

# ΑΝΑΠΑΡΑΣΤΑΣΗ ΓΝΩΣΗΣ ΣΤΟΝ ΠΑΓΚΟΣΜΙΟ ΙΣΤΟ

XML-XPath

Ι. Χατζηλυγερούδης

# Ερωτήματα σε έγγραφα XML

- Αντίστοιχα των ερωτημάτων (queries) σε βάσεις δεδομένων
- Γλώσσες ερωτημάτων (αντίστοιχες της SQL)
  - XQL (XML Query Language)
  - XML-QL (Query Language for XML)
  - **XQuery**
- Στηρίζεται σε εκφράσεις της Xpath, οι οποίες κάνουν πλοήγηση ανάμεσα στα στοιχεία και τα χαρακτηριστικά της XML

# XPath

- Η XPath χρησιμοποιεί «παραστάσεις διαδρομής» (path expressions).
- Κάθε παράσταση διαδρομής ορίζει τον τρόπο με τον οποίο μπορούμε να φτάσουμε σε/προσδιορίσουμε ένα κόμβο ή ένα σύνολο κόμβων στη **δενδρική αναπαράσταση** ενός εγγράφου XML.
- Η σύνταξη της XPath δεν βασίζεται στην XML.

# Χpath-Ορολογία

## ■ Κόμβοι (nodes)

- element (στοιχείου), attribute (χαρακτηριστικού), text (κειμένου), namespace (χώρου ονομάτων), processing-instruction (εντολής επεξεργασίας), comment (σχολίων), document (root) (εγγράφου ή ρίζας)

- Π.χ.

<library> (root node)

<author>Henry Wise</author> (element node)

location = "Bremen" (attribute node)

# Χρath-Ορολογία

- **Ατομικές τιμές (atomic values)**
  - Είναι κόμβοι χωρίς παιδιά ή γονέα
  - Π.χ. Henry Wise, “Bremen”
- **Συστατικά (Items)**
  - Είναι ατομικές τιμές ή κόμβοι
- **Σχέσεις μεταξύ κόμβων**
  - parent (γονέας), child (παιδί), sibling (αμφιθαλής), ancestor (πρόγονος), descendant (απόγονος)

# Xpath-Σύνταξη

- Εκφράσεις ή Παραστάσεις Διαδρομής (Path Expressions)
- Δύο τύποι
  - Απόλυτες (absolute)
    - Ξεκινούν με /, που αναφέρεται στη ρίζα του εγγράφου
  - Σχετικές (relative)
    - Ξεκινούν χωρίς το /, μ' ένα στοιχείο, αναφέρονται δηλ. σε κάποιο κόμβο αναφοράς
- Δομή
  - βήμα<sub>1</sub>/βήμα<sub>2</sub>/.../βήμα<sub>n</sub>
  - βήμα: <αναγνωριστικό> <έλεγχος κόμβου> <κατηγορία>

# Χpath-Σύνταξη

## ■ Επιλογή-Προσπέλαση κόμβων

- Αναγνωριστικά
  - όνομα-κόμβου, /, //, ., .., @
- Εκφράσεις διαδρομής
  - **library** (όλα τα παιδιά του στοιχείου)
  - **/library** (το στοιχείο-ρίζα library)
  - **/library/author** (όλα τα στοιχεία author που είναι παιδιά του library)
  - **/library/@location** (όλα τα χαρακτηριστικά location που είναι σε παιδιά του library)
  - **//author** (όλα τα στοιχεία author όπου κι' αν βρίσκονται)
  - **/library//author** (όλα τα στοιχεία author που είναι απόγονοι του library, όπου κι αν είναι)
  - **//@name** (όλα τα χαρακτηριστικά name)
  - **//book/@title="AI"** (όλοι οι κόμβοι title σε οποιοδήποτε κόμβο book με τιμή "AI")
  - **//book[@title="AI"]** (όλοι οι κόμβοι book που έχουν title με τιμή "AI")

```
<?xml version="1.0" encoding="UTF-16"?>  
<!DOCTYPE library PUBLIC "library.dtd">
```

```
<library location="Bremen">
```

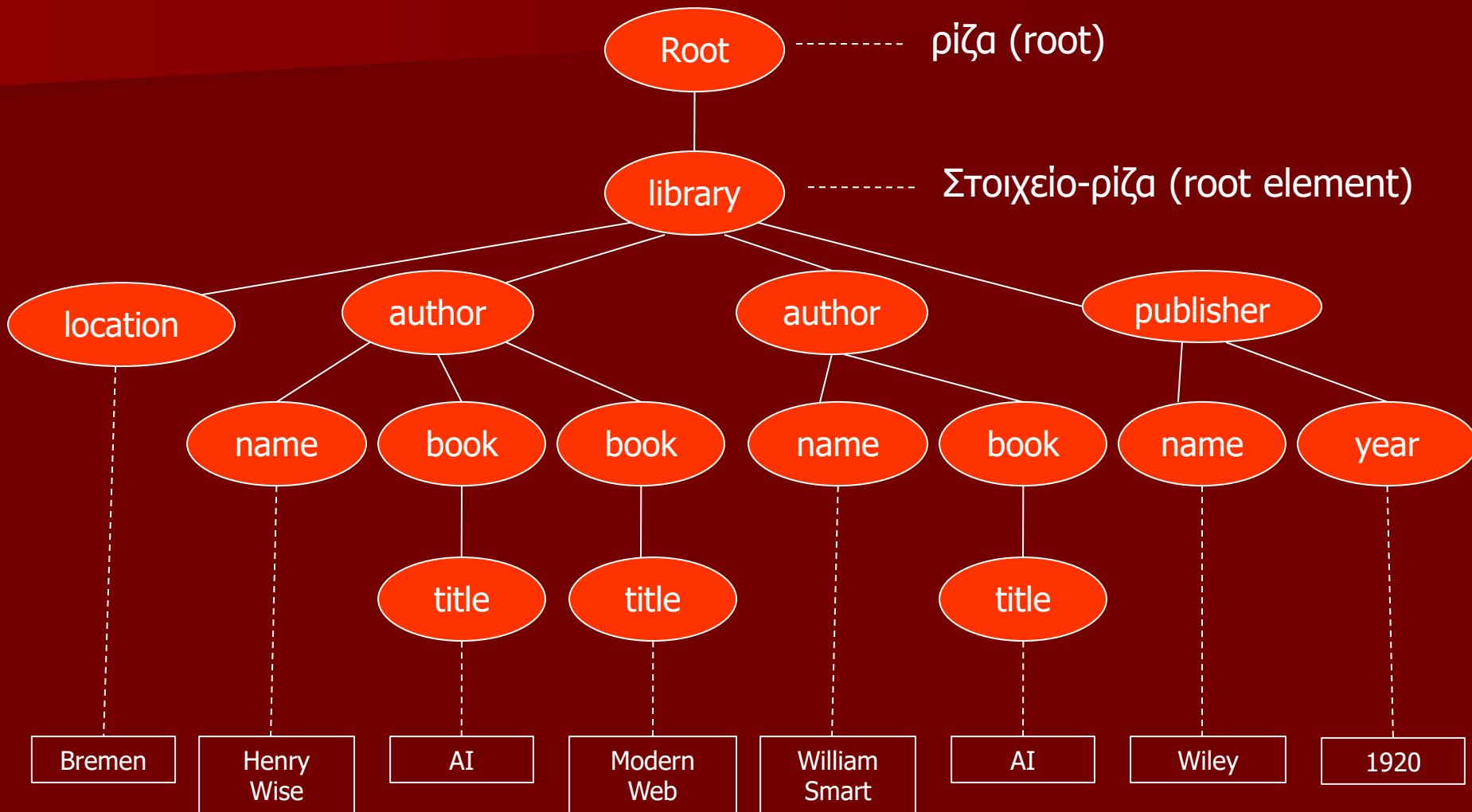
```
<author name="Henry Wise">  
  <book title="AI"/>  
  <book title="Modern Web"/>  
</author>
```

```
<author name="William Smart">  
  <book title="AI"/>  
</author>
```

```
<publisher>  
  <name>Wiley</name>  
  <year>1920</year>  
</publisher>
```

```
</library>
```

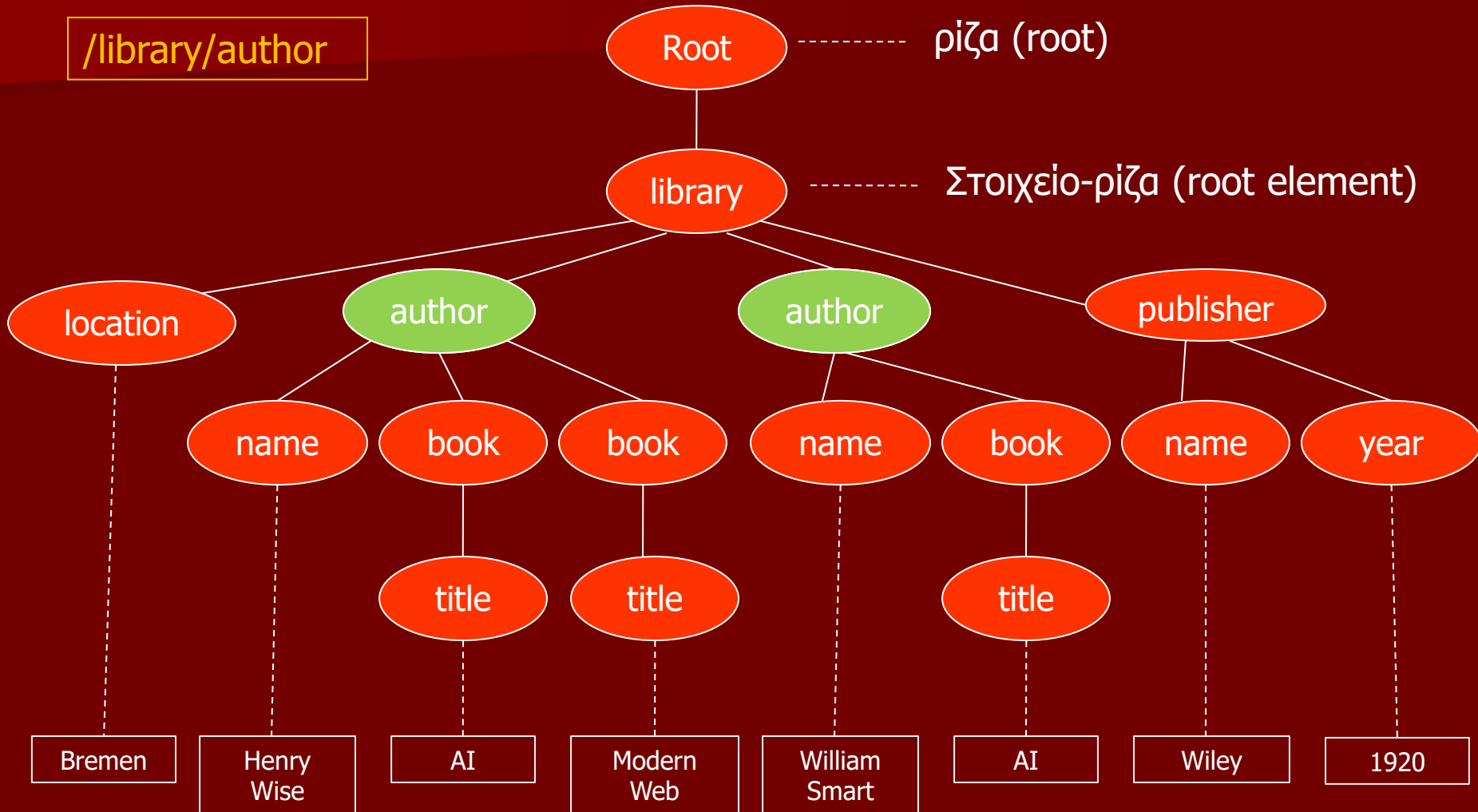
# Δενδρικό Μοντέλο Δόμησης





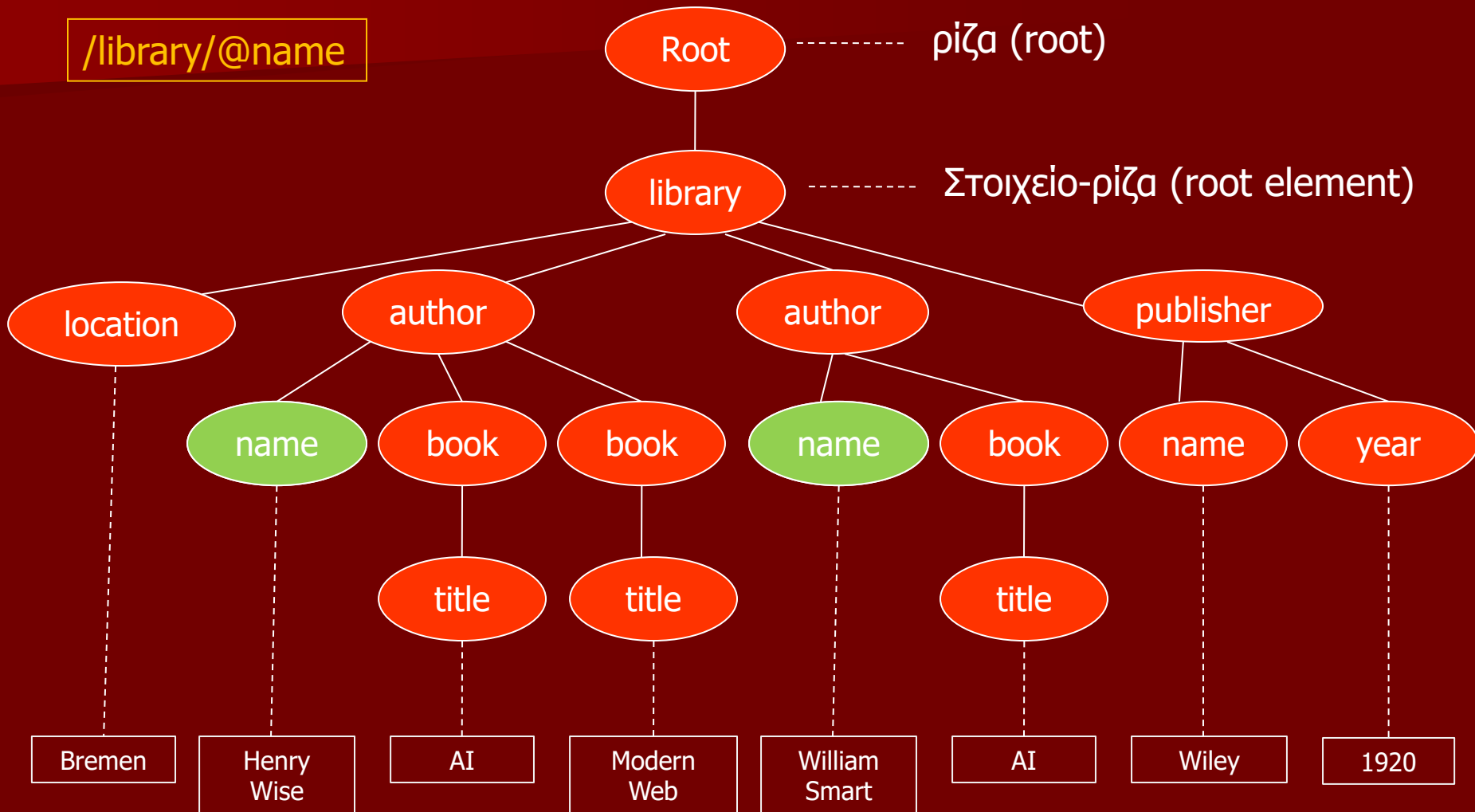
# Δενδρικό Μοντέλο-Παράδειγμα

`/library/author`



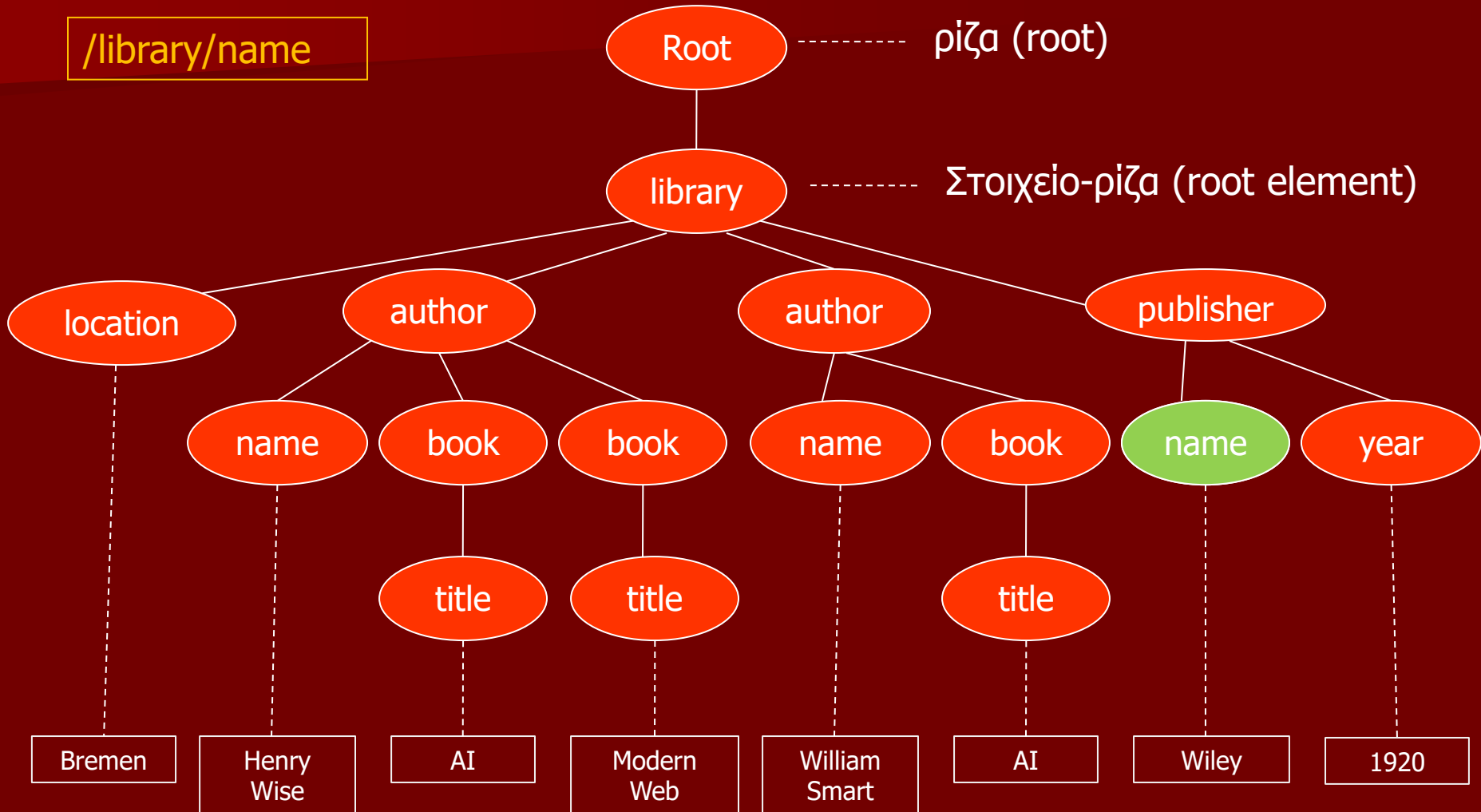
# Δενδρικό Μοντέλο-Παράδειγμα

`/library/@name`



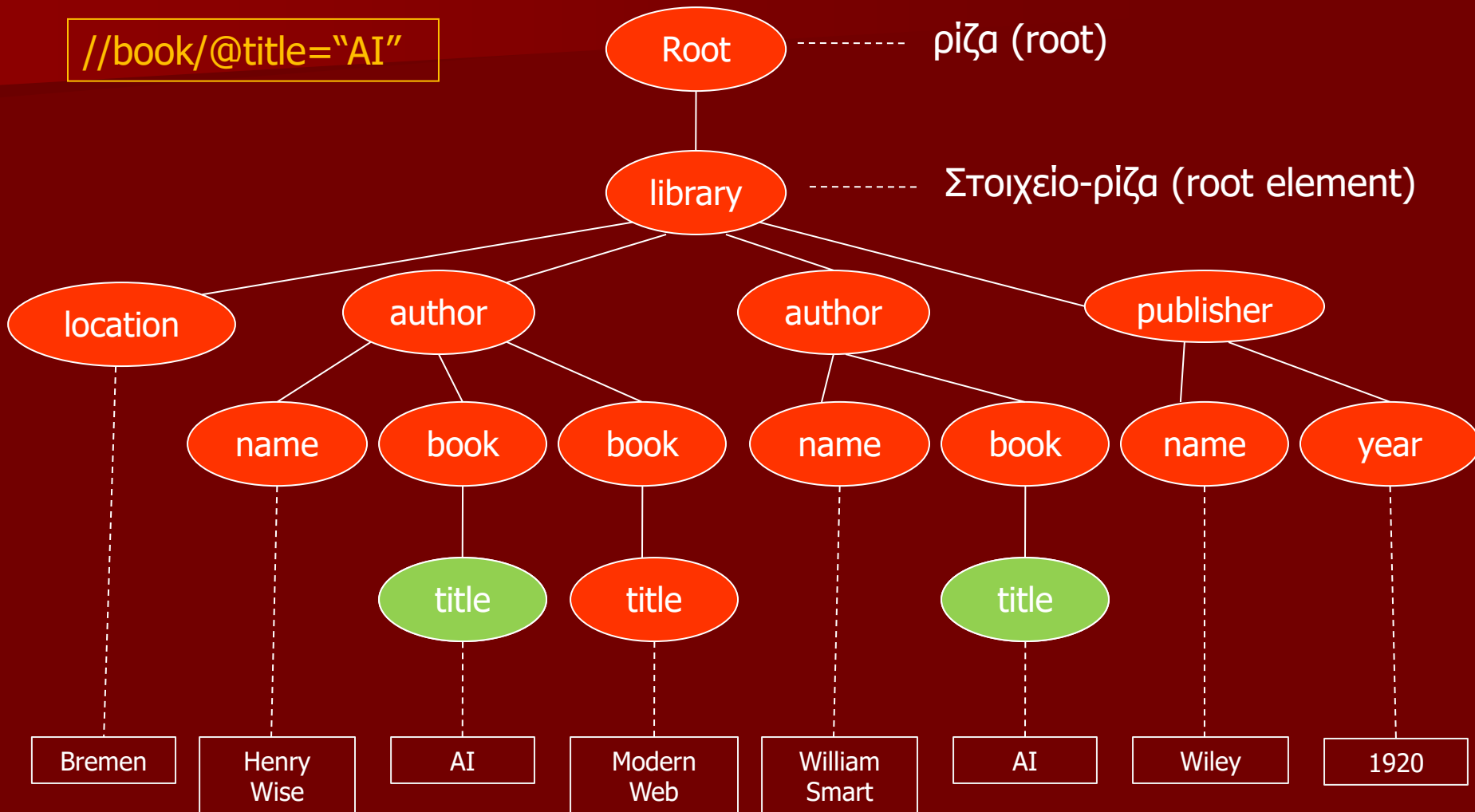
# Δενδρικό Μοντέλο-Παράδειγμα

`/library/name`



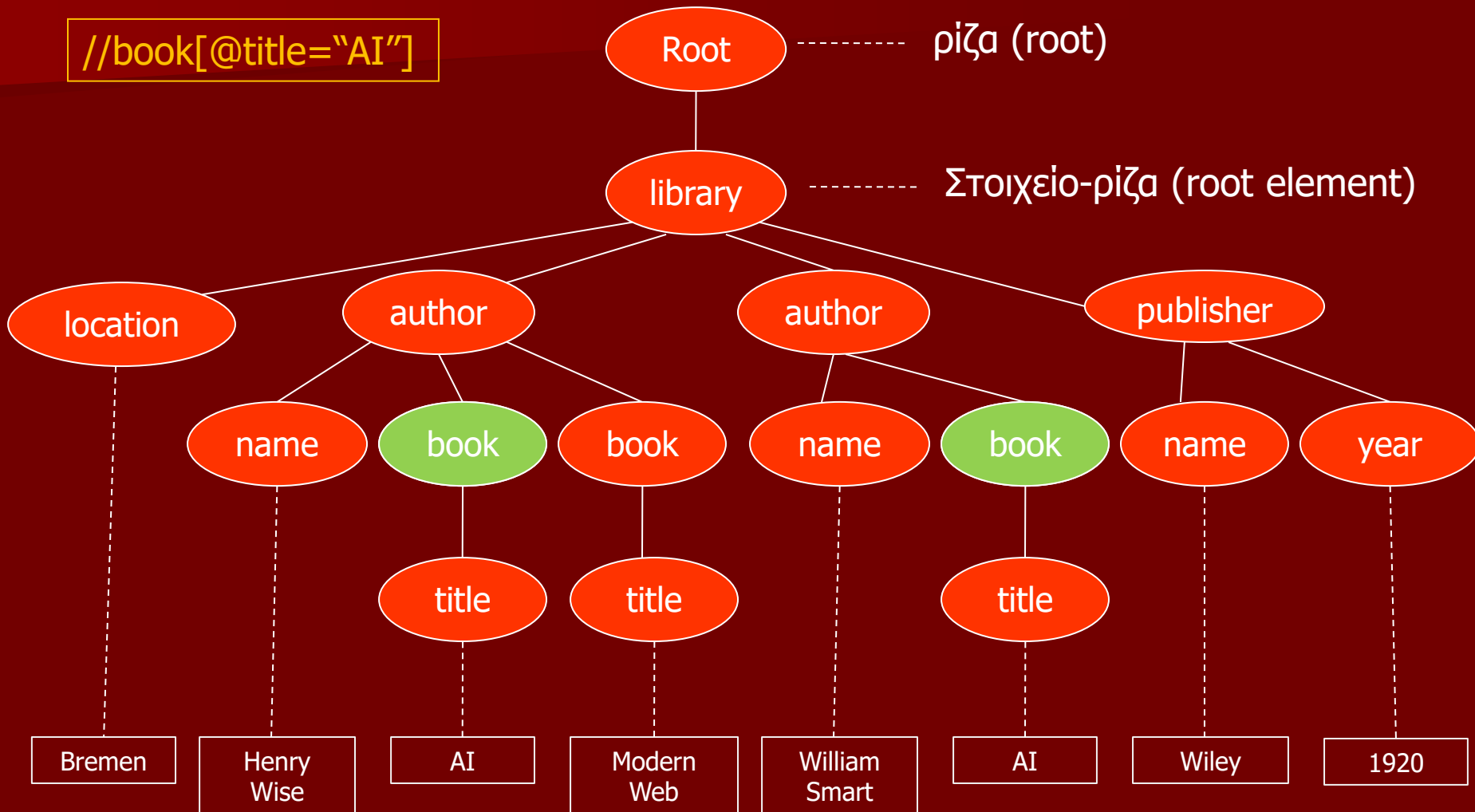
# Δενδρικό Μοντέλο-Παράδειγμα

`//book/@title="AI"`



# Δενδρικό Μοντέλο-Παράδειγμα

```
//book[@title="AI"]
```



# Χpath-Σύνταξη

## ■ Κατηγορήματα

(για εύρεση συγκεκριμένων κόμβων ή κόμβων που περιέχουν συγκεκριμένες τιμές-σε τετραγωνικές παρενθέσεις)

### – Εκφράσεις διαδρομής

- `/library/author[1]` (το πρώτο στοιχείο author που είναι παιδί του στοιχείου library)
- `/library/author[last()]` (το τελευταίο στοιχείο author που είναι παιδί του στοιχείου library)
- `/library/author[last()-1]` (το προτελευταίο στοιχείο author που είναι παιδί του στοιχείου library)
- `/library/author[position()<3]` (τα δύο πρώτα στοιχεία author που είναι παιδιά του στοιχείου library)
- `//book[@title]` (όλα τα βιβλία που έχουν χαρακτηριστικό title)
- `//book[@title="AI"]` (όλοι οι κόμβοι book που έχουν title με τιμή "AI")
- `/library/publisher[year>1900]` (όλα τα στοιχεία publisher που έχουν στοιχείο year με περιεχόμενο/τιμή μεγαλύτερο/η του 1900)

```
<?xml version="1.0" encoding="UTF-16"?>  
<!DOCTYPE library PUBLIC "library.dtd">
```

```
<library location="Bremen">
```

```
<author name="Henry Wise">  
  <book title="AI"/>  
  <book title="Modern Web"/>  
</author>
```

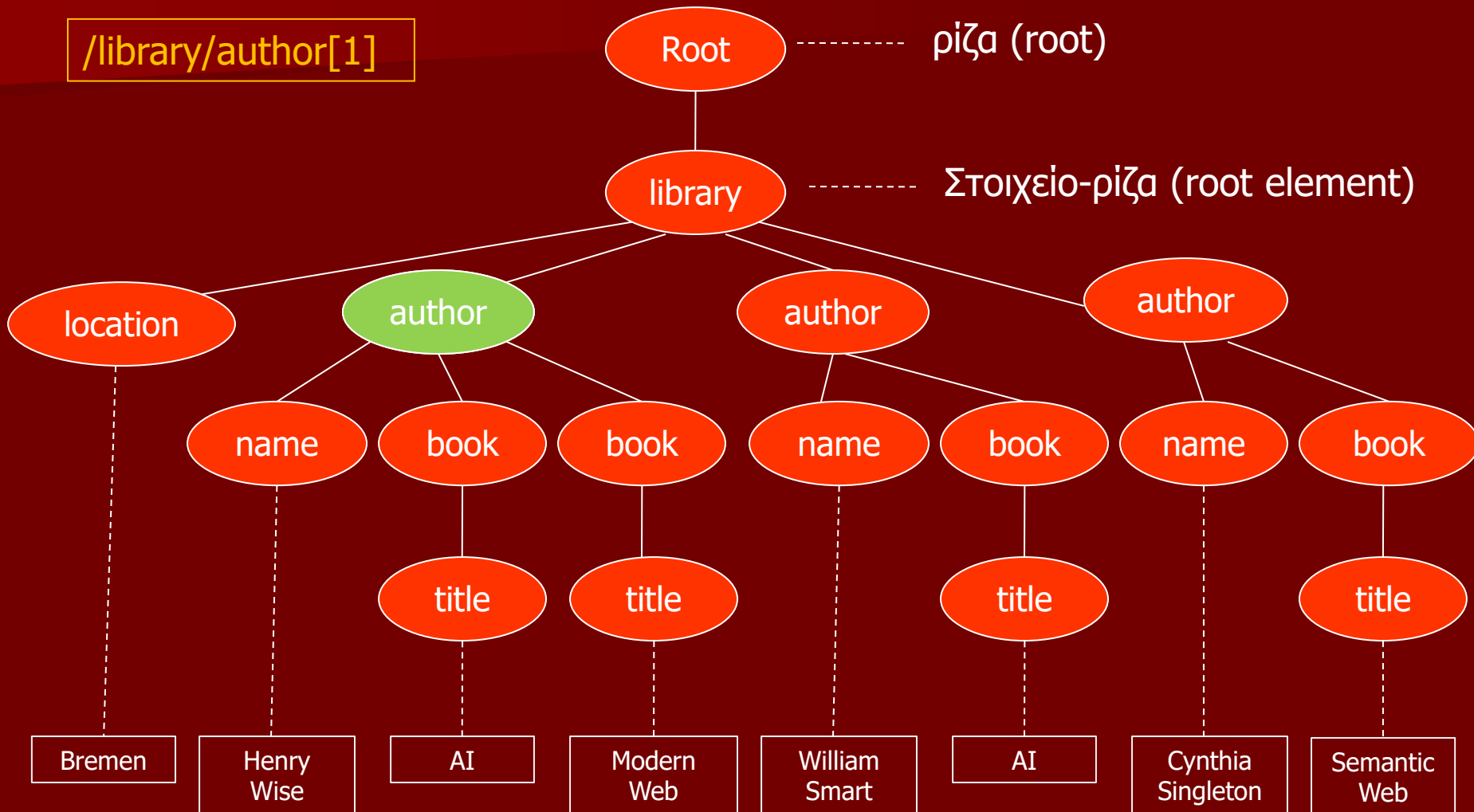
```
<author name="William Smart">  
  <book title="AI"/>  
</author>
```

```
<author name="Cynthia Singleton">  
  <book title="Semantic Web"/>  
</author>
```

```
</library>
```

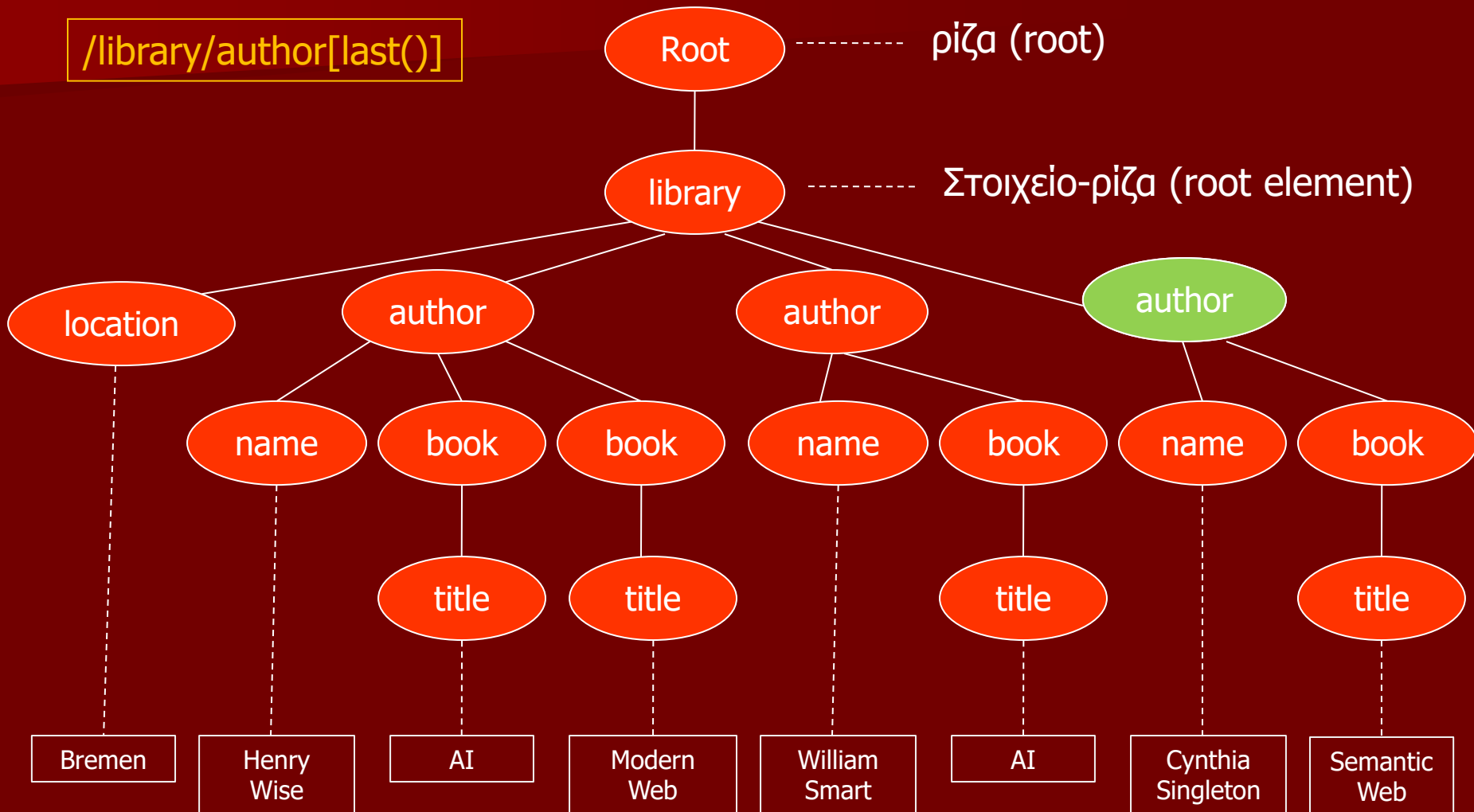
# Δενδρικό Μοντέλο-Παράδειγμα

`/library/author[1]`



# Δενδρικό Μοντέλο-Παράδειγμα

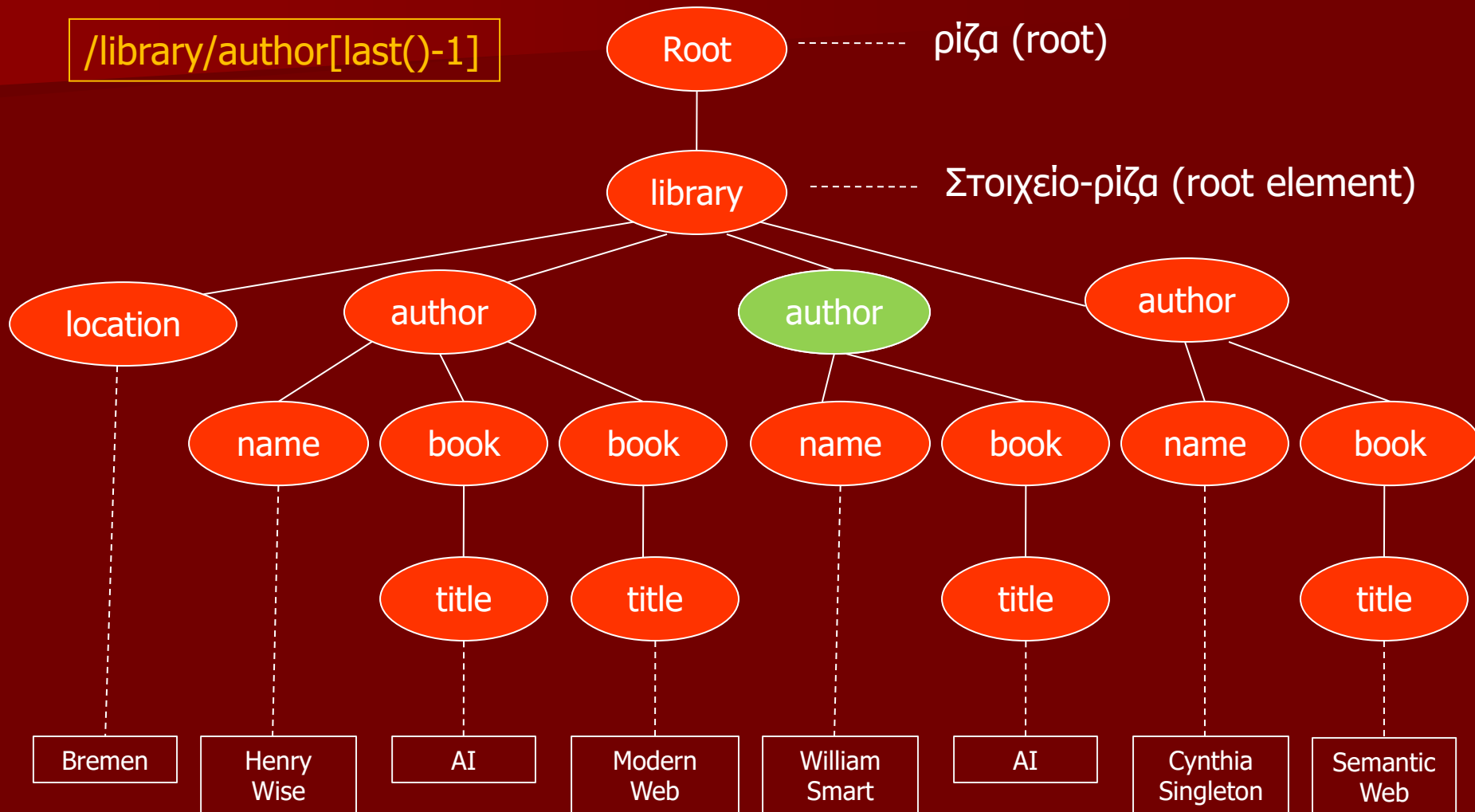
`/library/author[last()]`





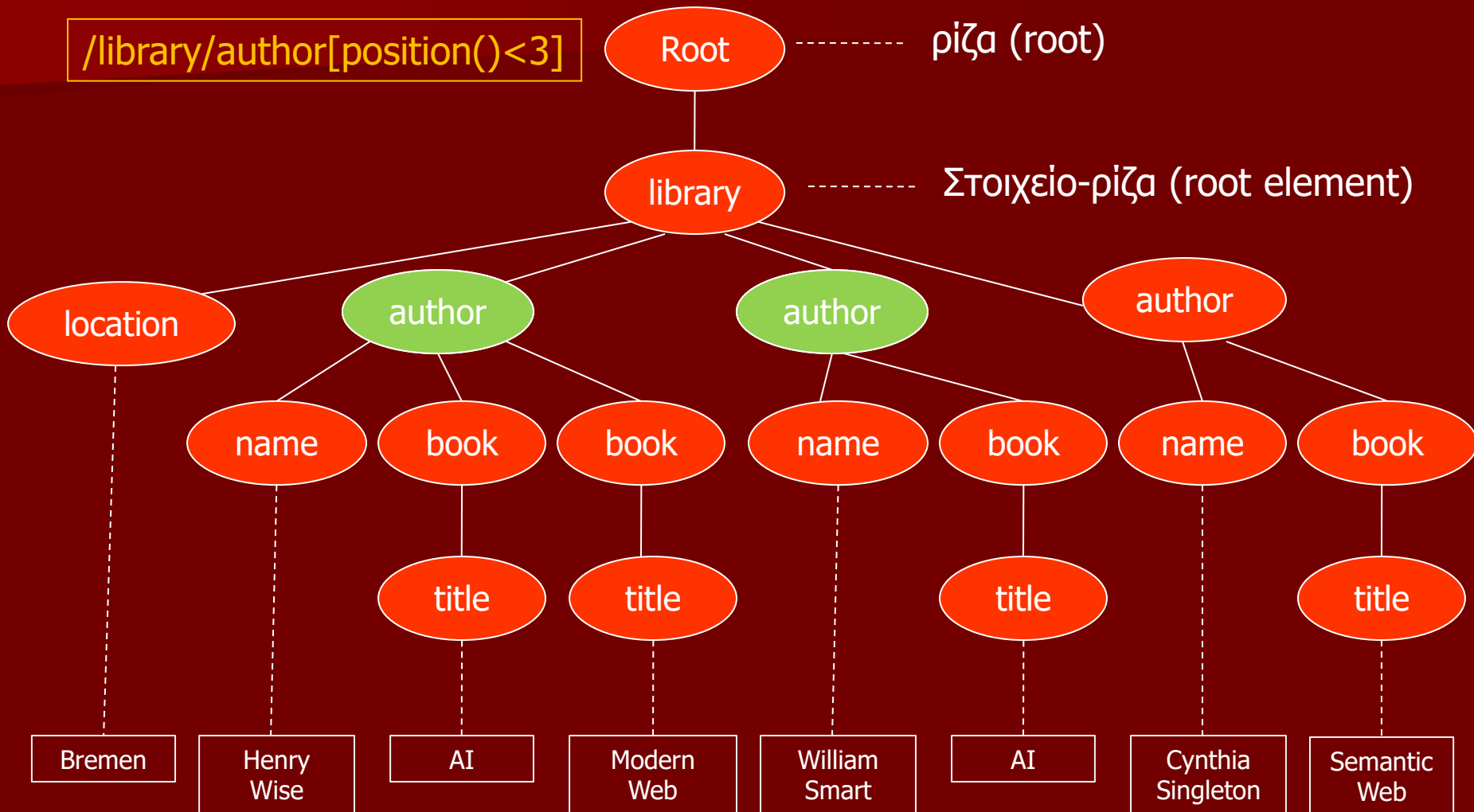
# Δενδρικό Μοντέλο-Παράδειγμα

`/library/author[last()-1]`



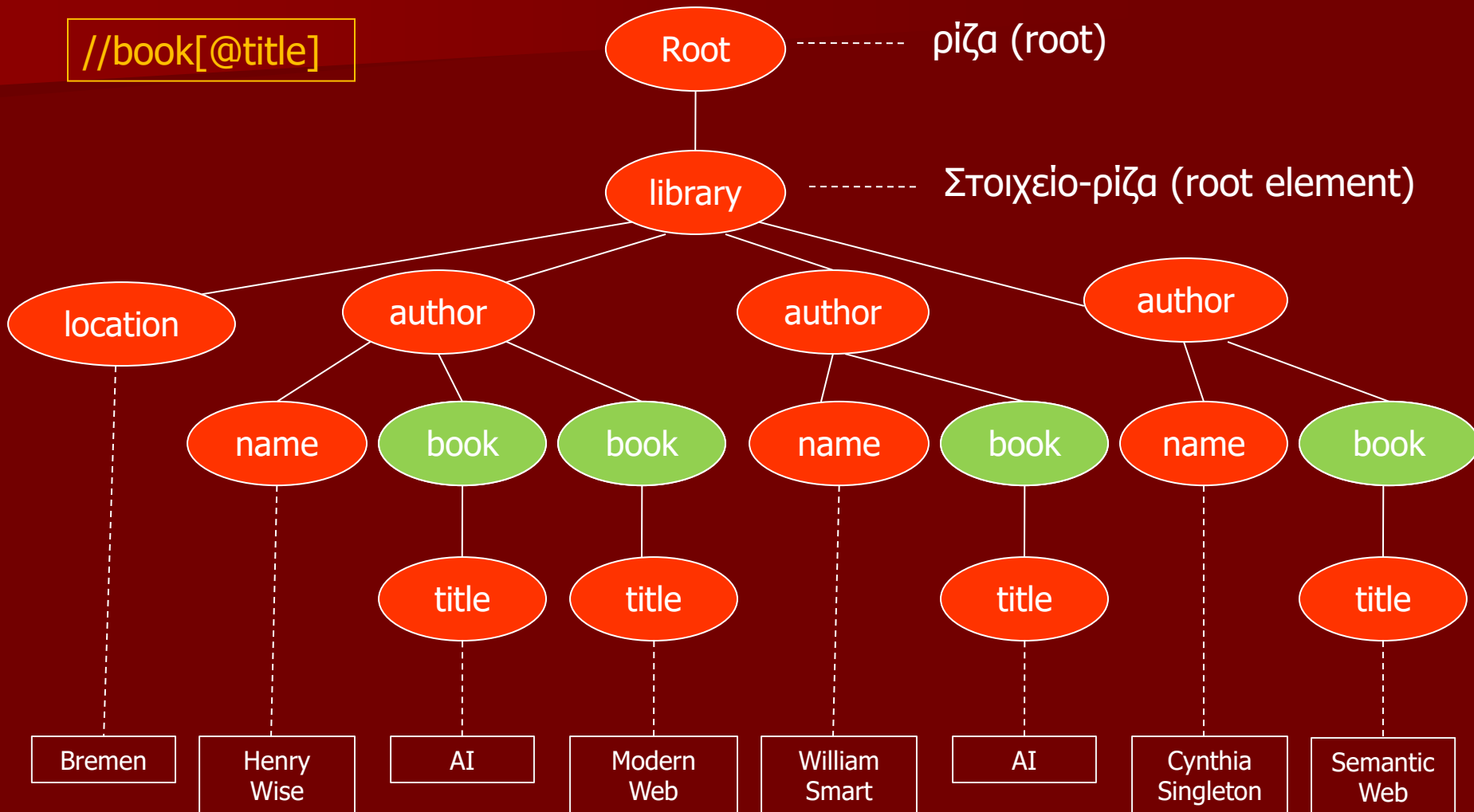
# Δενδρικό Μοντέλο-Παράδειγμα

`/library/author[position()<3]`



# Δενδρικό Μοντέλο-Παράδειγμα

`//book[@title]`



# Δενδρικό Μοντέλο-Παράδειγμα

```
//book[@title="AI"]
```

