

# XPath

# XPath izrazi

---

- izraz je namenjen za označavanje čvora ili skupa čvorova u dokumentu
- ovi izrazi liče na izraze za rad sa fajl-sistemom
  - fajl-sistem
    - C:\windows\fonts\arial.ttf
    - \windows\fonts\arial.ttf
    - ../../windows/fonts/arial.ttf

# Standardne XPath funkcije

---

- XPath sadrži oko 100 ugrađenih funkcija
  - operacije sa stringovima
  - operacije numeričkim vrednostima
  - operacije sa datumom i vremenom
  - manipulacija čvorovima
  - manipulacija sekvencama
  - ...

# XPath čvorovi

---

- XML dokument se, sa stanovišta XPath-a, posmatra kao stablo
- više tipova čvorova
  - element
  - atribut
  - tekst
  - namespace
  - procesna instrukcija
  - komentar
  - dokument (koren)

# XPath čvorovi

- primer

```
<?xml version="1.0"?>
```

```
<bookstore>
```

```
  <book>
```

```
    <title lang="en">Harry Potter</title>
```

```
    <author>J. K. Rowling</author>
```

```
    <year>2005</year>
```

```
    <price>29.99</price>
```

```
  </book>
```

```
</bookstore>
```

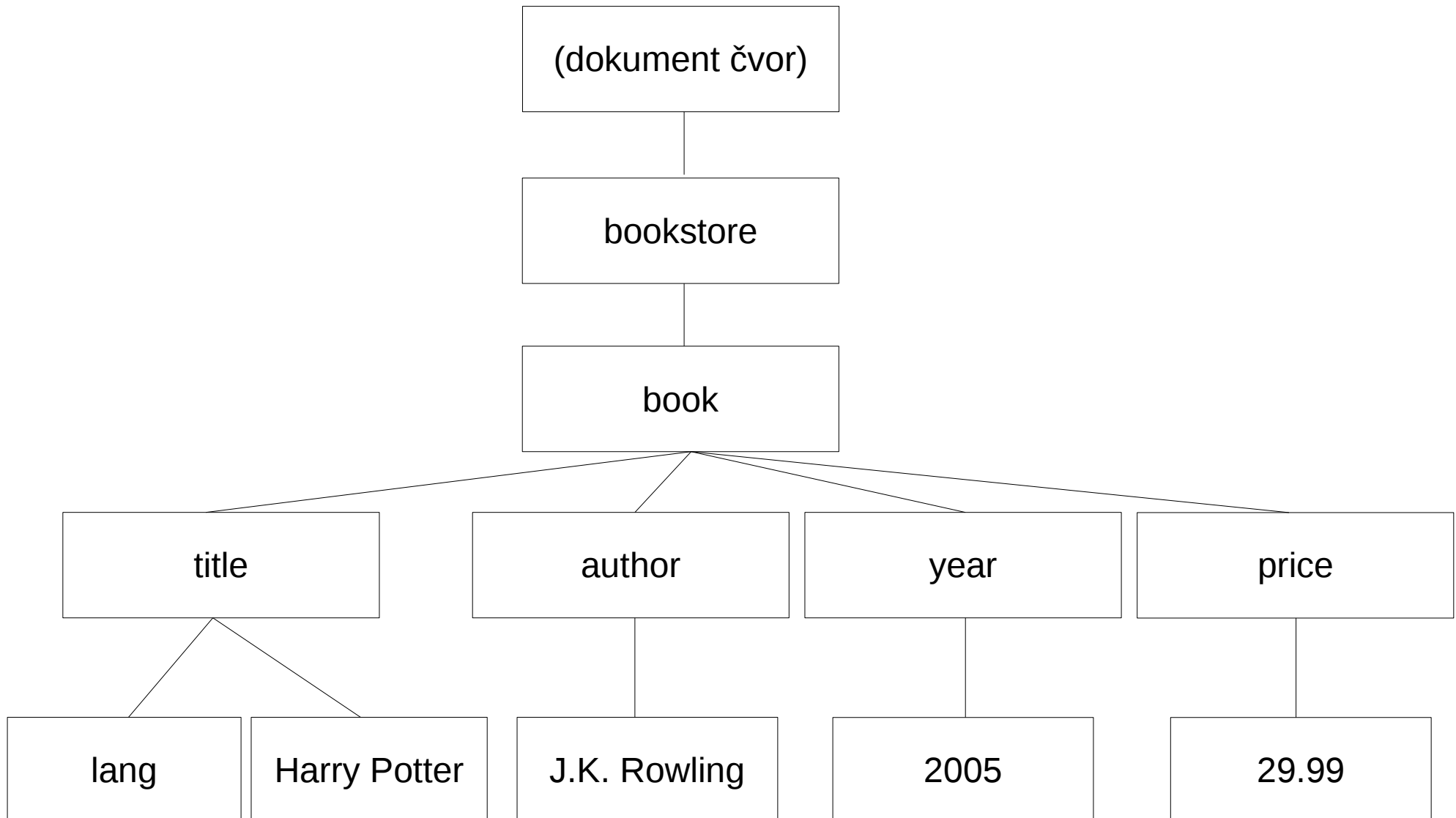
dokument čvor

atribut čvor

element čvor

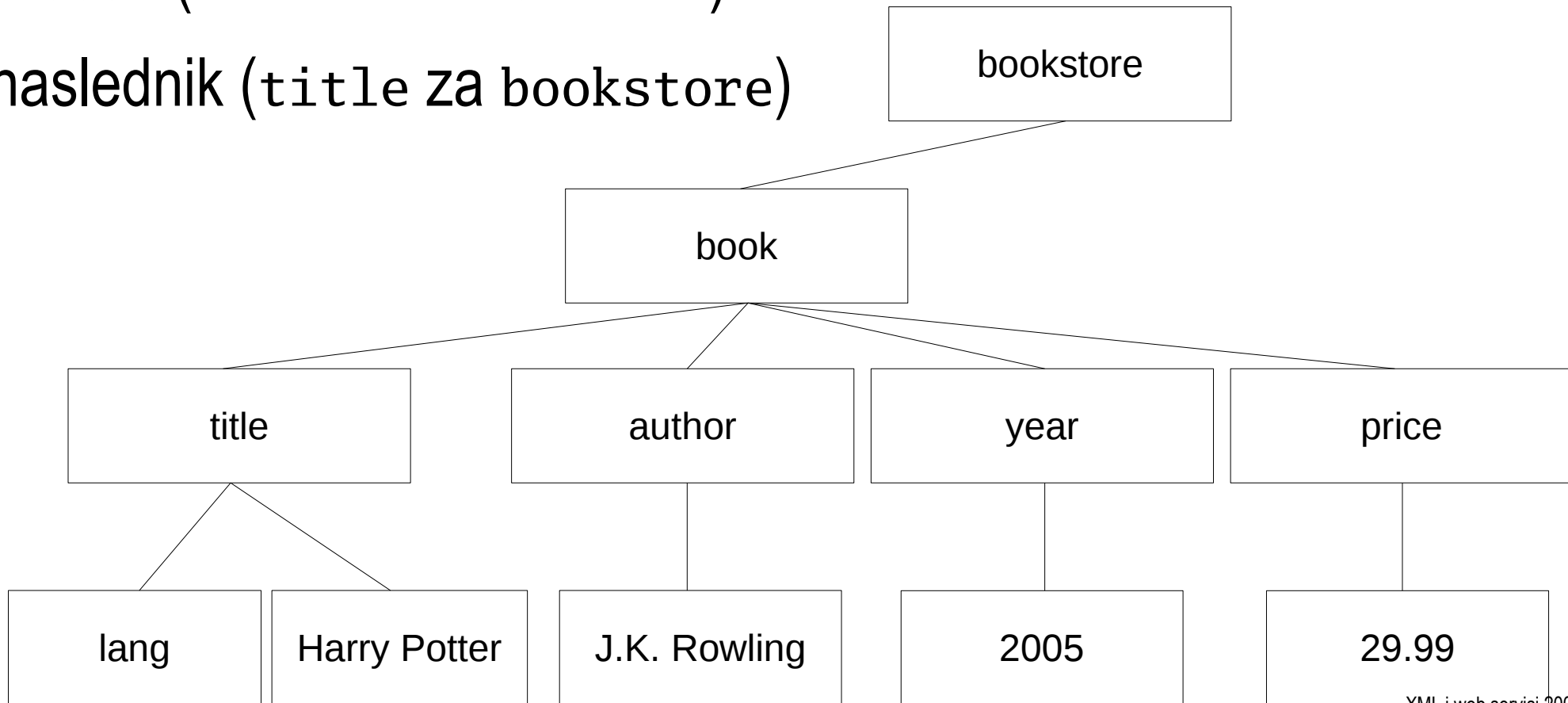
# XPath čvorovi

- primer



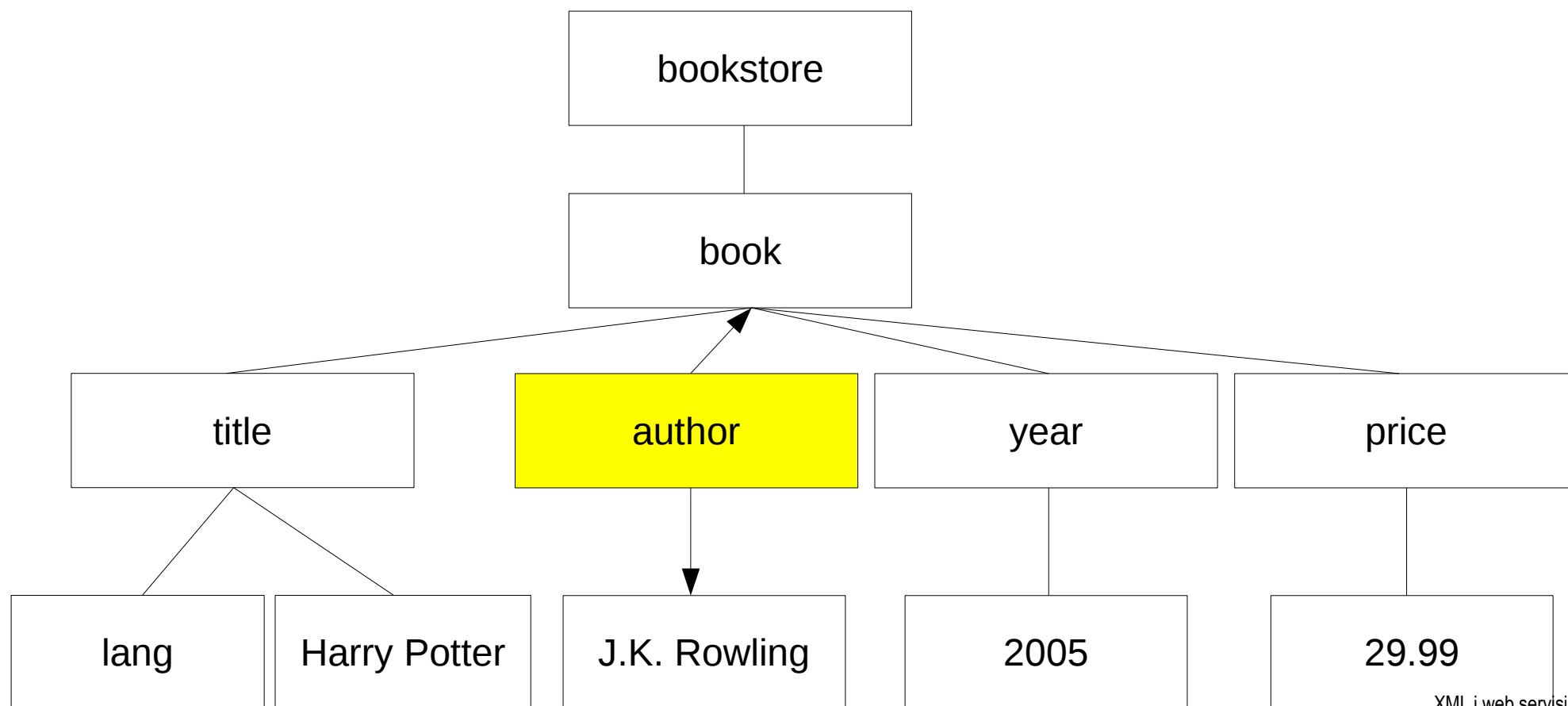
# Odnosi među čvorovima

- roditelj (book za title)
- dete (title za book)
- brat/sestra (title i author)
- predak (bookstore za title)
- naslednik (title za bookstore)



# Navigacija kroz stablo dokumenta

- koncept tekućeg čvora = “gde se trenutno nalazim”
- ako je tekući čvor author
  - u smeru roditelj nalazi se book
  - u smeru dete nalazi se text()=“J.K. Rowling”





# XPath putanja

---

- sastoji se iz više koraka razdvojenih znakom /
- apsolutna putanja: počinje znakom /
  - /korak/korak/...
  - kreće uvek od dokument-čvora
- relativna putanja: ne počinje znakom /
  - korak/korak/...
- svaki korak se izračunava u odnosu na čvorove u tekućem skupu čvorova (node-set)

# XPath putanja

---

- korak putanje se sastoji iz
  - ose kretanja
  - testa čvora
  - nula ili više predikata
- opšti oblik izraza za korak je

`osa::test[predikat]`

# XPath ose

---

- osa definiše skup čvorova u odnosu na tekući čvor
  - ancestor svi preci
  - ancestor-or-self svi preci ili sam čvor
  - attribute svi atributi
  - child sva deca
  - descendant svi potomci
  - descendant-or-self svi potomci ili sam čvor
  - following sve u dokumentu iza krajnjeg taga tekućeg
  - following-sibling sva braća posle tekućeg
  - namespace svi namespace čvorovi tekućeg
  - parent roditelj
  - preceding sve u dokumentu iza početnog taga tekućeg
  - preceding-sibling sva braća pre tekućeg
  - self tekući čvor

# XPath putanja

---

- skraćeno pisanje
  - ako se osa ne navede podrazumeva se `child::`
  - `self::` se može pisati kao tačka `.`
  - `parent::` se može pisati kao dve tačke `..`
  - `descendant::` se može pisati kao dve kose crte `//`
  - `attribute::` se može pisati kao `@`
  - `following-sibling::` se može pisati kao `../`

# Predikati

---

- koriste se za pronalaženje čvora koji zadovoljava dati uslov
  - osnovna primena: filtriranje čvorova
    - primer: od svih *title* čvorova, izdvoj one čiji je tekstualni sadržaj duži od 20 znakova
- uvek se pišu unutar uglastih zagrada [ . . . ]

# Predikati

- primeri
  - `//title[@lang]`
    - svi elementi *title* koji imaju atribut *lang*
  - `//title[@lang='eng']`
    - svi *title* elementi koji imaju atribut *lang* sa vrednošću 'eng'
  - `/bookstore/book[price > 35.00]`
    - svi *book* elementi deca *bookstore* korena koji imaju podelement *price* sa vrednošću većom od 35.00
  - `/bookstore/book[price > 35.00]/title`
    - svi *title* elementi koji pripadaju *book* elementima koji pripadaju *bookstore* korenu koji imaju *price* element sa vrednošću većom of 35.00
  - `/bookstore/book[price > 35.00]/price`
    - svi *price* elementi koji pripadaju *book* elementima koji pripadaju *bookstore* korenu koji imaju *price* element sa vrednošću većom of 35.00

# Predikati

---

- primeri
  - `/bookstore/book[1]`
    - prvi *book* element koji je dete *bookstore* korena
  - `/bookstore/book[last()]`
    - poslednji *book* element koji je dete *bookstore* korena
  - `/bookstore/book[last()-1]`
    - preposlednji *book* element koji je dete *bookstore* korena
  - `/bookstore/book[position()<3]`
    - prva dva *book* elementa koji su deca *bookstore* korena

# XPath operatori

Operator	Opis	Primer	Rezultat
	unija dva skupa čvorova	//book   //cd	skup čvorova sa svim book i cd elementima
+	sabiranje	6 + 4	10
-	oduzimanje	6 — 4	2
*	množenje	6*4	24
div	deljenje	6 div 4	1
mod	ostatak pri deljenju	6 mod 4	2
=	jednako	price=9.80	true ako je price=9.80, inače false
!=	nejednako	price!=9.80	
<	manje	price<9.80	
<=	manje ili jednako	price<=9.80	
>	veće	price>9.80	
>=	veće ili jednako	price>=9.80	
or	disjunkcija	price=9.80 or price=9.70	
and	konjunkcija	price>9.00 and price<9.90	



# Biblioteka standardnih funkcija

---

- core functions: 27 funkcija
  - `boolean()`
  - `ceiling()`
  - `concat()`
  - `contains()`
  - `count()`
  - `false()`
  - `floor()`
  - `id()`
  - `lang()`
  - `last()`
  - `local-name()`
  - `name()`
  - `namespace-uri()`
  - `normalize-space()`
  - `not()`
  - `number()`
  - `position()`
  - `round()`
  - `starts-with()`
  - `string()`
  - `string-length()`
  - `substring()`
  - `substring-after()`
  - `substring-before()`
  - `sum()`
  - `translate()`
  - `true()`

# Označavanje više putanja odjednom

---

- navođenjem više XPath izraza povezanih operatorom |
- rezultat je unija skupova čvorova dobijenih osnovnim izrazima
- primeri
  - `//book/title | //book/price`
    - svi *title* elementi deca *book*-a i svi *price* elementi deca *book*-a
  - `//title | //price`
    - svi *title* elementi i svi *price* elementi
  - `/bookstore/book/title | //price`
    - svi *title* elementi *book* elementa *bookstore* korena i svi *price* elementi