

XML i nowoczesne technologie zarządzania treścią

Egzamin, 2 lutego 2010 r.

Imię i nazwisko: _____

Pytania testowe są wielokrotnego wyboru, prawdziwa jest co najmniej jedna odpowiedź. Każde pytanie – 1 punkt. W przypadku pytań opisowych możliwe są także oceny ułamkowe.

1. Zaznacz stwierdzenia prawdziwe:

- a) Każdy dokument XML-owy musi być poprawny strukturalnie (ang. *valid*).
- b) Element pusty nie może zawierać atrybutów.
- c) Jeśli wartość atrybutu nie zawiera spacji, można pominąć otaczające ją cudzysłowy lub apostrofy.
- d) Dokument XML-owy nie musi być zapisany w kodowaniu UTF-8.

2. Które deklaracje elementów są poprawne?

- a) `<!ELEMENT test (#PCDATA)>`
- b) `<!ELEMENT test CDATA>`
- c) `<!ELEMENT test EMPTY>`
- d) `<!ELEMENT test (ANY)>`

3. Wskaż i objaśnij trzy błędy składni XML-owej w poniższym dokumencie. Zakładamy, że dokument jest zapisany w kodowaniu UTF-8.

```
<?xml version="1.0" ?>
<!-- $Id: -- egz-2010-02-02, v 1.2 <mog> Exp $ -->
<html><p>Ten dokument
<div align='center'>
<b>nie jest, <i><br>
<_tag-własny_>niestety</_tag-własny_>,<b/> </i></b>,
<i>poprawny składniowo</i>.</div></p>
<hr noshade>
</html>
```

4. Które stwierdzenia są prawdziwe?

- a) Normalizacja dla napisów typu `normalizedString` polega na usunięciu białych znaków z początku i końca napisu oraz zastąpieniu ciągów białych znaków przez pojedynczą spację.
- b) Element o wartości nieokreślonej (ang. *nil value*) musi mieć zawartość pustą niezależnie od definicji jego modelu zawartości w schemacie, natomiast jego atrybuty muszą być zawsze zgodne z modelem.
- c) Jedną z metod zapewniania odporności na zmianę struktury dokumentów w czasie jest parametryzacja schematem – użycie atrybutów opcjonalnych do przechowania etykiet pól formularza, odwzorowania elementów na tabele i pola w bazie danych itp.
- d) W składni XPointer napis `dokument.xml#element(test)` oznacza odwołanie do wszystkich elementów o nazwie `test` znajdujących się w pliku `dokument.xml`.

5. Które stwierdzenia nt. przestrzeni nazw w XML-u są prawdziwe?

- a) Nazwy kwalifikowane muszą być poprzedzone prefiksem skojarzonym z przestrzenią nazw, do której należą.
- b) W definicji schematu URI używane w deklaracji docelowej przestrzeni nazw nie musi wskazywać miejsca lokalizacji schematu.
- c) Nazwa atrybutu nie poprzedzona prefiksem przestrzeni nazw należy do tej samej przestrzeni nazw, co nazwa elementu, w którym atrybut się znajduje.
- d) Nazwy encji mogą używać niestandardowych prefiksów przestrzeni nazw.

6. Zapisz dany fragment DTD w składni XML Schema (jako fragment pełnego schematu – zdefiniuj globalne elementy **test** i **pierwszy** wraz z atrybutami, pominięciem definicje elementów **drugi**, **trzeci** i **czwarty**, ale odwołaj się do nich, jak gdyby były zdefiniowane globalnie).
Każdy element za 0,5 pkt.

```
<!ELEMENT test (pierwszy?, (drugi | trzeci+), czwarty*)>
<!ELEMENT pierwszy (#PCDATA | drugi)*>
```

7. XML Schema pozwala na:

- a) ograniczenie zawartości elementu do postaci listy 5 liczb naturalnych z przedziału 1-10 oddzielanych dowolną liczbą białych znaków,
- b) określenie, że wymienione podelementy danego elementu mogą wystąpić w dowolnej kolejności,
- c) określenie, że wartość tekstowa dwóch atrybutów danego elementu ma być identyczna,
- d) określenie, że element może zawierać dowolne atrybuty.

8. Które stwierdzenia nt. standardu XLink są prawdziwe?

- a) XLinki mogą łączyć nie więcej niż 2 zasoby.
- b) XLinki mogą występować wyłącznie w dokumentach poprawnych strukturalnie.
- c) XLinki umożliwiają nawigację w wielu kierunkach.
- d) XLinki nie muszą być zagnieżdżane w treści linkowanych dokumentów.

9. Zaznacz prawdziwe stwierdzenia dotyczące dokumentów XSL Formatting Objects.

- a) Dokument XSL-FO pełni rolę arkusza stylu.
- b) Formatowanie elementów danego rodzaju (np. tabel) opisuje się w jednym miejscu dokumentu.
- c) Formatowanie elementów zawierających treść opisuje się w ich atrybutach.
- d) Można raz podać zawartość, która będzie powtarzana na każdej stronie (np. w nagłówku).

10. Poniższy fragment arkusza XSLT

```
<xsl:variable name="x" select="/test/*"/>
<xsl:if test="$x">
  <xsl:variable name="x" select="/test/*[1]"/>
</xsl:if>
<xsl:value-of select="$x"/>
```

przy dokumencie wejściowym

```
<test><a>jeden</a><b>dwa</b><c>trzy</c></test>
```

może (w XSLT 1.0 lub 2.0) spowodować wypisanie na wyjście (przy metodzie serializacji xml):

- a) <a>jeden
- b) <a>jedendwa<c>trzy</c>
- c) jeden
- d) jeden dwa trzy

11. Zaznacz prawdziwe stwierdzenia dotyczące XQuery 1.0.

- a) Zapytanie XQuery jest dokumentem XML.
- b) Efektywna wartość logiczna sekwencji składającej się z jednego węzła tekstowego to prawda, niezależnie od tekstu zawartego w węźle.
- c) Wynik zapytania może zawierać elementy znajdujące się w dokumencie źródłowym.
- d) Wynik zapytania może zawierać elementy nie znajdujące się w dokumencie źródłowym.

12. Pewien dokument XML zawiera wiele elementów **klucz** z zawartością tekstową. Napisz ścieżkę XPath 1.0, która zwraca liczbę różnych wartości tekstowych elementów **klucz**.

13. Zaznacz prawdziwe stwierdzenia dotyczące korzystania z SAX2.

- a) Parsowanie nie wymaga wczytywania całego dokumentu na raz do pamięci.
- b) W oparciu o schemat dokumentu typy zawartości elementów i atrybutów są odzwierciedlane na odpowiednie typy w języku programowania.
- c) Aby rozpoznawać identyfikatory przestrzeni nazw elementów i atrybutów, należy zapamiętywać historię deklaracji przestrzeni nazw.
- d) Filtry mogą przekazywać do dalszego przetwarzania zdarzenia nie występujące na wejściu.

14. Podaj zalety i wady XML jako formatu elektronicznej wymiany danych w stosunku do tradycyjnego EDI (np. EDIFACT).

15. Wymień trzy czynności, które podsystem wyszukiwarki przetwarzający tekst zapytania powinien wykonać zanim zacznie dopasowywać go do zawartości indeksu.

16. W jaki sposób reprezentowane są zdania logiczne w standardzie RDF (na poziomie idei, nie składni)? Podaj przykład z objaśnieniem.