprawie nadania regulaminu organizacyjnego Urzędu miasta stołecznego Warszawy.

Na podstawie art. 33 ust. 2 ustawy z dnia 8 marca 1990 r. o samorządzie gminnym (tj. Dz. U. z 2001 r. Nr 142, poz. 1591 z późn. zm.) zarządza się, co następuje:

§ 1

Nadaje się regulamin organizacyjny Urzędu miasta stołecznego Warszawy, stanowiący załącznik do niniejszego zarządzenia.

§ 2

Traci moc zarządzenie Nr 162/2003 Prezydenta miasta stołecznego Warszawy z dnia 14 lutego 2003 r. w sprawie nadania Urzędowi miasta stołecznego Warszawy tymczasowego regulaminu organizacyjnego urzędu miasta, zmienione zarządzeniami Prezydenta m.st. Warszawy: Nr 326/2003 z dnia 30 kwietnia 2003 r., Nr 1757/2004 z dnia 16 sierpnia 2004 r., Nr 2097/2005 z dnia 4 stycznia 2005 r., Nr 2127/2005 z dnia 20 stycznia 2005 r., Nr 2512/2005 z dnia 28 czerwca 2005 r., Nr 3008/2005 z dnia 7 grudnia 2005 r., Nr 3079/2005 z dnia 22 grudnia 2005 r., Nr 3389/2006 z dnia 12 kwietnia 2006 r., Nr 3680/2006 z dnia 14 lipca 2006 r., Nr 3717/2006 z dnia 31 lipca 2006 r., Nr 3812/2006 z dnia 31 sierpnia 2006 r., Nr 4005/2006 z dnia 31 października 2006 r., Nr 18/2006 z dnia 29 grudnia 2006 r., Nr 100/2007 z dnia 30 stycznia 2007 r., Nr 130/2007 z dnia 9 lutego 2007 r., Nr 198/2007 z dnia 1 marca 2007 r., Nr 238/2007 z dnia 14 marca 2007 r.

§ 3

Wykonanie zarządzenia powierza się Zastępcom Prezydenta m.st. Warszawy, Sekretarzowi m.st. Warszawy, Skarbnikowi m.st. Warszawy, Zarządom Dzielnic m.st. Warszawy, Dyrektorom biur, Naczelnikom wydziałów dla Dzielnic oraz Kierownikowi Urzędu Stanu Cywilnego m.st. Warszawy.

¹ Tzw. spacja nierozdzielająca (Ctrl+Shift+Spacja) – wstawienie jej pomiędzy wyrazami spowoduje, że dwa (lub więcej) wyrazy będą zawsze wyświetlane w tej samej linii. Dzięki temu zmiana formatowania czy dodanie nowych wyrazów nie spowoduje problemów z tekstem

Dzięki wprowadzeniu „twardych spacji” pomiędzy wyrazami, które mają występować zawsze razem w jednej linii, jakakolwiek zmiana w tekście nie spowoduje ich rozdzielenia.

Identycznie sprawa wygląda z XMLem.

Jeżeli kopiowany tekst zamiast ustawień akapitów i twardej spacji, będzie zawierał znaki nowej linii i formatowanie uzyskiwane poprzez dodanie kilku znaków ENTER, jego interpretacja w XML może nie być do końca prawidłowa.

To z kolei może powodować kolejne problemy, w szczególności, gdy będziemy chcieli dokonywać nowelizacji takiego tekstu i automatycznie uzyskiwać tzw. „tekst ujednolicony” oraz „tekst porównawczy”. Dlatego też, należy zwracać baczną uwagę na to w jaki sposób jest sformatowany tekst, który zamierzamy „przekonwertować” na XML przy użyciu Edytora XML Legislators.

Dlatego też, czasami najprostszym i jednocześnie skutecznym rozwiązaniem może się okazać uprzednie skopiowanie tekstu do „notatnika” (automatycznie wszelkie formatowania zostaną usunięte) i dopiero z „notatnika” wklejenie do Edytora XML (jest to opisane na początku niniejszego opracowania).

Konsekwencje używania ENTER w obrębie jednostki redakcyjnej

Przykładowo przeklejając następującą zawartość:

Zarządzenie Nr 312/2007
Prezydenta miasta stołecznego Warszawy
z dnia 4 kwietnia 2007 r.

w sprawie nadania regulaminu organizacyjnego Urzędu miasta stołecznego Warszawy.

Na podstawie art. 33 ust. 2 ustawy z dnia 8 marca 1990 r. o samorządzie gminnym (tj. Dz. U. z 2001 r. Nr 142, poz. 1591 z późn. zm.) zarządza się, co następuje:

§ 1

Nadaje się regulamin organizacyjny Urzędu miasta stołecznego Warszawy, stanowiący załącznik do niniejszego zarządzenia.

§ 2

Traci moc zarządzenie Nr 162/2003 Prezydenta miasta stołecznego Warszawy z dnia 14 lutego 2003 r. w sprawie nadania Urzędowi miasta stołecznego Warszawy tymczasowego regulaminu organizacyjnego urzędu miasta, zmienione zarządzeniami Prezydenta m.st. Warszawy: Nr 326/2003 z dnia 30 kwietnia 2003 r., Nr 1757/2004 z dnia 16 sierpnia 2004 r., Nr 2097/2005 z dnia 4 stycznia 2005 r., Nr 2127/2005 z dnia 20 stycznia 2005 r., Nr 2502/2005 z dnia 28 czerwca 2005 r., Nr 3008/2005 z dnia 7 grudnia 2005 r., Nr 3079/2005 z dnia 22 grudnia 2005 r., Nr 3389/2006 z dnia 12 kwietnia 2006 r., Nr 3680/2006 z dnia 14 lipca 2006 r., Nr 3717/2006 z dnia 31 lipca 2006 r., Nr 3812/2006 z dnia 31 sierpnia 2006 r., Nr 4005/2006 z dnia 31 października 2006 r., Nr 18/2006 z dnia 29 grudnia 2006 r., Nr 100/2007 z dnia 30 stycznia 2007 r., Nr 130/2007 z dnia 9 lutego 2007 r., Nr 198/2007 z dnia 1 marca 2007 r., Nr 238/2007 z dnia 14 marca 2007 r.

§ 3

Wykonanie zarządzenia powierza się Zastępcom Prezydenta m.st. Warszawy, Sekretarzowi m.st. Warszawy, Skarbnikowi m.st. Warszawy, Zarządom Dzielnic m.st. Warszawy, Dyrektorom biur, Naczelnikom wydziałów dla Dzielnic oraz Kierownikowi Urzędu Stanu Cywilnego m.st. Warszawy.

ENTER - W XML INTERPRETOWANY
JAKO NOWY ELEMENT AKAPIT

Zasada jest następująca: treść jednej jednostki redakcyjnej (paragrafu, ustępu, punktu, litery) piszemy jako „tekst ciągły”.

Jeżeli w celu „poprawienia wyglądu” złamiemy tę zasadę i wprowadzamy znaki „nowej linii” (poprzez użycie klawisza ENTER), będzie to miało swoje odzwierciedlenie również w utworzonym w wyniku konwersji pliku XML.

W szczególności będzie to miało swoje konsekwencje przy tworzeniu aktów nowelizujących. Chodzi o to, że poprzez wstawienie ENTER, cała treść paragrafu 2 automatycznie została podzielona na dwa akapity, a tak przecież nie jest (merytorycznie).

Owszem, można używać ENTER – ale czynimy to tylko i wyłącznie wówczas gdy rzeczywiście chcemy wyodrębnić poszczególne bloki tekstu w osobne „akapity”, ale nigdy w celu „poprawienia wyglądu”.

Podsumowanie

Tworzenie aktów prawnych zgodnie z obowiązującymi zasadami techniki prawodawczej to nie tylko wyraz „umiejętności prawniczych” i „dobrej praktyki”, ale przede wszystkim konieczność i wymóg, który należy spełnić. Jeżeli akt w postaci elektronicznej nie będzie sporządzony prawidłowo, to może się okazać, że nie zostanie ogłoszony w dzienniku i tym samym nie wejdzie w życie (np. plan zagospodarowania przestrzennego).

Opracowany przez naszą firmę Edytor XML Legislador posiada szereg funkcjonalności wspierających prawidłowe tworzenie aktów prawnych – m.in. zostały zaimplementowane logiki pozwalające chociażby na automatyczne rozpoznawanie akapitów i konwertowanie ich zawartości na treść danej jednostki redakcyjnej, i szereg innych. Niemniej jednak, jeżeli akt prawny, w swojej „źródłowej” postaci jest tworzony w sposób nieprawidłowy, z rażącym naruszeniem „zasad techniki prawodawczej” wówczas nie można liczyć na uzyskanie pozytywnego efektu końcowego. W rezultacie może skutkować to odmową ogłoszenia aktu w dzienniku urzędowym, wraz ze wszystkimi tego negatywnymi konsekwencjami.