

XML i nowoczesne metody zarządzania treścią

Wykład 14: Studium przypadku: System SET

Władysław Baksza, Maciej Ogrodniczuk

MIMUW, 14 stycznia 2010

Przed rokiem 2005:

- system zarządzający metadanymi ustaw i projektów ustaw,
- dokumenty przechowywane w systemie plików w postaci plików MS Worda,
- ręczne wersjonowanie dokumentów,
- ręczne ujednolicanie tekstów.

Wyodrębnienie części procesowej i redakcyjnej:

- 1 moduł serwera obsługujący procesy legislacyjne:
 - rejestracja nowych tekstów,
 - kontrola nad stanem prac,
 - obsługa wariantów projektów nowelizujących,
 - uruchamianie akcji edycji/podglądu/ujednolicania,
 - ...
- 2 środowisko edycyjne (XMetaL):
 - zapis kompletu informacji w tekstach:
 - adresy publikacji,
 - daty: uchwalenia, wejścia w życie, obowiązywania zestawów przepisów,
 - powiązania nowelizacyjne (przywołania),
 - znaczniki definicji,
 - dostosowania edytora XML-owego (CSS/makra/skróty klawiaturowe),
 - integracja z repozytorium z użyciem protokołu WebDAV.

USTAWA

z dnia 30 czerwca 2000 r.

Prawo własności przemysłowej

Tytuł I

Przepisy ogólne

Art. 1.

1. Ustawa normuje:

- 1) stosunki w zakresie wynalazków, wzorów użytkowych, wzorów przemysłowych, znaków towarowych, oznaczeń geograficznych i topografii układów scalonych;
- 2) zasady, na jakich przedsiębiorcy mogą przyjmować projekty racjonalizatorskie i wynagradzać ich twórców;
- 3) zadania i organizację Urzędu Patentowego Rzeczypospolitej Polskiej, zwanego dalej „Urzędem Patentowym”.

2. Przepisy ustawy nie uchybiają ochronie przedmiotów, o których mowa w ust. 1 pkt 1, przewidzianej w innych ustawach.

Art. 2.

Zwalczanie nieuczciwej konkurencji reguluje odrębna ustawa.

Art. 3.

1. Ilekroć w ustawie jest mowa o:

- 1) osobie – rozumie się przez to osobę fizyczną lub prawną;
- 2) osobie zagranicznej – rozumie się przez to osobę niemającą obywatelstwa polskiego i odpowiednio miejsca zamieszkania albo siedziby bądź poważnego przedsiębiorstwa na obszarze Rzeczypospolitej Polskiej;

USTAWA
z dnia 7 grudnia 2006 r.

**o zmianie ustawy o ujawnianiu informacji o dokumentach organów
bezpieczeństwa państwa z lat 1944-1990 oraz treści tych dokumentów**

Art. 1.

W ustawie z dnia 18 października 2006 r. o ujawnianiu informacji o dokumentach organów bezpieczeństwa państwa z lat 1944-1990 oraz treści tych dokumentów (Dz.U. Nr 218, poz. 1592) wprowadza się następujące zmiany:

1) art. 1 otrzymuje brzmienie:

„Art. 1. Ustawa określa zasady i tryb ujawniania informacji o dokumentach organów bezpieczeństwa państwa oraz treści tych dokumentów, znajdujących się w archiwum Instytutu Pamięci Narodowej – Komisji Ścigania Zbrodni przeciwko Narodowi Polskiemu, zwanego dalej „Instytutem Pamięci Narodowej” oraz tryb składania oraz oceny zgodności z prawdą oświadczeń dotyczących pracy lub służby w organach bezpieczeństwa państwa lub współpracy z tymi organami w okresie od dnia 22 lipca 1944 r. do dnia 31 lipca 1990 r.”;

2) po art. 3 dodaje się art. 3a w brzmieniu:

„Art. 3a. 1. Współpracą w rozumieniu ustawy jest świadoma i tajna współpraca z ogniwami operacyjnymi lub śledczymi organów bezpieczeństwa państwa w charakterze tajnego informatora lub pomocnika przy operacyjnym zdobywaniu informacji.

2. Współpracą w rozumieniu ustawy jest również działanie, którego obowiązek wynikał z ustawy obowiązującej w czasie tego działania w związku z pełnioną funkcją, zajmowanym stanowiskiem, wykonywaną pracą lub pełnioną służbą, jeżeli informacje przekazywane organom bezpieczeństwa państwa stwarzały zagrożenie dla wolności i praw człowieka i

The screenshot shows the XMetaL Author editor window. The title bar reads "XMetaL Author - [http://localhost:8080/ksrp-webdav/Workspace/Ref_ed1d8310a0002f00b09321e219e30a.xml]". The menu bar includes "plik", "Edycja", "Widok", "Wstaw", "Narzędzia", "Okno", and "Pomoc". The toolbar contains various icons for file operations and editing. The main text area displays a document with the following content:

USTAWA
z dnia 29 sierpnia 2003 r.

o zmianie ustawy o Lasach Państwowych

Art. 1.

W ustawie z dnia 26 marca 1982 r. o Lasach Państwowych wprowadza się następujące zmiany:

1) w art. 2 w ust. 1 wyrazy „lub za popełnienie przestępstwa” zastępuje się wyrazami „za popełnienie przestępstwa lub przestępstwa skarbowego”

2) w art. 9 w ust. 9 wyrazy „uznaje się za nieprzedstawiony przez Komisję Odpowiedzialności Konstytucyjnej” zastępuje się wyrazami „nie podlega rozpatrzeniu”;

3) w art. 10:

- a) w ust. 1a po wyrazie „podjęcia” dodaje się wyraz „przez Zgromadzenie Narodowe”;
- b) w ust. 2 wyrazy „odpowiada warunkom wymaganym” zastępuje się wyrazami „spełniać warunki wymagane przez ustawę”;

4) w art. 11:

- a) w ust. 3 w zdaniu pierwszym po wyrazie „uchwałę” dodaje się wyraz „Sejmu”;
- b) w ust. 4 po wyrazie „uchwały” dodaje się wyraz „Sejmu”;

5) w art. 13:

- a) dodaje się ust. 2c w brzmieniu: „2c. Jeżeli Komisja Odpowiedzialności Konstytucyjnej występuje z wnioskiem o umorzenie w stosunku do Prezydenta postępowania z powodu przedawnienia lub śmierci, wówczas przewodniczący Zgromadzeniu Narodowemu stwierdza umorzenie postępowania w sprawie.”;
- b) dodaje się ust. 4 w brzmieniu: „4. Jeżeli Komisja Odpowiedzialności Konstytucyjnej występuje z wnioskiem o umorzenie w stosunku do osób, o których mowa w art. 1 ust. 1 pkt 2-7 albo ust. 2, postępowania z powodu przedawnienia lub śmierci osoby, której postępowanie dotyczy, wówczas Marszałek Sejmu stwierdza umorzenie postępowania w sprawie.”;

Art. 2.

Ustawa wchodzi w życie po upływie 14 dni od dnia ogłoszenia.

The status bar at the bottom shows "Rules Checking On" and "NUM".

Założenia ogólne:

- brak konstrukcji formatujących,
- reprezentacja struktury:
 - hierarchia jednostek: księga, dział, rozdział, artykuł, ...
 - zawartość tekstowa poszczególnych części aktu,
 - tytuły części tytułowanych,
 - przypisy końcowe,
- elementy dodatkowe:
 - uwagi do tekstu,
 - oznaczenia definicji,
 - dodatkowe oznaczenia orzeczeń TK,
 - elementy zewnętrzne w postaci wysp edycyjnych,
- znaczniki nowelizacyjne.

Jak wygląda ustawa?

```
<ustawa data-uchwalenia="20 grudnia 1990">
  <tytul>o jednostkach innowacyjno-wdrożeniowych</tytul>
  <adres-publicacji>
    <adres-dzu rok="1991" numer="2" pozycja="7"/>
  </adres-publicacji>
  <artykul nr="1">Urząd Postępu Naukowo-Technicznego
    i Wdrożeń na dotychczasowych zasadach:
    <punkt nr="1">dokonuje skreśleń w rejestrze jednostek
      innowacyjno-wdrożeniowych w okresie trzech lat
      od dnia wejścia w życie niniejszej ustawy,</punkt>
    <punkt nr="2">wpisuje jednostki do rejestru.</punkt>
  </artykul>
  <artykul nr="2">Ustawa wchodzi w życie po upływie
    14 dni od dnia ogłoszenia.</artykul>
</ustawa>
```

Obecnie:

- dokumenty tworzone w XML-u (zgodnie z tzw. „schematem sejmowym”),
- prezentacja tekstów w wielu formatach (PDF, HTML, DOCX) obsługiwana za pomocą XSLT,
- wersjonowanie przy każdym zapisie; teksty przechowywane w bazie danych (Oracle, CLOB),
- mechanizm automatycznego ujednolicania (przywołania, znaczniki nowelizacyjne, ścieżki nowelizacyjne),
- konwerter Word2XML – prawie automatyczny,
- architektura: Solaris (Windows)/Oracle/JBoss/XMetaL.

```
ustawa:before  
{ content: "USTAWA\Az dnia " attr(data-uchwalenia) " r.";  
  font-weight: bold;  
  display: block;  
  text-align: center; }
```

```
<xsl:template match="artykul">
  <xsl:choose>
    <xsl:when test="count(ancestor::*[name() = 'dodaj'
      or name() = 'zastap']) > 0">
      <div class="artykul-cytowany">
        <xsl:if test="position() = 1">„</xsl:if>
        Art. <xsl:value-of select = "@nr"/>
        <xsl:apply-templates/>
      </div>
    </xsl:when>
    <xsl:otherwise>
      <div class="artykul">
        <div class="naglowek-artykulu">Art.
          <xsl:value-of select = "@nr"/></div>
          <xsl:apply-templates/>
        </div>
      </xsl:otherwise>
    </xsl:choose>
  </xsl:template>
```

Projekty nowelizacji można nakładać na teksty obowiązujące:

- linki nowelizacyjne pomiędzy parami tekstów zapisane w treści nowelizacji:
 - OO: przywołania = linki między ustawą nowelizującą a nowelizowaną,
 - CC: znaczniki nowelizacyjne reprezentujące pojedyncze zmiany wraz z datami ich obowiązywania;
 - cele linków zapisane w postaci wyrażeń XPath,
- możliwość nałożenia wielu projektów nowelizacji na pojedynczy tekst obowiązujący,
- wykrywanie potencjalnych konfliktów nowelizacyjnych.

Przywołania przechowują informację o linku między fragmentem tekstu nowelizacji a ustawą bazową. Są wstawiane w miejscu odwołania do tytułu ustawy nowelizowanej i zastępują go; informacja o ustawie pobierana jest przy każdym otwarciu tekstu bezpośrednio z bazy danych.

Przykład:

```
<przywołanie ustawa-id="8ab282971830bb8a01183221c82c0010"
      typ="nowelizacyjne">
  <ustawa-info data-uchwalenia="6 grudnia 1996">
    <tytul>o zastawie rejestrowym
      i rejestrze zastawów</tytul>
    <adres-publicacji>
      <adres-dzu rok="1996" numer="149" pozycja="703"/>
    </adres-publicacji>
  </ustawa-info>
  <przypis nr="2"/>
</przywołanie>
```

Dwa wymiary znaczników:

- 1 rodzaj operacji:
 - <dodaj>,
 - <zmien>,
 - <usun>,
- 2 cel operacji:
 - element struktury,
 - fragment tekstu.

Tekst:

W art. 7 w ust. 1 dodaje się pkt 6a w brzmieniu:

„6a) analizę wykorzystania wynikających z kontroli wniosków dotyczących stanowienia lub stosowania prawa.”.

Tekst:

W art. 7 w ust. 1 dodaje się pkt 6a w brzmieniu:
„6a) analizę wykorzystania wynikających z kontroli wniosków dotyczących stanowienia lub stosowania prawa.”.

Reprezentacja XML-owa:

W art. 7 w ust. 1 dodaje się pkt 6a w brzmieniu:

```
<dodaj>  
  <po-elemencie id="//artykul[@nr='7']  
    /ustep[@nr='1']/punkt[@nr='6']"/>  
  <element>  
    <punkt nr="6a">analizę wykorzystania  
      wynikających z kontroli wniosków dotyczących  
      stanowienia lub stosowania prawa.</punkt>  
  </element>  
</dodaj>
```

Tekst:

W art. 2 ust. 3 wyrazy „armie” zastępuje się wyrazami „siły zbrojne” .

Tekst:

W art. 2 ust. 3 wyrazy „armie” zastępuje się wyrazami „siły zbrojne” .

Reprezentacja XML-owa:

W art. 2 ust. 3 wyrazy

```
<zastap>  
  <w-elemencie id="//artykul[@nr='2']  
                /ustep[@nr='3']"/>  
  <tekst>armie</tekst>  
  <akcja>zastępuje się wyrazami</akcja>  
  <tekstem>siły zbrojne</tekstem>  
</zastap>.
```

Model znacznika <zastap>

```
<xsd:element name="zastap">
  <xsd:complexType>
    <xsd:choice>
      <xsd:sequence>
        <xsd:choice maxOccurs="unbounded">
          <xsd:element ref="element"/>
          <xsd:element ref="elementy"/>
        </xsd:choice>
        <xsd:element ref="elementem" minOccurs="0"/>
      </xsd:sequence>
      <xsd:sequence>
        <xsd:element ref="w-elemencie"
          minOccurs="0" maxOccurs="unbounded"/>
        <xsd:element ref="tekst"/>
        <xsd:element ref="akcja" minOccurs="0"/>
        <xsd:element ref="tekstem"/>
      </xsd:sequence>
    </xsd:choice>
    <xsd:attributeGroup ref="atr-nowelizacyjne"/>
  </xsd:complexType>
</xsd:element>
```

5 modeli zawartości – do reprezentacji:

- 1 zmiany brzmienia pojedynczego elementu struktury:
 - `<element id="ścieżka">` – element pusty, ze ścieżką do elementu, którego zawartość ma zostać wymieniona,
 - `<elementem>` – nowe brzmienie elementu struktury (zapisane w treści znacznika),
- 2 zmiany brzmienia ciągłego zestawu elementów struktury:
 - `<elementy od="ścieżka1" do="ścieżka2">` – ze ścieżkami wskazującymi początek i koniec ciągłego zestawu elementów o zmienianym brzmieniu,
 - `<elementem>` – nowe brzmieniem zestawu,
- 3 zmiany brzmienia tekstu:
 - `<w-elemencie id="ścieżka">` – ze ścieżką do elementu, w którym ma zostać zmieniony tekst,
 - `<tekst>` – tekst, który ma zostać zastąpiony nowym,
 - `<akcja>` – tekst łączący (np. „zastępuje się wyrazy”),
 - `<tekstem>` – nowe brzmienie zastępowanych wyrazów,

- 4 obudowania tekstu elementem struktury:
 - `<element id="ścieżka">` – ze ścieżką do elementu, którego tekst ma zostać obudowany elementem struktury,
 - `<elementem poziom="element">` – ze wskazaniem nazwy elementu reprezentującego poziomu struktury, którego należy użyć do obudowania tekstu (np. *ustęp*),
- 5 obniżenia tekstu do poziomu tekstu podstawowego:
 - `<element id="ścieżka">` – ze ścieżką do elementu, którego zawartość ma zostać obniżona.

2. Kontrola działalności samorządu terytorialnego przeprowadzana jest pod względem legalności, gospodarności i rzetelności.

3. Kontrola działalności jednostek organizacyjnych i przedsiębiorców, o których mowa w art. 2 ust. 3, jest przeprowadzana pod względem legalności i gospodarności.

Art. 6.

1. Najwyższa Izba Kontroli podejmuje kontrole na zlecenie Sejmu lub jego organów, na wniosek Prezydenta Rzeczypospolitej Polskiej, Prezesa Rady Ministrów oraz z własnej inicjatywy.

2. Najwyższa Izba Kontroli wykonuje swoje zadania na podstawie **okresowych planów⁽¹⁾ planu⁽²⁾ pracy, które⁽³⁾ który⁽⁴⁾** przedkłada Sejmowi; może też przeprowadzać kontrole doraźne.

Art. 7.

1. Najwyższa Izba Kontroli przedkłada Sejmowi:

- 1) analizę wykonania budżetu państwa i założeń polityki pieniężnej;
- 2) opinię w przedmiocie absolutorium dla Rady Ministrów;
- 3) informacje o wynikach kontroli zleconych przez Sejm lub jego organy;
- 4) informacje o wynikach kontroli przeprowadzonych na wniosek Prezydenta Rzeczypospolitej Polskiej, Prezesa Rady Ministrów oraz innych ważniejszych kontroli;
- 5) wnioski w sprawie rozpatrzenia przez Sejm określonych problemów związanych z działalnością organów wykonujących zadania publiczne;
- 6) wystąpienia zawierające wynikające z kontroli zarzuty, dotyczące działalności osób wchodzących w skład Rady Ministrów, kierujących urzędami centralnymi, Prezesa Narodowego Banku Polskiego i osób kierujących instytucjami, o których mowa w art. 4 ust. 1;

^(6a) ⁽⁷⁾ [analizę wykorzystania wynikających z kontroli wniosków dotyczących stanowienia lub stosowania prawa.](#)

⁽⁷⁾ ⁽¹⁰⁾ ~~sprawozdanie ze swej działalności w roku ubiegłym.~~

1a ⁽¹¹⁾ [Najwyższa Izba Kontroli przedstawia Sejmowi coroczne sprawozdanie ze swojej działalności.](#)

2. Tryb wykonywania przez Najwyższą Izbę Kontroli obowiązków wobec Sejmu i jego organów określa regulamin Sejmu.

Art. 8.

1. Najwyższa Izba Kontroli przedkłada Prezydentowi Rzeczypospolitej Polskiej informacje o wynikach kontroli przeprowadzonych na

Gotowe Zaufane wtrymy 100%

Większość tekstów jest wciąż tworzonych w formatach MS Worda. Muszą zostać skonwertowane do formatu XML-owego:

- pliki Worda zapisywane w formacie DOCX,
- proces konwersji zaimplementowany w Javie,
- oprócz struktury weryfikowane dodatkowe własności tekstu (np. interpunkcja),
- błędy konwersji prezentowane w postaci komentarzy Worda,
- dokument wynikowy tworzony na bieżąco,
- wyrażenia regularne dostępne jako parametr – mogą być modyfikowane przez klienta.

USTAWA z dnia 3 lipca 2002 r.

Prawo lotnicze

Rozdział 1 Zakres regulacji i definicje

1. Przepisy Prawa lotniczego regulują stosunki prawne z zakresu lotnictwa cywilnego.)
2. Przepisy Prawa lotniczego stosuje się do polskiego lotnictwa cywilnego oraz, w zakresie ustalonym przez to prawo, również do obcego lotnictwa cywilnego.
3. Lotnictwo cywilne obejmuje wszystkie rodzaje lotnictwa, z wyjątkiem lotnictwa państwowego, to jest państwowych statków powietrznych oraz lotnisk państwowych wykorzystywanych wyłącznie do startów i lądowań państwowych statków powietrznych.)
4. Przepisów prawa lotniczego nie stosuje się do lotnictwa państwowego, z wyjątkiem przepisów art. 1 ust. 6-7²⁾, art. 2, 3-10, 11 ust. 1, art. 12, 14, 33, 35 ust. 2, art. 43, 44, 60, 74, 75, art. 76 ust. 2, art. 89, 92, 119-125, 128, 130, 133, 140, 149 i 150 ustawy, oraz z zastrzeżeniem ust. 5,
5. W przypadku wykorzystywania:
- 1) polskich i obcych państwowych statków powietrznych do prowadzenia działalności innej niż służba publiczna;
 - 2) lotnisk lotnictwa państwowego do startów i lądowań cywilnych statków powietrznych oraz statków, o których mowa w pkt 1
- przepisy ustawy stosuje się odpowiednio.

[6. (uchylony).]

Komentarz [D1]: Znaczenie elementu 'statki', s spodzielano się jednemu z następujących elementów: '(oddział, średyna, artykuł, uwaga, wyspa)'.
(Dashed line connects to item 1)

Komentarz [D2]: Element ma za duzo wcięcia
(Dashed line connects to item 3)

Komentarz [D3]: Niektórzy mają zdania elementu zmuszającego? Okazanie średnika.
(Red line connects to item 1.1)

Czyli czego nie ma, a co można:

- prezentacja każdego tekstu w wersji „na dany dzień” ,
- wyszukiwanie pełnotekstowe (i ew. bardziej zaawansowane),
- obsługa prac komisji,
- ???