

KML, XML & Auszeichnungssprachen

Christoph Rinne*

04. November 2021

Inhaltsverzeichnis

1	Allgemeine Information	1
2	KML	2
2.1	Punktgeometrien	2
2.2	Liniengeometrien	4

1 Allgemeine Information

Der Titel ist etwas hochtrabend und es wird nicht eine Einführung in [XML](#) werden, sondern es geht darum, Ihnen die Organisation von [Auszeichnungssprachen](#) etwas näher zu bringen. Ich nutze dazu [KML](#)-Dateien, die Sie ggf. schon von [Google Earth](#) kennen.

Auszeichnungssprachen sollen Text gliedern, strukturieren, formatieren oder auch dem Text in einem definierten Kontext Bedeutung geben. Dies erfolgt, indem Textabschnitt von sogenannte *tags* eingeschlossen werden. Auszeichnungssprachen sind vorrangig maschinenlesbar und sollen plattformunabhängig den Datenaustausch zwischen Computern ermöglichen. Daten in einer allgemeingültig definierten XML-Struktur sind also eine gute Grundlage für nachhaltige Interoperabilität. Ein Beispiel das sie sicher schon mal gehört und gesehen haben ist HTML (Hypertext Markup Language), die genuine Sprache des Internets.

```
<!--Beispiel für HTML-->
<p>
  <h2>Die Macht der Computer und die Ohnmacht der Vernunft</h2>
  <i>by <b>J. Weizenbaum</b></i>
  Frankfurt am Main 1977
</p>

<!--Beispiel für inhaltliche Strukturierung (XML)-->
<book>
  <title>Die Macht der Computer und die Ohnmacht der Vernunft</title>
  <author>J. Weizenbaum</author>
  <place>Frankfurt am Main</place>
  <year>1977</year>
</book>
```

HTML zielt auf das Layout von Text, XML hingegen zielt allgemein auf die Definition einer Auszeichnungssprache, auf deren Grundlage eine Strukturierung von Daten in einzelnen **Elemente** erfolgt. Diese Elemente sollen dem beinhalteten Text eine Bedeutung zuweisen, z.B. als Autor kenntlich machen. Die Strukturdefinition einer XML basierten Datei wird in einem frei zugänglichem Dokument abgelegt, der

*Christian-Albrechts Universität zu Kiel, crinne@ufg.uni-kiel.de

Document Type Description (DTD) und auf dieses am Anfang der Datei verwiesen oder aber innerhalb des XML-basierten Dokumentes abgelegt.

Ziel dieser Übung wird es nicht, XML-Dateien von Hand zu schreiben, sondern eben grundlegende Strukturen zu erkennen und ggf. manipulieren oder nachnutzen zu können. Als Beispiel verwende ich KML.

2 KML

Es handelt sich um eine Auszeichnungssprache für Geodaten. Aktuell finden Sie überwiegend KMZ-Dateien, es sind komprimierte (zip) Projektordner mit KML-Dateien. KML ist als Standard vom Open Geospatial Consortium (OGC) akzeptiert und umfassend dokumentiert ([OGC Dokument](#)). Das zuvor referenzierte Dokument thematisiert mit didaktischen Bildern zugleich Fragen der Projektion. Etwas lesefreundlicher ist die Dokumentation der Erweiterungen bei [Google](#). Insgesamt ist KML für die Archäologie durchaus sehr interessant, denn es bietet Raum für ergänzende Sachinformationen. Hier unterscheidet es sich von Well-Known-Text [WKT](#), hier werden ausschließlich Geometrien beschrieben.

2.1 Punktgeometrien

Öffnen Sie die Datei “BBG_Siedlungen” mit Notepad++ und wählen Sie bei den Sprachen XML. Dadurch erhalten sie eine farbliche Differenzierung der Syntax und am linken Rand mit den Linien, [-] und [+] die Möglichkeit, die geschachtelten Elemente auf- und zuzuklappen. Sie erkennen die Zusammenhänge zudem durch die Einrückungen und den *start tag* `<kml>` und den *end tag* `</kml>`. Natürlich gibt es weitere Bedingungen und Ausnahmen, so z.B. *tags* ohne *end tag*.

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2" xmlns:gx="http://www.google.com/kml/ext/2.2" xmlns:kml="http://www.opengis.net/kml/2.2">
<Document>
  <name>BBG Siedlungen, Höhen- und Talrandlage</name>
  <open>1</open>
  <Style id="sn_bbg_sdl">
  <Style id="sh_bbg_sdl">
  <StyleMap id="ms_bbg_sdl">
  <Placemark>
  ...
</Document>
</kml>
```

Das erste *tag* verweist auf die xml-Version und die verwendete Kodierung. Und wenn Sie bis zum Ende der Datei springen finden Sie kein *end tag*.

Das zweite *tag* `<kml>` hat ein *end tag*, führt aber innerhalb des *start tags* eine lange Liste von Verweisen auf die verwendeten DTD (s.o.), die essentielle steht am Anfang, die anderen erweitern die Syntax. Diese Verweise dienen der automatisierten Verarbeitung zur Validierung der Syntax.

Das dritte Element ist ein Dokument innerhalb des kml-Elementes, es entspricht einem Ordner mit Geometrieobjekten bei Google Earth. Entsprechend sehen sie als erstes Element innerhalb dieses Ordners einen *name tag* gefolgt von der Eigenschaft ob aufgeklappt oder geschlossen. Der weitere Abschnitt innerhalb des *document tag* wirkt erst etwas unübersichtlich, deshalb ist die obige Darstellung “zusammengeklappt”.

Für die Darstellung der Geometrien werden als erstes Darstellungsstile definiert, innerhalb des *start tag* stet stets eine ID, diese Bezeichnung ist frei wählbar, minimalistische Lesbarkeit schadet aber nicht. “sn_bbg_sdl”: Symbol für normal dargestellte Bernburger Siedlung, “sh_bbg_sdl”: Symbol für *highlighted* Bernburger Siedlung und “ms_bbg_sdl”: *mapsymbol* für Bernburger Siedlung. Das letztere wird für die entsprechende Geometrie zu Kartierung gewählt und verweist auf die beiden vorgenannten je nachdem, ob das Element nun markiert wird oder eben nicht. Die Syntax innerhalb dieser Stile ist im vorliegen Fall recht simpel.

```
<Style id="sn_bbg_sdl">
```

```

<IconStyle>
  <color>ff0000aa</color>
  <scale>0.6</scale>
  <Icon>
    <href>http://maps.google.com/mapfiles/kml/shapes/homegardenbusiness.png</href>
  </Icon>
  <hotSpot x="0.5" y="0" xunits="fraction" yunits="fraction"/>
</IconStyle>
<LabelStyle>
  <scale>0</scale>
</LabelStyle>
</Style>

```

Der Stil hat einen Abschnitt für das Icon und einen für die Beschriftung. Die Eigenschaften *tags* sind weitgehend selbst erklärend und mal abgesehen von der im Internet üblichen Farbkodierung durchaus verständlich. Das *tag* darf erklärt werden. Durch den Schrägstrich am Ende wird übrigens klar, es gibt kein *end tag*. Dieses *tag* beschreibt die Lage des Symbols zur angegebenen Koordinate. Sollten Sie den Link auf das Icon angeklickt haben, erkennen Sie, dass das Haus in der Karte mittig auf dem Punkt dargestellt wird. Die zum Stil gehörende Beschriftung wird mit der Größe 0, also nicht dargestellt.

Die Definition für den hervorgehobenen Stil wiederzugeben lohnt sich nicht wirklich, das Symbol wird lediglich größer dargestellt `<scale>1</scale>` und die Beschriftung wird mit der selben Anweisung in normaler Größe angezeigt.

```

<StyleMap id="ms_bbg_sdl">
  <Pair>
    <key>normal</key>
    <styleUrl>#sn_bbg_sdl</styleUrl>
  </Pair>
  <Pair>
    <key>highlight</key>
    <styleUrl>#sh_bbg_sdl</styleUrl>
  </Pair>
</StyleMap>

```

Auch dies kann nicht viel weiter erklärt werden. Zentral ist das Element *key* das genau zwei Werte haben kann *normal* oder *highlight*. Daraus ergeben sich zwei Darstellungspaare mit dem jeweiligen Verweis auf einen zuvor definierten Stil.

Innerhalb des *Document tag* folgt nun eine lange Liste von Fundplätzen, die mit dem Element *Placemark* beschrieben werden.

```

<Placemark>
  <name>Burgörner Weinberg</name>
  <LookAt>
    <longitude>11.50428795600097</longitude>
    <latitude>51.64067006920091</latitude>
    <altitude>0</altitude>
    <heading>0.0002321020520841536</heading>
    <tilt>0</tilt>
    <range>171.1583632529837</range>
    <altitudeMode>relativeToGround</altitudeMode>
  </LookAt>
  <styleUrl>#ms_bbg_sdl</styleUrl>
  <ExtendedData>
    <Data name="Literatur">
      <value>Müller 1990</value>
    </Data>
  </ExtendedData>
</Placemark>

```

```

        </Data>
        <Data name="Fundart">
            <value>Siedlung</value>
        </Data>
    </ExtendedData>
    <Point>
        <coordinates>11.50428795600097,51.64067006920091,0</coordinates>
    </Point>
</Placemark>

```

Die zugehörigen Elemente haben alle sprechende Namen. Das Element *LookAt* ist optional (kann also entfallen) und beschreibt detailliert die Ansicht auf die Geometrie. Im vorliegenden Fall ist es eine Reminiszenz der Datenzusammenstellung in Google Earth. Interessant ist das Element *ExtendedData*. Es erlaubt das Anhängen von Sachdaten in einer Liste aus Feldname und Feldwert. Die spätere Darstellung erfolgt als zweiseitige Tabelle mit dem Variablenamen in der ersten und dem zugehörigen Wert in der zweiten Spalte.

Damit ist die Syntax für eine KML-Datei mit Fundpunkten und zugehörigen Sachdaten grundlegend beschrieben. Das ist nicht wirklich kompliziert und kann aus jeder Datenbank generiert werden.

Wo ist der Unterschied zu einer CSV-Tabelle (UTF-8) mit Koordinaten als WKT? In vielen wichtigen Punkten gibt es keinen. Beides ist maschinenlesbar und dennoch auch für Menschen lesbar. Es werden Geometrien mit anhängenden Sachdaten verbunden. Beide Formate liegen als Text vor und sind damit einfach zu archivieren. Soweit kein Unterschied für die zentralen Informationen. KML ist per Definition in UTF-8 kodiert bei einer CSV-Datei ist das nicht der Fall, hier bestimmt das Betriebssystem die Kodierung (Windows: ANSI). CSV speichert die Daten als Tabelle in der uns geläufigen Form: eine Spalte je Attribut die bei fehlender Information einfach leer bleibt. KML ist hier viel effizienter und kompakter, es gibt definierte Attribute und nur die belegten werden am Objekt gelistet. Ein durchaus übliches Verfahren in der Datenorganisation, es setzt aber einheitliche Definition und Schreibweise der Attribute voraus. KML bietet mit der Darstellung und weiteren Definitionen natürlich mehr Optionen. Damit ist es aber abhängig von der Umgebung, z.B. liegt das Icon bei unserer Datei im Web, was sich natürlich ändern ließe. Die KML zielt also ergänzend auf eine visuelle als auch interaktive Darstellung und geht damit über die CSV deutlich hinaus.

2.2 Liniengeometrien

Nutzen Sie nachfolgend bitte die Datei: Hochbruecke.kml. Für die einfache Visualisierung sind in der Struktur der Syntax keine Unterschiede vorhanden, bei den Elementen zur Beschreibung und Interaktion aber schon. Das Ganze ist mit der vorangehenden Erläuterung zu den Punkten aber verständlich und kleine Variationen mit 0 und 1 für die Elemente *<extrude>* und *<tessellate>* erschließen sich bei der Betrachtung in Google Earth unmittelbar. Das Element *<altitudeMode>* benötigt andere Werte: *relativeToGround*, *absolute* oder *clampToGround* und bedingt ebenfalls die Darstellung der Linie.

```

<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2" xmlns:gx="http://www.google.com/kml/ext/2.2" xmlns:kml="http://www.opengis.net/kml/2.2">
<Document>
    <name>Hochbruecke</name>
    <Style id="route-hl">
        <LineStyle>
            <color>ff0000aa</color>
            <width>8</width>
        </LineStyle>
        <PolyStyle>
            <color></color>
        </PolyStyle>
    </Style>
    <Style id="route-n">
        <LineStyle>

```

```

        <color>ff0000aa</color>
        <width>2</width>
    </LineStyle>
    <PolyStyle>
        <color>ff7df5ff</color>
    </PolyStyle>
</Style>
<StyleMap id="route">
    <Pair>
        <key>normal</key>
        <styleUrl>#route-n</styleUrl>
    </Pair>
    <Pair>
        <key>highlight</key>
        <styleUrl>#route-hl</styleUrl>
    </Pair>
</StyleMap>
<Placemark>
    <name>Holtenau</name>
    <styleUrl>#route</styleUrl>
    <LineString>
        <extrude>1</extrude>
        <tessellate>1</tessellate>
        <altitudeMode>absolute</altitudeMode>
        <coordinates>
            10.12093796366777,54.36590694049225,39.99999999999999
            10.12109123715073,54.36643494457174,39.99999999999999
            10.12127580261095,54.36689379308591,39.99999999999999
            10.12146501336677,54.36735153011876,39.99999999999999
            10.12172279167329,54.36776565000539,39.99999999999999
            10.12187831244731,54.36838201838961,39.99999999999999
            10.12217427037541,54.36892432558026,39.99999999999999
            10.12231905470356,54.36943118392424,39.99999999999999
            10.12260799672135,54.36996726586428,39.99999999999999
            10.12286854180568,54.37067130582465,39.99999999999999
            10.12304780654306,54.37097445808439,39.99999999999999
            10.12335429706786,54.37135390598432,39.99999999999999
        </coordinates>
    </LineString>
</Placemark>
</Document>
</kml>

```