

Grundlagen von Datenbanken

Probeklausur

XML, XPath

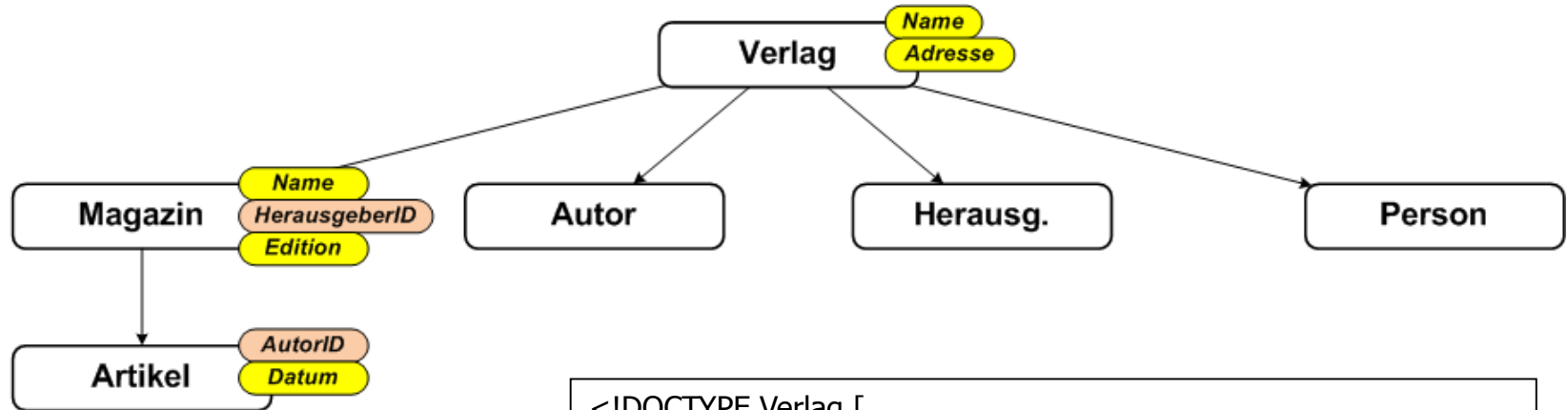


XML und XPath

Gegeben: DTD eines XML-Dokuments.

Tipp: Zunächst einen *Knotenbaum* mit der Struktur des XML-Dokuments erstellen

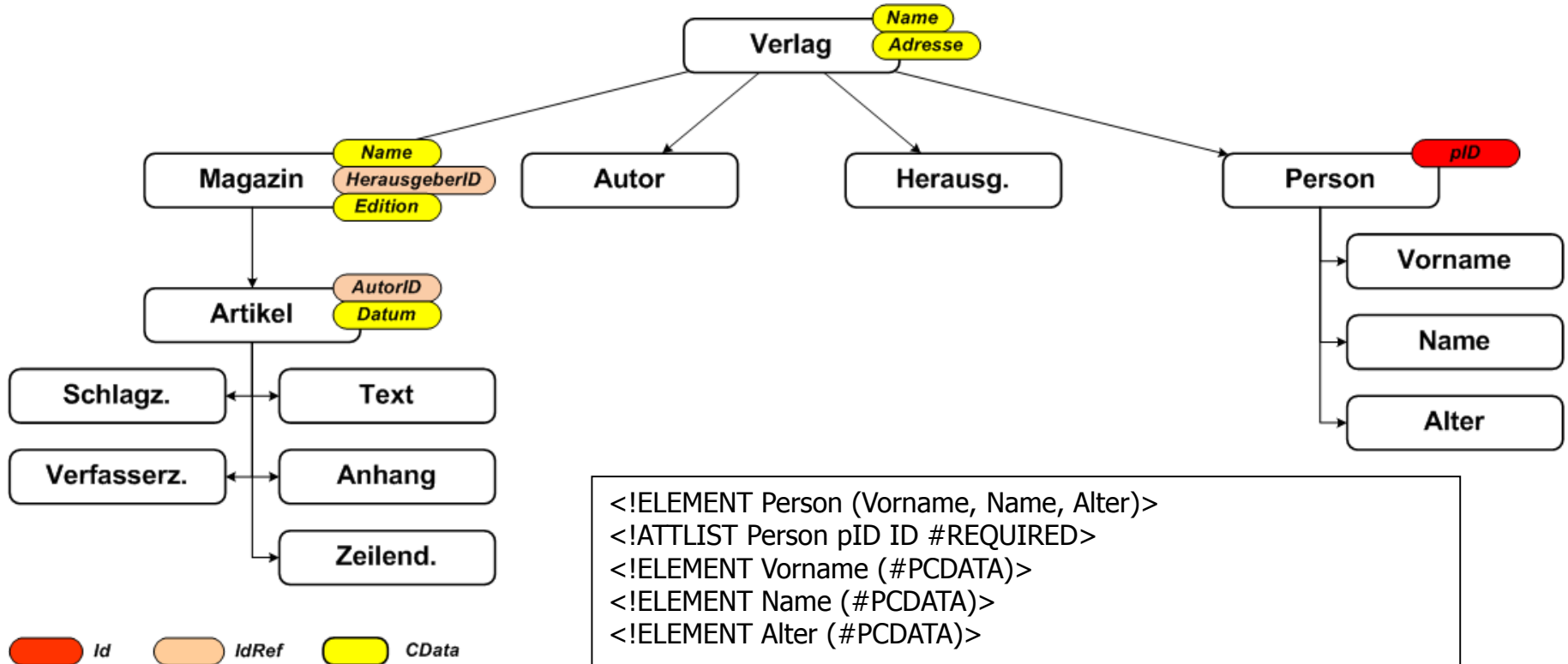
XML und XPath



```
<!DOCTYPE Verlag [  
<!ELEMENT Verlag (Magazin+, Autor+, Herausgeber+, Person+ )>  
<!ATTLIST Verlag Name CDATA #REQUIRED>  
<!ATTLIST Verlag Adresse CDATA #REQUIRED>  
  
<!ELEMENT Magazin (Artikel+)>  
<!ATTLIST Magazin Name CDATA #REQUIRED>  
<!ATTLIST Magazin HerausgeberID IDREF #IMPLIED>  
<!ATTLIST Magazin Edition CDATA #IMPLIED>  
  
<!ELEMENT Artikel (Schlagzeile, Verfasserzeile,  
Zeilendurchschuss, Text, Anhang)>  
<!ATTLIST Artikel AutorID IDREF #REQUIRED>  
<!ATTLIST Artikel Datum CDATA #IMPLIED>
```

 *Id*  *IdRef*  *CData*

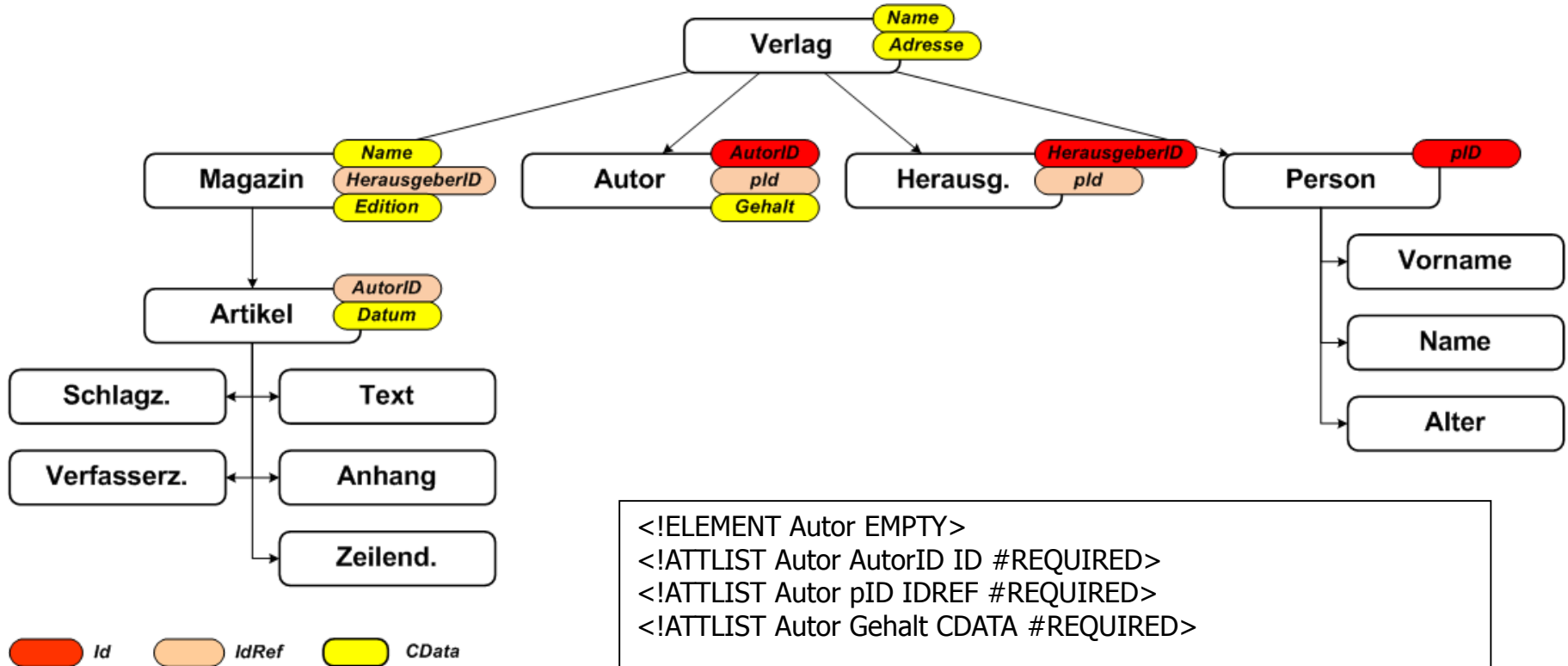
XML und XPath



```
<!ELEMENT Person (Vorname, Name, Alter)>  
<!ATTLIST Person pID ID #REQUIRED>  
<!ELEMENT Vorname (#PCDATA)>  
<!ELEMENT Name (#PCDATA)>  
<!ELEMENT Alter (#PCDATA)>
```

```
<!ELEMENT Schlagzeile (#PCDATA)>  
<!ELEMENT Verfasserzeile (#PCDATA)>  
<!ELEMENT Zeilendurchschuss (#PCDATA)>  
<!ELEMENT Text (#PCDATA)>  
<!ELEMENT Anhang (#PCDATA)>
```

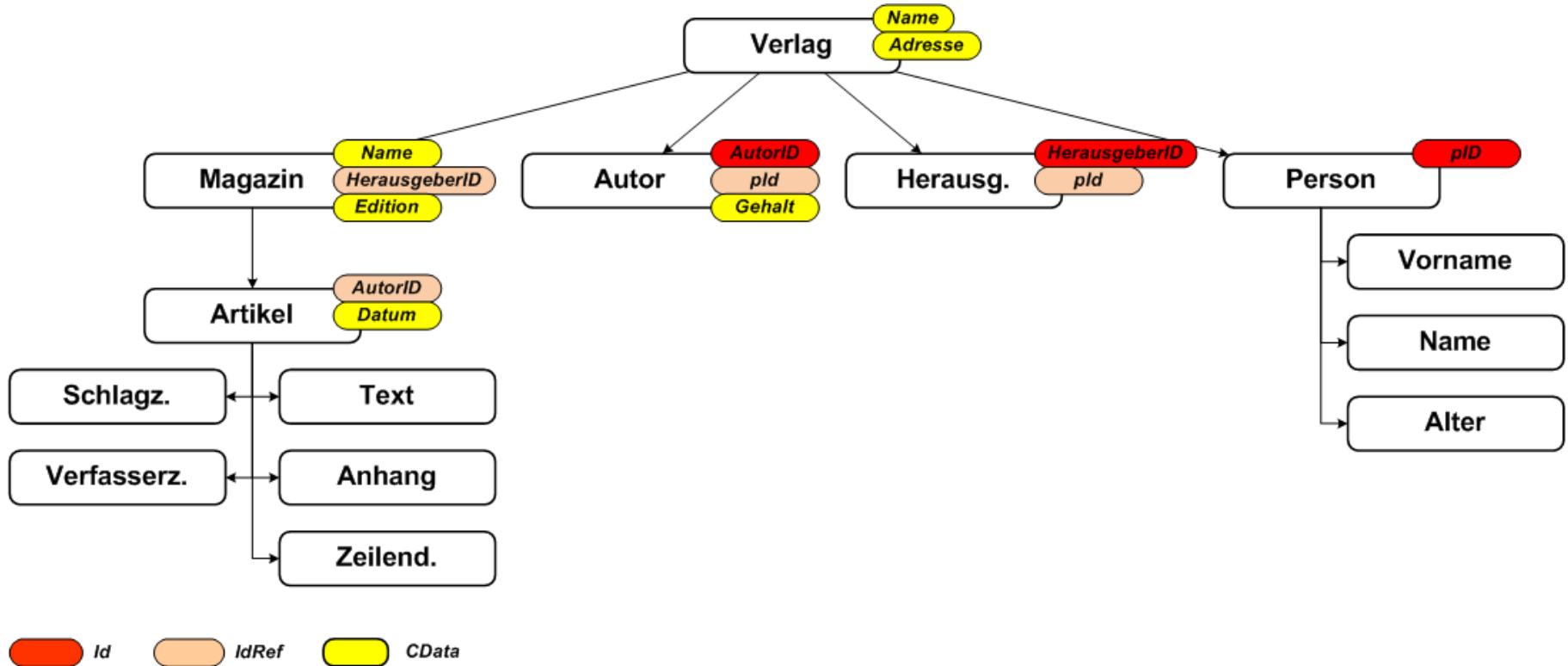
XML und XPath



```
<!ELEMENT Autor EMPTY>
<!ATTLIST Autor AutorID ID #REQUIRED>
<!ATTLIST Autor pID IDREF #REQUIRED>
<!ATTLIST Autor Gehalt CDATA #REQUIRED>
```

```
<!ELEMENT Herausgeber EMPTY>
<!ATTLIST Herausgeber HerausgeberID ID #REQUIRED>
<!ATTLIST Herausgeber pID IDREF #REQUIRED>
]>
```

XML und XPath



XML und XPath

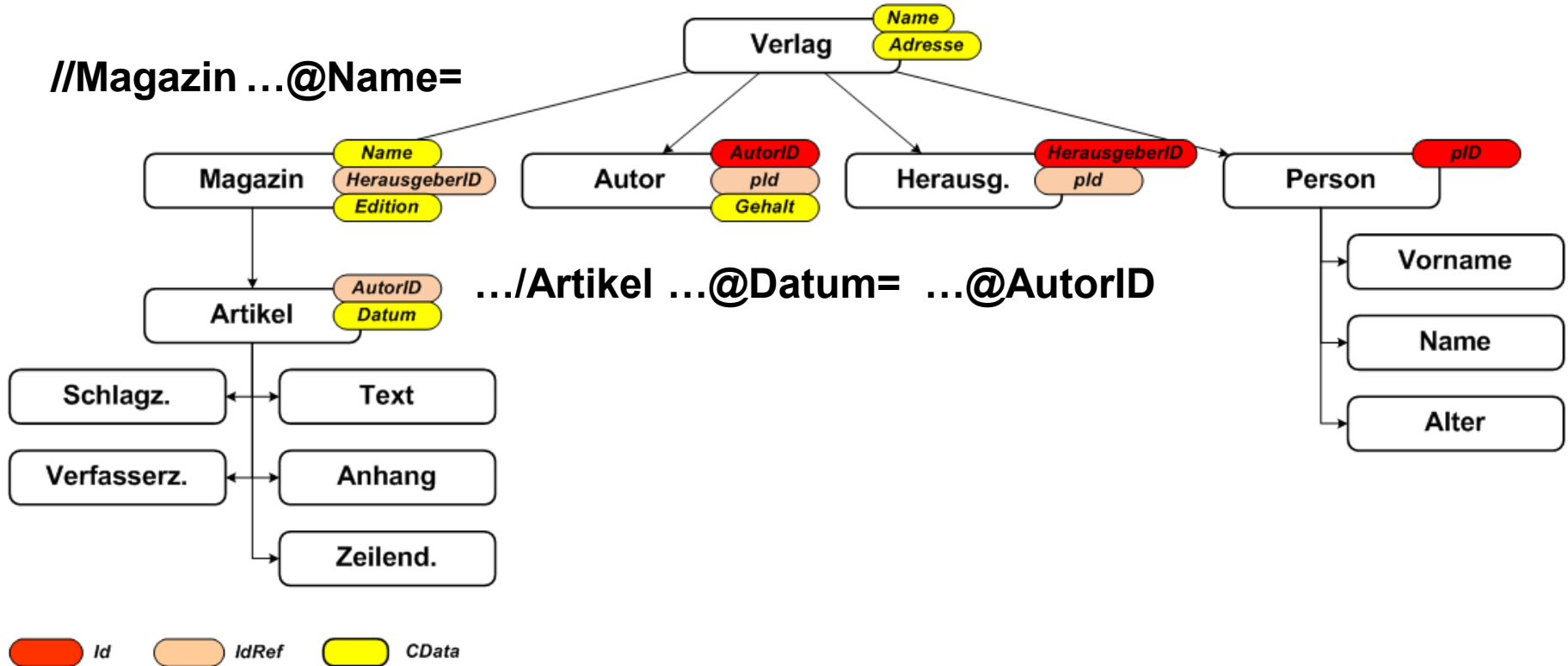
- Pfadausdruck adressiert (selektiert) eine Sequenz von Knoten in einem Dokument
 - besteht aus Schritten, durch "/" voneinander getrennt
 - Bsp.: Namen aller Personen
`/child::Verlag/child::Magazin[attribute::Edition='spezial']/attribute::Name`
 - Verkürzte Schreibweisen für häufig genutzte Konstrukte ("@"= attribute::, "/"= /descendant-or-self::node()/, Achse fehlt= child::, etc.)
 - Bsp.: Namen aller Personen
`//Magazin[@Edition='spezial']/@Name`
 - wird sukzessive, von links nach rechts ausgewertet
 - Pfadanfang: Dokumentwurzel oder von außen vorgegebener Kontext
 - Jeder Schritt geht von Knotensequenz aus
 - Sucht für jeden Knoten in der Sequenz weitere Knoten auf
 - Prädikate (z.B. `[@Edition='spezial']`) sorgen dabei für eine Knotenselektion
 - Leere Resultate führen nicht zu Fehlern

XML und XPath

- Nutzung von Funktionen
 - fn:id = löst IDREF-Referenzen auf und gibt referenzierte Elemente zurück
 - fn:count = zählt die selektierten Elemente

Bsp1: //magazin[@Name = 'Spiegel']/Artikel[@Datum = '2010-03-22']/@AutorID

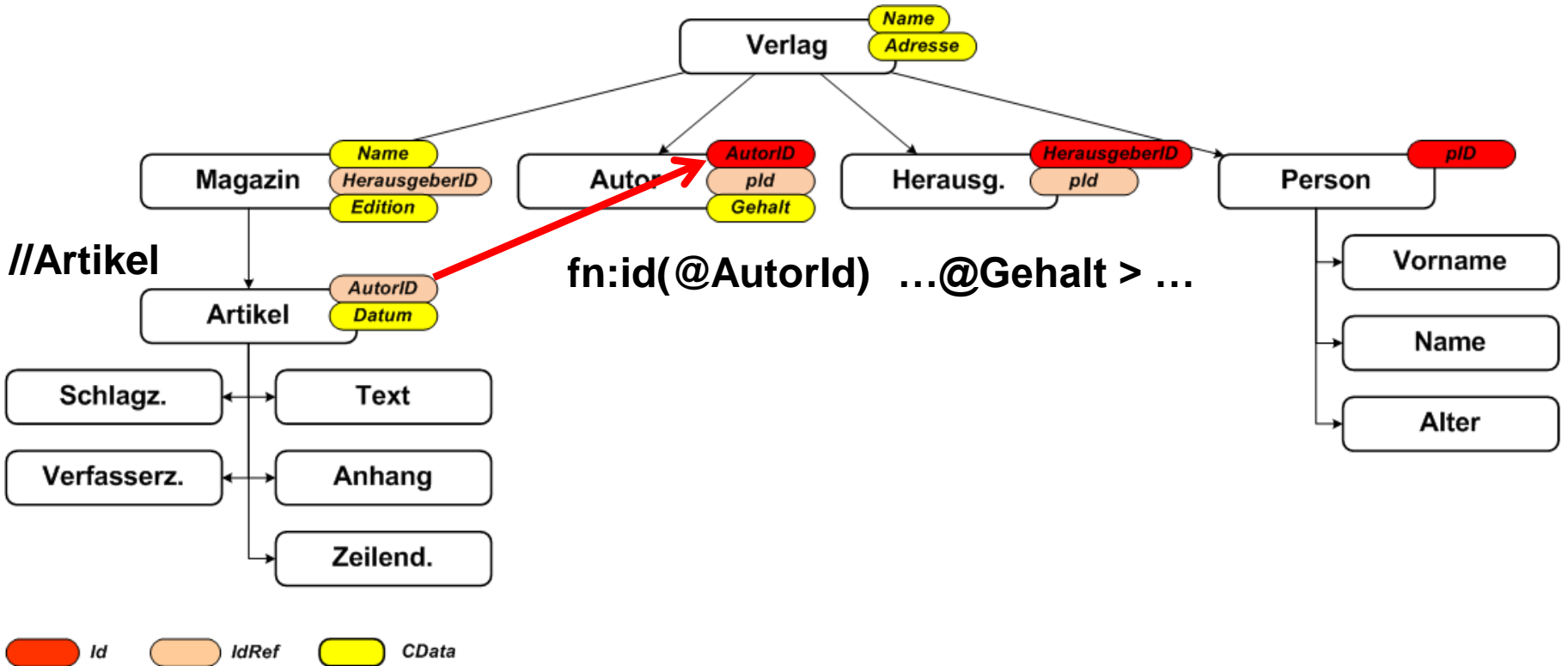
//Magazin ...@Name=



.../Artikel ...@Datum= ...@AutorID

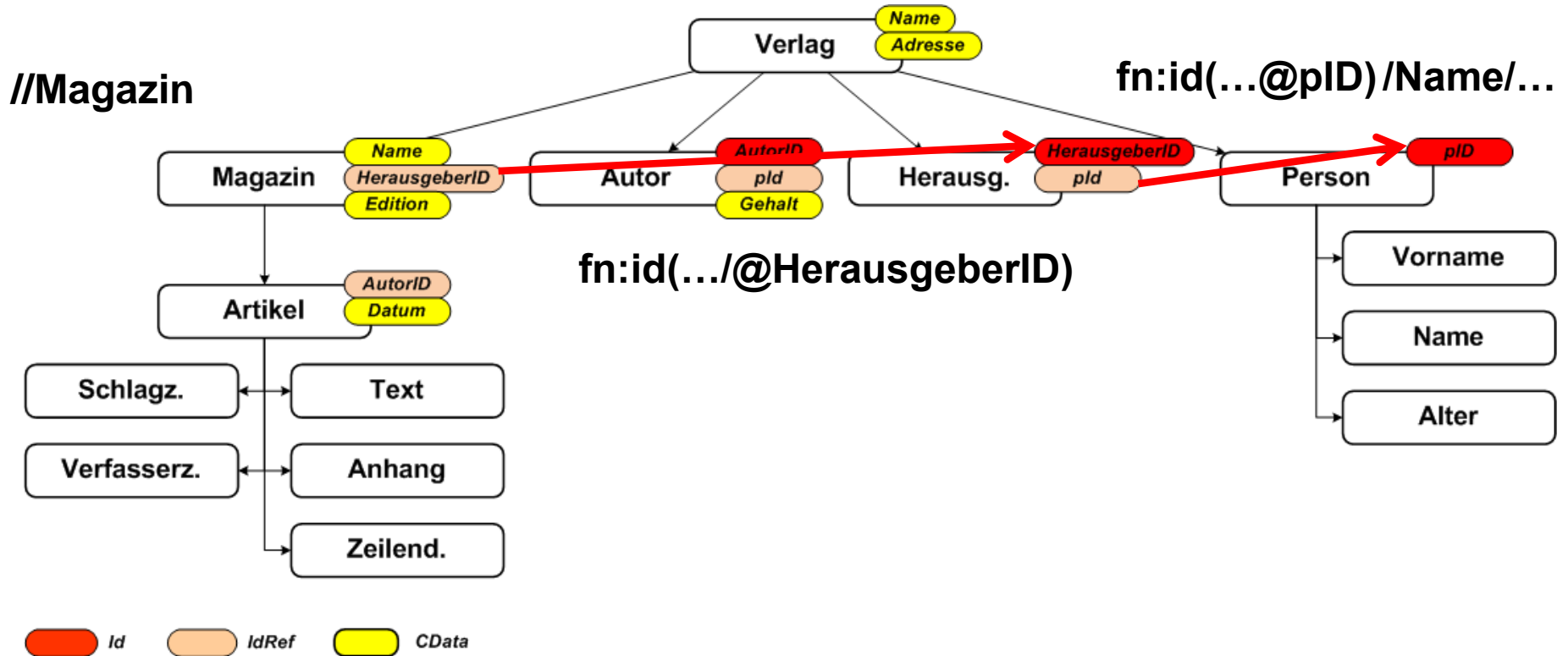
Die IDs aller Autoren, die für das Magazin 'Spiegel' am 22. März einen Artikel verfasst haben.

Bsp2: `fn:count(//Artikel[fn:id(@AutorID)/@Gehalt > 100000])`



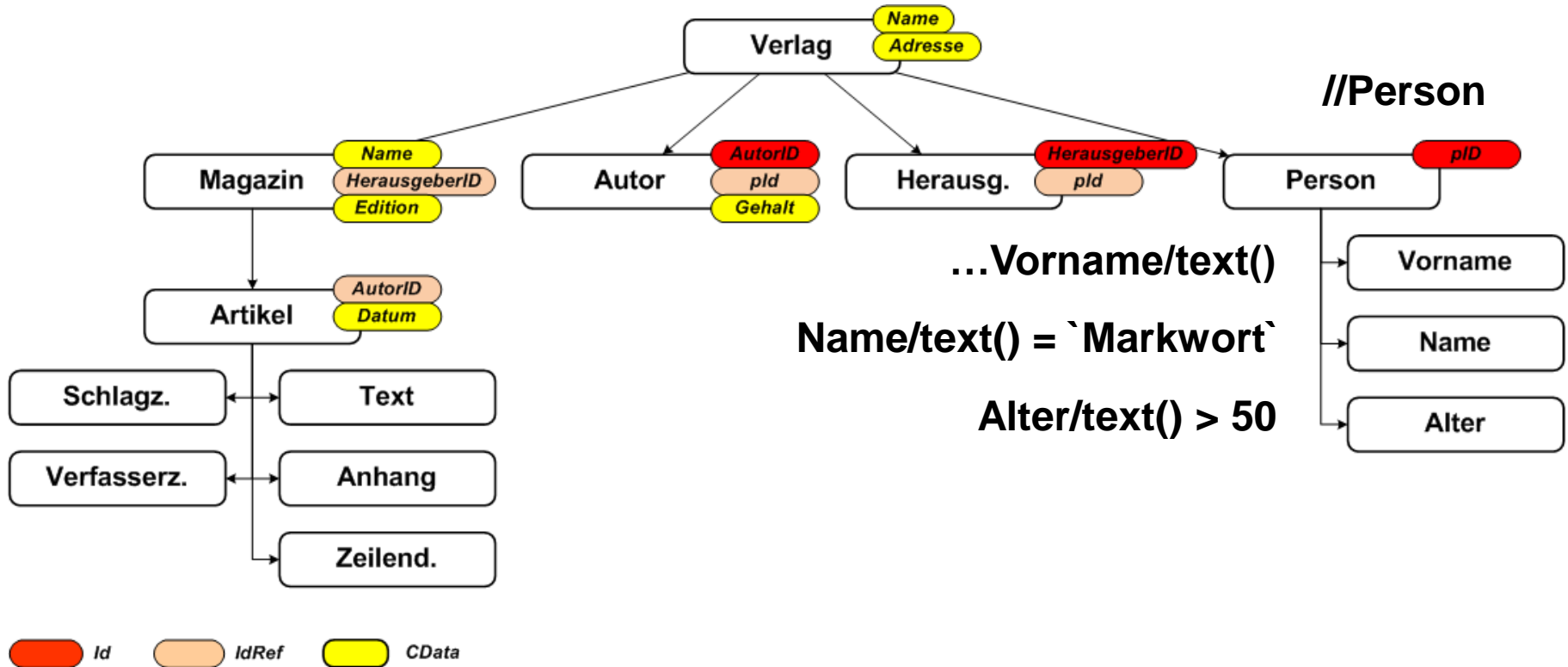
Die Anzahl der Artikel, die von Autoren mit mehr als 100000 Euro Gehalt geschrieben wurden.

Bsp3: `fn:id(fn:id(//Magazin/@HerausgeberID)/@pID)/Name/text()`



Die Namen der Herausgeber von Magazinen.

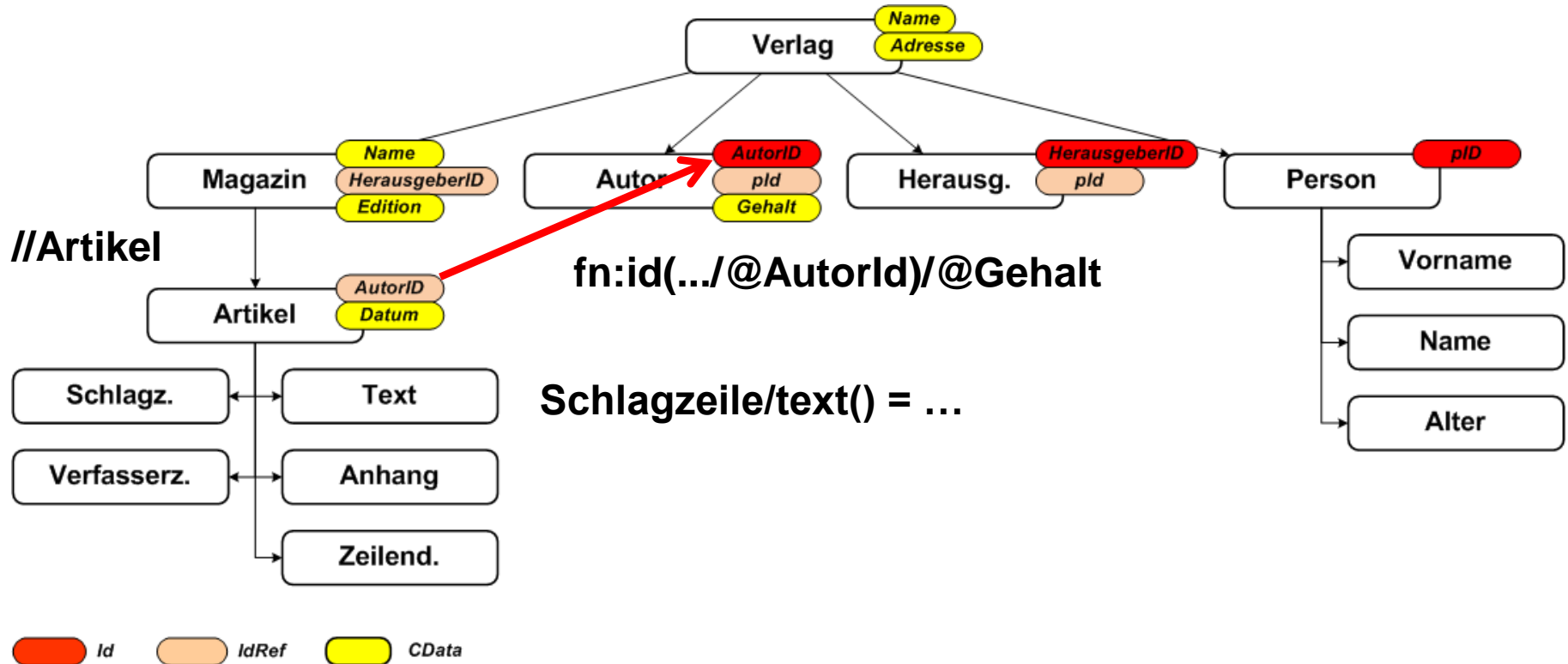
Bsp4: Die Vornamen aller Personen mit dem Namen 'Markwort' älter als 50 Jahre.



...Vorname/text()
 Name/text() = 'Markwort'
 Alter/text() > 50

`//Person[Alter/text() > 50 and Name/text() = 'Markwort']/Vorname/text()`

Bsp5: Die Gehälter von Autoren, die einen Artikel mit der Schlagzeile 'Spätromische Dekadenz!' geschrieben haben.

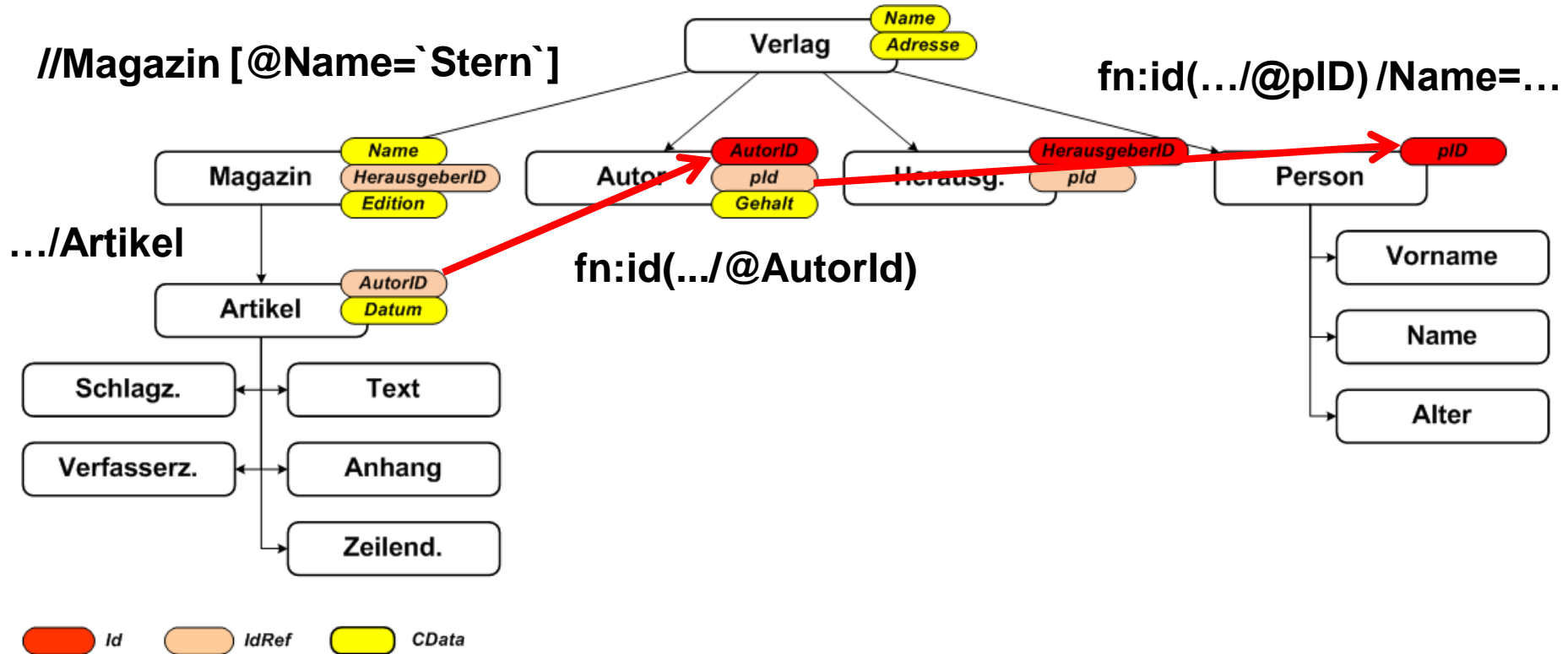


`fn:id(.../@AutorID)/@Gehalt`

`Schlagzeile/text() = ...`

`fn:id(//Artikel[Schlagzeile/text() = 'Spätromische Dekadenz']/@AutorID)/@Gehalt`

Bsp6: Alle Artikel im Magazin „Stern“, die vom Autor mit dem Namen `Thomsen` geschrieben wurden



```
//Magazin[@Name='Stern']/Artikel[ fn:id(fn:id(@AutorID)/@pID)/Name/text() = 'Thomsen' ]
```

Organisatorisches

Klausurtermine

Klausur 1: **Freitag, 21.02.2014**, 10.30-12.30 Uhr, ESA A und ESA B

Klausur 2: Dienstag, 11.03.2014, 10:30-12:30 Uhr, ESA B

Besprechung Probeklausur

Die Besprechung der Probeklausur findet am **31.01.2014** im **Hörsaal A der Chemie** statt.