

DAIMLERCHRYSLER

XML-Namensräume Idee und Umsetzung

Mario Jeckle

DaimlerChrysler Forschungszentrum Ulm

mario.jeckle@daimlerchrysler.com

mario@jeckle.de

www.jeckle.de

Gliederung

I XML-Namensräume: Grundidee

II W3C Namespace Recommendation

III Verwendung von XML-Namensräumen in der Praxis

Grundidee der XML-Namensräume

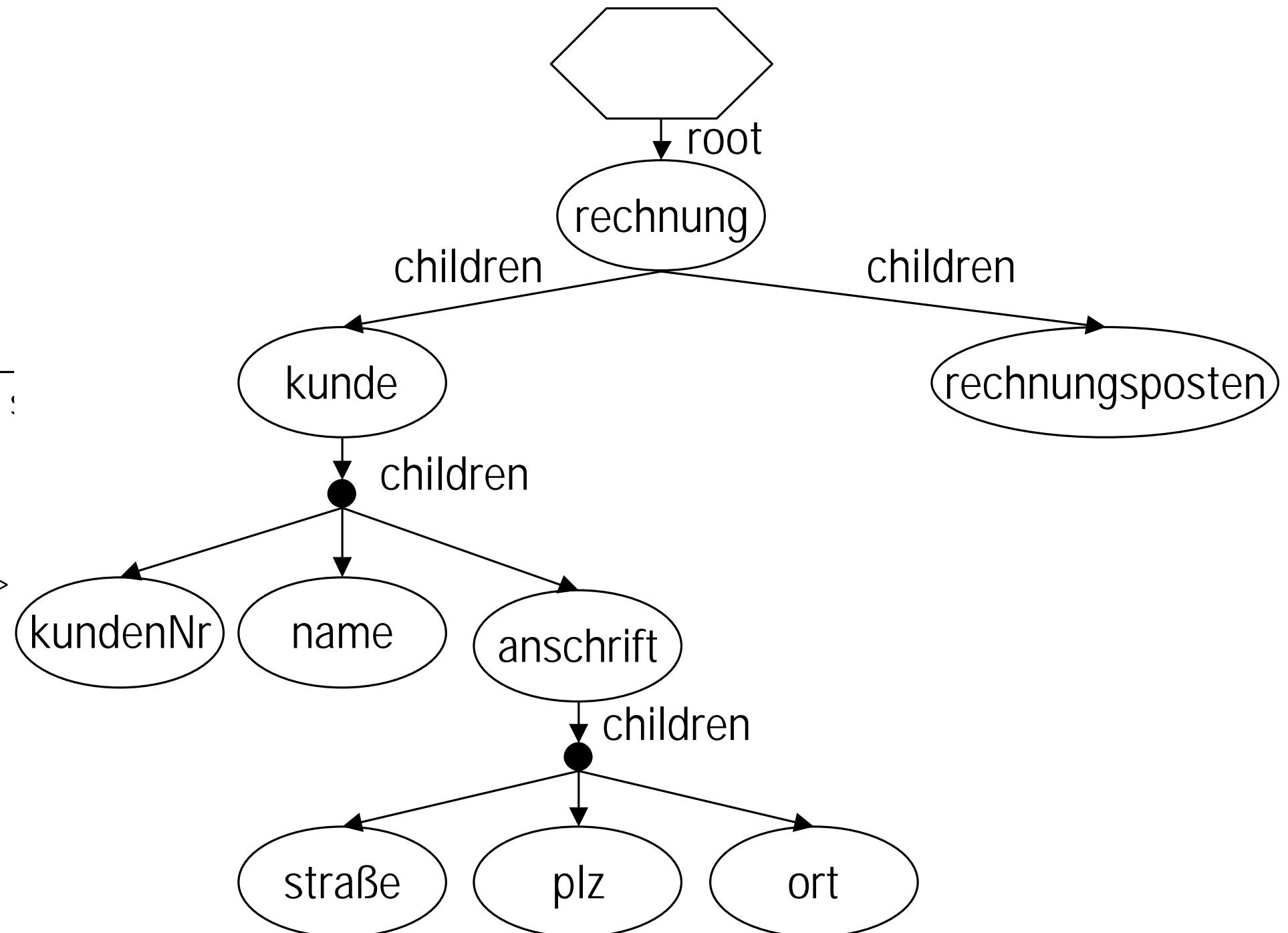
- **Lösung** der prinzipiellen Mehrdeutigkeit von Element- und Attributnamen in verschiedenen Kontexten
- Beibehaltung des **Freiheitsgrades** bei der Erstellung eigener Vokabulare hinsichtlich der Element- und Attributbenennung
- **Ziel:** Möglichst flexibler und einheitlicher Mechanismus zur uneindeutigen Bezeichnung von Elementen- und Attributen ohne zentrale Registrierung

Grundidee der XML-Namensräume

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<rechnung>
  <kunde>
    <kundenNr>4711</kundenNr>
    <name>Max Mustermann</name>
    <anschrift>
      <straße>Musterplatz 1</straße>
      <plz>12345</plz>
      <ort>Musterstadt</ort>
    </anschrift>
  </kunde>
  <rechnungsposten>
    ...
  </rechnung>
```

Grundidee der XML-Namensräume

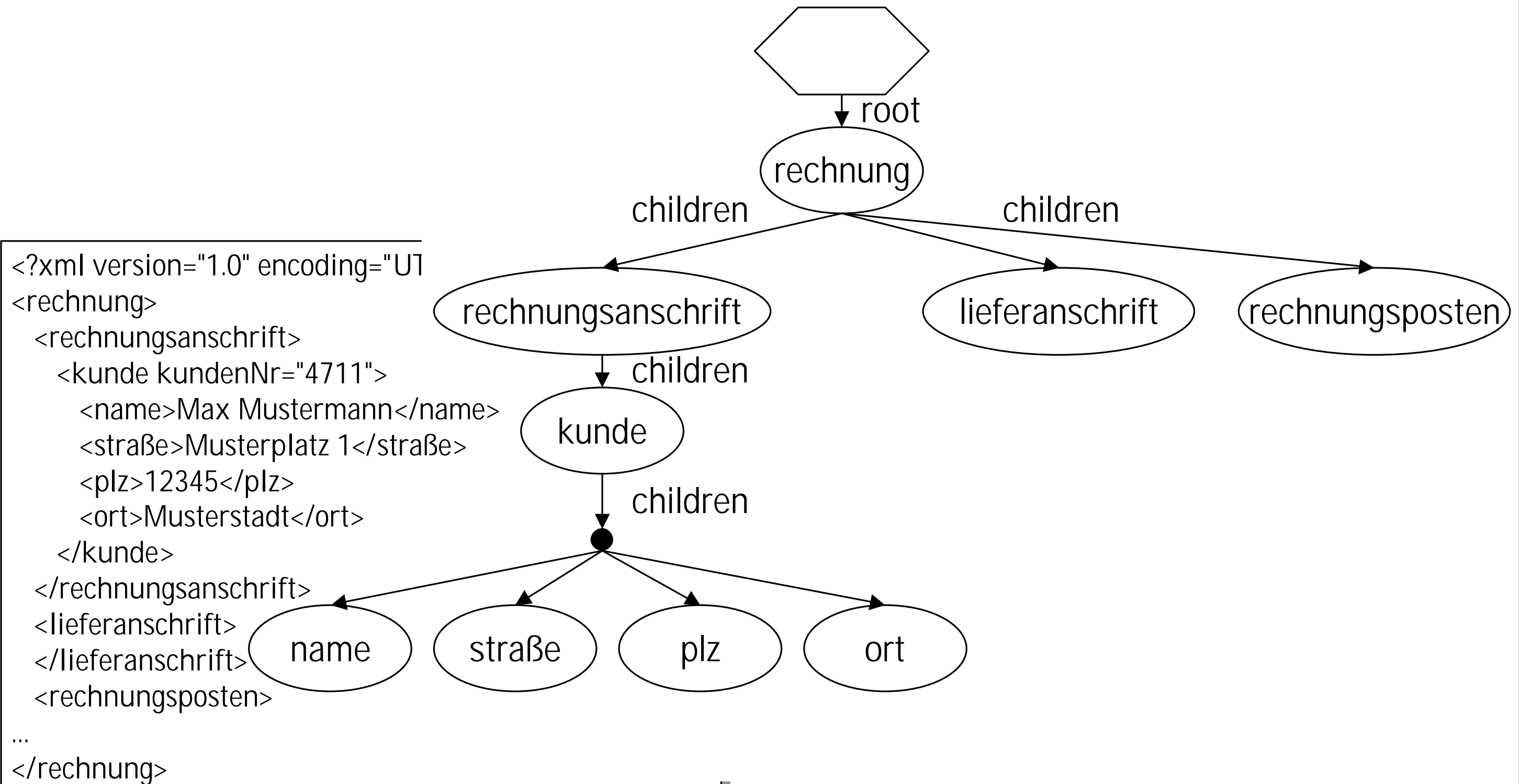
```
<?xml version="1.0" encoding="UTF-8" ?>
<rechnung>
  <kunde>
    <kundenNr>4711</kundenNr>
    <name>Max Mustermann</name>
    <anschrift>
      <straße>Musterplatz 1</straße>
      <plz>12345</plz>
      <ort>Musterstadt</ort>
    </anschrift>
  </kunde>
  <rechnungsposten>
...
</rechnung>
```



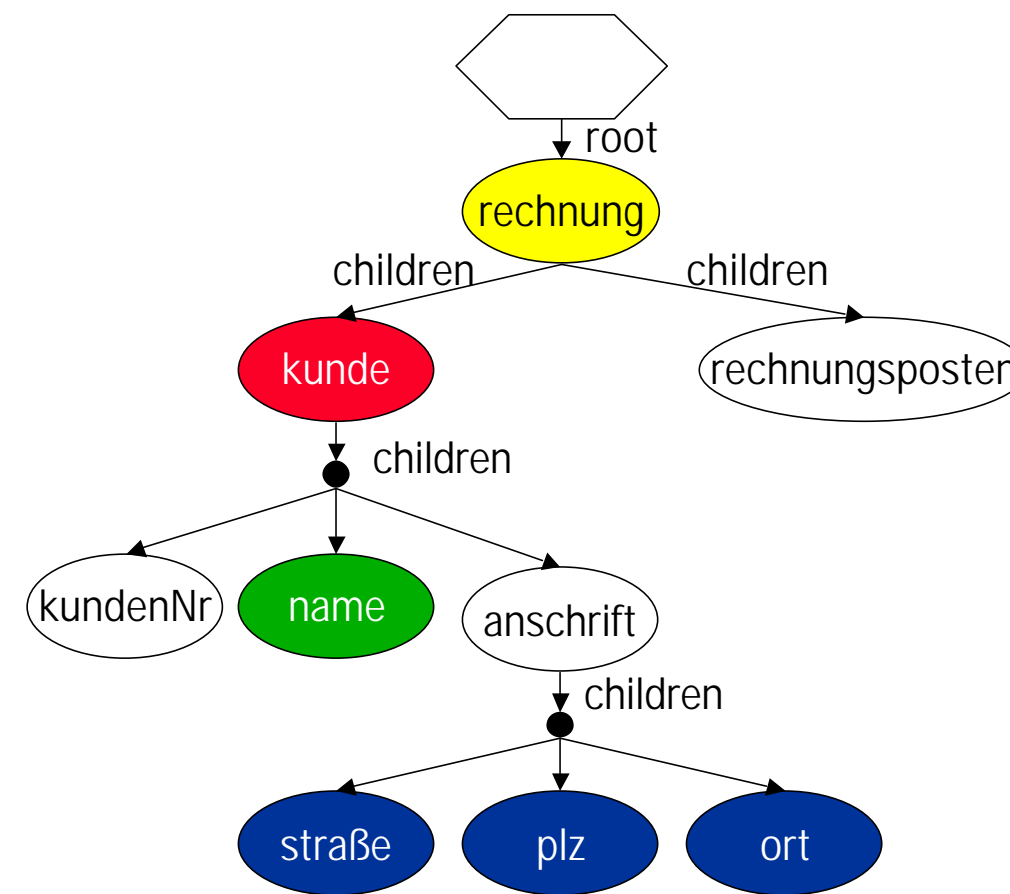
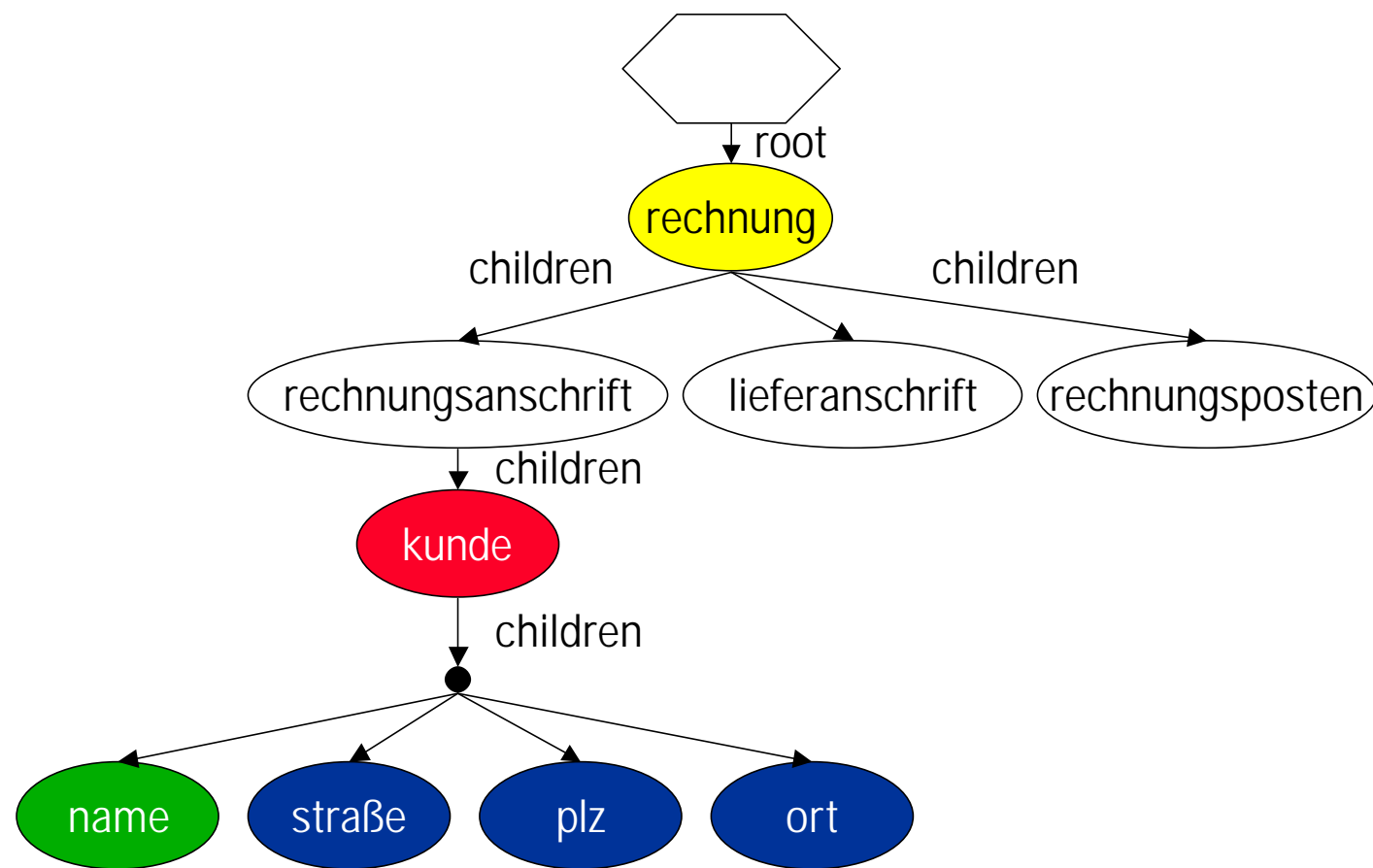
Grundidee der XML-Namensräume

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<rechnung>
  <rechnungsanschrift>
    <kunde kundenNr="4711">
      <name>Max Mustermann</name>
      <straße>Musterplatz 1</straße>
      <plz>12345</plz>
      <ort>Musterstadt</ort>
    </kunde>
  </rechnungsanschrift>
  <lieferanschrift>
  </lieferanschrift>
  <rechnungsposten>
  ...
</rechnung>
```

Grundidee der XML-Namensräume



Grundidee der XML-Namensräume



Grundidee der XML-Namensräume

- Lösungsmöglichkeit 1: (*klassisches* Vorgehen in EDI, STEP, ...) Standardisierung (d. h. normative Festlegung) der Bezeichner
 - Praktisch unmöglich, eine globale Einigung hinsichtlich Semantik und Inhalt zu erzielen
- Lösungsmöglichkeit 2: (ähnlich DNS) Prinzipiell freie Vergabe, bei zwingender zentraler Registrierung
 - Nicht praktikabel hinsichtlich des abzuschätzenden Aufwandes sowie der abzusehenden (Rechts-)Streitigkeiten

W3C Namespace Recommendation

*Namespaces in XML
Working Draft*

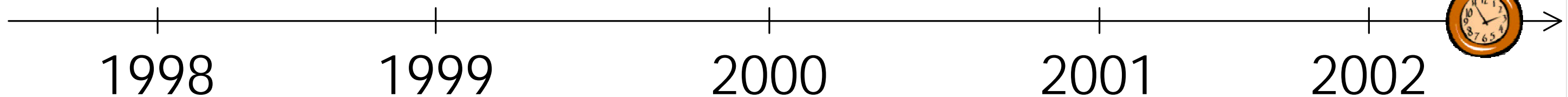
- W3C-Standard; verabschiedet 1999-01-14

*XML Schema
Recommendation*

*XML v1.0 2nd edition
Recommendation*

*Namespaces in XML
Recommendation*

*XML v1.0
Recommendation*



W3C Namespace Recommendation

- W3C-Standard; verabschiedet 1999-01-14
- Namespaceidentifikation durch RFC 2396-konforme URI

Namespaces in XML v1.0:

[Definition:] An XML namespace is a collection of names, identified by a URI reference [RFC2396], which are used in XML documents as element types and attribute names.

[Definition:] URI references which identify namespaces are considered identical when they are exactly the same character-for-character.

W3C Namespace Recommendation

- W3C-Standard; verabschiedet 1999-01-14
- Namespaceidentifikation durch RFC 2396-konforme URI
- Namespaceidentifikatoren nicht notwendigerweise dereferenzierbar

— *Namespaces in XML v1.0:* —

[Definition:] [...] It is not a goal that it be directly usable for retrieval of a schema (if any exists).

W3C Namespace Recommendation

- W3C-Standard; verabschiedet 1999-01-14
- Namespaceidentifikation durch RFC 2396-konforme URI
- Namespaceidentifikatoren nicht notwendigerweise dereferenzierbar
- Namespaceidentifikatoren sollten langfristig existent sein

Namespaces in XML v1.0:

[Definition:] The attribute's value, a URI reference, is the namespace name identifying the namespace. The namespace name, to serve its intended purpose, should have the characteristics of uniqueness and persistence.

W3C Namespace Recommendation

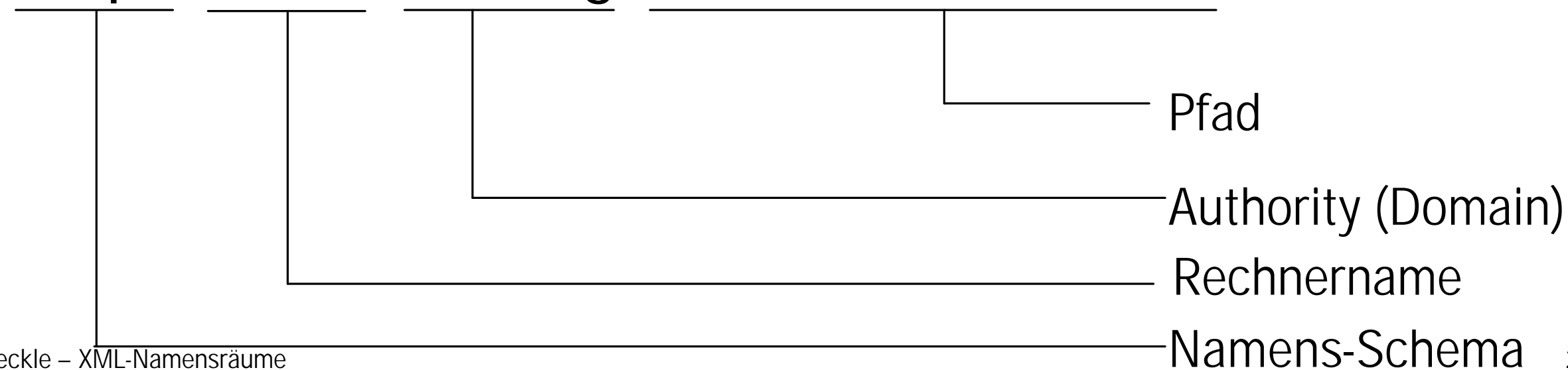
- W3C-Standard; verabschiedet 1999-01-14
- Namespaceidentifikation durch RFC 2396-konforme URI
- Namespaceidentifikatoren nicht notwendigerweise dereferenzierbar
- Namespaceidentifikatoren sollten langfristig existent sein
- In XML 2nd edition und XML v1.1 berücksichtigt
(bis dato nicht integriert)
- Verwendet in allen Vokabularen der *zweiten Generation*
 - XHTML, SVG, XMI
 - XML Schema, XSL-FO, XSLT
 - XLink, MathML, RDF, ...
- Unterstützt durch Parser und XML-APIs
 - SAX, DOM
 - Pull-Parser, Data Binding

W3C Namespace Recommendation

Technische Eigenschaften der Lösung:

- (Wiederbe-)Nutzung der URI als Namenszusatz der Attribut- und Elementbezeichner
- Zentral verwalteter Domain-Anteil garantiert generell Eindeutigkeit
- Dezentral verwalteter Pfad-Anteil garantiert größtmögliche Flexibilität

<http://www.w3.org/2001/XMLSchema>



W3C Namespace Recommendation

Naive Lösung: Erweiterung der Elementnamen

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<http://www.example.com/sales:rechnung>
  <http://www.example.com/sales:kunde>
    <http://www.example.com/sales:kundenNr>4711</http://www.example.com/sales:kundenNr>
    <http://www.example.com/sales:name>Max Mustermann</http://www.example.com/sales:name>
    <http://www.example.com/sales:anschrift>
      <http://www.example.com/sales:straße>Musterplatz 1</http://www.example.com/sales:straße>
      <http://www.example.com/sales:plz>12345</http://www.example.com/sales:plz>
      <http://www.example.com/sales:ort>Musterstadt</http://www.example.com/sales:ort>
    </http://www.example.com/sales:anschrift>
  </http://www.example.com/sales:kunde>
  <http://www.example.com/sales:rechnungsposten>
  ...
</http://www.example.com/sales:rechnung>
```


W3C Namespace Recommendation

Naive Lösung: Erweiterung der Elementnamen

- Umständliche Notation, senkt Lesbarkeit dramatisch
- Vergrößert Speicherplatzaufwand signifikant
(im Beispiel rund 150%!)
- Problematisch hinsichtlich Domain- oder Pfadänderungen
- Verstoß gegen XML-Syntax;
Pfadtrenner „/“ (U+002F) ist innerhalb Elementnamen nicht zugelassen!

Erweiterung von Elementnamen durch URIs

Verstoß gegen XML-Syntax;

Pfadtrenner „/“ (U+002F) ist innerhalb Elementnamen nicht zugelassen!

XML v1.1 Candidate Recommendation 2002-10-15:

[5]	Name	::=	NameStartChar NameChar*	<i>XML v1.0 2nd ed:</i>
[4a]	NameChar	:=	NameStartChar "-" "." [0-9] #xB7 [#x0300-#x036F] [#x203F-#x2040]	} > #x002F
[4]	NameStartChar	:=	":" [A-Z] "_" [a-z] [#xC0-#x2FF] [#x370-#x37D] [#x37F-#x1FFF] [#x200C-#x200D] [#x2070-#x218F] [#x2C00-#x2FEF] [#x3001-#xD7FF] [#xF900-#xEFFFF]	

Erweiterung von Elementnamen durch URIs

Verstoß gegen XML-Syntax;

Pfadtrenner „/“ (U+002F) ist innerhalb Elementnamen nicht zugelassen!

XML v1.0 2nd edition:

[5]	Name	::=	(Letter '_' ':') (NameChar)*	<i>XML v1.1:</i>
[4]	NameChar	::=	Letter Digit '.' '-' '_' ':' CombiningChar Extender	
[84]	Letter	::=	BaseChar Ideographic	
[85]	BaseChar	::=	[#x0041-#x005A] ...	} > #x002F
[86]	Ideographic	::=	[#x4E00-#x9FA5] ...	
[88]	Digit	::=	[#x0030-#x0039] ...	
[87]	CombiningChar	::=	[#x0300-#x0345] ...	
[89]	Extender	::=	#x00B7 ...	

W3C Namespace Recommendation

Standardisierte Lösung: Zweistufiges Vorgehen:

1. Bindung einer URI an ein frei wählbares Präfix durch ein standardisiertes Attribut

```
<rechnung xmlns:rch="http://www.example.com">
```

2. Qualifizierung der Element- oder Attributnamen mit dem definierten Präfix

```
<rch:rechnung xmlns:rch="http://www.example.com">
```

```
<rch:kunde>
```

```
<rch:kundenNr>4711</rch:kundenNr>
```

```
</rch:kunde>
```

...

W3C Namespace Recommendation

Standardisierte Lösung: Zweistufiges Vorgehen:

1. Oder: Überschreiben des Vorgabenschemas

```
<rechnung xmlns="http://www.example.com">
```

2. Propagierung an alle Kindelemente ohne explizites Namensraumpräfix oder redeklariertem Namensraum

```
<rechnung xmlns="http://www.example.com">
```

```
<kunde>
```

```
<kundenNr>4711</kundenNr>
```

```
</kunde>
```

```
...
```

Alle Elemente sind im
Namensraum
<http://www.example.com>

Verwendung von XML-Namensräumen in der Praxis

- URLs (häufigst genutzte Variante)
 - <http://www.w3.org/2001/XMLSchema>
 - <http://www.w3.org/1999/XSL/Transform>
 - <http://www.w3.org/2000/09/xmlidsig#>
- URNs (selten genutzte Variante)
 - <urn:schemas-microsoft-com:office:word>
 - <urn:nbn:fi-fe19991055>
 - <urn:nbn:fi-fea-5c5875e6e49ae649cad63e5ee4f6c346>
 - <urn:oasis:names:specification:docbook:dtd:xml:4.1.2>

Verwendung von XML-Namensräumen in der Praxis

Alternative Identifikationsschemata:

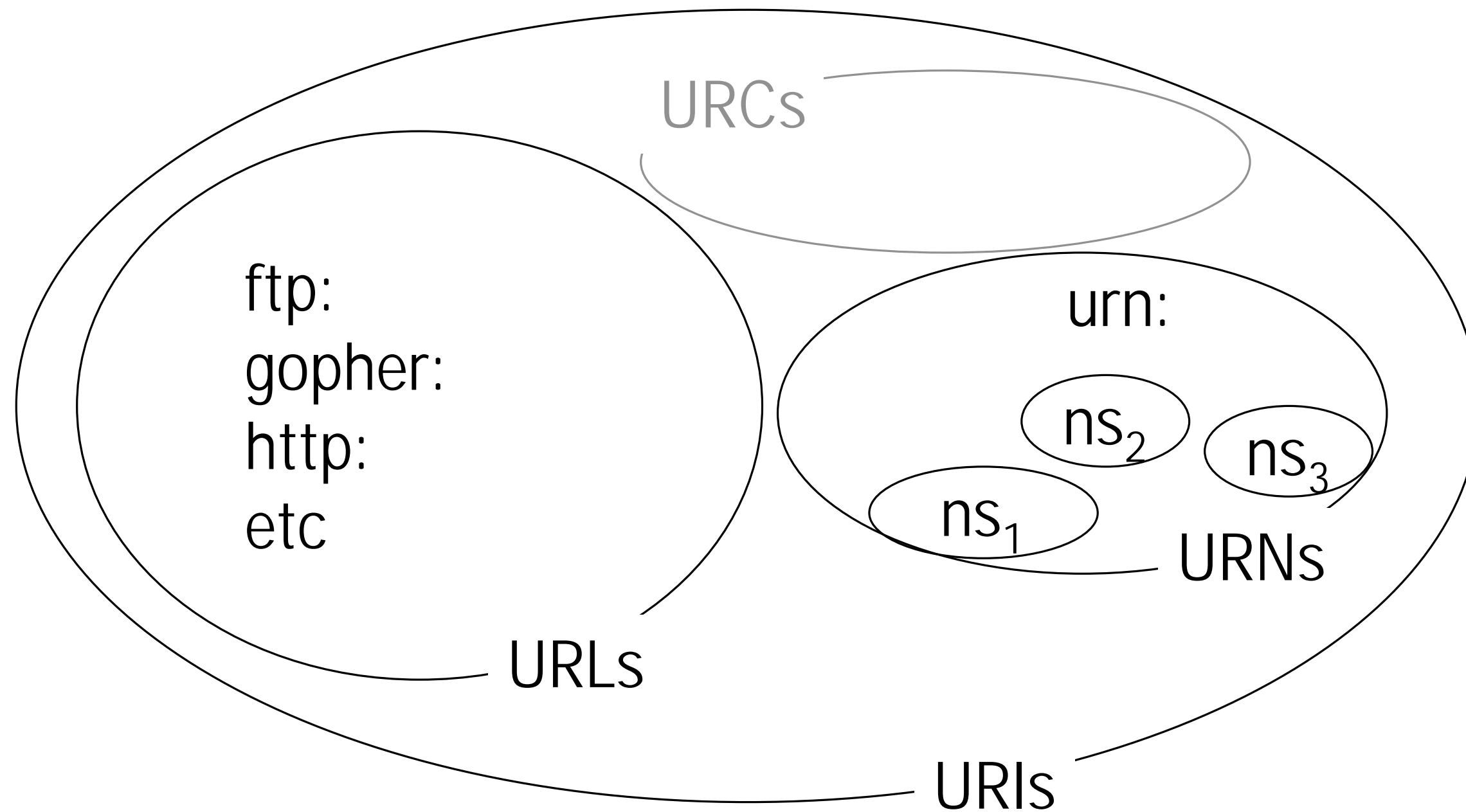
- **OASIS** (IETF RFC 3121)
 - urn:oasis:names:specification:docbook:dtd:xml:4.1.2
 - urn:oasis:names:tc:docbook:dtd:xml:docbook:5.0b1
 - urn:oasis:names:technical:memo:9502:1995
 - urn:oasis:member:A00024:x
- **Personal Internet Name** (IETF RFC 3043)
 - URN:pin:bs4321234
 - URN:pin:324kj5hkj45
 - URN:pin:mm2136
- **IETF** (IETF RFC 2648)
 - urn:ietf:rfc:2141
 - urn:ietf:std:50
 - urn:ietf:id:ietf-urn-ietf-06
 - urn:ietf:mtg:41-urn

Verwendung von XML-Namensräumen in der Praxis

Alternative Identifikationsschemata:

- NewsML (IETF RFC 3085)
 - urn:newsml:iptc.org:20001006:NewsMLv1.0:1
 - urn:newsml:reuters.com:20000206:IIMFFH05643_2000-02-06_17-54-01_L06156584:1U
- International Serial Standard Number (IETF RFC 3044)
 - urn:ISSN:1046-8188
- Object Management Group (UML, XMI)
 - org.omg/standards/UML
 - <http://www.omg.org/XMI>

Exkurs: URIs, URLs, URNs ...



Verwendung von XML-Namensräumen in der Praxis

- Warum URLs?
 - Ubiquitär => Gewisser Vertrautheitsgrad
 - Leicht zu erstellen
 - Vorbildfunktion des W3C

Namespaces in XML v1.0:

[Definition:] The attribute's value, a URI reference, is the namespace name identifying the namespace. The namespace name, to serve its intended purpose, should have the characteristics of uniqueness and persistence. [...]

An example of a syntax that is designed with these goals in mind is that for Uniform Resource Names [RFC2141].

However, it should be noted that ordinary URLs can be managed in such a way as to achieve these same goals.

Verwendung von XML-Namensräumen in der Praxis

Namespaceidentifikator \Leftrightarrow URL?!

Pro:

- URL-Syntax und Semantik ist vertraut
- URL-Syntax wird verstanden
- (inzwischen) quasi-intuitive Pragmatik
- Schema wäre lokalisierbar

Kontra:

- URL-Syntax wird (miß-)verstanden
- URL-Semantik ist nicht identisch mit der Absicht der XML-Namensräume
- Pragmatik nicht durch XML-Namensraumspezifikation gedeckt
- Schema muß nicht zwingend existieren, oder eineindeutig an einen Namensraum gekoppelt sein (zu XHTML existieren drei normative Schemata!)
- URL-Dereferenzierung von XML-Namensräumen gegenwärtig höchst inkonsistent umgesetzt

Verwendung von XML-Namensräumen in der Praxis

Anforderungen an einen XML-Namensraum-adäquaten Identifikationsmechanismus:

- Einfach zu benutzen wie URLs
 - Einfach zu gewährleistende weltweite Eindeutigkeit
 - Dezentrale Vergabe
- Keine semantischen und pragmatischen Konflikte
- Unabhängigkeit von physischer Repräsentation und Lokation

Verwendung von XML-Namensräumen in der Praxis

Vorschlag einer Initiative von BMW, SAP, Siemens, Mozquito und DaimlerChrysler:

Syntax:

`nsd://domain/date/[version/]freePath/`

Beispiel:

`nsd://w3.org/2003-01-01/1.0/foo/bar/...`

`nsd://daimlerchrysler.com/baz/...`

Anmerkungen:

- *domain* enthält RFC 2396-konform nur die second und top level domain (z.B. example.com nicht: www.example.com)
- *date* gemäß ISO 8601
- *freePath* sollte in absteigender Signifikanz organisiert sein und keine Organisationsstrukturen o.ä. abbilden