



# Bundestags-Plenar-Protokolle im XML-Format: Aufbau der Strukturdefinition – DTD

## Inhaltsübersicht:

1.	Im Prozess beteiligte Informationen/Dateien . . . . .	2
2.	Das Wurzelement <dbtplenarprotokoll>. . . . .	3
3.	Der Hauptabschnitt <vorspann> . . . . .	4
3.1	Das Element <kopfdaten> . . . . .	4
3.2	Element <inhaltsverzeichnis>. . . . .	5
4.	Elemente im Sitzungsverlauf . . . . .	6
4.1	Element <sitzungsverlauf> . . . . .	6
4.2	Die Elemente <sitzungsbeginn> und <sitzungsende> . . . . .	6
4.3	Das Element <tagesordnungspunkt> . . . . .	6
5.	Anlagen zum Plenarprotokoll . . . . .	7
5.1	Element <anlagen> . . . . .	7
5.2	Element <anlage> . . . . .	7
6.	Rednerliste, Redner, Namen und Reden . . . . .	8
6.1	Element <rednerliste>. . . . .	8
6.2	Element <redner>. . . . .	8
6.3	Das Element <name>. . . . .	9
6.4	Das Element <rede>. . . . .	10
7.	Elemente für Textstrukturen . . . . .	11
7.1	Das Element <p> . . . . .	11
7.2	Das Element <zitat> . . . . .	11
7.3	Das Element <kommentar> . . . . .	11
7.4	Tabellen-Elemente . . . . .	12
8.	Die Inline-Elemente . . . . .	13
8.1	Element <a> . . . . .	13
8.2	Element <xref> . . . . .	14
8.3	Element <sub> . . . . .	14
8.4	Element <sup> . . . . .	14
8.5	Element <fussnote> . . . . .	14

## 1. Im Prozess beteiligte Informationen/Dateien

Mit dem Ziel, möglichst genau die Informationen aus den Plenarprotokollen zu extrahieren und sie anschließend differenziert zuordnen zu können, fließen in jede XML-Version eines Plenarprotokolls Informationen aus folgenden Quellen ein:

1. Plenarprotokolle mit dem gedruckten Inhalt.
2. Liste der Redner. Sie enthält aktuelle Informationen zu allen Rednern mit eindeutiger ID für:
  - Abgeordnete
  - Bundesminister
  - Staatssekretäre
  - Minister der Bundesländer
  - ▶ Vom Referat IT-Systementwicklung – IT4 – wird die Liste der aktuellen Abgeordneten zur Verfügung gestellt. Diese Datei enthält Titel, Namen, Ortszusatz, Fraktions-Zugehörigkeit und eine vom System erzeugte, eindeutige ID für jeden Abgeordneten.
  - ▶ Vom Referat Stenografischer Dienst – PD3 – werden Änderungsmitteilungen geliefert, anhand derer die Rednerliste aktualisiert wird.
  - ▶ Die Abgeordneten-Daten werden in eine XML-Datei eingepflegt und um die Angaben zu Bundesministern, Staatssekretären und den Ministern der Bundesländer ergänzt. Eindeutige IDs für diese Personen werden bei der Erfassung direkt zugeordnet.
  - ▶ Bei einigen Personen – den Bundestagspräsidenten und Bundesministern – werden zusätzlich noch Informationen zu ihrer Rolle ergänzt und gepflegt.
  - ▶ Siehe auch “Element <redner>” auf Seite 8.



### Fehlende Angaben zu Rednern und Reden

Wenn Personen vor dem Plenum reden, die als Redner nicht in der Redner-Liste enthalten sind, kann aktuell dafür keine Rede getagged werden.

- ▶ Für die Verarbeitung der xml-basierten Plenarprotokolle ist die Aktualität der Liste mit den Informationen zu den Rednern wichtig.
-

## 2. Das Wurzelement <dbtplenarprotokoll>

### a.) Element-Definition:

```
<!ELEMENT dbtplenarprotokoll (vorspann, sitzungsverlauf, anlagen, rednerliste) >
<!ATTLIST dbtplenarprotokoll
wahlperiode          CDATA          #REQUIRED
sitzung-nr           CDATA          #REQUIRED
sitzung-datum        CDATA          #REQUIRED
sitzung-start-uhrzeit CDATA          #REQUIRED
sitzung-ende-uhrzeit CDATA          #REQUIRED
sitzung-naechste-datum CDATA          #REQUIRED
start-seitennr       CDATA          #REQUIRED

sitzung-ort          CDATA          #FIXED      "Berlin"
herausgeber          CDATA          #FIXED      "Deutscher Bundestag"
herstellung          CDATA          #FIXED
"H. Heenemann GmbH Co., Buch- und Offsetdruckerei, Bessemerstraße 83-91,
12103 Berlin, www.heenemann-druck.de"
vertrieb             CDATA          #FIXED
"Bundesanzeiger Verlagsgesellschaft mbH, Postfach 1 0 05 34, 50445 Köln,
Telefon (02 21) 97 66 83 40, Fax (02 21) 97 66 83 44, www.betrifft-gesetze.de"
issn                 CDATA          #FIXED      "0722-7980"

version              CDATA          #IMPLIED
status               (erfasst | korrektur | freigegeben) #IMPLIED
kommentare           CDATA          #IMPLIED >
```

### Beispiel:

```
<dbtplenarprotokoll
    wahlperiode="18"
    sitzung-nr="97"
    sitzung-datum="26.03.2015"
    sitzung-start-uhrzeit="9:01"
    sitzung-ende-uhrzeit="20:37"
    sitzung-naechste-datum="27.03.2015"
    start-seitennr="9153">
```

**Bedeutung:** Im Element wurden Attribute hinterlegt, die drei Kategorien zuzuordnen sind:

- Attribute zur eindeutigen Identifikation eines Protokolls
- Attribute, die in jeder gedruckten Ausgabe als unveränderlicher Text erscheinen
- Attribute, die für eine mögliche Weiterverarbeitung den Prozess-Status und eine Version angeben
- Attribut zum Hinterlegen eines Kommentars

Das Inhaltsmodell des Wurzelements ist feststehend. Alle als #REQUIRED angegebenen Elemente müssen vorhanden sein, um die Validität zu sichern.

Zu Attributen zur eindeutigen Identifikation eines Protokolls:

wahlperiode	Beschreibt die aktuelle Wahlperiode
sitzung-nr	Beschreibt die Nummer der Plenarsitzung innerhalb der Wahlperiode
sitzung-datum	Beschreibt das Datum, an dem die Sitzung stattfand
sitzung-start-uhrzeit	Beschreibt die Uhrzeit, an dem die Sitzung begann
sitzung-ende-uhrzeit	Beschreibt die Uhrzeit, an dem die Sitzung endete
sitzung-naechste-datum	Beschreibt die in der Sitzung bekanntgegebene Datum der nächsten Sitzung
start-seitennr	Beschreibt die Seitennummer, mit der das gedruckte Protokoll beginnt.. Innerhalb einer Wahlperiode werden zur eindeutigen Identifikation jeder Seite eines Protokolls die Seiten für alle Protokolle fortlaufend nummeriert.

Zu Attributen, die in jeder gedruckten Ausgabe als unveränderlicher Text erscheinen:

```
sitzung-ort
herausgeber
herstellung
vertrieb
issn
```

sind zwar definiert, werden jedoch beim XML-Export aktuell nicht ausgegeben.

Zu Attributen, die für eine mögliche Weiterverarbeitung den Prozess-Status und eine Version angeben:

Die Attribute `version`, `status` und `kommentare` wurden nur zu dem Zweck angelegt, für weiterverarbeitende Prozesse zur Verfügung zu stehen. Sie werden bei der XML-Erzeugung deshalb nicht gefüllt.

### 3. Der Hauptabschnitt <vorspann>

#### a.) Element-Definition:

```
<!ELEMENT vorspann (kopfdaten, inhaltsverzeichnis) >
Keine Attribut-Definition
```

Beispiel: <vorspann>

**Bedeutung:** Das Element <vorspann> enthält Verwaltungsangaben und das Inhaltsverzeichnis. Beide Elemente müssen enthalten sein. Sie umfassen im Vergleich zum gedruckten Protokoll das Deckblatt und alle Seiten des Inhaltsverzeichnisses. Der Vorspann endet vor der Seite, die den ersten stenografierten Text enthält.

#### 3.1 Das Element <kopfdaten>

#### a.) Element-Definition:

```
<!ELEMENT kopfdaten (plenarprotokoll-nummer, herausgeber,
berichtart, sitzungstitel, veranstaltungsdaten) >
Keine Attribut-Definition
```

Beispiel: <kopfdaten>

**Bedeutung:** Enthält die Verwaltungsangaben des Protokolls, wie sie in Klartext in der gedruckten Fassung vorkommen. Inhaltlich gibt es keinen Unterschied zwischen den Kopfdaten und den Attributwerten des Wurzelements.

#### b.) Element-Definition:

```
<!ELEMENT plenarprotokoll-nummer (#PCDATA | wahlperiode | sitzungsnr)* >
Keine Attribut-Definition
```

Beispiel: <plenarprotokoll-nummer>Plenarprotokoll <wahlperiode>18</wahlperiode>/  
<sitzungsnr>97</sitzungsnr></plenarprotokoll-nummer>

**Bedeutung:** Enthält den gedruckten Inhalt der Überschrift des Protokolls, z. B. „Plenarprotokoll 18/97“. Speziell gekennzeichnet sind die Elemente <wahlperiode> und <sitzungsnummer>.

#### c.) Element-Definition:

```
<!ELEMENT herausgeber (#PCDATA) >
Keine Attribut-Definition
```

Beispiel: <herausgeber>Deutscher Bundestag</herausgeber>

**Bedeutung:** Enthält den gedruckten Inhalt der Überschrift „Deutscher Bundestag“

#### d.) Element-Definition:

```
<!ELEMENT berichtart (#PCDATA) >
Keine Attribut-Definition
```

Beispiel: <berichtart>Stenografischer Bericht</berichtart>

**Bedeutung:** Enthält den gedruckten Inhalt der Überschrift „Stenografischer Bericht“

#### e.) Element-Definition:

```
<!ELEMENT sitzungstitel (#PCDATA | sitzungsnr)* >
Keine Attribut-Definition
```

Beispiel: <sitzungstitel><sitzungsnr>97</sitzungsnr>. Sitzung</sitzungstitel>

**Bedeutung:** Enthält den gedruckten Inhalt der Überschrift z. B. „97. Sitzung“

#### f.) Element-Definition:

```
<!ELEMENT veranstaltungsdaten (#PCDATA | ort | datum)* >
Keine Attribut-Definition
```

Beispiel: <veranstaltungsdaten><ort>Berlin</ort>, <datum date="26.03.2015">Donnerstag,  
den 26. März 2015</datum></veranstaltungsdaten>

**Bedeutung:** Enthält den gedruckten Inhalt der Überschrift z. B. „Berlin, Donnerstag den 26. März 2015“  
Speziell gekennzeichnet sind die Elemente <ort> und <datum>.

Zur leichten Auswertung wird das Datum im Format „TT.MM.JJJJ“ noch im Attribute date wiederholt.

### 3.2 Element <inhaltsverzeichnis>

#### a.) Element-Definition:

```
<!ELEMENT      ivt-verzeichnis      (ivz-titel, (ivz-block | ivz-eintrag)+)>
Keine Attribut-Definition
```

**Beispiel:** <inhaltsverzeichnis>

**Bedeutung:** Das Element umfasst alle Angaben des Inhaltsverzeichnisses. Die wesentlichen Bausteine sind die Elemente <ivz-eintrag> für einzelne Angaben in der Tagesordnung und <ivz-block> für alle Einträge, die zu einem Tagesordnungspunkt oder einer Anlage gehören.

#### b.) Element-Definition:

```
<!ELEMENT      ivz-titel              (#PCDATA) >
Keine Attribut-Definition
```

**Beispiel:** <ivz-titel>Inhalt:</ivz-titel>

**Bedeutung:** Enthält den gedruckten Inhalt der unveränderlichen Überschrift „Inhalt.“

#### c.) Element-Definition:

```
<!ELEMENT      ivz-block      (ivz-block-titel?, (ivz-eintrag | ivz-block | p)+) >
Keine Attribut-Definition
```

**Beispiel:**

```
<ivz-block>
  <ivz-block-titel>Tagesordnungspunkt 26:</ivz-block-titel>
  <ivz-eintrag>
    <ivz-eintrag-inhalt>a) Erste Beratung des von der Bundesregierung eingebrachten
    Entwurfs eines Gesetzes zur Verbesserung der internationalen Rechtshilfe bei der
    Vollstreckung von freiheitsentziehenden Sanktionen und bei der Überwachung von
    Bewährungsmaßnahmen Drucksache 18/4347 </ivz-eintrag-inhalt>
    <a href="S9195" typ="druckseitennummer">
      <seite>9195</seite> <seitenbereich>D</seitenbereich></a>
    </ivz-eintrag>
  ...
</ivz-block>
```

**Bedeutung:** Enthält alle Einträge, die zu einem Tagesordnungspunkt oder einer Anlage gehören und deshalb zu einem Block zusammengefasst werden.

#### d.) Element-Definition:

```
<!ELEMENT      ivz-eintrag      (#PCDATA | ivz-eintrag-inhalt | xref | a)* >
Keine Attribut-Definition
```

**Beispiel:**

```
<ivz-eintrag>
  <ivz-eintrag-inhalt>Erweiterung und Abwicklung der Tagesordnung </ivz-eintrag-inhalt>
  <a href="S9153" typ="druckseitennummer">
    <seite>9153</seite> <seitenbereich>A</seitenbereich></a>
</ivz-eintrag>
```

**Bedeutung:** In einem Element <ivz-eintrag> wird innerhalb des Elements <ivz-eintrag-inhalt> der gedruckte Text und mit <a> der Link zum Inhalt innerhalb des Elements <sitzungsverlauf> abgebildet. Im Element <a> wird mit dem Attribut href="SXXXX" auf die Seitennummer im gedruckten Protokoll verwiesen.



#### Im Element <a> ist die Angabe von href auf Druckseiten nicht eindeutig

Der Wert vom Attribut href ist nicht als einmalige ID zu verstehen, denn er entspricht der Seitennummer der gedruckten Protokoll-Ausgabe. Auch bei Ergänzung des Attributwertes mit den Inhalt vom Element <seitenbereich> – der Seitenbereich entspricht den Quadranten A, B, C, D auf jeder gedruckten A4-Seite – wird die Angabe nicht eindeutig, da innerhalb eines Quadranten mehrere Verweisziele Platz finden können.

- ▶ Soll mit <a> ein eindeutiger Verweis umgesetzt werden, kann id als zusätzliches Attribut mit eindeutigem Wert ergänzt werden. Genaue Details zum Element <a> lesen Sie auf Seite 13

## 4. Elemente im Sitzungsverlauf

### 4.1 Element <sitzungsverlauf>

#### a.) Element-Definition:

```
<!ELEMENT      sitzungsverlauf      (sitzungsbeginn?, rede*,
                                     tagesordnungspunkt+, sitzungsende?) >
```

Keine Attribut-Definition

**Beispiel:** <sitzungsverlauf>

**Bedeutung:** Im Gegensatz zu den Elementen <vorspann> und <anlagen> enthält das Element <sitzungsverlauf> den stenografierten Inhalt der Plenarsitzung mit allen darin differenzierten Elementen. Im Vergleich zum gedruckten Protokoll beginnt der Sitzungsverlauf auf der Seite, auf der der Sitzungsbeginn angegeben wurde und endet auf der Seite, auf der das Sitzungsende vermerkt wurde.

### 4.2 Die Elemente <sitzungsbeginn> und <sitzungsende>

Die Elemente <sitzungsbeginn> und <sitzungsende> sind das erste bzw. das letzte Kind-Element im Element <sitzungsverlauf>.

#### a.) Element-Definition:

```
<!ELEMENT      sitzungsbeginn      (#PCDATA | p | zitat | kommentar | a | name)* >
<!ATTLIST     sitzungsbeginn
               sitzung-start-uhrzeit      CDATA      #REQUIRED >
```

**Beispiel:** <sitzungsbeginn sitzung-start-uhrzeit="9:01">

**Bedeutung:** Der Attributwert im Element gibt die genaue Uhrzeit für den Sitzungsbeginn an.

#### b.) Element-Definition:

```
<!ELEMENT      sitzungsende      (#PCDATA | p | a)* >
<!ATTLIST     sitzungsende
               sitzung-ende-uhrzeit      CDATA      #REQUIRED >
```

**Beispiel:** <sitzungsende sitzung-ende-uhrzeit="20:37"/>

**Bedeutung:** Der Attributwert im Element gibt die genaue Uhrzeit für das Sitzungsende an.

### 4.3 Das Element <tagesordnungspunkt>

#### a.) Element-Definition:

```
<!ELEMENT      tagesordnungspunkt  (top-titel | name | rede
                                     | p | zitat | kommentar | a)* >
<!ATTLIST     tagesordnungspunkt
               top-id      CDATA      #IMPLIED >
```

**Beispiel:** <tagesordnungspunkt top-id="Tagesordnungspunkt 20">

**Bedeutung:** Innerhalb des Sitzungsverlaufs nehmen die Tagesordnungspunkte eine Schlüsselrolle für die Inhaltsgliederung ein. Zu einem Tagesordnungspunkt können einleitende Absätze, Namen, Kommentare und Reden zugeordnet werden.

#### b.) Element-Definition:

```
<!ELEMENT top-titel      (#PCDATA | p | a)* >
Keine Attribut-Definition
```

**Bedeutung:** Sollte ein Tagesordnungspunkt (top) zusätzlich einen einzeln und konkret aufgeführten Titel haben, kann dieser mit diesem Element hervorgehoben werden.

Da üblicherweise vom Bundestagspräsidenten die Tagesordnungspunkte aus einem anderen Kontext heraus aufgerufen und genannt werden, kommt dieses Element im Prinzip nicht zum Einsatz.

## 5. Anlagen zum Plenarprotokoll

Als Anlagen zum Plenarprotokoll werden verschiedene Informationen angefügt. Die erste Anlage ist immer die „Liste der entschuldigten Abgeordneten“.

### 5.1 Element <anlagen>

#### a.) Element-Definition:

```
<!ELEMENT      anlagen      (anlagen-titel?, anlage+) >
Keine Attribut-Definition
```

#### Beispiel:

```
<anlagen>
  <anlagen-titel>
    <a typ="druckseitennummer" id="S9291" name="S9291"/>
      Anlagen zum Stenografischen Bericht
  </anlagen-titel>
  ...
</anlagen>
```

**Bedeutung:** Im Element <anlagen> werden <anlage>-Elemente aufgereiht. Sie beginnt im Vergleich zum gedruckten Protokoll mit der Seite, auf der die Liste der entschuldigten Abgeordneten steht und endet mit der letzten gedruckten Seite des Protokolls.

### 5.2 Element <anlage>

#### a.) Element-Definition:

```
<!ELEMENT      anlage      (anlagen-titel?, anlagen-text+) >
Keine Attribut-Definition
```

#### Beispiel:

```
<anlage>
  <anlagen-titel><a typ="druckseitennummer" id="S9293" name="S9293"/>Anlage 3
</anlagen-titel>
  <anlagen-text anlagen-typ="Erklärung nach § 31 GO">
    <p klasse="Anlage_2">Erklärung nach § 31 GO </p>
  ...
</anlage>
```

**Bedeutung:** Jede einzeln genannte Anlage wird in einem Element <anlage> gekapselt.

#### b.) Element-Definition:

```
<!ELEMENT      anlagen-titel  (#PCDATA | p | %inline-elemente;)* >
Keine Attribut-Definition
```

**Bedeutung:** Enthält den Titel der Anlage. Steht der Titel am Beginn einer neuen Seite, wird diese im Titel über das Element <a> angezeigt.

#### c.) Element-Definition:

```
<!ELEMENT      anlagen-text    (#PCDATA | p | kommentar | rede | table |
                               zitat | %inline-elemente;)* >
<!ATTLIST      anlagen-text    anlagen-typ      CDATA      #IMPLIED >
```

**Bedeutung:** Enthält den Anlagen-Inhalt, eine Tabelle, eine Rede, usw.

## 6. Rednerliste, Redner, Namen und Reden

### 6.1 Element <rednerliste>

#### a.) Element-Definition:

```
<!ELEMENT      rednerliste      (redner+) >
<!ATTLIST     rednerliste
      sitzung-datum      CDATA      #REQUIRED>
```

**Beispiel:** <rednerliste sitzung-datum="26.03.2015">

**Bedeutung:** Die Rednerliste wird aus einer externen Datenquelle – einer XML-Datei – bei der Konvertierung nach XML zusätzlich mit am Ende eingefügt. Aus der Rednerliste werden alle konkreten Informationen zu allen Rednern aus dem Sitzungsverlauf gewonnen.

Die Rednerliste enthält aktuelle Informationen zu allen Rednern mit eindeutiger ID. Aufgelistet werden:

- Abgeordnete
- Bundesminister
- Staatssekretäre
- Minister der Bundesländer

### 6.2 Element <redner>

#### a.) Element-Definition:

```
<!ELEMENT      redner      (name) >
<!ATTLIST     redner
      id      CDATA      #REQUIRED >
```

#### Beispiel:

```
<redner id="11002624">
  <name>
    <vorname>Marieluise</vorname>
    <nachname>Beck</nachname>
    <ortszusatz>(Bremen)</ortszusatz>
    <fraktion>BÜNDNIS 90/DIE GRÜNEN</fraktion>
  </name>
</redner>
<redner id="999990001">
  <name>
    <vorname>Heiko</vorname>
    <nachname>Maas</nachname>
    <rolle>
      <rolle_lang>Bundesminister der Justiz und
        für Verbraucherschutz</rolle_lang>
    </rolle>
  </name>
</redner>
```

**Bedeutung:** Jedem Redner ist eine eindeutige ID zugeordnet.

Vom Referat IT-Systementwicklung – IT4 – wird die Liste der aktuellen Abgeordneten zur Verfügung gestellt. Diese Datei enthält Titel, Namen, Ortszusatz, Fraktions-Zugehörigkeit und eine eindeutige ID für jeden Abgeordneten. Die Abgeordneten-Daten werden in eine XML-Datei eingepflegt und um die Angaben zu Bundesministern, Staatssekretären und den Ministern der Bundesländer ergänzt. Eindeutige IDs für diese Personen werden bei der Erfassung direkt zugeordnet und unterscheiden sich von den IDs der Abgeordneten. Siehe auch "Im Prozess beteiligte Informationen/Dateien" auf Seite 2.



#### Ziffernfolgen in den Redner-IDs

- ▶ Abgeordneten-IDs sind in der Regel 8-stellig und beginnen mit den Ziffern '1100'
- ▶ IDs für alle anderen Redner, die nicht Abgeordnete sind, sind 9-stellig und beginnen mit den Ziffern '9999'.



## 6.3 Das Element <name>

### a.) Element-Definition:

```

<!ELEMENT      name      (#PCDATA | titel | vorname | nachname | namenszusatz |
                           ortszusatz | fraktion | rolle | bdland)* >

<!ELEMENT      titel     (#PCDATA) >
<!ELEMENT      vorname   (#PCDATA) >
<!ELEMENT      namenszusatz (#PCDATA) >
<!ELEMENT      nachname  (#PCDATA) >
<!ELEMENT      ortszusatz (#PCDATA) >
<!ELEMENT      fraktion  (#PCDATA) >
<!ELEMENT      rolle     (rolle_lang | rolle_kurz)* >
<!ELEMENT      rolle_lang (#PCDATA) >
<!ELEMENT      rolle_kurz (#PCDATA) >
<!ELEMENT      bdland    (#PCDATA) >
Keine Attribut-Definitionen für die genannten Elemente

```

### Beispiel:

```

<name>
    <titel>Dr.</titel>
    <vorname>Angela</vorname>
    <nachname>Merkel</nachname>
    <fraktion>CDU/CSU</fraktion>
    <rolle>
        <rolle_lang>Bundeskanzlerin</rolle_lang>
        <rolle_kurz>BK</rolle_kurz>
    </rolle>
</name>
oder
<name>
    <titel>Dr.</titel>
    <vorname>Beate</vorname>
    <nachname>Merk</nachname>
    <rolle>
        <rolle_lang>Staatsministerin</rolle_lang>
    </rolle>
    <bdland>Bayern</bdland>
</name>

```

**Bedeutung:** Im Element <name> werden die Informationen zu den Akteuren der Plenarsitzungen zusammengefasst.

Die Bedeutung der Kind-Elemente im Namen eines Redners:

titel	Enthält alle akademischen Titel eines Redners.	optional
vorname	Enthält jeweils einen Vornamen, kann mehrfach auftreten.	erforderlich
namenszusatz	Enthält Namenszusätze, z.B. „von der“ oder „Freiherr von“	optional
nachname	Enthält den Nachnamen	erforderlich
ortszusatz	Enthält einen Ortszusatz, um Abgeordnete mit gleichem Nachnamen ohne Vornamen-Nennung zu unterscheiden.	optional
fraktion	Enthält die Fraktion in einer definierten Schreibweise.	bei Abgeord. erforderlich
rolle	Enthält <rolle_lang> oder <rolle_kurz>.	optional
rolle_lang	Angabe der Rolle ausgeschrieben.	optional
rolle_kurz	Angabe der Rolle in Kurzform.	optional
bdland	Angabe des Bundeslandes für Minister der Bundesländer.	optional

## 6.4 Das Element <rede>

### a.) Element-Definition:

```
<!ELEMENT rede (name | p | kommentar | zitat | a)+ >
<!ATTLIST rede
  id ID #REQUIRED
  redeart (plenum | zu_protokoll | zu_protokoll_anlage) #IMPLIED
>
```

### Beispiel:

```
<rede id="ID1809710100">
  <p klasse="redner"><redner id="11004190">
    <name><titel>Dr.</titel><vorname>Valerie</vorname><nachname>Wilms</nachname>
    <fraktion>BÜNDNIS 90/DIE GRÜNEN</fraktion></name></redner>
    Dr. Valerie Wilms (BÜNDNIS 90/DIE GRÜNEN):
  </p>
  <p klasse="J_1">Frau Präsidentin! Werte Kolleginnen und Kollegen! ... </p>
  <kommentar>(Beifall beim BÜNDNIS 90/DIE GRÜNEN)</kommentar>
  <a typ="druckseitennummer" id="S9287" name="S9287"/>
  <p klasse="J">Danke. </p>
  <name>Vizepräsidentin Ulla Schmidt: </name>
  <p klasse="J_1">Vielen Dank. </p>
</rede>
```

**Bedeutung:** Im Element <rede> wird der gesamte Inhalt einer Rede strukturiert.

Da die Reden innerhalb eines Protokolls anhand von Formaten im Vergleich mit den Einträgen im Inhaltsverzeichnis identifiziert werden, Beginnen die Reden mit der Nennung des Redners und enden – wenn sie im Sitzungsverlauf gehalten wurden – mit den Worten eines Bundestagspräsidenten bzw. einer Bundestagspräsidentin.

Die ID im Element <rede id="ID1809710100"> wird zuerst im <inhaltsverzeichnis> verwendet und im <sitzungsverlauf> dann referenziert. **Die Zuordnung der Rede-ID ist eineindeutig.**

Der Ursprung der Redner-ID wurde bereits beschrieben unter "Element <redner>" auf Seite 8.

Innerhalb von <rede> sind Kind-Elemente mit folgender Bedeutung erlaubt:

name	Angabe eines Namens, der nicht zum Redner gehört.
p	Absatz, der wiederum sowohl Redetext enthält als auch Angaben zum Redner.
kommentar	Enthält Kommentare mit der Reaktion der Zuhörer.
zitat	Enthält zitierten Text
a	Enthält nach Seitenwechsel der Druckvorlage die Angabe der Druck-Seitennummer.

Das Attribut `id` wird für jede Rede eindeutig generiert, die einen direkten Verweis erlaubt.

Der Attributwert von `rid` setzt sich zusammen aus:

ID	feststehende Zeichenfolge
18	aktuelle Wahlperiode
097	Sitzungsnummer
098	laufende Nummer der verknüpften Rede
00	Reserve für etwaige Korrekturen

## 7. Elemente für Textstrukturen

### 7.1 Das Element <p>

#### a.) Element-Definition:

```
<!ELEMENT      p      (#PCDATA | redner | table | %inline-elemente;)* >
<!ATTLIST     p
      klasse      CDATA      #IMPLIED>
```

**Beispiel:** `<p klasse="0">Er sagte weiter: </p>`

**Bedeutung:** In Elementen vom Typ `<p>` steht der hauptsächlichliche Inhalt des Protokolls. Die Bezeichnung ist in Analogie zum Element `<p>` in der HTML-Strukturdefinition für Paragraph=Absatz zu sehen.

In jedem Absatz-Element wird das Attribut `klasse` angegeben, welches einen beliebigen Wert haben kann. Der Attributwert wird aus der Bezeichnung der Formatvorlage abgeleitet, die dem Text bei der Protokoll-Erfassung zugewiesen wird. Diese Bezeichnung kann als zusätzliche Information verstanden werden, die nicht verloren gehen soll. Die Bezeichnung der Formatvorlagen ist jedoch unabhängig von der hier vorliegenden Strukturdefinition. Im Erstellungsprozess können Formatvorlagen ergänzt und ihre Bezeichnung geändert werden. Deshalb sind diese Werte in der DTD nicht als Auswahlliste hinterlegt.

In den bisherigen XML-Protokollen können die häufig vorkommenden Werte des Attributs `klasse` folgender Verwendung zugeordnet werden:

Normaler Text	J, J_1, O
Tagesordnungspunkt	A_TOP, T_Beratung, T_Drs, T_E_Drs, T_E_E_Drs, T_E_fett, T_NaS, T_NaS_NaS, T_ZP_NaS, T_ZP_NaS_NaS, T_ZP_NaS_NaS_Strich, T_Ueberweisung, T_fett, T_ohne_NaS
Abstimmungslisten	AL_Ja-Nein-Enth, AL_Namen, AL_Partei, AL_Vorspann_1, AL_Vorspann_2, AL_Vorspann_3
Anlagen	Anlage_2, Anlage_3
Redner	redner
Namen	N
Zitat	Z

### 7.2 Das Element <zitat>

#### a.) Element-Definition:

```
<!ELEMENT      zitat  (#PCDATA | p | %inline-elemente;)* >
Keine Attribut-Definition
```

#### Beispiel:

`<p klasse="J">Das Bild vom „gemeinsamen Haus Europa“ geht auf eine Rede Michail Gorbatschows aus dem Jahre 1987 zurück, in der er wörtlich sagte: </p>`

`<zitat>`

`<p klasse="Z">Europa ist ... ein gemeinsames Haus, wo Geografie und Geschichte die Geschehnisse von Dutzenden von Ländern eng miteinander verwoben haben. </p>`

`</zitat>`

`<p klasse="0">Er sagte weiter: </p>`

`<zitat>`

`<p klasse="Z">Doch nur zusammen, gemeinschaftlich und indem sie die vernünftigen Regeln der Koexistenz befolgen, können die Europäer ihr Haus bewahren ... es besser und sicherer machen ... </p>`

`</zitat>`

**Bedeutung:** Enthält zitierte Aussagen

### 7.3 Das Element <kommentar>

#### a.) Element-Definition:

```
<!ELEMENT      kommentar  (#PCDATA | p | %inline-elemente;)* >
Keine Attribut-Definition
```

#### Beispiel:

`<kommentar>(Beifall beim BÜNDNIS 90/DIE GRÜNEN) </kommentar>`

**Bedeutung:** Kommentare enthalten in der Regel Vermerke zur Reaktion der Zuhörer im Plenum.

## 7.4 Tabellen-Elemente

### a.) Element-Definition:

```
<!ELEMENT table (caption?, (col*|colgroup*), thead?, tfoot?, (tbody+|tr+))>
```

Beispiel: `<table>`

```

    <thead>
      <tr>
        <th colspan="2"> Abgeordnete(r)</th>
        <th>entschuldigt bis einschließlich</th>
      </tr>
    </thead>
    <tbody>
      <tr>
        <td>Auernhammer, Artur</td>
        <td>CDU/CSU</td>
        <td>26.03.2015</td>
      </tr>
    </tbody>
  </table>
```

**Bedeutung:** Tabellen kommen im Plenarprotokoll in der Regel nur in den Anhängen vor.

Zur Darstellung wurde ein allgemeines Tabellen-Modell genutzt, dass an das CALS-Tabellenmodell angelehnt ist. Auf die Nennung und genaue Erklärung der Attribut-Definition der Tabellen-Elemente wird verzichtet, da sich diese aus allgemein verfügbaren Quellen ermitteln lassen.

### b.) Element-Definition:

```
<!ELEMENT caption (#PCDATA)>
```

Beispiel: `<caption>Übersicht</caption>`

**Bedeutung:** Enthält den Tabellentitel

### c.) Element-Definition:

```
<!ELEMENT thead (tr)+>
```

**Bedeutung:** Tabellen-Kopfbereich, enthält die Kopfreihe mit Kopf-Tabellenzellen `<th>`

### d.) Element-Definition:

```
<!ELEMENT tfoot (tr)+>
```

**Bedeutung:** Tabellen-Fußbereich, enthält die Reihen mit Fuß-Tabellenzellen `<th>`  
Kommt selten vor.

### e.) Element-Definition:

```
<!ELEMENT tbody (tr)+>
```

**Bedeutung:** Tabellen-Bereich für die Inhalts-Reihen

### f.) Element-Definition:

```
<!ELEMENT tr (th|td)+>
```

**Bedeutung:** Tabellenreihe; enthält in tbody Tabellenzellen `<td>` und im Kopf-/Fußbereich Tabellenzellen `<th>`.

### g.) Element-Definition:

```
<!ELEMENT th (#PCDATA | %inline-elemente;)*>
```

Beispiel: `<th colspan="2"> Abgeordnete(r)</th>`

**Bedeutung:** Tabellenzelle im Kopf- oder Fußbereich einer Tabelle

### h.) Element-Definition:

```
<!ELEMENT td (#PCDATA | %inline-elemente;)*>
```

**Bedeutung:** Tabellenzelle in den Standardreihen einer Tabelle.

## 8. Die Inline-Elemente

### 8.1 Element <a>

#### a.) Element-Definition:

```
<!ELEMENT a      #PCDATA | seite | seitenbereich)* >
<!ATTLIST a
  typ (druckseitennummer | bild | tabelle)          #IMPLIED
  id   CDATA                                         #IMPLIED
  name CDATA                                         #IMPLIED
  href CDATA                                         #IMPLIED>
```

#### Beispiel:

Definition eines Verweisziels:

```
<a typ="druckseitennummer" id="S9153" name="S9153"/>
```

Definition eines Verweises auf ein entsprechend festgelegtes Verweisziel:

```
<ivz-eintrag>
  <ivz-eintrag-inhalt>Erweiterung der Tagesordnung </ivz-eintrag-inhalt>
  <a href="S9153" typ="druckseitennummer"><seite>9153</seite>
    <seitenbereich>A</seitenbereich>
  </a>
</ivz-eintrag>
```

**Bedeutung:** Analog zur Verwendung des Elements <a> innerhalb der HTML-Struktur-Definition wurde dieses Element entworfen. Es dient dazu, Verweise und Verweisziele festzulegen.

Innerhalb des Textes im gedruckten Protokoll besteht außer für Reden, deren Anfang durch Textvergleich und Namenserkennung ermittelt wird, keine Möglichkeit, Texte im Sitzungsverlauf über eine eindeutige ID zu identifizieren. Deshalb wird im Text an jedem Druckseiten-Anfang ein Verweisziel für die gerade begonnene Seite eingefügt. Um darzustellen, dass die ID nicht als eindeutige ID auf eine bestimmte Struktur oder zur Identifikation eines Elements gedacht ist, wird im Attribut `id` die Seitennummer mit einem führenden `S=Seite` eingefügt. Das Attribut `name` ist historisch zu betrachten und kann als Label des Verweisziels verstanden werden.

Im Gegensatz zur Kennzeichnung des Beginns einer Druckseite werden Reden und Redner über eindeutige IDs identifiziert.



#### Im Element <a> ist die Angabe von `href` auf Druckseiten nicht eindeutig

Der Wert vom Attribut `href` ist nicht als einmalige ID zu verstehen, denn er entspricht der Seitennummer der gedruckten Protokoll-Ausgabe. Auch bei Ergänzung des Attributwertes mit den Inhalt vom Element `<seitenbereich>` – der Seitenbereich entspricht den Quadranten A, B, C, D auf jeder gedruckten A4-Seite – wird die Angabe nicht eindeutig, da innerhalb eines Quadranten mehrere Verweisziele Platz finden können.

- ▶ Soll mit <a> ein eindeutiger Verweis umgesetzt werden, kann `id` als zusätzliches Attribut mit eindeutiger ID ergänzt werden.

## 8.2 Element <xref>

### a.) Element-Definition:

```
<!ELEMENT      xref (#PCDATA | a)* >
<!ATTLIST     xref
ref-type      (rede | kommentar | anlage)      #REQUIRED
rid           CDATA                            #REQUIRED
pnr           CDATA                            #REQUIRED
div           CDATA                            #IMPLIED>
```

### Beispiel:

```
<xref ref-type="rede" rid="ID1809709800" pnr="9282" div="B">
  <a href="S9282" typ="druckseitennummer">
    <seite>9282</seite>
    <seitenbereich>B</seitenbereich></a>
</xref>
```

**Bedeutung:** Das Element <xref> hat die Funktion eines Querverweises. Es kombiniert angezeigten Text mit einem Verweis. Der Querverweis wird z. B. genutzt, um eine eindeutige Verlinkung vom Inhaltsverzeichnis zu den Reden im Protokoll herzustellen.

Das Attribut `rid` ist eine Abkürzung für „Rede-ID“. In jedem Protokoll wird für jede Rede eine eindeutige ID generiert, die einen direkten Verweis erlaubt.

Der Attributwert von `rid` setzt sich zusammen aus:

ID	feststehende Zeichenfolge
18	aktuelle Wahlperiode
097	Sitzungsnummer
098	laufende Nummer der verknüpften Rede
00	Reserve für etwaige Korrekturen

## 8.3 Element <sub>

### a.) Element-Definition:

```
<!ELEMENT      sub      #PCDATA)>
Keine Attribut-Definition
```

**Beispiel:** <p klasse="J"> ... Wasser ... H<sub>2</sub>O ... </p>

**Bedeutung:** Kennzeichnen von tiefgestellten Zeichen

## 8.4 Element <sup>

### a.) Element-Definition:

```
<!ELEMENT      sup      #PCDATA)>
Keine Attribut-Definition
```

**Beispiel:** <sup>1</sup>

**Bedeutung:** Kennzeichnen von hochgestellten Zeichen

## 8.5 Element <fussnote>

### a.) Element-Definition:

```
<!ELEMENT      fussnote (#PCDATA | p)*>
Keine Attribut-Definition
```

**Beispiel:** <p klasse="J">Wir kommen zur Abstimmung ... <sup>1</sup><fussnote>1)Anlage2</fussnote></p>

**Bedeutung:** Kennzeichnen einer Fußnote innerhalb eines Absatzes

Ende des Dokuments