

# KNA



## **Dokumentation**

**Änderungen in der Auslieferung  
von NewsML-G2 Text-Nachrichten  
vom Dezember 2015**

**Version 3 – vom 2015-12-23**

## Einführung

NewsML-G2 ist ein flexibles Datenmodell mit dem laufend das Format von Nachrichteninhalten aber auch von deren Beschreibung, den Metadaten, verändert, verbessert oder verfeinert werden kann.

Im Januar 2016 setzt KNA diesen Schritt, um die Qualität der gelieferten Nachrichten zu erhöhen.

Dieses Dokument beschreibt die Änderungen gegenüber der seit September 2013 ausgelieferten Version von KNA-NewsML-G2. Die bisher ausgelieferte Version von KNA-NewsML-G2 ist in einem Dokument vom September 2013 mit dem Dateinamen NewsML-G2@KNA-Dok\_2013-09-18.pdf beschrieben, Hinweise auf Veränderungen beziehen sich auf diese Dokument.

(Diese vollständige Dokumentation wird demnächst in einer aktualisierten Version erscheinen.)

Anmerkung: wenn hier von einer neuen Version des KNA-NewsML-G2 gesprochen wird, dann wird damit die spezielle Nutzung dieses Formats im Rahmen der Auslieferung durch KNA gemeint. Davon unabhängig wird NewsML-G2 als Format von der IPTC laufend weiter entwickelt, dies hat jedoch keinen unmittelbaren Einfluss auf die KNA-Auslieferung. In dieser KNA-NewsML-G2-Version werden jedoch auch neuere Fähigkeiten des Formats genutzt.

Bei Fragen oder Anmerkungen zu diesem Dokument und zur neuen Version KNA-NewsML-G2 kontaktieren Sie bitte:

KNA-Produktentwicklung Heinrich-Brüning-Str. 9, 53113 Bonn

Tel. +49 (228) 26 00 01 12 E-Mail: [NewsMLG2@kna.de](mailto:NewsMLG2@kna.de)

Version dieses Dokuments: 3 vom 2015-12-23 (mws/bmi)

## Änderungen von KNA-NewsML-G2

In den folgenden Abschnitten werden einzelne Änderungen beschrieben.

### Neues XHTML Format

Die von KNA gelieferten Texte werden jetzt in einem an HTML5 orientierten XHTML formatiert.

Dazu hier ein Beispiel:

```
<inlineXML contentType="application/xhtml+xml" wordcount="164">
  <html xmlns="http://www.w3.org/1999/xhtml">
    <head>
      <title>Studie: Flüchtlingsaufnahme kostet 17 Milliarden Euro jährlich </title>
      <meta charset="utf-8"/>
    </head>
    <body>
      <header>
        <ul class="slugline">
          <li class="subject" data-qcode="knasubject:644">Flüchtlinge</li>
          <li class="subject" data-qcode="knasubject:624">Steuern</li>
          <li class="subject" data-qcode="knasubject:540">Arbeitsmarkt</li>
          <li class="geo" data-qcode="knacountry:1">Deutschland</li>
        </ul>
        <time class="publicationDate" data-datetime="2015-11-25T17:02:47+01:00">25.11.2015 17:02</time>
        <h1>Studie: Flüchtlingsaufnahme kostet 17 Milliarden Euro jährlich </h1>
      </header>
      <section class="main kgenre-0001">
        <p>
          <span class="dateline">Berlin <span class="credit">(KNA) </span>
          </span>Die Flüchtlingsaufnahme kostet die Steuerzahler in Deutschland einer Studie zufolge jährlich
          dauerhaft 17 Milliarden Euro. Das geht aus einer Berechnung der Stiftung Marktwirtschaft hervor, die am Dienstag in
          Berlin vorgestellt wurde. Entscheidend sei jedoch die Geschwindigkeit der Integration: Je schneller Flüchtlinge arbeiten
          könnten, desto niedriger seien die Kosten.</p>
          <p>Bei den Berechnungen wurde den Angaben zufolge davon ausgegangen, dass bis zum Jahr 2018 zwei
          Millionen Flüchtlinge nach Deutschland kommen und eine Integration in den Arbeitsmarkt innerhalb eines Zeitraums
          von sechs Jahren gelingt. Zur Finanzierung müssten die Ausgaben dauerhaft um etwa ein Prozent gesenkt oder
          Steuern und Abgaben um 1,5 Prozent erhöht werden.</p>
          <p>Bernd Raffelhüschen von der Stiftung Marktwirtschaft forderte angesichts der Berechnung eine Reform der
          Zuwanderungspolitik. "Die Lösung könnte ein Einwanderungsgesetz sein, mit dessen Hilfe die Zuwanderung nach
          Deutschland besser gesteuert und stärker an den Bedürfnissen des Arbeitsmarktes ausgerichtet werden kann." Zudem
          könnten qualifizierte Zuwanderer, die schnell auf dem Arbeitsmarkt und in der Gesellschaft integriert würden, dazu
          beitragen, die steuerlichen Herausforderungen zu lösen. </p>
        </section>
      </body>
    </html>
  </inlineXML>
```

### Was ändert sich:

- Im <head> wurde <meta charset="utf-8"/> ergänzt
- Der <body> verwendet jetzt eine HTML5-nahe Struktur:
  - body: umfasst den gesamten Text
  - header: umfasst Daten über den Meldungstext
    - ul mit class="slugline": je li wird ein Eintrag der bisherigen Slugline gezeigt
    - time mit class="publicationDate": Datum und Zeit zu dem die Nachricht von KNA veröffentlicht wurde.
    - h1, h2, h3: die Überschriften der Meldung
    - by: die Autorenzeile
  - section mit class="main ...": umfasst den Meldungstext, so wie bisher direkt unter <body>
- Anmerkungen zur Verarbeitung:
  - Die Überschriften sind jetzt innerhalb des body-Elements vorhanden
  - wenn die Textabsätze über einen XPath wie ../inlineXML/h:html/h:body/h:p abgegriffen wurden, muss dieser angepasst werden ../inlineXML/h:html/h:body/h:section[starts-with(@class,"main")]/h:p

## Kenntlichmachung von Hinweisen an die Redaktion

Siehe auch vollständige KNA-NewsML-G2 Dokumentation vom September 2013: Abschnitt „KNA Abläufe für Textnachrichten“ Seite 28

Für Hinweise an die empfangende Redaktion wird allgemein das Element <edNote> verwendet. Welcher konkrete Hinweis das ist, wird über den Wert des @role Attributs kenntlich gemacht. Bisher wurden für viele unterschiedliche Anwendungsfälle jeweils eigene @role-Werte definiert, siehe die Liste in der September 2013-Dokumentation.

### Was ändert sich:

Jetzt wird für alle Hinweise, die sich auf die Verwendung des <signal>-Elements beziehen, einheitlich die @role-Werte

- knaednoterole:edWarningLong ... für ausführliche Hinweise
- knaednoterole:edWarningShort ... für kurze Hinweise

verwendet.

Die @role-Werte knaednoterole:correction, knaednoterole:correctionshort und knaednoterole:updatenote werden nicht mehr verwendet.

## Information zu Sende-Wiederholungen

Zum Beispiel

```
<edNote role="knaednoterole:edWarningLong">Hinweis zur Wiederholung, lang</edNote>
<edNote role="knaednoterole:edWarningShort">Hinweis zur Wiederholung, kurz</edNote>
...
<signal qcode="sig:repeat">
  <name xml:lang="en">Repeat</name>
</signal>
```

## Information zu Korrekturen

Zum Beispiel

```
<edNote role="knaednoterole:edWarningLong">Hinweis zur Berichtigung, lang</edNote>
<edNote role="knaednoterole:edWarningShort">Hinweis zur Berichtigung, kurz</edNote>
...
<signal qcode="sig:correction">
  <name xml:lang="en">Correction</name>
</signal>
```

## Information zum Zurückziehen einer Meldung

Zum Beispiel

```
<pubStatus qcode="stat:canceled">
  <name xml:lang="en">Cancelled</name>
</pubStatus>
...
<edNote role="knaednoterole:edWarningLong">Hinweis zum Zurückziehen, lang</edNote>
<edNote role="knaednoterole:edWarningShort">Hinweis zum Zurückziehen, kurz</edNote>
```

## Verwendung des KNA-Personen-Vokabulars

Bisher wurden zu Personen nur deren Name als Stichwort verarbeitet und mit dem <keyword> Element in NewsML-G2 abgebildet.

Zum Beispiel:

```
<keyword rank="1">Papst Franziskus</keyword>
```

### Was ändert sich:

KNA hat ein Vokabular von häufig vorkommenden Personen erstellt. Dieses wird nun verwendet und Personen, die in diesem Vokabular definiert sind, werden mittels <subject>-Element dargestellt:

```
<subject qcode="knaperson:339" type="cpnat:person">
  <name>Papst Franziskus</name>
</subject>
```

Anmerkung: jede Person wird mit einem QCode aus dem knaperson-Vokabular identifiziert. Dieses Vokabular soll Anfang 2016 auch allen KNA-Kunden als Referenz zur Verfügung stehen.

## Andere Vokabulare bei Ländern, darunter ISO-Ländercodes

In der bisherigen KNA-NewsML-G2-Version wurden Angaben zu Ländern so festgehalten:

```
<subject qcode="knageo:35" rank="1" type="cpnat:geoArea">
  <name>Deutschland</name>
  <sameAs qcode="knacountry:1">
    <name role="nrol:display" xml:lang="de">Deutschland</name>
    <name role="nrol:short" xml:lang="de">DEU</name>
  </sameAs>
  <broader qcode="wldreg:r150">
    <name xml:lang="en-GB">Europe</name>
  </broader>
</subject>
```

### Was ändert sich:

Nun werden die Länderangaben so festgehalten

```
<subject qcode="knacountry:1" rank="1" type="cpnat:geoArea">
  <name role="nrol:display" xml:lang="de">Deutschland</name>
  <name role="nrol:short" xml:lang="de">DEU</name>
  <sameAs qcode="dpacountry:1">
    <name role="nrol:display" xml:lang="de">Deutschland</name>
    <name role="nrol:short" xml:lang="de">DEU</name>
  </sameAs>
  <sameAs qcode="iso3166-1a2:DE">
    <name role="nrol:mnemonic">DE</name>
  </sameAs>
  <sameAs qcode="knageo:35">
    <name>Deutschland</name>
  </sameAs>
  <sameAs qcode="iso3166-1a3:DEU">
    <name role="nrol:mnemonic">DEU</name>
  </sameAs>
  <broader qcode="wldreg:r150">
    <name xml:lang="en">Europe</name>
  </broader>
</subject>
```

Konkrete Änderungen:

- der @qcode-Wert wird nun bei Ländern dem KNA-Vokabular knacountry: entnommen
- Beim unmittelbaren name-Kinderelement des subject-Elements werden nun @role-Attribute verwendet. Die bisherige Anzeige entspricht dem name-Element mit der @role="nrol:display"
- Bei den sameAs-Werten werden und auch die ISO-Landesnamen-Codes im 2-Zeichen (iso3166-1a2:) und im 3-Zeichen (iso3166-1a3:) Format verwendet

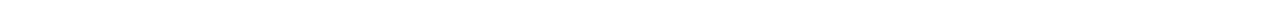
# KNA

## **NewsML-G2 Dokumentation**

**NewsML**  **G2**

---

Diese Seite bleibt geplanter Weise frei.



# Inhalt

<b>NewsML-G2 Allgemein</b>	<b>1</b>
Über diese Dokumentation.....	1
Zielgruppen.....	1
Version der Dokumentation.....	1
Sprachliche Aspekte .....	1
Technische Notationen .....	1
Urheberrechtliches .....	2
Kontakt .....	2
Was NewsML-G2 ist und kann.....	2
Anforderungen an NewsML-G2 Empfänger .....	4
Das NewsML-G2 Datenmodell.....	5
Das NewsML-G2 Datenformat in XML .....	6
Die Items .....	6
Die News Message .....	9
XML-Elemente für den Inhalt eines Items .....	9
NewsML-G2 und betriebliche Abläufe .....	10
Metadaten in NewsML-G2 .....	11
Ausdrucksformen von RDF-Objekten .....	12
Begriffe - ihre IDs und QCodes.....	13
Begriffe/Concepts in NewsML-G2.....	14
Kennung eines Begriffs - Concept-Id .....	15
Kurzfassung von Kennungen = QCodes .....	16
<b>NewsML-G2 von KNA</b>	<b>19</b>
Übersicht.....	19
News Message .....	19
Text-Nachrichten .....	20
KNA-spezifische Daten .....	20
KNA Abläufe für Text-Nachrichten .....	28
Begriff-Vokabularien .....	31
KNA-eigene Begriff-Vokabularien .....	31
Andere Begriff-Vokabularien.....	33
<b>Wo finde ich Information über ...</b>	<b>37</b>
<b>Referenzen</b>	<b>39</b>
Technologien und Standards .....	39
IPTC NewsML-G2 .....	39
IPTC NewsCodes .....	39
IPTC QCodes .....	39
Resource Description Framework – RDF .....	39
Semantic Web.....	40
Extensible Markup Language – XML.....	40
<b>Index</b>	<b>41</b>





# NewsML-G2 Allgemein

---

## Über diese Dokumentation

Diese Dokumentation beschreibt das IPTC Nachrichtenformat [NewsML-G2](#) in seiner weltweit allgemeingültigen Form als auch besonders detailliert die spezifische Form der Nachrichtenagentur KNA.

Diese Dokumentation ersetzt jedoch nicht die Originaldokumentation, insbesondere die formalen Spezifikationen der IPTC.

## Zielgruppen

Die Zielgruppen dieser Dokumentation sind:

- Ganz allgemein: die Empfänger von Nachrichtenströmen im NewsML-G2 Format
- Die IT-Mitarbeiter von NewsML-G2 Empfängern
- Die Mitarbeiter der Redaktion von NewsML-G2 Empfängern die für die technische Organisation zuständig sind.

## Version der Dokumentation

Das ist die Version 0.9  
mit dem Ausgabedatum 2013-09-18

## Sprachliche Aspekte

Alle originalen Dokumentationen und Spezifikationen von NewsML-G2 durch die IPTC liegen nur in Englisch vor. Diese Dokumentation verwendet ausschließlich die originalen Begriffe in Englisch wenn es sich um technisch-formale Bezeichnungen handelt. Bei allen Beschreibungen und Erläuterungen wird möglichst weitgehend die deutsche Sprache verwendet.

## Technische Notationen

Für XML Elemente wird oft deren Darstellung in XML Dokumenten verwendet, also der Namen des Elements in Spitzklammern: <elementName>

Für XML Attribute wird die weit verbreitete aus XPath entlehene Form des dem Attributnamen vorangestellten @-Zeichens verwendet: @attributName

## Urheberrechtliches

Diese Dokumentation wurde von NewsIT-M.W.Steidl – [www.newsit.biz](http://www.newsit.biz) – erstellt.

Der Abschnitt NewsML-G2 von KNA wurde im Auftrag und in Zusammenarbeit mit der Katholischen Nachrichten-Agentur GmbH, Bonn, erarbeitet.

Copyright 2013, KNA - Katholische Nachrichtenagentur GmbH, Bonn – [www.kna.de](http://www.kna.de)

Die vorliegende Dokumentation ist urheberrechtlich geschützt. Alle Rechte vorbehalten.

Die Verwendung der Texte und Abbildungen, auch auszugsweise, ist ohne die schriftliche Zustimmung der KNA urheberrechtswidrig. Dies gilt insbesondere für die Vervielfältigung, Übersetzung oder die Verwendung in elektronischen Systemen.

Es wird darauf hingewiesen, dass die in der Dokumentation genannten Produktbezeichnungen und Markennamen der jeweiligen Organisationen oder Firmen internationalem warenzeichen-, marken- und patentrechtlichem Schutz unterliegen.

Diese Dokumentation liegt als PDF oder als auf HTML-Seiten basierendes Web-Help-System vor.

## Kontakt

Bei Fragen oder Anmerkungen zu dieser Dokumentation kontaktieren Sie bitte:

KNA-Produktentwicklung  
Heinrich-Brüning-Str. 9, 53113 Bonn

Tel. +49 (228) 26 00 01 12

Email: [NewsMLG2@kna.de](mailto:NewsMLG2@kna.de)

---

## Was NewsML-G2 ist und kann

NewsML-G2 ist ein umfassendes Datenformat für den Austausch von Nachrichten und anderen Informationen im Nachrichtengewerbe und ist ein Produkt der IPTC (International Press Telecommunications Council – [www.iptc.org](http://www.iptc.org)), dem weltweiten Verband von Nachrichtenagenturen, anderen Nachrichtenherausgebern und Systemherstellern für die Medienbranche

**Das Ziel von NewsML-G2** ist es, alles übermitteln zu können, was von professionellen Herausgebern von Nachrichten verbreitet wird:

- Journalistische Nachrichten in den gängigen Medientypen Text, Bild, Grafik, Audio und Video
- Selektierte und strukturierte Sammlungen von solchen journalistischen Nachrichten, in NewsML-G2 als Pakete bezeichnet.
- Strukturierte Daten über Ereignisse/Termine.
- Daten über die redaktionelle Planung der aussendenden Nachrichtenredaktion („Welche Art von Nachricht in welchem Umfang bis wann auf welchem Dienst geliefert wird.“)
- Die Definition und weltweit eindeutige Kennung von Begriffen, die als Stichworte, Klassifikationen oder sachliche Referenzen des Nachrichteninhalts verwendet werden.
- Ganze Sammlungen (auch Vokabulare genannt) solcher Begriffe.

**Zur Erreichung dieses Zieles definiert NewsML-G2**

- Ein [Datenmodell](#), das sowohl die jeweiligen journalistischen oder begrifflichen Inhalte als auch begleitende Beschreibungen (Metadaten) umfasst
- Ein [Datenformat](#) unter Verwendung der XML Technologie
- Einige Regeln für [betriebliche Abläufe](#)

**Hervorzuhebende Fähigkeiten** von NewsML-G2 sind:

- ***Gemeinsamkeiten vereinfachen den Umgang mit im Detail Unterschiedlichem:***  
Das Datenmodell von NewsML-G2 und die damit verbundenen Strukturen im XML Format bildet alles Gemeinsame der oben beschriebenen übermittelten Inhalte in einer einheitlichen Weise ab. Nur die ganz spezifischen Eigenschaften eines bestimmten Typs von übermittelten Inhalten sind unterschiedlich gestaltet. So kann auf der Seite von NewsML-G2 Empfängern etwa eine Weiterleitung von Inhalten nach Stichworten sehr einfach umgesetzt werden, da diese Stichworte immer an derselben Stelle der Datenstruktur zu finden ist, unabhängig davon, ob es eine Text- oder Bildnachricht oder ein Ereignis/Termin ist.
- ***NewsML-G2 ist präzise – und doch flexibel:***
  - Es stehen viele unterschiedliche den Inhalt beschreibende Daten zur Verfügung
  - Die als beschreibende Daten übergebenen Werte sind jedoch genau definiert: als Zahlenwert, als Datum, als Datum mit Zeit und Zeitzone, oder als Begriff mit einer weltweit einmaligen Kennung. So sollten Unsicherheiten beim empfangenen Stichwort „König“, ob es sich um einen monarchischer Herrscher oder eine Person dieses Namens handelt, nicht mehr auftreten.
- ***NewsML-G2 unterstützt den journalistischen und publizistischen Arbeitsablauf:***
  - Überarbeitungen derselben Nachricht können genau nachverfolgt werden, da alle Änderungen einer Nachricht immer dieselbe Basis-Kennung, aber eine ansteigende Versionsnummer ausweisen.
  - Bei Korrekturen bietet NewsML-G2 eine saubere Trennung zwischen dem geänderten Text – der so wie er ist verwendet werden kann – und einer redaktionellen Erläuterung dessen, was geändert wurde.
  - Die Publikationsfähigkeit einer Meldung ist ein eindeutiges Datenfeld: so kann eine zurückgezogene Meldung sofort entsprechend warnend im empfangenden Redaktionssystem angezeigt werden.
  - Zu jeglichem übermittelten Inhalt können sauber davon getrennt redaktionelle Kommentare und Hinweise übermittelt werden.
  - Eine besondere Neuerung ist jedoch die Verwendung von redaktionellen Planungsdaten: Ankündigungen, dass zu einem Thema Nachrichten gesendet werden, können nun präzise ergänzt werden: bis wann folgen Textnachrichten in welchen Genres, bis wann folgen circa wie viele Fotos oder Videos. Und wenn diese angekündigten Nachrichten ausgesendet werden, dann können sie sich auf die Ankündigung beziehen, d.h. ein empfangendes Redaktionssystem kann anzeigen: „die angekündigten Fotos und die Textmeldung sind eingetroffen, der Volltext des Interviews fehlt noch.“
- ***NewsML-G2 verwaltet Ereignisse/Termine in einer wohlorganisierten Weise.***  
Daher können nun Nachrichten über Ereignisse mit diesen im Voraus zugesendeten Ereignisinformationen direkt verbunden werden. Ein empfangendes Redaktionssystem kann so anzeigen „diese Meldung ist über die angekündigte Pressekonferenz XY“.
- ***NewsML-G2 kann nicht nur einzelne Nachrichten oder Ereignisse/Termine, sondern auch Pakete davon übermitteln:***
  - Zu einer Textmeldung kann ein strikt damit verbundenes Paket von Fotos mitgeliefert werden
  - Nachrichtensammlungen wie „die Top-Meldungen Politik vom heutigen Tag“, oder „die fünf derzeit wichtigsten Meldungen“ können als ein Paket mit einer klar definierten Reihenfolge zugestellt werden.
- ***NewsML-G2 bietet ein einfaches System zur Organisation von Wissen:***  
Viele Herausgeber von journalistischen Inhalten erstellen parallel dazu Verzeichnisse von

Themen, Personen, Organisationen, Unternehmen, Orten usw. und verwenden diese als Referenzen für die Beschreibung des Inhalts. Das Verwalten von entsprechenden Begriffen und die Übermittlung von erstellten Listen an Bezieher wird von NewsML-G2 im Einklang mit gängigen Modellen unterstützt.

---

## Anforderungen an NewsML-G2 Empfänger

NewsML-G2 ist eine Technik der IPTC zur Übermittlung von Nachrichten und verwandten Inhalten, wie sie schon von IPTC 7901, dem IIM oder dem NITF Format bereitgestellt wurde. NewsML-G2 greift allerdings auf neuere Technologien zurück

Der Empfang und die weitere Nutzung von NewsML-G2 stellen folgende Anforderungen, die teilweise über bisherige hinausgehen:

- Kenntnisse rund um NewsML-G2
  - Kenntnisse der grundlegenden Fähigkeiten von NewsML-G2 und der organisatorischen Anforderungen an die empfangende Redaktion und Technik. Diese Anforderung sollte sowohl von Mitarbeitern der Technik als auch der Redaktion wahrgenommen werden.
  - Kenntnisse der mit NewsML-G2 verbundenen betrieblichen Abläufe. Diese Anforderung sollte sowohl von Mitarbeitern der Technik als auch der Redaktion wahrgenommen werden.
  - Kenntnisse des Datenmodells und der eingesetzten Datenformate von NewsML-G2. Diese Anforderung sollte vor allem von Mitarbeitern der Technik wahrgenommen werden.
- Zustand von Übermittlungswegen
  - Es können sowohl Texte als auch bildliche und audible Inhalte übermittelt werden, daher müssen die Übermittlungswege XML Dateien und Binärdateien unterstützen.
  - NewsML-G2 greift auf die Übermittlung von Daten mittels des http Web-Protokolls zurück, daher ist für die weitere Verarbeitung ein Internetzugriff de facto erforderlich.
- Anbindung von NewsML-G2 an eigene Systeme des Empfängers
  - Die Mindestanforderung auf der Empfängerseite von NewsML-G2 ist, dass XML-Dateien ausgewertet und damit verbundene Binär-Dateien weiterverarbeitet werden können.
  - Das empfangende System sollte das das grundlegende [Datenmodell von NewsML-G2](#), insbesondere die Darstellung der den Inhalt beschreibenden Daten, möglichst vollständig in das eigene Datenmodell überführen können.
  - Wesentlich ist, dass die mittels NewsML-G2 übermittelten betrieblichen Abläufe, wie etwa Wiederholungen, Korrekturen oder das Zurückziehen von Inhalten, entsprechend in das empfangende System übernommen werden können.
- Technologien
  - [XML](#) als Format der Dokumente mit denen jedenfalls die Inhalte beschreibende Daten aber nur bei Texten auch die Inhalte selbst übermittelt werden.
  - Kenntnis des [W3C Resource Description Frameworks](#) zum Verständnis der Ausdrucksformen der die Inhalte beschreibenden Daten.
  - Für ein gutes Verständnis der Handhabung von die Inhalte beschreibenden Daten sind Grundkenntnisse des [Semantic Web](#) von Vorteil.

Diese Dokumentation greift alle NewsML-G2 unmittelbar betreffenden Themen und Fakten auf, in darüber hinausgehende Themen wird kurz eingeführt und auf weitergehende Informationen vor allem im Web verwiesen.

---

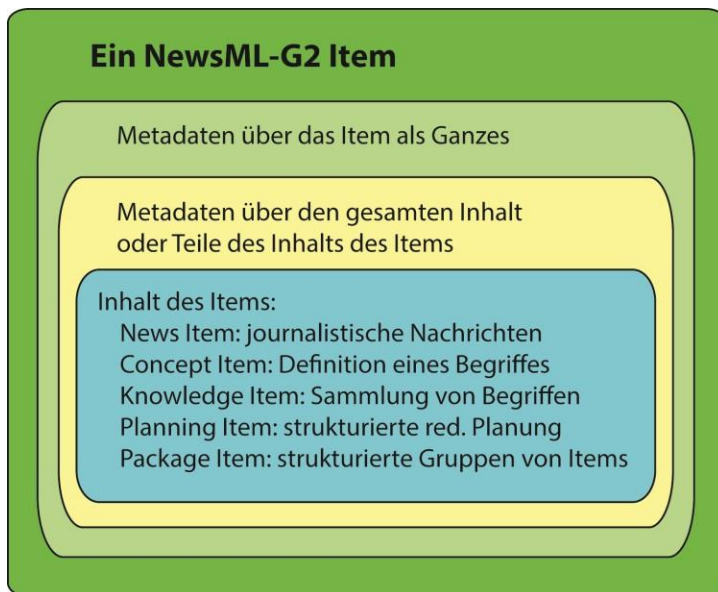
# Das NewsML-G2 Datenmodell

Hinter den konkreten [Formatdefinitionen von NewsML-G2 in XML](#) steht ein umfassendes Datenmodell.

Die Eckpunkte des Datenmodells lauten:

- Jede einzelne Nachricht, jedes einzelne Ereignis, jeder einzelne Begriff wird als eigenständiges Objekt (engl.: Item) verwaltet
- Nachrichtenobjekte können jedoch dieselbe Nachricht in unterschiedlichen technischen Formaten enthalten. Text z.B. als unformatierter Text oder als HTML, Fotos in unterschiedlichen Größen, Videos in unterschiedlichen Codierungen usw.
- Jedes dieser Objekte besitzt eine weltweit einmalige individuelle Kennung (engl: Globally Unique Identifier – GUID oder guid)
- Jedes dieser Objekte kann jedoch über seinen Lebenslauf hinweg in mehreren Versionen vorliegen: als Erstfassung, Überarbeitung, Ausweitung oder Korrektur. Jegliche Veränderung an den Daten eines Objekts muss mit einer Erhöhung der Versionsnummer einhergehen.
- Ein Objekt besitzt folgenden grundlegenden Aufbau:
  - o Dem eigentlichen **Inhalt** des Objekts. Dieser hängt von der Art des Objekts ab
    - Text-Nachricht: ein Meldungstext
    - Bild-Nachricht: ein digitales Bild
    - Infografik-Nachricht: eine digitale Grafik
    - Audio-Nachricht: ein digitalisierter Ton
    - Video-Nachricht: ein digitalisiertes Bewegtbild
    - Ereignis/Termin: strukturierte digitalisierte Daten davon
    - Redaktionelle Planung: strukturierte digitalisierte Daten dazu
    - Begriff: eine weltweit einmalige Kennung, eine Bezeichnung, eine Definition des Begriffs und allenfalls weitere, ergänzende Informationen dazu
    - Sammlung von Begriffen: eine Liste von Begriffen die in der oben beschriebenen Struktur vorliegen
    - Ein Paket an Objekten: eine in Gruppen strukturierte Liste der weltweit einmaligen Kennungen (GUID) der in dieses Paket aufgenommenen Objekte.
  - o **Metadaten, die den eigentlichen Inhalt beschreiben** bzw. betreffen (z.B. Stichwörter, Überschrift, zusammenfassende Beschreibung, der Autor der Meldung, dem Datum der Aufnahme des Bildes, der im Text verwendeten Sprachen.)
  - o **Metadaten, die den Inhalt und dessen Metadaten gesamt beschreiben** (z.B. die weltweit einmalige Kennung und die Versionsnummer, das Datum der Erstellung des Objektes, den Ersteller des Objektes, die Publikationsfähigkeit des Objektes, Urheberrechtinformationen zum Objekt usw.)

Das Diagramm zeigt diesen Aufbau von Inhalt und Metadaten:



- Das Datenmodell der Metadaten folgt den Ideen der Semantischen Technik: aus NewsML-G2 Objekten können Aussagen nach dem Resource Description Framework (RDF) abgeleitet und weiterverarbeitet werden.
- Das von der IPTC definierte Datenmodell hat die Übermittlung von Daten zwischen der ausgehenden Schnittstelle eines Senders (Nachrichtenagentur) und der eingehenden Schnittstelle eines Empfängers im Blick. Das heißt, dass dieses Datenmodell nur sehr begrenzt Informationen zur Verwaltung hausinterner Abläufe anbietet – aber ...
- ... das Datenmodell ist erweiterbar: an einigen Strukturpunkten des Modells ist die Erweiterung des von IPTC definierten Modells durch hauseigene Untermodelle erlaubt und technisch ermöglicht.

---

## Das NewsML-G2 Datenformat in XML

Zur technischen Darstellung des Datenmodells – Serialisierung – wird die XML Technologie verwendet:

- XML 1.0 als Norm für die Syntax
  - XML Schema 1.0 zur Definition von Datentypen und Strukturen.
  - IPTC bietet die laufend weiterentwickelten XML Schemas und eine dazugehörige Spezifikation und Dokumentation von NewsML-G2 auf ihrem Webserver an ([www.newsml-g2.org](http://www.newsml-g2.org))
- Die Nachrichtenagentur KNA verwendet derzeit die Version 2.14 des Standards.

### Die Items

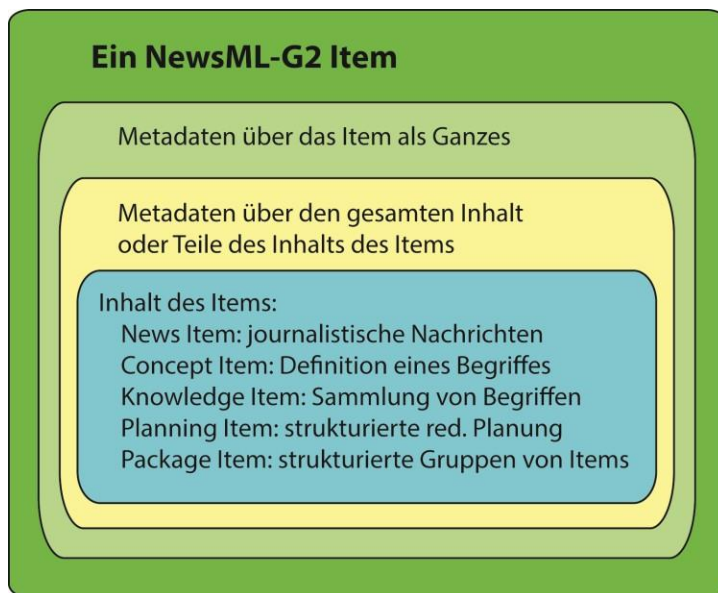
Die im Datenmodell beschriebenen Objekte werden in NewsML-G2 als Item (das „(einzelne) Stück“) bezeichnet, folgende Typen von Items sind als XML Dokumente definiert:

- **<newsItem>** ... für alle Arten von journalistischen Nachrichten
- **<conceptItem>** ... für Concepts, das sind konkret Ereignisse/Termine und alle Arten von Begriffen
- **<knowledgeItem>** ... für eine Sammlung von Concepts.
- **<planningItem>** ... für Daten zur redaktionellen Planung

- **<packagelitem>** ... ein strukturiertes Paket von Referenzen auf bestehende andere Items, das sind in erster Linie newsItems, können aber auch andere, selbst andere packagelItems, sein.
- (In den IPTC Spezifikationen ist auch ein <catalogItem> zu finden, das jedoch für die Übermittlung zwischen Nachrichtenagenturen und ihren Empfängern keine praktische Bedeutung hat.)

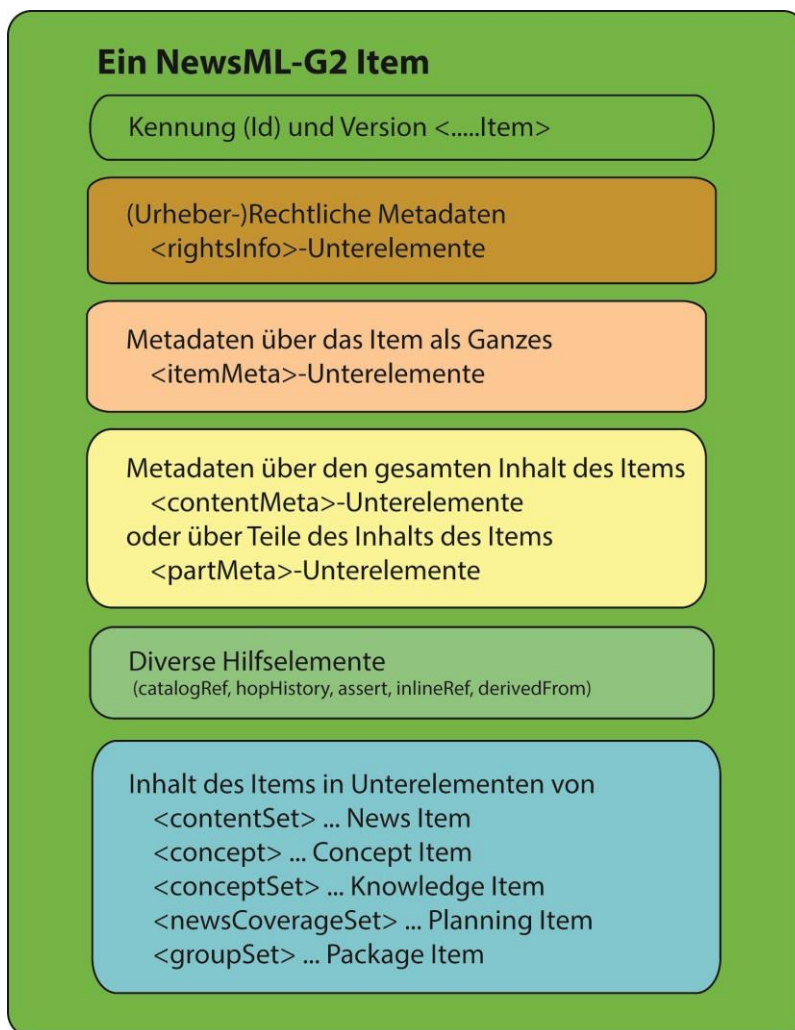
Das allen Items Gemeinsame ist in dieser Grafik dargestellt:

- Der Kern jedes Items ist der Inhalt. Das ist der einzige in der Struktur stark variable Teil eines Item-Typs im Vergleich aller Typen.
- Zum Inhalt eines Items werden diesen beschreibende Daten (= Metadaten) bereitgestellt.
- In einer weiteren äußeren Schicht werden Metadaten bereitgestellt, die das Item als Ganzes beschreiben, also den Inhalt samt den ihn beschreibenden Metadaten. (Beispiel: ein Copyright kann beides umfassen, den Inhalt und dessen Metadaten.)





Dieser geordnete Aufbau des Datenmodells bildet sich auch im Aufbau der XML-Struktur ab, die folgende Grafik gibt einen Überblick:



Den markanten Unterschied zwischen NewsML-G2 Items setzt das Element, das den Inhalt des Items enthält. Siehe obenstehende Grafik, mehr dazu im Abschnitt [XML-Elemente für den Inhalt eines Items](#).

Das folgende XML zeigt als Beispiel eine minimale Struktur eines News Items:

```
<newsItem guid="..." version="1" standard="NewsML-G2" standardversion="2.14" xml:lang="de">
  <catalogRef href="http://www.iptc.org/std/catalog/catalog.IPTC-G2-Standards_22.xml"/>
  <catalogRef href=".....Katalog der Nachrichtenagentur...."/>
  <rightsInfo>
<!-- Urheberrechtlich relevante Informationen -->
  </rightsInfo>
  <itemMeta>
<!-- itemMeta enthält Metadaten die für den Inhalt (content) und dessen Metadaten (contentMeta) gelten -->
    <itemClass qcode="ninat:text">
      <name xml:lang="en-GB">Text Item(s)</name>
    </itemClass>
    <provider qcode="nprov:nag">
      <name xml:lang="de">Eine Nachrichtenagentur</name>
    </provider>
    <versionCreated>2013-10-12T12:10:57+02:00</versionCreated>
    <pubStatus qcode="stat:usable">
      <name xml:lang="en-GB">Usable</name>
    </pubStatus>
  </itemMeta>
  <contentMeta>
<!-- contentMeta enthält nur Metadaten über den Inhalt (content) -->
```

```

    </contentMeta>
    <contentSet>
<!-- contentSet enthält den Inhalt des News Items -->
    <inlineXML contentType="application/xhtml+xml" wordcount="246">
        <html xmlns="http://www.w3.org/1999/xhtml">
            <head>
                <title>Überschrift der Meldung</title>
            </head>
            <body>
                <p>Text der Meldung</p>
            </body>
        </html>
    </inlineXML>
</contentSet>
</newsItem>

```

Der detaillierte Aufbau der Items ist im agenturspezifischen Kapitel der Dokumentation [NewsML-G2 von KNA](#) zu finden.

## Die News Message

Laut Datenmodell sind die NewsML-G2 Items alleinstehende Objekte die keinerlei Information über die Übermittlung zu Kunden besitzen.

Diese Transportinformation ist in NewsML-G2 der `<newsMessage>` vorbehalten. Diese kann als Hülle um ein oder mehrere Items gesehen werden. Dementsprechend werden alle transportierten Items in die XML Struktur einer News Message aufgenommen.

Eine News Message hat folgenden grundlegenden Aufbau:

```

<?xml version="1.0" encoding="UTF-8"?>
<newsMessage xmlns="http://iptc.org/std/nar/2006-10-01/">
    <header>
<!-- Angaben zur konkreten Übermittlung: wann an welchen Dienst gesendet wurde -->
    </header>
    <itemSet>
<!-- In das itemSet können ein bis beliebig viele Items, auch unterschiedlicher Art, eingeklinkt werden. -->
        <newsItem guid="..." version="1" standard="NewsML-G2" standardversion="2.14" xml:lang="de">
<!-- Inhalt eines News Items -->
        </newsItem>
    </itemSet>
</newsMessage>

```

---

## XML-Elemente für den Inhalt eines Items

Wie im Datenmodell beschrieben, unterscheiden sich die NewsML-G2 Items vor allem dadurch, dass sie unterschiedliche Arten von Inhalten transportieren.

Generell gilt für die Struktur eines Items: das Element, das den Inhalt transportiert ist das letzte Unterelement des Wurzelementes. Siehe auch die Übersicht im Abschnitt [Die Items](#)

Daraus ergibt sich folgender spezifischer Aufbau der NewsML-G2 Items:

```

<newsItem>
    ... die gemeinsamen Elemente aller Items: <catalogRef>, <rightsInfo>, <itemMeta>, <contentMeta>
    ... sowie: <partMeta>, <assert>, <inlineRef>, <derivedFrom>
    ... und jetzt das den Nachrichteninhalte liefernde Element
    <contentSet>
        ... mit den Nachrichteninhalten ...
    </contentSet>
</newsItem>

<conceptItem>

```

... die gemeinsamen Elemente aller Items: `<catalogRef>`, `<rightsInfo>`, `<itemMeta>`, `<contentMeta>`  
 ... sowie: `<assert>`, `<inlineRef>`, `<derivedFrom>`  
 ... und jetzt das den Inhalt liefernde Element  
`<concept>`  
 ... mit dem Concept/Begriff ...  
`</concept>`  
`</conceptItem>`

`<knowledgeItem>`  
 ... die gemeinsamen Elemente aller Items: `<catalogRef>`, `<rightsInfo>`, `<itemMeta>`, `<contentMeta>`  
 ... sowie: `<partMeta>`, `<assert>`, `<inlineRef>`, `<derivedFrom>`  
 ... und jetzt das den Inhalt liefernde Element  
`<conceptSet>`  
 ... mit der Sammlung von Concepts/Begriffen ...  
`</conceptSet>`  
`</knowledgeItem>`

`<planningItem>`  
 ... die gemeinsamen Elemente aller Items: `<catalogRef>`, `<rightsInfo>`, `<itemMeta>`, `<contentMeta>`  
 ... sowie: `<partMeta>`, `<assert>`, `<inlineRef>`, `<derivedFrom>`  
 ... und jetzt das den Inhalt liefernde Element  
`<newsCoverageSet>`  
 ... mit den Planungsinhalten ...  
`</newsCoverageSet >`  
`</planningItem>`

`<packageItem>`  
 ... die gemeinsamen Elemente aller Items: `<catalogRef>`, `<rightsInfo>`, `<itemMeta>`, `<contentMeta>`  
 ... sowie: `<partMeta>`, `<assert>`, `<inlineRef>`, `<derivedFrom>`  
 ... und jetzt das den Paketinhalt liefernde Element  
`<groupSet>`  
 ... mit den gruppierten Referenzen auf andere Items ...  
`</groupSet>`  
`</packageItem>`

In welcher Form genau diese Inhalte ausgeliefert werden, ist im Abschnitt [„NewsML-G2 von KNA“](#) zu finden.

---

## NewsML-G2 und betriebliche Abläufe

Journalistische Nachrichten, Daten über Ereignisse/Termine und Mitteilungen der sendenden Redaktion über redaktionelle Planungen werden immer im Kontext von vereinbarten betrieblichen Abläufen übermittelt, in NewsML-G2 wird dies so abgebildet:

- Die üblichen formalen Mitteilungen werden in genau definierten Datenstrukturen übermittelt:
  - Der Publikationsstatus einer Meldung („kann verwendet werden“, „derzeit nicht verwenden“, „wird zurückgezogen“)
  - Die Sperrfrist
  - Ein klarer formaler Hinweis auf eine Korrektur mit Erläuterungen der Redaktion
    - Die journalistische Dringlichkeit einer Meldung („Alarm“, „Blitz“ „Eil“ ...)
- Jedes in Rahmen von NewsML-G2 übermittelte Objekt hat eine weltweit einmalige Kennung. Das tägliche Wiederholen von Laufnummern als Meldungskennung und damit die täglich auftretende Nichteindeutigkeit kann daher kein Problem mehr.

- Von übermittelten Objekten können Versionen verbreitet werden. Damit ist es einfach, Überarbeitungen, Korrekturen und auch Zurückziehungen von Objekten exakt nachzuverfolgen – sie alle besitzen dieselbe Kennung, aber mit ansteigenden Versionsnummern.
- Mitteilungen über redaktionelle Planungen der sendenden Redaktion werden nicht mehr als Freitext, sondern als detaillierte Angaben in Form von strukturierten Daten übergeben. So kann als Planung zu einem Ereignis oder Thema zum Beispiel festgestellt werden:
  - eine Textmeldung wird bis 15:00 auf dem Basisdienst verbreitet
  - mindestens 5 Fotos werden bis 15:10 auf dem Bildkanal verbreitet
  - sowohl Textmeldung als auch Fotos besitzen die Ereignis/Termin-Kennung „ET4711“

---

## Metadaten in NewsML-G2

Wie im [Datenmodell](#) angesprochen, sind Metadaten all jene Datenstrukturen, mit denen eine Aussage oder Feststellung über den Inhalt eines Objekts (Items) oder ein Item als Ganzes getroffen wird.

NewsML-G2 folgt darin dem Datenmodell des [Resource Description Frameworks \(RDF\)](#) der W3C.

Das RDF Datenmodell folgt dem Aufbau eines einfachen Satzes:

Subjekt (S) - Prädikat (P) – Objekt (O)

z.B: Diese Meldung (= S) – ist über (= P) – den Eiffelturm (= O)

(Hinweis: im Rahmen von RDF wird diese dreiteilige Struktur als Triple bezeichnet.) NewsML-

G2 folgt diesem Modell in folgender Weise:

- Für das Subjekt der Aussage gibt es drei Bezugspunkte:
  - Alle Metadaten, die in der Item-Struktur als Unterelemente von <contentMeta> stehen, machen eine Aussage über den gesamten Inhalt des Items. Das konkrete Subjekt einer Aussage ist
    - in einem News Item immer implizit „diese Meldung“, „dieses Foto“, „diese Grafik“, „dieses Video“ ....
    - in einem Package Item immer explizit „dieses Paket von Items“;
  - Alle Metadaten, die in der Item-Struktur als Unterelemente von <partMeta> stehen, machen eine Aussage über einen bestimmten Teil des Inhalts. Welcher Teil des Inhalts dies ist, wird entweder über das Attribut @contentrefs des <partMeta> Elements definiert, oder über die Unterelemente <timeDelim> oder <regionDelim>.
  - Alle Metadaten, die in der Item-Struktur als Unterelemente von <itemMeta> stehen, machen eine Aussage über das NewsML-G2 Item als Ganzes. (Damit letztlich auch über die Metadaten unter <itemMeta>, aber das stellt kein sachliches Problem dar.)
- Für das Prädikat der Aussage gilt: dieses wird durch die inhaltliche Definition des Metadaten-Elements festgelegt – siehe die folgenden Beispiele.
- Für das Objekt der Aussage gilt: dieses wird durch den Wert des Metadaten-Elements festgelegt – siehe die folgenden Beispiele.
- **Beispiele** von Metadatenelementen und ihren Werten:
  - <versionCreated> (unter <itemMeta>) ist definiert als „das Datum (und optional die Zeit) an dem diese Version des Items erstellt wurde“. Wenn dieses Element den Wert „2013-10-01T13:42:28+02:00“ ausweist, lautet die entsprechende RDF-Aussage:  
Von diesem Item (= S) – wurde diese Version erstellt (= P) – am 1. Oktober 2013 um 13:42:28 Uhr Mitteleuropäischer Sommerzeit.
  - <language> (unter <contentMeta>) ist definiert als „Sprache, die im Nachrichteninhalt verwendet wird“. Wenn dieses Element den Wert (im @tag Attribut) „de“ ausweist,

lautet die entsprechende RDF-Aussage:

Dieser Nachrichteninhalt (= S) – verwendet die Sprache (= P) – Deutsch (= O).

- <subject> (unter <contentMeta>) ist definiert als „ein wichtiges Thema des Inhalts bzw. etwas, wovon der Inhalt handelt“. Wenn dieses Element den Wert (im Unterelement <name>) „Eiffelturm“ ausweist, lautet die entsprechende RDF-Aussage:

Dieser Inhalt (= S) – handelt vom (= P) – Eiffelturm (= O).

- **Hinweis:** um die konkrete Bedeutung eines Metadaten-Elements kennenzulernen, sollte nicht nur dessen Name, sondern auch dessen Beschreibung in den NewsML-G2 Spezifikationen nachgeschlagen werden.
- Verfeinerung der Bedeutung eines Metadaten-Elements: in einigen Fällen ist die Definition der Bedeutung eines Elements so breit gefasst, dass zur Präzisierung eine verfeinernde Einschränkung durchgeführt werden kann und manchmal auch als notwendig erscheint. Beispiel: <creator> = Ersteller des Inhalts. Schon bei einem Text können das mehrere Personen sein, die unterschiedliche Rollen bei der Erstellung ausgeübt haben, und bei einem Video kann es einen Reporter, einen Kameramann und einen Tonmeister geben. Um solche Rollen ausdrücken zu können, stellt NewsML-G2 bei vielen Metadaten-Elementen ein @role Attribut bereit. In dieses kann dann die spezielle Rolle einer Person im Rahmen der allgemeinen Definition „Erstellerin des Inhalts“ eingetragen werden.

## Ausdrucksformen von RDF-Objekten

Die Ausdrucksformen für die Subjekte und Prädikate der RDF-Metadatenaussage sind relativ einfach:

- Das Subjekt wird aus der Position des Metadaten-Elements in der Gesamtstruktur abgeleitet. Siehe [Metadaten in NewsML-G2](#).
- Ein Prädikat entspricht der Bedeutung eines Metadaten-XML-Elements, falls vorhanden zusammen mit einem @role Attribut.
- Ein Objekt entspricht einem Wert der durch ein Metadaten-XML-Element ausgedrückt wird.

In welcher Form ein Objekt-Wert ausgedrückt wird, hängt von der Type des Wertes ab:

- **Datum, Datum mit Uhrzeit und Zeitzone:** werden immer im von XML adoptierten ISO 8601-Format festgehalten und erscheinen als Zeichenkette im Textknoten des XML Elements:
  - nur Datum: JJJJ-MM-TT
  - Datum mit Uhrzeit und Zeitzone: JJJJ-MM-TT“T“hh:mm:ss.sss±hh:mm**Hinweis:** bei einer Zeitangabe ist die Angabe einer Zeitzone verpflichtend, denn ohne sie ist die angegebene Zeit nicht mit anderen vergleichbar.
- **Zeichenkette** (z.B. Überschrift = <headline>): erscheint im Textknoten des XML Elements
- **Zahlenwert** (z.B. Dringlichkeit = <urgency>): erscheint als Ziffer oder Ziffernfolge im Textknoten des XML Elements.
- **Begriff:** wird mittels einer etwas komplexeren aber sehr leistungsfähigen XML-Struktur ausgedrückt. Sie gibt eine Sammlung von Faktischem und Wissenswertem über ein Ding, eine Sache, ein Person, eine Organisation, einen Ort, eine Kategorie, ein Ereignis, usw. wieder. Ein solcher Begriff wird in NewsML-G2 als Concept bezeichnet.

Begriffe/Concepts spielen in NewsML-G2 eine wichtige Rolle, da sie den Umgang mit Daten in der Art und Weise des Semantischen Webs in diesem Standard abbilden:

- solche Begriffe/Concepts haben in der Regeln eine weltweit einmalige Kennung
- solche Begriffe/Concepts können eine Vielfalt an Detaildaten über das beschriebene Ding haben (bei einer Person etwa Geburtsdatum, Größe, Augenfarbe)
- solche Begriffe/Concepts können auch eine Vielfalt an Beziehungen zu anderen Begriffen/Concepts haben, typischer Weise sowohl hierarchische Beziehungen (Oberbegriffe, Unterbegriffe) als auch freie, sachliche Beziehungen (ein Person hat z.B.

einen Arbeitgeber – also eine Beziehung zu einer Organisation – oder eine Stadt, in der sie lebt – also eine Beziehung zu einem geografischen Ort)

Die Ausdrucksform eines Begriffes hat in NewsML-G2 zwei Varianten:

- Eine Datenstruktur als Unterstruktur des Metadaten-Elements.  
Typische Elemente dieser Unterstruktur sind:  
<name> = die natürlichsprachliche Bezeichnung des Begriffs  
<definition> = eine präzise Beschreibung des Begriffs  
Mehr dazu in Abschnitt über [Begriffe](#).
- Eine weltweit einmalige Kennung für den Begriff die als URI ausgedrückt ist.  
Dieser URI kann in zwei Formaten festgehalten werden:  
- in unverkürzter Form als Wert des @uri Attributs des Metadaten-XML-Elements  
- in der verkürzten Form eines QCodes als Wert des @qcode Attributes des Metadaten-XML-Elements. Mehr dazu in Abschnitt über [Begriffe](#).

---

## Begriffe - ihre IDs und QCodes

Begriffe sind für die semantische Erschließung von Nachrichteninhalten wesentlich, insbesondere als eine der beiden Alternativen für RDF-Objekte in Metadaten-Aussagen:

- 1. Alternative: Man kann Inhalte mit konkreten Datums- (z.B. Datum der Erstellung) oder Zahlenwerten (z.B. Länge des Textes in Wörtern) beschreiben.  
Man kann Inhalten textliche Beschreibungen zuordnen (z.B. eine Kurzbeschreibung, Zusammenfassung)  
Diese Art von Werten wird als Literal bezeichnet.
- 2. Alternative: Man kann Inhalte aber auch mit „Dingen dieser Welt“ in Zusammenhang bringen: mit Themen, Personen, Orten, Ereignissen, Objekten, usw. (z.B. „dieser Artikel handelt von dem Maler Franz Mayer und der Stadt Paris“)
- Während Datumswerte, Zahlen oder beschreibender Text für sich sprechen, ist die Darstellung von „Dingen dieser Welt“ etwas komplizierter: man muss andere wissen lassen, welches „Ding dieser Welt“ man genau meint. Das scheint auf ersten Blick einfach, kann aber bei genauerer Recherche mehrdeutig werden: z.B. der Stadtname Paris steht für die Metropole von Frankreich, aber auch für mindestens 9 Städte in den USA, das heißt die Angabe der Zeichenkette „Paris“ als Metadatenwert ist mehrdeutig.
- Die große Herausforderung für den Umgang mit „Dingen dieser Welt“ besteht nun darin, diese als RDF-Objekte eindeutig darzustellen.

Ein wesentlicher Schritt zur Bewältigung dieser Herausforderung ist es, ein „Ding dieser Welt“ so zu beschreiben, dass möglichst jeder sich davon einen Begriff machen kann, das heißt, er genau weiß, um welches Ding es dabei geht. Daher wird eine solche Beschreibung eines „Dings dieser Welt“ als Begriff (engl. Concept) bezeichnet.

Wenn man nun so ein „Ding dieser Welt“ mit Nachrichteninhalten in Zusammenhang bringt, kann man nun diesen ganzen niedergeschriebenen Begriff zur unmittelbaren Definition des Dings verwenden.

Es wäre jedoch sehr aufwändig, wenn zum Beispiel von jeder in einer Nachricht vorkommenden Person eine umfassende Beschreibung in die Metadaten aufgenommen werden müsste, um zu wissen, um wen es genau gegangen ist.

Darum ist ein zweiter wesentlicher Schritt, diesem Begriff von einem „Ding dieser Welt“ eine weltweit einmalige Kennung zuzuordnen und nur diese Kennung zur Definition des in Zusammenhang gebrachten „Dings dieser Welt“ zu verwenden.

Wenn nun diese Kennung als Metadatenwert vorgefunden wird, dann muss man sich an die diese Kennung ausgebende Stelle (Institution, Unternehmen; ...) wenden um den damit gekennzeichneten Begriff abzufragen und damit festzustellen, um wen es sich genau handelt. Oder die Kennung liegt in



Form eines URL vor, was von der IPTC und NewsML-G2 empfohlen wird, und wenn dieser aufgerufen wird, dann wird eine mehr oder minder lange Beschreibung des Dings ausgeliefert. Der Umfang dieser Beschreibung kann mit wirtschaftlichen Überlegungen verbunden sein: kostenlos wird nur der Name des Dings geliefert, zahlende Kunden erhalten umfangreiche Beschreibungen.

Zusammenfassung mit terminologischen Festlegungen:

- Es geht um ein ganz bestimmtes „Ding dieser Welt“ (engl. Thing) das im Zusammenhang mit einem (Nachrichten-)Inhalt steht. Ein „Ding dieser Welt“ kann ein Thema, eine Person, ein Ort, ein Ereignis, ein materielles Objekt, usw. sein.
- Eine im Wissensmanagement tätige Organisation oder Unternehmen kann eine genaue und möglichst eindeutige Beschreibung dieses Dings festschreiben. Diese Beschreibung wird in NewsML-G2 als Concept (dt. Begriff) bezeichnet.
- Diesem Concept bzw. Begriff kann auch eine weltweit einmalige Kennung (in NewsML-G2: Concept-Id) zugeordnet werden. NewsML-G2 schlägt vor, dazu einen URL zu verwenden.
- Wenn nun von einem (Nachrichten-)Inhalt ein Zusammenhang mit einem solchen Ding hergestellt werden soll, dann kann man
  - o entweder dazu die ganze Beschreibung (= Concept bzw. Begriff) heranziehen
  - o oder man kann dazu deren weltweit einmalige Kennung (Concept-Id) verwenden
  - o oder man kann eine Mischung aus beidem verwenden:
    - die Menschen oft nicht viel sagende Kennung (Concept-Id)
    - und dazu einige wesentlich Teile der Beschreibung des Dings, etwa dessen Name.
- **Hinweis:** Es ist möglich, dass mehrere unterschiedliche Personen/Organisationen/Unternehmen ein „Ding dieser Welt“ eigenständig beschreiben und dazu jeweils eigene, das heißt auch unterschiedliche, weltweit einmalige Kennungen ihrer eigenen Beschreibung zuordnen. (Das ist vergleichbar mit den Beschreibungen desselben Dings in unterschiedlichen Lexika.) Darum ist es wichtig zu verstehen, dass eine solche Kennung nur für eine Beschreibung des „Dings dieser Welt“ steht und nicht für das „Ding dieser Welt“ selbst! Mit anderen Worten: für ein und dasselbe „Ding dieser Welt“ kann es mehrere unterschiedliche begriffliche Beschreibungen und damit unterschiedliche Kennungen dieser Beschreibungen geben.

## Begriffe/Concepts in NewsML-G2

Die Beschreibungen eines „Dings dieser Welt“, also ein Concept (dt. Begriff) dieses Dings hat in NewsML-G2 eine klar definierte Struktur.

Das diese Struktur umfassende Element ist das <concept>, es hat folgende Unterelemente:

- **<conceptId>** enthält die weltweit einmalige Kennung dieses Concepts. Mehr dazu im nächsten Abschnitt.
- **<type>** zeigt den Typus des Dings an: Person, Organisation, Ort, materielles Objekt, Ereignis ...
- **<name>** ist die natürlichsprachliche Bezeichnung des Dings in einer bestimmten Sprache (und Schrift). (Die Namen von Personen können noch feiner unterteilt werden)
- **<definition>** ist die möglichst eindeutige natürlichsprachliche Beschreibung des Dings in einer bestimmten Sprache (und Schrift) durch den Ersteller dieses Concepts/Begriffs.
- **<note>** ist eine Anmerkung zu diesem Ding in einer bestimmten Sprache (und Schrift).
- **<remoteInfo>** ist ein Verweis auf andere Beschreibungen dieses Dings. Diese Beschreibungen können auch anders als durch Text erfolgen: z.B. Fotos, Videos, Grafiken ...
- **<personDetails>** eine Gruppe von 4 Metadaten-Elementen die typisch für eine Person sind

- **<organisationDetails>** eine Gruppe von 5 Metadaten-Elementen die typisch für eine Organisation/ein Unternehmen sind
- **<geoAreaDetails>** eine Gruppe von 6 Metadaten-Elementen die typisch für einen geografischen Punkt oder ein geografisches Gebiet sind
- **<POIDetails>** eine Gruppe von 9 Metadaten-Elementen die typisch für einen Ort Von Interesse (Point Of Interest) sind
- **<objectDetails>** eine Gruppe von 4 Metadaten-Elementen die typisch für ein materielles Objekt sind
- **<eventDetails>** eine Gruppe von 13 Metadaten-Elementen die typisch für ein Ereignis (einen Termin) sind

(Anmerkung zu den ...Details: die große Anzahl an Unterelementen wird bei konkreter Verwendung im agenturspezifischen Teil der Dokumentation erläutert um hier nicht die Übersichtlichkeit einzuschränken.)

... darunter eine Gruppe von Elementen zur Definition von Beziehungen dieses Concepts bzw. dieses Dings zu anderen:

- **<sameAs>** definiert, dass der im <sameAs> Element angegebene Begriff dasselbe Ding beschreibt wie dieses Concept/Begriff. sameAs wird vor allem dazu verwendet, von einem Begriff eines Vokabulars auf den de facto identischen Begriff in einem anderen Vokabular zu verweisen.
- **<broader>** definiert, dass der im <broader> Element angegebene Begriff einen weiter gefassten Begriff darstellt. (Beispiel: Zum Bundesland Bayern ist ein <broader> Begriff der Staat Deutschland. )
- **<narrower>** definiert, dass der im <broader> Element angegebene Begriff einen enger gefassten Begriff darstellt. (Beispiel: Zum Bundesland Bayern ist ein <narrower> Begriff die Stadt München. )
- **<related>** definiert die Art der Beziehung, die dieses Ding zu einem anderen Ding hat. Dieses Element kann sehr flexibel eingesetzt werden, weil es ermöglicht, die Art der Beziehung frei zu definieren, während <sameAs>, <broader> und <narrower> nur eine einzige klar definierte Beziehung ausdrücken.

Viele dieser Unterelemente eines <concept>s können auch als Unterelemente von Metadaten-Elementen verwendet werden.

Beispiele:

```
<subject>
  <type qcode="cpnat:person"/>
  <name>Franz Mayer</name>
</subject>

<located>
  <type qcode="cpnat:geoArea"/>
  <name>Bonn</name>
  <broader>
    <name>Nordrhein-Westfalen</name>
  </broader>
</located>
```

## Kennung eines Begriffs - Concept-Id

Nach den Spezifikationen von NewsML-G2 muss die weltweit einmalige Kennung eines Concepts/Begriffes ein Uniform Resource Identifier (URI) sein.

Für die Praxis empfehlen die NewsML-G2 Spezifikationen die Verwendung eines Web-URLs, also eines URL aus dem Schema http. Dies hat den Vorteil, dass im URL die Internet Domain des Unternehmens, das ein solches Concept erstellt hat, verwendet werden kann: z.B.

<http://www.dienachrichtenagentur.com/begriffe/orte/berlinde> für ein Concept/Begriff über die Hauptstadt von Deutschland, Berlin. Und wenn dieser URL aufgerufen wird, dann könnte der



Besitzer der Kennung zumindest Minimalinformationen über das Ding, wie etwa dessen Namen, ausliefern.

In der NewsML-G2 Struktur eines <conceptId> kann außerdem angegeben werden:

- Wann dieser conceptId vergeben wurde (@created Attribut)
- Ab wann dieser conceptId nicht mehr aktiv verwendet werden sollte (@retired Attribut). Ein Grund für diese nicht-aktive Verwendung ist etwa, dass dieses Ding, das der Begriff beschreibt, nicht mehr existiert. (Beispiel: der Staat Tschechoslowakei existiert seit Ende 1992 nicht mehr, so dass der ISO Code für den Landesnamen, cs, heute nicht mehr aktiv verwendet werden sollte.)

Die ungekürzte Concept-Id eines Concepts/Begriffs kann zur Herstellung eines Zusammenhangs zwischen einem (Nachrichten-)Inhalt eines Items und einem Concept in das @uri Attribut des Metadaten-Elements eingefügt werden.

Beispiele:

```
<located uri="http://www.dienachrichtenagentur.com/begriffe/orte/berlinde"/>  
<subject uri="http://www.dienachrichtenagentur.com/begriffe/personen/fmayer1383" />
```

## Kurzfassung von Kennungen = QCodes

Um die in vielen Fällen langen URIs von Concept-Ids zu verkürzen hat die IPTC in enger Anlehnung an die Qnames von XML das URI verkürzende Format QCodes definiert – hier ist eine vollständige Beschreibung (in Englisch): [www.qcodes.org](http://www.qcodes.org)

**Hintergrund:** bei der Spezifikation von NewsML-G2 wurde die Anforderung zur Verkürzung von URIs diskutiert. Eine sehr bekannte Methodik sind die Qnames von XML: der auch oft sehr lange URI eines XML Namespace wird durch einen kurzen Präfix ersetzt.

Beispiel:

Namespace von NewsML-G2: <http://iptc.org/std/nar/2006-10-01/>

Bevorzugter Prefix dieses Namespace: nar

Der Qname des Elements <creator> lautet <nar:creator>

Diese Lösung war jedoch nicht für Concept-Ids zu verwenden, da es bei allen größeren Nachrichtenagenturen Begriffsverzeichnisse für die Klassifizierung von Nachrichten, für Unternehmen, für Personen usw. gab und die verzeichnisinternen Kennungen für die Begriff oft mit einer Ziffer oder einem Sonderzeichen (z.B. \$) beginnen – diese Zeichen sind jedoch am Beginn eines lokalen Namens in XML verboten.

Daher hier **das Wichtigste** zusammengefasst:

- QCodes gehen davon aus, dass Concepts immer im Rahmen eines Vokabulars (Verzeichnisses) definiert werden – also z.B. ein Vokabular von Personen, eines von Organisationen, eines von Orten, eines von Ereignissen ...
- Jedes dieser Vokabulare bekommt einen eigenen URI/URL als Kennung. Im Falle eines URL muss dieser mit einem / oder einem # enden.
- Jedes Concept/Begriff bekommt einen Code, der innerhalb des Vokabulars einmalig sein muss. Für diesen Code dürfen alle Zeichen verwendet werden, die in URIs erlaubt sind – verboten sind etwa Zeilenvorschub (line feed), Wagenrücklauf (carriage return), oder Tabulator.
- Die Zusammenfügung von Vokabular-URI (links) und Code (rechts) ergeben die vollständige Concept-Id, die wiederum eine URI sein muss, bzw. ein URL sein kann.
- Die Kurzform QCode wird so gebildet:
  - der Vokabular URI wird durch seine Kurzform namens Alias ersetzt
  - der Alias wird wie in XML Qname links von einem Doppelpunkt gesetzt
  - und rechts von den Doppelpunkt wird der Code gesetzt.

Beispiel mit weiteren Details:

Die Zuordnung des Alias zum Vokabular erfolgt in einem NewsML-G2 Catalog der so aussieht:

```
<catalog xmlns="http://iptc.org/std/nar/2006-10-01/">  
  <scheme alias="medtop" uri="http://cv.iptc.org/newscodes/mediatopic/" />  
</catalog>
```

Hinweis: dieser Catalog wird in der Regel auf einem Webserver als Datei hinterlegt und dann von NewsML-G2 Items aus so aufgerufen:

```
<catalogRef href="http://www.iptc.org/std/catalog/catalog.IPTC-G2-Standards_19.xml"/>
```

Das Concept des Dings "Massenmedien" erhält den Code 20000045

Daraus ergibt sich:

Die Concept-Id als vollständiger URL: <http://cv.iptc.org/newscodes/mediatopic/20000045>

Der kurze QCode: medtop:20000045

### ***Die Auflösung eines QCodes zu einem Concept-Id URI***

In einem NewsItem findet sich dieses Metadaten-Element: <subject qcode="medtop:20000045">

Daraus ergibt sich:

Der Alias lautet: medtop

Der Code lautet: 20000045

Nun müssen die Cataloge des Items überprüft werden:

- alle über <catalogref href=http://... /> referenzierten Catalog-Dateien aufrufen und öffnen
- Suche nach einem <scheme> Element, das im Attribut @alias den Wert „medtop“ hat:

```
<scheme alias="medtop" uri="http://cv.iptc.org/newscodes/mediatopic"/>
```

- der im Attribut @uri angegebene Wert ist der gesuchte Vokabular-URI

Nun Vokabular-URI und Code aneinanderfügen ergibt die Concept-Id:

<http://cv.iptc.org/newscodes/mediatopic/20000045>



# NewsML-G2 von KNA

---

## Übersicht

NewsML-G2 hat bei allen Nachrichtenagenturen, die es verwenden, eine spezifische Ausformung, so auch bei KNA.

Diese agenturspezifischen Ausformungen werden in diesem Abschnitt dokumentiert, die allgemeinen Datenstrukturen wurden bereits im Abschnitt [NewsML-G2 Datenformat](#) erläutert.

Die Notwendigkeit von Agenturspezifika hat weniger mit technischen Aspekten oder Beschränkungen von NewsML-G2 zu tun, als mit den lokalen Eigenheiten der Nachrichtenbranche, die eine weltweit einheitliche Definition nicht möglich macht:

- Die Aufgliederung von Redaktionen in Ressorts
- Jede Nachrichtenagentur hat ihre eigenen Begriff-Vokabulare, die deren Klassifizierungsgeschichte und –form reflektieren. Zur besseren weltweiten Austauschbarkeit von Nachrichten können diese mit den international gebräuchlichen Vokabularien der IPTC (den NewsCodes) in Zusammenhang gebracht werden.
- Details des Nachrichtenflusses sind regional unterschiedlich. Für Deutschland gibt es spezifische Abläufe und darüber hinaus KNA-spezifische Details.

---

## News Message

Das ist ein Beispiel für das XML einer NewsML-G2 News Message der KNA mit einer kurz Erläuterung über den wichtigsten Elementen, weitere Details dazu werden unterhalb dieses XML Beispiels erläutert.

```
<newsMessage xmlns="http://iptc.org/std/nar/2006-10-01/" xmlns:h="http://www.w3.org/1999/xhtml">
  <header>
<!-- Es folgt: die Zeit, zu der diese News Message ausgesendet wurde -->
    <sent>2013-08-16T11:34:58+02:00</sent>
<!-- Es folgen: die für die News Message notwendigen Catalog-Referenzen -->
    <catalogRef href="http://www.iptc.org/std/catalog/catalog.IPTC-G2-Standards_19.xml"/>
    <catalogRef href="http://www.kna.de/rtech/catalog001.xml"/>
    <catalogRef href="http://g2.dpa.com/catalog/catalog001.xml"/>
<!-- Es folgt: die Quelle der Verbreitung, das KNA Redaktionssystem namens "ines" -->
    <origin>ines</origin>
<!-- Es folgt: die Kennungen des Dienstes/der Dienste, über die diese News Message verbreitet wurde -->
    <destination qcode="knaprod:BD">Basisdienst</destination>
  </header>
  <itemSet>
<!--hier befindet sich eines oder mehrere NewsML-G2 items -->
  </itemSet>
</newsMessage>
```

## Erläuterungen

Die Zeit, zu der die News Message ausgesendet wurde, findet sich im `<sent>` Element. Diese Zeit ist nicht unbedingt identisch mit dem Zeitpunkt, an dem die Bearbeitung des Inhaltes abgeschlossen wurde. Diese Zeit findet sich im jeweiligen NewsML-G2 Item.

Die KNA-Dienste, über die diese News Message verbreitet wurde, sind als einzelne `<destination>` Elemente ausgeführt. Der Text-Wert des Elements gibt in der Regel eine gut erkennbare Kurzbezeichnung des Dienstes. Fragen Sie eventuell im KNA-Marketing nach, mit welchen Kennungen und Kurzbezeichnungen die aktuellen KNA-Dienste ausgestattet sind.

---

## Text-Nachrichten

### KNA-spezifische Daten

Das ist ein Beispiel für das XML einer NewsML-G2 Text-News Item der KNA mit einer kurz Erläuterung über den wichtigsten Elementen, weitere Details dazu werden unterhalb dieses XML Beispiels erläutert.

```
<newsItem xmlns="http://iptc.org/std/nar/2006-10-01/" xmlns:h="http://www.w3.org/1999/xhtml"
          conformance="power"
          guid="urn:newsml:kna.de:20130101:130816-89-00031" version="4" standard="NewsML-
          G2" standardversion="2.14"
          xml:lang="de">
<!-- Oben: guid gibt die weltweit einmalige Kennung und version die Version des Items an -->
<!-- Es folgen: die für das News Item notwendigen Catalog-Referenzen -->
  <catalogRef href="http://www.iptc.org/std/catalog/catalog.IPTC-G2-Standards_19.xml"/>
  <catalogRef href="http://www.kna.de/rtech/catalog001.xml"/>
  <catalogRef href="http://g2.dpa.com/catalog/catalog001.xml"/>
<!-- Es folgt: ein Block an Information zum rechtlichen Aspekten -->
  <rightsInfo>
<!-- Es folgt: die im Sinne des Presserechts bei KNA verantwortliche Person -->
    <accountable literal="Iri">
      <definition>Chefredakteur Ludwig Ring-Eifel</definition>
      <note>Verantwortlicher im Sinne des deutschen Presserechts</note>
    </accountable>
<!-- Es folgt: der aktuelle Inhaber der Nutzungsrechte am Inhalt dieses News Items -->
    <copyrightHolder qcode="knacprh:kna">
      <name xml:lang="de">KNA</name>
      <definition>
        KNA - Katholische Nachrichten-Agentur GmbH
      </definition>
      <copyrightHolder>
        Registergericht Bonn, HRB 5605
      </copyrightHolder>
<!-- Es folgt: der international übliche Copyright Vermerk -->
    <copyrightNotice>Copyright 2013, KNA (www.kna.de). Alle Rechte vorbehalten</copyrightNotice>
<!-- Es folgt: eine Aussage über die auf den Inhalt dieses News Items zutreffenden Nutzungsrechte -->
    <usageTerms>Nutzung nur nach schriftlicher Vereinbarung mit der KNA</usageTerms>
  </rightsInfo>
  <itemMeta>
<!-- Es folgt: die Angabe über den Medientype des News Items -->
    <itemClass qcode="ninat:text">
      <name xml:lang="en-GB">Text Item(s)</name>
    </itemClass>
<!-- Es folgt: die Angabe, welches Unternehmen das News Item erstellt und verbreitet hat -->
    <provider qcode="knaprov:kna">
      <name xml:lang="de">KNA</name>
    </provider>
<!-- Es folgt: der Zeitpunkt, zu dem diese Version des News Items fertiggestellt wurde -->
    <versionCreated>2013-08-16T11:34:58+02:00</versionCreated>
```

```

<!-- Es folgt: der Publikationsstatus dieses News Items -->
  <pubStatus qcode="stat:usable">
    <name xml:lang="en-GB">Usable</name>
  </pubStatus>
<!-- Es folgt: eine KNA-interne Angabe über die Erstellung des News Items -->
  <generator versioninfo="5.19.5.15975">ines-g2-exporter</generator>
<!-- Es folgt: eine Angabe über den strukturellen Aufbau des News Items -->
  <profile versioninfo="2.0.0">ines/text</profile>
<!-- Es folgt: die Kennungen des Dienstes/der Dienste, über die dieses News Item verbreitet werden soll -->
  <service qcode="knaprod:BD">
    <name role="nrol:display" xml:lang="de">Basisdienst</name>
    <name role="nrol:mnemonic">BD</name>
  </service>
<!-- Es folgt: eine Folge von Kürzeln der KNA-Redaktionsmitglieder, die am Inhalt mitgearbeitet haben -->
  <edNote role="knaednoterole:closingline">sky/cas</edNote>
<!-- Es folgt: eine KNA-interne Klassifikation -->
  <instanceOf qcode="krubr:c082">
    <name xml:lang="de">APP Inland</name>
  </instanceOf>
<!-- Es folgt: ein Link auf das Bildpaket zu diesem Text-News Item -->
  <link href="http://kna-bild.de/paket/130816-89-00037" rank="1" rel="irel:associatedWith" residref="kna-
    picturepackage:urn:newsml:kna.de:20130101:130816-89-00037">
    <itemClass qcode="ninat:composite"/>
    <channel chnlid="knatype:picture-package"/>
  </link>
</itemMeta>
<contentMeta>
<!-- Es folgt: die redaktionelle Dringlichkeit des Textes -->
  <urgency>4</urgency>
<!-- Es folgt: der redaktionelle Bezugsort für des Textes -->
  <located how="howextr:tool" why="why:inferred">
    <name>Bonn</name>
  </located>
<!-- Es folgt: die KNA-interne Kennung des Redaktionsmitglieds, das den Text erstellt hat -->
  <creator>
    <name>skleyb01</name>
  </creator>
<!-- Es folgt/folgen: eine oder mehrere KNA-interne Kennungen von Redaktionsmitgliedern, die am Text mitgewirkt
haben -->
  <contributor>
    <name>skleyb01</name>
  </contributor>
  <contributor>
    <name>carens01</name>
  </contributor>
<!-- Es folgt: eine KNA-interne Kennung zur Verbreitung des Textes. Sie kann als Referenz in Problemfällen genutzt
werden -->
  <altId environment="knaprod:BD"
type="kna7901rendition:iptc7901Id">bkn:014:KNA:0021:KNA:4:pl:277:vvvvb:20130816113505:130816-89-
00031#4</altId>
  <altId type="knatype:originator">KNA</altId>
<!-- Es folgt: das journalistische Genre des Textinhalts -->
  <genre qcode="kgenre:0001">
    <name xml:lang="de">MLD</name>
  </genre>
<!-- Es folgt: eine KNA-Klassifizierung des Inhalts mit Zuordnung zu den IPTC Subject Codes und -Mediatopics -->
  <subject qcode="knasrs:12000000">
    <name role="nrol:full" xml:lang="de">Religion, Weltanschauung</name>
    <name role="nrol:mnemonic" xml:lang="de">REL</name>
    <sameAs qcode="subj:12000000">
      <name xml:lang="de">Religion, Weltanschauung</name>
    </sameAs>
    <sameAs qcode="medtop:12000000">
      <name xml:lang="de">Religion, Weltanschauung</name>
    </sameAs>
  </subject>
<!-- Es folgt: die Zuordnung geografischer Begriffe aus geografischen Vokabularen -->

```

```

<!-- Es folgt: Staaten, auf die im Inhalt Bezug genommen wird. In der Regel mit Bezug zu einer Weltregion -->
  <subject qcode="knageo:35" rank="1" type="cpnat:geoArea">
    <name>Deutschland</name>
    <sameAs qcode="knacountry:1">
      <name role="nrol:display" xml:lang="de">Deutschland</name>
      <name role="nrol:short" xml:lang="de">DEU</name>
    </sameAs>
    <broader qcode="wldreg:r150">
      <name xml:lang="en-GB">Europe</name>
    </broader>
  </subject>
  <subject qcode="knageo:2" rank="2" type="cpnat:geoArea">
    <name>Ägypten</name>
    <sameAs qcode="knacountry:5">
      <name role="nrol:display" xml:lang="de">Ägypten</name>
    </sameAs>
    <broader qcode="wldreg:r002">
      <name xml:lang="en-GB">Africa</name>
    </broader>
  </subject>
<!-- Es folgt: eine Klassifizierung des Inhalts mit Begriffen aus den KNA Subject Vokabular mit Bezügen zu den
semantisch gleichen Begriffen in anderen Vokabularien -->
  <subject qcode="knasubject:698" rank="1" type="knatype:knasubject">
    <name xml:lang="de">Kirche</name>
    <sameAs qcode="dpasubject:664">
      <name xml:lang="de">Kirche</name>
    </sameAs>
    <sameAs qcode="knakeyword:133">
      <name xml:lang="de">Kirche</name>
    </sameAs>
    <sameAs qcode="medtop:20000698">
      <name xml:lang="de">Kirche</name>
    </sameAs>
  </subject>
<!-- Es folgt: weitere Klassifizierung des Inhalts mit Begriffen, die nicht zwingend Bestandteil des KNA Subject
Vokabulars sind -->
  <keyword rank="1">Konflikte</keyword>
  <keyword rank="2">Ägypten</keyword>
<!-- Es folgt: Zuordnung des Inhalts zu einem redaktionellen Ressort -->
  <subject qcode="knacat:pl" type="knatype:category">
    <name role="nrol:mnemonic" xml:lang="de">pl</name>
    <sameAs qcode="dpacat:pl">
      <name xml:lang="de">Politik</name>
      <name role="nrol:mnemonic" xml:lang="de">pl</name>
    </sameAs>
  </subject>
<!-- Es folgen: die Überschriften/Titel, @rank="1" = Haupttitel, @rank="2" = Untertitel -->
  <headline rank="1">Bischofskonferenz fordert Ende der Gewalt in Ägypten </headline>
<!-- Es folgen: die Autorenzeile -->
  <by>Von Sabine Mayer (KNA)</by>
<!-- Es folgt: Ortsmarke und Agenturkennung, traditionell als "Datumszeile" bezeichnet -->
  <dateline>Bonn (KNA) </dateline>
<!-- Es folgt: die Creditline = mit dieser Kennung müssen weiterverbreitete KNA-Texte gezeichnet (zitiert) werden -->
  <creditline>KNA</creditline>
<!-- Es folgt: ein Text, der als Web-Teaser verwendet werden kann (hier leer) -->
  <description role="drol:teaser"/>
<!-- Es folgt/folgen: die im Inhalt verwendete(n) Sprache(n) -->
  <language tag="de"/>
</contentMeta>
<contentSet>
<!-- Es folgt: der Meldungstext im XHTML Format. Im @wordcount die Textlänge in Wörtern -->
  <inlineXML contentType="application/xhtml+xml" wordcount="262">
    <html xmlns="http://www.w3.org/1999/xhtml">
      <head>
        <title>Bischofskonferenz fordert Ende der Gewalt in Ägypten </title>
      </head>

```

```

    <body>
      <p>
        <span class="dateline">Bonn (KNA) </span>Die Deutsche Bischofskonferenz fordert
        ein Ende der blutigen Auseinandersetzungen in Ägypten. (...)
      </p>
    </body>
  </html>
</inlineXML>
</contentSet>
</newsItem>

```

Bei einem Vergleich des Beispiels mit der IPTC Spezifikation von NewsML-G2 kann man feststellen, dass die KNA nicht alle der möglichen Elemente verwendet. Das zeigt die Flexibilität von NewsML-G2 auf und ist nicht als (freiwillige) Beschränkung des KNA-spezifischen Formats zu verstehen.

## Detaillierte Beschreibung der wichtigsten XML Elemente für Empfänger des KNA NewsML-G2 Feeds

Die erläuterten Elemente oder Attribute erscheinen in **Rot**, die besonders relevanten sind **fett** hervorgehoben.

### Das **newsItem** Element

Das **<newsItem>** Element ist das sogenannte Wurzelement (root element) des News Item XML Dokuments.

Dieses Element und alle Unterelemente sind einem speziellen XML Namespace zugeordnet:

**xmlns="http://iptc.org/std/nar/2006-10-01/"**

Im **<newsItem>** Element ist auch die Kennung des verwendeten Standards – NewsML-G2 – und die Version von dessen verwendeter Spezifikation angeführt. (Anmerkung: in allen Arten von Items des NewsML-G2 Standards, die ja unterschiedliche Tag-Namen tragen, ist das Attribut **standard="NewsML-G2"** zu finden, um damit klarzustellen, dass es sich hier um ein Item gemäß dieses IPTC Standards handelt.)

### Die weltweit einmalige Kennung und die Version eines Items

Die weltweit einmalige Kennung eines Items ist im **@guid** Attribut des **<newsItem>** Elements zu finden.

Bei KNA basiert diese Kennung auf dem newsml Namespace von URNs wie er im RFC3085 definiert ist.

Diese Kennung hat folgenden Aufbau:

Vollständiges Beispiel: **urn:newsml:kna.de:20130101:131003-89-00031**

Der ständig gleichbleibende Teil davon lautet: **urn:newsml:kna.de:20130101:**

Der variable Teil davon lautet – in diesem Beispiel: **131003-89-00031**

Die Version des Items mit dieser Kennung ist im **@version** Attribut des **<newsItem>** Elements zu finden.

Dafür gelten folgende Regeln:

- der niedrigste Wert ist 1
- der Wert wird in späteren Versionen immer erhöht
- die an Kunden ausgelieferten Versionswerte müssen sich jedoch nicht fortlaufend erhöhen: eine Abfolge der Versionsnummern 2, 3, 6,8 ist durchaus möglich.

### (Urheber-)Rechtliche Information

Rechtliche Informationen sind in **<rightsInfo>** Elementen zu finden, in der Regel ist bei KNA nur ein einziger vorhanden.

Die nach dem deutschen Presserecht für den Inhalt verantwortliche Person ist im **<accountable>** Element zu finden.



Die KNA wird als Besitzer der urheberrechtlichen Nutzungsrechte am Inhalt im `<copyrightHolder>` Element genannt.

Die `<copyrightNotice>` besagt in der Sache dasselbe, jedoch in einem für die Veröffentlichung geeigneten Format.

Die `<usageTerms>` geben einen kurzen natürlichsprachlichen Hinweis auf die urheberrechtlichen Nutzungsbedingungen, in der Regel wird hier auf den mit KNA abzuschließenden Vertrag verwiesen.

## Metadaten über das Item als Ganzes

Diese Metadaten sind als Unterelemente des Elements `<itemMeta>` zu finden. Hier werden nur jene Elemente angeführt, die für die aktive Auswertung durch einen Empfänger eine Bedeutung haben.

Das Element `<itemClass qcode="ninat:text">` besagt, dass es sich hier um ein News Item mit Textinhalt handelt. Dieser Wert ist anders für Bilder/Fotos, Grafiken, Audio und Video.

Das Element `<provider>` führt die KNA als Ersteller des News Items an. Damit ist nichts darüber gesagt, wer der Ersteller des journalistischen Inhaltes des News Items ist.

Das Element `<versionCreated>` gibt die Zeit an, zu der die Arbeit an dieser Version des Items abgeschlossen wurde. Diese Zeit darf nicht jünger sein, als die Zeit im `<sent>` Element der News Message.

Das Element `<embargoed>` ist optional und zeigt eine bestehende Sperrfrist an. Dazu wird als Wert ein Datum, eine Tageszeit und die zutreffende Zeitzone angegeben, bis zu dem der Inhalt dieses Nachrichten-Items nicht veröffentlicht werden darf.

Das Element `<pubStatus>` ist von wesentlicher Bedeutung für die weitere Verarbeitung eines Items, der Wert ist im `@qcode` Attribut festgehalten:

@qcode Wert	... bedeutet für Sie als Empfänger
stat:usable	Sie können das Item und seinen Inhalt publizieren.
stat:withheld	Sie dürfen das Item vorerst nicht publizieren. Es kann eine Version des Items mit dem pubStatus stat:usable oder stat:canceled folgen.
stat:canceled	Die KNA zieht das News Item zurück, es darf von Ihnen NICHT publiziert werden. Und wenn es schon publiziert wurde, dann ist vor allem auf elektronischen Systemen (Web, Mobil-Apps, ...) die Publikation zu stoppen.

Das oder die Elemente `<service>` halten fest, in welchen KNA-Diensten dieses News Item publiziert wurde. . Das `<name>` Unterelement gibt in der Regel eine gute Kurzbezeichnung des Dienstes, das `@qcode` Attribute eine eindeutige Kennung. Fragen Sie eventuell im KNA-Marketing nach, mit welchen Kennungen die aktuellen KNA-Dienste ausgestattet sind.

### Die `<edNote>` Elemente

Die `<edNote>` Elemente transportieren generell Information der KNA-Redaktion an die Redaktion der Empfänger, also auch die Redaktion in Ihrem Haus.

Mehr Details dazu im Abschnitt [Redaktionelle Hinweise](#)

### `<link>` auf ein Bildpaket

KNA liefert zu zahlreichen Text-Nachrichten Items eine zum Inhalt passende Sammlung an Fotos an, diese Sammlung wird als Bildpaket bezeichnet.

Diese Fotos werden am Bildauslieferungssystem der KNA auf einem Webserver von KNA-Bild dargestellt.

Zur Darstellung der Fotos eines Bildpaketes muss ein Web URL aufgerufen werden, z.B. <http://kna-bild.de/paket/130816-89-00037>

Dieser URL wird in einem `<link>` Element übergeben:

- Der URL ist im @href Attribut zu finden
- Das @rel Attribut hat immer den Wert ired:associatedWith
- Im Unterelement <channel> hat das @channelid Attribut immer den Wert knatype:picture-package

### **Metadaten über den Nachrichten-Inhalt**

Diese Metadaten sind als Unterelemente des Elements <contentMeta> zu finden. Hier werden nur jene Elemente angeführt, die für die aktive Auswertung durch einen Empfänger eine Bedeutung haben.

Das Element **<urgency>** enthält die journalistische Dringlichkeit des Inhalts als Zahlenwert:

- Wert 6: geringe Dringlichkeit
- Wert 1: höchste Dringlichkeit – „Blitz“ Meldung

Das Element **<located>** gibt den Ort an, der den redaktionellen Bezug für die Erstellung des Textes darstellt: Sitz eines Korrespondenten oder Büros der KNA, Ort über den Zugang zum Inhalt möglich war.

Das oder die **<creator>** Elemente und die **<contributor>** Elemente nennen die KNA-Redaktionsmitglieder, die an der Erstellung des Inhalts mitgearbeitet haben. Während eine <edNote> Kürzel anführt wie sie in der Publikation der Meldung verwendet werden können (siehe [Betriebsabläufe](#)), stellen diese Elemente eine präzise Dokumentation vor allem für interne Zwecke dar. KNA empfiehlt, diese Elemente nicht weiter auszuwerten.

Das Element **<genre>** ordnet dem Textinhalt ein journalistisches Genre zu, der Wert ist im @qcode Attribut festgehalten:

@qcode Wert	Journalistisches Genre
kgenre:0001	MLD – Meldung
kgenre:0002	BIO – Biographie, Porträt
kgenre:0003	CHR – Chronologie
kgenre:0004	DOK – Dokumentation, Wortlaut
kgenre:0005	FEA – Feature
kgenre:0006	HIG – Hintergrund
kgenre:0007	INT – Interview
kgenre:0009	KOM – Kommentar
kgenre:0010	KORR – Korrespondentenbericht, Reportage
kgenre:0011	PRO – Tagesvorschau/Programm
kgenre:0012	INFOBOX
kgenre:0014	STICHWORT
kgenre:0015	TER – Terminkalender, Wochenterminliste
kgenre:0019	ZIT – Zitat, Zitat der Woche
kgenre:0030	RED - Achtung
kgenre:0031	ZFA - Zusammenfassung
kgenre:0034	Zur Person

## <subject> Elemente

Diese Elemente repräsentieren aus Vokabularien entnommene Begriffe um die es im Inhalt geht. Die wichtigsten Facetten dieser Begriffe sind:

- Die (weltweit einmalige) Kennung des Begriffes wird im **@qcode** Attribut übergeben.
- Welche Art von Begriff dies ist, ist im **@type** Attribut definiert.

@type Wert	Art des Begriffs bzw. des damit beschriebenen „Dings der Welt“
cpnat:abstract	Abstrakter Begriff, Thema
knatype:knasubject	Abstrakter Begriff aus einem KNA-Vokabular
cpnat:person	Person
cpnat:organisation	Organisation
cpnat:geoArea	Geographisches Gebiet
cpnat:poi	Ort von Interesse (Point Of Interest)
cpnat:object	Materielles Objekt
cpnat:event	Ereignis
knatype:category	Redaktionsressort

- Ein natürlichsprachlicher Name des Begriffs ist im <name> Element zu finden.

## Beziehungen zu anderen Begriffen

Diese Beziehungen werden durch folgende Unterelemente ausgedrückt:

**<sameAs qcode=“...Begriff-URI...”>**: der im @qcode Attribute angegeben Begriff hat dieselbe Bedeutung wie der der vom <subject> Element angegebene. In der Regel ist dies eine Begriffsdefinition über dasselbe Ding, nur aus einen andere Vokabular.

**<broader qcode=“...Begriff-URI...”>**: der im @qcode Attribute angegeben Begriff hat eine breitere Bedeutung wie der der vom <subject> Element angegebene.

## Redaktionelles Ressort

Dieses wird wie oben beschrieben durch ein <subject> Element mit einem **@type=“knatype:category“** Attribut ausgedrückt, der Wert für das Ressort ist im @qcode Attribut festgehalten:

@qcode Wert	Ressort
knacat:pl	Politik
knacat:ku	Kultur
knacat:vm	Vermischtes
knacat:wi	Wirtschaft
knacat:sp	Sport
knacat:rs	Redaktioneller Service

Die Elemente **<keyword>** stellen auch einen Begriff dar, nur ist dieser Begriff nicht aus einem Vokabular entnommen und nur durch einen meist natürlichsprachlichen Ausdruck definiert.

**Achtung:** das von KNA gelieferte NewsML-G2 enthält KEINE in IPTC 7901 übliche Stichwortzeile. Die Gründe dafür sind:

- Die „echten“ Stichworte sind Begriffe in <subject> oder allenfalls <keyword> Elementen.

- Die Beschränkung der Länge der Stichwortzeile hat auch die Anzahl der Stichwörter begrenzt.
- Die Stichwortzeile hat auch anderes als Stichwörter enthalten, wie etwa besondere journalistische Genres oder betriebliche Hinweise, z.B. auf eine Korrektur. Diese Dinge wurden in klar definierte Elemente ausgelagert.

Die Elemente **<headline>** stellen die Überschriften/Titel der Textnachricht dar. Das @rank Attribut gibt die Reihenfolge an:

- rank="1" ... Hauptüberschrift
- rank="2" ... Zweite Überschrift
- rank="3" ... Dritte Überschrift

Das Element **<by>** gibt eine natürlichsprachliche Autorenzeile wieder.

Das oder die Element(e) **<description>** geben eine Beschreibung des Inhaltes der Textnachricht wieder. Grundsätzlich ist es möglich, dass diese Beschreibungen in unterschiedlichen Sprachen vorliegen, die über das @xml:lang Attribut unterschieden werden.

Das Element **<creditline>** gibt an, mit welcher Kennung publizierte oder zitierte Texte gezeichnet werden sollen. Bei KNA-Meldungen ist dies in der Regel „KNA“.

Das Element **<dateline>** gibt eine Ortsmarke und die Agenturkennung wieder, dies wird oft an den Beginn des Meldungstextes gestellt. Dieser Text ist in der Regel eine Mischung aus den Werten des <located> und des <creditline> Elements.

Das oder die Element(e) **<language>** geben die im Text verwendeten Sprachen an, dazu werden die Language Identifier nach IETF BCP 47 im @tag Attribut verwendet.

## Verweise ins Web

Neben dem eigentlich Text werden zu vielen Meldungen auch ergänzende Links in das Web geliefert, wie etwa die Website einer Organisation, die Gegenstand der Meldung war, oder der Link auf ein Video, das in der Meldung angesprochen wurde.

Ein solcher Link kann in zwei Formen geliefert werden:

- eingebettet in den Meldungstext
- in einem speziellen **<link>** Element. Dies wird vor allem dann verwendet, wenn der Link als Ergänzung zum Artikeltext, etwa in der Art Fußnote, erscheinen soll.

Zu dieser Art von <link> Element hier ein Beispiel:

```
<link href="http://www.youtube.com/watch?v=m09Sr0f6iWg&amp;feature=c4-overview&amp;list=UUUXi6fGFfGU5s0t5tXTfC7g"
  rank="1"
  rel="irel:seeAlso">
  <title>Bayern evangelisch</title>
</link>
```

Zu beachten ist das @rel Attribut: für diese ergänzenden Hinweise wird der Wert „irel:seeAlso“ angewendet.

## Der Text der Meldung

Wie im der allgemeinen Beschreibung der NewsML-G2 [News Item Struktur](#) angeführt, wird der Inhalt eines News Items in Unterelementen des **<contentSet>** Elements übermittelt; in mehreren, falls der Inhalt in unterschiedlichen Formaten vorliegt.

Die KNA übermittelt Texte ausschließlich in XHTML.

Dazu wird unterhalb des <contentSet> Elements ein <inlineXML> Element angeordnet und unterhalb dessen ein <html> Element. Achtung: das <html> Element und alle seine Unterelement haben einen anderen XML Namespace als die NewsML-G2 Elemente: `xmlns="http://www.w3.org/1999/xhtml"`

Das <inlineXML> Element besitzt auch ein @wordcount Attribut, das die Anzahl der Wörter im Text wiedergibt und ein @contenttype Attribut mit dem XHTML entsprechenden MIME Type "application/xhtml+xml"

Das XHTML Format ist sehr einfach ausgeführt:

- Die Hauptüberschrift des Textes ist neben dem <headline rank="1"> Element der Metadaten auch im <title> Element des XHTML zu finden.
- Der <body> enthält nur den unmittelbaren Meldungstext und keinerlei redaktionelle Anmerkungen, diese sind ausschließlich in den <edNote> Elementen der Metadaten zu finden.
- Die Absätze des Meldungstextes sind in <p> Elemente gefasst.
- Im ersten Absatz steht am Beginn der Inhalt des <dateline> Metadaten Elements in ein <span class="dateline">Element gefasst.

## KNA Abläufe für Text-Nachrichten

Gegenüber dem Nachrichtenformat IPTC 7901 stehen mehr Metadaten zur Verfügung, sodass in NewsML-G2 Betriebsabläufe viel eindeutiger und granularer dokumentiert werden können.

### Der Normalfall

Als Normalfall gilt die Übermittlung eines Textnachricht-Items, das sofort veröffentlicht werden darf.

Voraussetzungen dafür sind:

- Dass kein <embargoed> Element vorhanden ist.
- Dass das <pubStatus> Element den qcode Wert „stat:usable“ ausweist.

Wenn der Wert von zumindest einem dieser Elemente von dieser Vorgabe abweicht, dann muss eine dieser Abweichung entsprechende Aktion gesetzt werden. Mehr zu diesen Elementen ist im Abschnitt [Metadaten über das Item als Ganzes](#) zu finden.

### Dringlichkeit

Die Dringlichkeit einer Meldung wird über das <urgency> Element ausgedrückt. Das Element enthält einen Zahlenwert zwischen 1 und 6. 1 drückt die höchste Dringlichkeit aus, 6 die geringste.

### Redaktionelle Hinweise

Eine charakteristische Eigenschaft von NewsML-G2 ist eine klare Trennung von journalistischem Inhalt und Hinweisen dazu.

Alle Arten dieser Hinweise sind in <edNote> Elementen zu finden, die sich aber in unterschiedlichem Ausmaß für eine Veröffentlichung eignen.

Welche Art des Hinweises es sich genau handelt, wird über das @role Attribut angegeben. Die folgende Tabelle gibt eine Übersicht aller derzeit verwendeten @role Attribut Werte, deren Bedeutung für den redaktionellen Gebrauch, jeweils ein Beispiel des Inhaltes des Elements und eine Klassifizierung der KNA ob der Inhalt veröffentlicht werden darf.

@role Wert knaednoterole: ....	Redaktionelle Bedeutung	Beispiel für den Inhalt des <edNote> Elements	Veröffent- lichung
closingline	Durch /-Zeichen getrennte	vwa/joh	möglich

	Kürzel von KNA-Redaktionsmitgliedern, die am Inhalt des Items mitgearbeitet haben, aufgeführt.		
contacts	Ansprechpartner einer in der Meldung angesprochenen Einrichtung	Die Dienstleistende GmbH, Franz Müller, fmueller@dienstleistende.biz	möglich
correction	Begründung der Korrektur oder Hinweis auf den Grund des Zurückziehens der Meldung	Der Name Veranstaltungsort, lautet Siebershofen und nicht Seibershofen.	nein
correctionsort	Kurzfassung des obigen	Korrektur des Ortsnamens	nein
editorialnote	Allgemeine Hinweis an die empfangende Redaktion	Es folgt KNA-Korrespondentenbericht bis 15.00 Uhr	nein
edservice	Service-Hinweis: Ergänzender Hinweis zum Meldungsinhalt	Das besprochene Buch: „Caritas hilft allen“ von Fritz Müller, Abc Verlag	möglich
genrenote	Hinweis in Zusammenhang mit dem Genre der Meldung	(aus einer Zusammenfassung) Zusammenfassung 11.00 - Neu: Stellungnahme CDU	nein
knacontacts	Kontaktperson(en) in der KNA-Redaktion zu dieser Meldung	CvD: Thomas Winkel, Tel. 02 28 / 26 00 01 70	nein
localities	Adressen mit Bezug auf den Meldungsinhalt	Veranstaltungsort: Erika-Mayer-Halle, 12345 Schöndorf	möglich
picture	Bild/Grafik-Hinweis		möglich
updatenote	Hinweis über die Aktualisierung der Meldung.	Aussagen von Ursula Sintshofen ergänzt.	nein

Zu Veröffentlichung:

- möglich = es liegt im Ermessen der empfangenden Redaktion, ob dieser Inhalt veröffentlicht wird
- nein = die KNA geht davon aus, dass dieser Inhalt nicht veröffentlicht, sondern als interner Hinweis zwischen sendender und empfangender Redaktion behandelt wird.

**Hinweis:** das <edNote> Element kann auch das Attribut pubconstraint mit folgendem Wert enthalten:

`pubconstraint="knapconstraint:nonpublic"`

Dieser Wert zeigt an, dass der Inhalt der <edNote> ausdrücklich nicht zur Publikation in Ihrem Medium zugelassen ist, sondern nur zur Information der Redaktion.

### **Aktualisierung**

Im Rahmen der Aktualisierung eines Meldungstextes werden nur geringe Anteile an journalistischem Inhalt hinzugefügt oder es werden Fakten auf den aktuellen Stand gebracht.

Die Aktualisierung einer Meldung hat im NewsML-G2 Ablauf und XML folgende Merkmale:

- Die Kennung (@guid) des News Items bleibt erhalten, die Versionsnummer in @version ist höher als in der zuletzt erhaltenen Version.
- Das Element <signal> weist den qcode Wert „sig:update“ aus.
- Es gibt ein Element <edNote role="knaednoterole:updatenote"> das die durchgeführte Aktualisierung erläutert.

- Der Meldungstext liegt in der geänderten Fassung vor – und enthält keinerlei Hinweise darauf, was geändert wurde. Das heißt, er kann so verwendet werden, wie er geliefert wurde.

Hier ein Beispiel der XML Elemente, die unter dem <itemMeta> Element zu finden sind:

```
<signal qcode="sig:update" />
....
<edNote role="knaednoterole:updatenote"> Aussagen von Ursula Sintshofen ergänzt.</edNote>
```

## **Sperrfrist**

Das Setzen einer Sperrfrist für eine Meldung hat im NewsML-G2 Ablauf und XML folgendes Merkmal:

- Das Element <embargoed> ist vorhanden ...
- ... und weist Datum, Uhrzeit und Zeitzone aus, bis zu dem die Sperrfrist gilt.

Hier ein Beispiel des XML Elements, das unter dem <itemMeta> Element zu finden ist:

```
<embargoed>2013-10-03T14:00:00+02:00</embargoed>
```

Dieses Beispiel teilt konkret mit: die Sperrfrist endet am 3. Oktober 2013 um 14:00 Uhr Mitteleuropäischer Sommerzeit.

## **Die Korrektur einer Meldung**

Die Korrektur einer Meldung hat im NewsML-G2 Ablauf und XML folgende Merkmale:

- Die Kennung (@guid) des News Items bleibt erhalten, die Versionsnummer in @version ist höher als in der zuletzt erhaltenen Version.
- Das Element <signal> weist den qcode Wert „sig:correction“ aus.
- Es gibt ein Element <edNote role="knaednoterole:correction"> das die durchgeführte Korrektur erläutert.
- Der Meldungstext liegt in der geänderten Fassung vor – und enthält keinerlei Hinweise darauf, was geändert wurde. Das heißt, er kann so verwendet werden, wie er geliefert wurde.

Hier ein Beispiel der XML Elemente, die unter dem <itemMeta> Element zu finden sind:

```
<signal qcode="sig:correction" />
....
<edNote role="knaednoterole:correction">Der Name des bedrohten Mannes lautet Mayer und nicht Meier.</edNote>
```

## **Das Zurückziehen einer Meldung**

Das Zurückziehen einer Meldung hat im NewsML-G2 Ablauf und XML folgende Merkmale:

- Die Kennung (@guid) des News Items bleibt erhalten, die Versionsnummer in @version ist höher als in der zuletzt erhaltenen Version.
- Das Element <pubStatus> weist den qcode Wert „stat:canceled“ aus.
- Es gibt ein Element <edNote role="knaednoterole:correction"> das die Gründe für das Zurückziehen der Meldung erläutert.

Hier ein Beispiel der XML Elemente, die unter dem <itemMeta> Element zu finden sind:

```
<pubStatus qcode="stat:canceled">
....
<edNote role="knaednoterole:correction">Bitte stellen Sie sicher, dass dieser satirisch gemeinte Text nicht verbreitet wird.</edNote>
<edNote role="knaednoterole:correctionshort">Bitte diesen Text nicht verwenden </edNote>
```



---

## Begriff-Vokabularien

Begriff-Vokabularien im Sinne von NewsML-G2 sind verwaltete Sammlungen von Begriffen, so wie im allgemeinen Abschnitt über [Begriffe](#) erläutert.

Solche Sammlungen können nach unterschiedlichen Kriterien aufgeteilt werden, KNA folgt dabei folgenden Überlegungen:

- Für Begriffe die im Inhalt enthaltene oder angesprochene Themen, Personen, Organisationen, Unternehmen, Orte, Objekte u.ä. darstellen, gibt es jeweils ein eigenes Vokabular: Ein KNA-Themenvokabular, ein Personenvokabular, ...
- Für organisatorische Einheiten im redaktionellen Bereich gibt es eigenständige Vokabulare, z.B. für die Ressorts von Redaktionen.
- Für NewsML-G2-Funktionalitäten gibt es je nach Verwendungszweck eigenständige Vokabulare, z.B. für die Rolle der <edNote> Elemente, für die Rolle eines Namens, ...

KNA verwendet in ihren News Items Vokabularien, die von folgenden Organisationen/Unternehmen verwaltet werden:

- Der KNA selbst
- Dem internationalen Verband von Nachrichtenagenturen IPTC
- Anderen internationalen Organisationen wie etwa der ISO für Staatennamen oder Währungen.
- Und der dpa als de-facto Marktführer in Deutschland

Grundsätzlich sollten zu allen Begriffen, die in diesen Vokabularien enthalten sind, neben deren weltweit einmaliger Kennung zumindest der Name des Begriffs, wenn möglich in Deutsch, zur Verfügung stehen. Insbesondere über alle Vokabularien, die nicht von der KNA verwaltet werden, kann die KNA keine weiteren Zusagen machen.

### KNA-eigene Begriff-Vokabularien

Das sind die von der KNA ihren Kunden zur Verfügung gestellten Vokabularien. Die Tabellen sind lexikalisch nach dem von der KNA verwendeten Scheme Alias sortiert.

Scheme Alias der KNA	kgenre
Bezeichnung	Genre des journalistischen Inhalts
Scheme URI	<a href="http://mds.kna.de/mds/genre/">http://mds.kna.de/mds/genre/</a>

Scheme Alias der KNA	knaarea
Bezeichnung	Konkrete geopolitische Regionen von Ländern.
Scheme URI	<a href="http://mds.kna.de/mds/area/">http://mds.kna.de/mds/area/</a>

Scheme Alias der KNA	knaarearole
Bezeichnung	Typisierung von geopolitischen Regionen von Ländern
Nähere Beschreibung	Für Deutschland z.B. Bundesland und Kreis, für die USA State und County.
Scheme URI	<a href="http://mds.kna.de/mds/arearole/">http://mds.kna.de/mds/arearole/</a>



Scheme Alias der KNA	knacat
Bezeichnung	Redaktionsressort
Scheme URI	<a href="http://mds.kna.de/mds/category/">http://mds.kna.de/mds/category/</a>

Scheme Alias der KNA	knachannel
Bezeichnung	Verbreitungskanal
Scheme URI	<a href="http://mds.kna.de/mds/channel/">http://mds.kna.de/mds/channel/</a>

Scheme Alias der KNA	knacountry
Bezeichnung	Staat
Nähere Beschreibung	Die Liste der Staaten folgt deren Aufnahme in die UN
Scheme URI	<a href="http://mds.kna.de/mds/country/">http://mds.kna.de/mds/country/</a>

Scheme Alias der KNA	knacprh
Bezeichnung	Copyright-Inhaber
Scheme URI	<a href="http://mds.kna.de/mds/copyrightholder/">http://mds.kna.de/mds/copyrightholder/</a>

Scheme Alias der KNA	knaednoterole
Bezeichnung	Rolle eines edNote Inhalts
Scheme URI	<a href="http://mds.kna.de/mds/ednoterole/">http://mds.kna.de/mds/ednoterole/</a>

Scheme Alias der KNA	knageo
Bezeichnung	Geographischer oder geopolitischer Begriff
Anmerkung	Wenn es sich bei dem Begriff um einen Staat handelt, der Mitglied der UN ist, dann wird auch mittels sameAs der Landesnamencode der ISO angezeigt.
Scheme URI	<a href="http://mds.kna.de/mds/geosubject/">http://mds.kna.de/mds/geosubject/</a>

Scheme Alias der KNA	knapconstraint
Bezeichnung	Publikationsbeschränkung
Scheme URI	<a href="http://mds.kna.de/mds/pubconstraint/">http://mds.kna.de/mds/pubconstraint/</a>

Scheme Alias der KNA	knaprod
Bezeichnung	Redaktionelles Produkt der KNA
Scheme URI	<a href="http://mds.kna.de/mds/redprod/">http://mds.kna.de/mds/redprod/</a>

Scheme Alias der KNA	knaprovider
Bezeichnung	Nachrichtenherausgeber

Scheme URI	<a href="http://mds.kna.de/mds/provider/">http://mds.kna.de/mds/provider/</a>
------------	---

Scheme Alias der KNA	knascn
Bezeichnung	Fotoszenerie
Nähere Beschreibung	Beschreibung des Aufbaus oder der Charakteristik eines Fotos.
Scheme URI	<a href="http://mds.kna.de/mds/knascene/">http://mds.kna.de/mds/knascene/</a>

Scheme Alias der KNA	knasubject
Bezeichnung	KNA Thema
Nähere Beschreibung	Thematischer Begriff wie von KNA verwendet.
Scheme URI	<a href="http://mds.kna.de/mds/subject/">http://mds.kna.de/mds/subject/</a>

Scheme Alias der KNA	knatype
Bezeichnung	Natur eines Begriffes
Nähere Beschreibung	Wie von KNA definiert.
Scheme URI	<a href="http://mds.kna.de/mds/type/">http://mds.kna.de/mds/type/</a>

Scheme Alias der KNA	krubr
Bezeichnung	KNA Rubrik
Nähere Beschreibung	Spezielle Charakterisierungen des Inhalts durch die KNA
Scheme URI	<a href="http://mds.kna.de/mds/rubrik/">http://mds.kna.de/mds/rubrik/</a>

## Andere Begriff-Vokabularien

Das sind von der KNA verwendete Vokabularien die allerdings nicht von der KNA verwaltet oder besessen werden. Die Tabellen sind lexikalisch nach dem von der KNA verwendeten Scheme Alias sortiert.

Von KNA übernommener Schema Alias	cpnat
Verwalter/Eigentümer	IPTC
Bezeichnung	Natur, Type eines Begriffs
Nähere Beschreibung	Wie von IPTC für die grundlegenden Concepts Abstrakter Begriff, Person, Organisation, Geographischer Raum, Ort von Interesse, materielles Objekt definiert.
Scheme URI	<a href="http://cv.iptc.org/newscodes/cpnature/">http://cv.iptc.org/newscodes/cpnature/</a>

Von KNA übernommener Schema Alias	drol
-----------------------------------	------

Verwalter/Eigentümer	IPTC
Bezeichnung	Art (Rolle) der Inhaltsbeschreibung
Nähere Beschreibung	
Scheme URI	<a href="http://cv.iptc.org/newscodes/descriptionrole/">http://cv.iptc.org/newscodes/descriptionrole/</a>

Von KNA übernommener Schema Alias	irel
Verwalter/Eigentümer	IPTC
Bezeichnung	Beziehung zum Ziel
Nähere Beschreibung	Beschreibt die Beziehung dieses Items zur Ziel-Ressource, wird in der Regel zusammen mit einem href Attribut verwendet, das die Ziel-Ressource definiert.
Scheme URI	<a href="http://cv.iptc.org/newscodes/itemrelation/">http://cv.iptc.org/newscodes/itemrelation/</a>

Von KNA übernommener Schema Alias	medtop
Verwalter/Eigentümer	IPTC
Bezeichnung	IPTC Media Topics
Nähere Beschreibung	Die neuere Variante einer Themensammlung der IPTC
Scheme URI	<a href="http://cv.iptc.org/newscodes/mediatopic/">http://cv.iptc.org/newscodes/mediatopic/</a>

Von KNA übernommener Schema Alias	ninat
Verwalter/Eigentümer	IPTC
Bezeichnung	Natur eines News Items
Nähere Beschreibung	Beschreibt die grundsätzliche Art/Variante eines News Items
Scheme URI	<a href="http://cv.iptc.org/newscodes/ninature/">http://cv.iptc.org/newscodes/ninature/</a>

Von KNA übernommener Schema Alias	nrol
Verwalter/Eigentümer	IPTC
Bezeichnung	Rolle/Variante eines Namens
Scheme URI	<a href="http://cv.iptc.org/newscodes/namerole/">http://cv.iptc.org/newscodes/namerole/</a>

Von KNA übernommener Schema Alias	sig
Verwalter/Eigentümer	IPTC

Bezeichnung	Signal
Nähere Beschreibung	Hinweise an den Empfänger, dass das Item in einer speziellen Weise verarbeitet werden soll.
Scheme URI	<a href="http://cv.iptc.org/newscodes/signal/">http://cv.iptc.org/newscodes/signal/</a>

Von KNA übernommener Schema Alias	stat
Verwalter/Eigentümer	IPTC
Bezeichnung	Publikationsstatus
Scheme URI	<a href="http://cv.iptc.org/newscodes/pubstatusg2/">http://cv.iptc.org/newscodes/pubstatusg2/</a>

Von KNA übernommener Schema Alias	subj
Verwalter/Eigentümer	IPTC
Bezeichnung	IPTC Subject NewsCode
Nähere Beschreibung	Die etwas ältere Variante einer Themensammlung der IPTC
Scheme URI	<a href="http://cv.iptc.org/newscodes/subjectcode/">http://cv.iptc.org/newscodes/subjectcode/</a>



# Wo finde ich Information über ...

Die Hinweise sind in Datenstrukturen und Betriebsabläufe gruppiert und in diesem alphabetisch sortiert.

## Datenstrukturen

- Autorenzeile einer Text-Nachricht: siehe <by> Element in [KNA / Text-Nachrichten / Daten / Metadaten / Inhalt](#)
- **Bildpaket**, Link auf: siehe [KNA / Text-Nachrichten / Metadaten über das Item als Ganzes](#)
- **Datum** einer Meldung: siehe <versionCreated> Element in siehe [KNA / Text-Nachrichten / Metadaten über das Item als Ganzes](#)
- **Genre** einer Text-Nachricht: siehe <genre> Element in [KNA / Text-Nachrichten / Metadaten über den Nachrichten-Inhalt](#)
- **Kennung** (und Version) einer Meldung: [KNA / Text-Nachrichten / Daten / Kennung](#)
- **Ortsmarke** einer Meldung: siehe <located> Element in [KNA / Text-Nachrichten / Daten / Metadaten / Inhalt](#)
- **Rechtliche Informationen**: siehe [KNA / Text-Nachrichten / Daten / \(Urheber-\)Rechtliche Information](#)
- **Redaktionelle Hinweise**: siehe <edNote> Element in siehe [KNA / Text-Nachrichten / Metadaten über das Item als Ganzes](#) oder siehe [KNA / Text-Nachrichten / Betriebsabläufe / Redaktionelle Hinweise](#)
- **Ressort** für eine Nachricht: siehe „Redaktionelles Ressort“ in in [KNA / Text-Nachrichten / Daten / Metadaten / Inhalt](#)
- **Sperrfrist**: siehe <embargoed> Element in [KNA / Text-Nachrichten / Metadaten über das Item als Ganzes](#)
- **Status** einer Nachricht (verwendbar, nicht verwendbar): siehe <pubStatus> Element in siehe [KNA / Text-Nachrichten / Metadaten über das Item als Ganzes](#)
- **Stichwörter**: siehe <subject> und <keyword> Elemente in [KNA / Text-Nachrichten / Daten / Metadaten / Inhalt](#)
- **Text** einer Meldung: siehe [KNA / Text-Nachrichten / Daten / Text einer Meldung](#)
- **Überschrift** einer Meldung: siehe <headline> Element in [KNA / Text-Nachrichten / Daten / Metadaten / Inhalt](#)
- **Urheberrechtliche Informationen**: [KNA / Text-Nachrichten / Daten / \(Urheber-\)Rechtliche Informationen](#)
- **Versandinformationen**: siehe [KNA / NewsMessage](#)

- **Verantwortliche** Person im Sinne des Presserechts: [KNA / Text-Nachrichten / Daten / \(Urheber-\)Rechtliche Informationen](#)

## Betriebsabläufe

- **Aktualisierung** einer Textmeldung: [KNA / Text-Nachrichten / Betriebsabläufe / Aktualisierung](#)
- **Dringlichkeit** einer Textmeldung: [KNA / Text-Nachrichten / Betriebsabläufe / Dringlichkeit](#)
- **Korrektur** einer Textmeldung: siehe [KNA / Text-Nachrichten / Betriebsabläufe / Korrektur einer Meldung](#)
- **Normalfall** der Meldungsübermittlung: siehe [KNA / Text-Nachrichten / Betriebsabläufe / Normalfall](#)
- **Redaktionelle Hinweise**: siehe [KNA / Text-Nachrichten / Betriebsabläufe / Redaktionelle Hinweise](#)
- **Sperrfrist**: siehe [KNA / Text-Nachrichten / Betriebsabläufe / Sperrfrist](#)**Zurückziehen** einer Meldung: siehe [KNA / Text-Nachrichten / Betriebsabläufe / Zurückziehen einer Meldung](#)

# Referenzen

---

## Technologien und Standards

In der folgenden Liste finden Sie Referenzen/Verweise auf eine umfassende Dokumentation bzw. Spezifikationen von Technologien und Standards.

Die Aufnahme in diese Liste erhebt keinerlei Anspruch, dass diese Standards von der dieses Dokument herausgebenden Nachrichtenagentur erstellt wurden oder dass ein Urheber- oder Markenrecht an diesen Standards und Dokumentationen oder Spezifikationen besteht.

### IPTC NewsML-G2

Allgemeine Informationen: [www.newsml-g2.org](http://www.newsml-g2.org)

Verzeichnis der Spezifikationen: auf der **Specification** Seite von NewsML-G2

Die Dateien liegen in den Unterverzeichnissen von [www.iptc.org/std/NewsML-G2/](http://www.iptc.org/std/NewsML-G2/) vor.

Die derzeit aktuelle Standardversion 2.15

ist zu finden unter [www.iptc.org/std/NewsML-G2/2.15/](http://www.iptc.org/std/NewsML-G2/2.15/)

### IPTC NewsCodes

Die IPTC NewsCodes sind die von der IPTC entwickelten und weiter gepflegten Vokabularien, mehr dazu auf [www.newscodes.org](http://www.newscodes.org)

Die Art der Begriffe teilt sich in zwei große Gruppen auf:

- Begriffe zur Klassifizierung von Nachrichteninhalten: folgen Sie dem Verweis auf **Descriptive NewsCodes**
- Begriffe zur Unterstützung eines IPTC Nachrichtenformats, dies umfasst auch die Vokabularien zur Unterstützung von NewsML-G2: folgen Sie dem Verweis auf **NewsML-G2 NewsCodes**

### IPTC QCodes

Die QCodes sind eine Technologie um lange URIs, die als Concept-Id verwendet werden, abzukürzen.

Mehr dazu auf [www.qcodes.org](http://www.qcodes.org)

### Resource Description Framework – RDF

Ist ein Standard des W3C, mehr dazu auf <http://www.w3.org/RDF/>



## **Semantic Web**

Das Semantic Web ist eine Technologie, die zuerst von dem W3C entwickelt wurde, das aber heute weite Nutzung, Unterstützung und Weiterentwicklung genießt, ein Startpunkt dazu auf der W3C Website ist <http://www.w3.org/standards/semanticweb/>

## **Extensible Markup Language – XML**

Ist ein Standard des W3C zur Notation von strukturierten Daten, mehr dazu auf <http://www.w3.org/XML/>

# Index

Keine Indexeinträge gefunden.