

Kerndatenquelle v1.3

Autor: Philip Helger, BRZ

Historie

Version	Datum	Autor	Änderung
V0.1 (intern)	2017-12-05	Philip Helger	Initiale Version
V0.2 (intern)	2017-12-18	Philip Helger	XSD-Adaption, Beispiel hinzugefügt
V0.3 (intern)	2017-12-21	Philip Helger	Änderungen basierend auf eVS-Betreiber Feedback
V1.0	2018-01-23	Philip Helger	Spezialfälle beschrieben
V1.1	2018-06-28	Philip Helger	Unterstützung für den Umzug von Kerndaten
V1.2	2018-09-12	Philip Helger	Unterstützung für Hashwerte von Kerndaten
V1.3	2018-12-03	Philip Helger	Verbesserung des XML Schemas

Inhaltsverzeichnis

Historie	1
1 Einleitung.....	1
2 Struktur.....	1
3 Technische Umsetzung.....	3
3.1 Abfrage	3
Appendix A – XML Schema	4
Appendix B - Beispieldatei.....	8

1 Einleitung

Ein Metadatensatz von data.gv.at enthält u.a. einen Link auf eine „Kerndatenquelle“ (KDQ). Dieses Spezifikationsdokument beschreibt die Felder, die in der Kerndatenquelle enthalten sein können.

Die Kerndatenquellen sind für die Aufbereitung aller Ausschreibungen im USP essentiell, da sie u.a. die eigentlichen Referenzen auf die Kerndaten des jeweiligen eVS-Betreibers enthalten. Sie dient somit als Input für den USP-Ausschreibungs-Indexer. Eine Kerndatenquelle kann als Inhaltsverzeichnis aller Kerndaten betrachtet werden.

2 Struktur

Eine Kerndatenquelle besteht aus folgenden Feldern:

- Einem allgemeinen Header, der hauptsächlich für die rasche Problembehebung im Fehlerfall hilfreich ist, bestehend aus:
 - Der Name der Publizierenden Stelle (z.B. Name des eVS Betreibers) (Pflichtfeld)
 - Die technische Ansprechperson in der publizierenden Stelle (optional)

- Die technische Kontakt-E-Mail-Adresse der publizierenden Stelle, falls abweichend von den Daten in data.gv.at (optional)
- Ein optionales „Extension“-Element soll vorhanden sein, um zukünftige Anforderungen ggf. ohne Anpassung des Schemas vornehmen zu können (optional)
- Einer Liste von 0 bis n Referenzen auf Kerndaten, wobei jeder Eintrag wie folgt aussieht:
 - Eine eindeutige ID der referenzierten Ausschreibung (Pflichtfeld)
 - Jede ID darf pro Kerndatenquelle nur einmal vorkommen
 - Eine ID darf maximal 80 Zeichen haben
 - Eine ID muss mindestens aus 1 Zeichen bestehen
 - Es wird empfohlen, keine Sonderzeichen zu verwenden
 - Die ID darf sich nie ändern
 - Über diese ID werden Kerndaten-Nachrichten wiedererkannt
 - Das ist eine vom Betreiber zu vergebende eindeutige „Dokumenten-ID“
 - Diese ID muss keinerlei Ausschreibungs-Semantik haben
 - Diese ID ermöglicht, dass sich die URL einer referenzierten Kerndaten-Nachricht ändern lässt, ohne den Bezug zur Kerndaten-Nachricht zu verlieren
 - Die systemweit eindeutige ID ergibt sich aus der zusammengesetzten ID des Metadaten-Eintrags von data.gv.at und dieser ID
 - Eine Gruppierung über 1-n Ausschreibungs-Dokumente (weder technisch noch semantisch) ist mit dieser ID nicht möglich
 - Eine URL zum Kerndaten-XML (Pflichtfeld)
 - Es sind nur die Protokolle „http“ und „https“ erlaubt
 - Es dürfen keine Anker (#) in den URLs vorhanden sein
 - Es dürfen keine Credentials (Benutzername/Passwort, Client-Zertifikat etc.) notwendig sein, um auf die Daten zugreifen zu können
 - Die referenzierten XMLs müssen einem vom Projekt veröffentlichten XML Schema für Kerndaten entsprechen
 - Datum und Uhrzeit der letzten Änderung der referenzierten Kerndaten (Pflichtfeld)
 - Wenn die Kerndaten noch nie geändert wurden (beim erstmaligen Erstellen), dann muss dieses Feld das erstmalige Veröffentlichungsdatum enthalten
 - Bei jeder Änderung der referenzierten Kerndaten-Nachricht muss dieses Feld sekundengenau aktualisiert werden
 - Es wird empfohlen Datum und Uhrzeit in UTC (GMT) anzugeben
 - Zwei optionale Datenfelder, die auf Daten aus der Original-Kerndatenquelle verweisen (seit V1.1)
 - Diese Felder sind nur in dem Spezialfall auszufüllen, wenn ein Kerndatensatz von einer Kerndatenquelle in eine andere Kerndatenquelle transferiert wurde.
 - Es müssen beide Felder oder keines von beiden ausgefüllt sein.
 - Es muss sowohl die URL der originalen Kerndatenquelle (Achtung – nicht des Kerndatensatzes) als auch die in der originalen Kerndatenquelle verwendete ID des Kerndatensatzes angegeben werden.
 - Falls ein Kerndatensatz mehr als einmal die Kerndatenquelle wechselt, so müssen immer nur die Daten der allerersten Kerndatenquelle angegeben werden.

- Hinweis: Falls ein Kerndatensatz nach einem Umzug wieder aus der originalen Kerndatenquelle referenziert wird, dann muss die alte ID als primäre ID verwendet werden und die Angabe dieser optionalen Datenfelder ist hinfällig.
- Zwei optionale Datenfelder, mit denen der Hash-Algorithmus und der Hash-Wert einer Kerndatennachricht angegeben werden können (seit V1.2)
 - Diese Felder können ausgefüllt werden, um die Integrität der Kerndaten zu erhöhen. Damit kann beim Einlesen eines Kerndatensatzes sichergestellt werden, dass dieser während der Übertragung nicht verändert wurde.
 - In einem Feld kann der Hash-Algorithmus angegeben werden und im anderen Feld der Hash-Wert des Kerndatensatzes.
 - Es kann maximal ein Hash-Wert pro Kerndatensatz angegeben werden.
 - Es müssen beide Felder oder keines von beiden ausgefüllt sein.
 - Es müssen mindestens die Hash-Algorithmen „SHA-256“ und „SHA-512“ unterstützt werden. Die Hash-Algorithmen „MD5“ und „SHA1“ dürfen nicht verwendet werden¹.
 - Der Name der Hash-Algorithmen ist nicht case-sensitiv. D.h. „sha-256“ und „SHA-256“ sind als gleichwertig zu interpretieren.
 - Der Hash-Wert soll Base64-kodiert angegeben werden.
 - Wenn der bei der Abholung berechnete Hash-Wert von einem Kerndatensatz nicht mit dem in der Kerndatenquelle angegebenen Hash-Wert übereinstimmt, dann wird der Kerndatensatz nicht verarbeitet.
- Ein optionales „Extension“-Element soll vorhanden sein, um zukünftige Anforderungen ggf. ohne Anpassung des Schemas vornehmen zu können (optional)

3 Technische Umsetzung

Die Definition der Struktur der KDQ wurde in XML umgesetzt, da sowohl die OGD-Metadaten auf data.gv.at als XML spezifiziert sind, als auch die Kerndaten an sich in XML spezifiziert werden müssen. Das XML Schema (XSD) ist auch in Appendix A – XML Schema abgebildet.

Es wird empfohlen, für sämtliche XML-Dokumente ausschließlich den UTF-8-Zeichensatz zu verwenden.

3.1 Abfrage

Die technische Umsetzung kann entweder als statische XML-Datei ausgeführt werden oder als dynamische REST-basierte Schnittstelle, die die Daten im gewünschten Format zurückliefert. Die Aufruf-URL ist in den Metadaten von data.gv.at hinterlegt und wird immer per HTTP GET abgefragt.

Alle HTTP Statuscodes ≥ 400 werden als Fehler interpretiert und die Kerndatenquelle wird nicht ausgewertet. Falls es technisch möglich ist, sollen automatisierte E-Mails an die Verantwortlichen gesendet werden (falls vorhanden an die letzte bekannte E-Mail-Adresse aus der Kerndatenquelle, sonst an die in den Metadaten angegebene E-Mail-Adresse).

Hinweis: eine Kerndatenquelle wird geplanter Weise mindestens einmal am Tag abgerufen und die URL ist grundsätzlich für alle sichtbar.

¹ Siehe z.B. https://developer.mozilla.org/en-US/docs/Web/Security/Weak_Signature_Algorithm

111 Auch bei Nicht-Vorhandensein von Kerndaten muss ein valides KDQ-XML zurückgeliefert werden. Die
 112 Verwendung eines HTTP-Statuscodes ist für diesen Fall nicht zulässig.

113 Appendix A – XML Schema

114 Das zu verwendende XML-Schema für Kerndatenquellen schaut wie folgt aus:

```

115 <?xml version="1.0" encoding="UTF-8"?>
116 <schema xmlns="http://www.w3.org/2001/XMLSchema"
117         targetNamespace="http://www.brz.gv.at/eproc/kdq/20180626"
118         xmlns:tns="http://www.brz.gv.at/eproc/kdq/20180626"
119         elementFormDefault="qualified">
120   <annotation>
121     <documentation>
122       Official XML schema for "Kerndatenquellen" (KDQ).
123       An XML instance based on this XSD must be referenced from data.gv.at
124 metadata
125
126       Status: v1.3
127       Version: 2018-12-03
128       History
129         v1.3 Limited the length of the ID to be between 1 and 80 as stated in the
130 specification (no change in targetNamespace)
131         v1.2 Added support for hash values (no change in targetNamespace)
132         v1.1 Added original ID and original KDQ URL
133         v1.0 Initial version
134     </documentation>
135   </annotation>
136
137   <simpleType name="ShortToken">
138     <restriction base="token">
139       <minLength value="1" />
140       <maxLength value="80" />
141     </restriction>
142   </simpleType>
143
144   <complexType name="ExtensionType">
145     <sequence>
146       <any processContents="lax">
147         <annotation>
148           <documentation>
149             The extension content.
150           </documentation>
151         </annotation>
152       </any>
153     </sequence>
154     <attribute name="key" type="string" use="required">
155       <annotation>
156         <documentation>
157           Required key to separate multiple extensions.
158           This extension key must be unique per extension usage block (list of
159 cohesive extension elements).
160         </documentation>
161       </annotation>
162     </attribute>
163   </complexType>
164
165   <complexType name="OriginalItemType">
166     <annotation>

```

```

167     <documentation>
168         This type encapsulates the necessary information required to identify a
169 Kerndaten record that
170         was moved from one KDQ to another KDQ.
171         If a reference was moved from KDQ1 to KDQ2, the URL of KDQ1 and the ID
172 used there must be provided.
173         If a reference was moved from KDQ1 to KDQ2 to KDQ3, also the URL of KDQ1
174 and the ID used there must be provided.
175     </documentation>
176 </annotation>
177 <sequence>
178     <element name="kdq-url" type="anyURI" minOccurs="1" maxOccurs="1">
179         <annotation>
180             <documentation>
181                 The URL of the original KDQ where the Kerndaten record was published
182 first.
183                 Only "http" and "https" URLs are supported.
184             </documentation>
185         </annotation>
186     </element>
187 </sequence>
188 <attribute name="id" type="token" use="required">
189     <annotation>
190         <documentation>
191             The original unique identifier of the Kerndaten record from the original
192 URL.
193             The length of this ID field is limited to 80 characters by the
194 specification!
195         </documentation>
196     </annotation>
197 </attribute>
198 </complexType>
199
200 <complexType name="HashValueType">
201     <annotation>
202         <documentation>
203             This type combines the hash algorithm and the hash value of a single
204 Kerndaten message.
205         </documentation>
206     </annotation>
207     <simpleContent>
208         <extension base="base64Binary">
209             <attribute name="algorithm" type="string" use="required">
210                 <annotation>
211                     <documentation>
212                         The hash algorithm to use.
213                         The name of the hash algorithm is case insensitive.
214                         At least "SHA-256" and "SHA-512" must be supported.
215                         The algorithms "MD5" and "SHA1" must not be used.
216                     </documentation>
217                 </annotation>
218             </attribute>
219         </extension>
220     </simpleContent>
221 </complexType>
222
223 <complexType name="ItemType">
224     <annotation>
225         <documentation>
226             This type contains the reference to a single "Kerndaten"-XML
227         </documentation>

```

```

228     </annotation>
229     <sequence>
230         <element name="url" type="anyURI" minOccurs="1" maxOccurs="1">
231             <annotation>
232                 <documentation>
233                     The URL where the "Kerndaten"-XML is publicly available.
234                     Only "http" and "https" URLs are supported.
235                 </documentation>
236             </annotation>
237         </element>
238         <element name="hash-value" type="tns:HashValueType" minOccurs="0"
239 maxOccurs="1">
240             <annotation>
241                 <documentation>
242                     The optional hash value of this Kerndaten record.
243                     Added in V1.1.
244                 </documentation>
245             </annotation>
246         </element>
247         <element name="original-reference" type="tns:OriginalItemType" minOccurs="0"
248 maxOccurs="1">
249             <annotation>
250                 <documentation>
251                     The optional information about the original reference of this record,
252 in case
253                     the referenced Kerndaten record was moved from one KDQ to another KDQ.
254                     Added in V1.2.
255                 </documentation>
256             </annotation>
257         </element>
258         <element name="extension" type="tns:ExtensionType" minOccurs="0"
259 maxOccurs="unbounded">
260             <annotation>
261                 <documentation>
262                     Optional custom extensions.
263                 </documentation>
264             </annotation>
265         </element>
266     </sequence>
267     <attribute name="id" type="tns:ShortToken" use="required">
268         <annotation>
269             <documentation>
270                 The unique identifier of the tender, referenced by the URL below.
271                 This ID might not change.
272                 The length of this ID field is limited to 80 characters by the
273 specification!
274             </documentation>
275         </annotation>
276     </attribute>
277     <attribute name="Lastmod" type="dateTime" use="required">
278         <annotation>
279             <documentation>
280                 The last modification date and time (incl. time zone offset) of the
281 referenced "Kerndaten"-XML file.
282                 Every time the referenced XML changes, this entry must be updated.
283                 It is recommended to use UTC timestamps here.
284             </documentation>
285         </annotation>
286     </attribute>
287 </complexType>
288

```

```

289 <complexType name="HeaderType">
290   <annotation>
291     <documentation>
292       The type of the static header information.
293     </documentation>
294   </annotation>
295   <sequence>
296     <element name="publisher" type="string" minOccurs="1" maxOccurs="1">
297       <annotation>
298         <documentation>
299           Entity responsible for publishing this file.
300         </documentation>
301       </annotation>
302     </element>
303     <element name="contact-person" type="string" minOccurs="0" maxOccurs="1">
304       <annotation>
305         <documentation>
306           Optional technical contact person name for this file.
307         </documentation>
308       </annotation>
309     </element>
310     <element name="contact-email" type="string" minOccurs="0" maxOccurs="1">
311       <annotation>
312         <documentation>
313           Optional technical contact email address for this file.
314           If this element is not present, the email address from data.gv.at is
315 used.
316         </documentation>
317       </annotation>
318     </element>
319     <element name="extension" type="tns:ExtensionType" minOccurs="0"
320 maxOccurs="unbounded">
321       <annotation>
322         <documentation>
323           Optional custom extensions.
324         </documentation>
325       </annotation>
326     </element>
327   </sequence>
328 </complexType>
329
330 <complexType name="KDQType">
331   <annotation>
332     <documentation>
333       The root type.
334       It contains one header and 0-n items.
335     </documentation>
336   </annotation>
337   <sequence>
338     <element name="header" type="tns:HeaderType" minOccurs="1" maxOccurs="1" />
339     <element name="item" type="tns:ItemType" minOccurs="0" maxOccurs="unbounded"
340 />
341   </sequence>
342 </complexType>
343
344 <element name="kdq" type="tns:KDQType">
345   <annotation>
346     <documentation>
347       The root element of this XSD.
348     </documentation>
349   </annotation>

```

```
350     </element>
351 </schema>
```

352 Appendix B - Beispieldatei

353 Beispiel-Kerndatenquelle im XML Format mit drei referenzierten Kerndaten:

```
354 <?xml version="1.0" encoding="UTF-8"?>
355 <kdq xmlns="http://www.brz.gv.at/eproc/kdq/20180626">
356   <header>
357     <publisher>BRZ GmbH</publisher>
358     <contact-person>Philip Helger</contact-person>
359     <contact-email>philip.helger@brz.gv.at</contact-email>
360   </header>
361   <item id="12345677" lastmod="2017-12-18T14:13:00Z">
362     <url>https://test.evs.at/kerndaten/613c0b76-e3f5-11e7-80c1-9a214cf093ae</url>
363   </item>
364   <item id="12345678" lastmod="2017-12-18T14:11:00Z">
365     <url>https://test.evs.at/kerndaten/613c0e14-e3f5-11e7-80c1-9a214cf093ae</url>
366   </item>
367   <item id="12345679" lastmod="2017-12-18T14:12:00Z">
368     <url>https://test.evs.at/kerndaten/613c124c-e3f5-11e7-80c1-9a214cf093ae</url>
369   </item>
370 </kdq>
371
```

372 Beispiel-Kerndatenquelle im XML Format ohne referenzierte Kerndaten:

```
373 <?xml version="1.0" encoding="UTF-8"?>
374 <kdq xmlns="http://www.brz.gv.at/eproc/kdq/20180626">
375   <header>
376     <publisher>BRZ GmbH</publisher>
377     <contact-person>Philip Helger</contact-person>
378     <contact-email>philip.helger@brz.gv.at</contact-email>
379   </header>
380 </kdq>
381
```

382 Beispiel-Kerndatenquelle im XML Format mit zwei referenzierten Kerndaten, wovon ein
383 referenzierter Kerndatensatz aus einer anderen Kerndatenquelle stammt:

```
384 <?xml version="1.0" encoding="UTF-8"?>
385 <kdq xmlns="http://www.brz.gv.at/eproc/kdq/20180626">
386   <header>
387     <publisher>BRZ GmbH</publisher>
388     <contact-person>Philip Helger</contact-person>
389     <contact-email>philip.helger@brz.gv.at</contact-email>
390   </header>
391   <item id="12345677" lastmod="2017-12-18T14:13:00Z">
392     <url>https://test.evs.at/kerndaten/613c0b76-e3f5-11e7-80c1-9a214cf093ae</url>
393     <original-reference id="1234">
394       <kdq-url>http://www.example.org/kdq</kdq-url>
395     </original-reference>
396   </item>
397   <item id="12345678" lastmod="2017-12-18T14:11:00Z">
398     <url>https://test.evs.at/kerndaten/613c0e14-e3f5-11e7-80c1-9a214cf093ae</url>
399   </item>
400 </kdq>
```


401

402 Beispiel-Kerndatenquelle im XML Format mit zwei referenzierten Kerndaten, wovon ein
403 referenzierter Kerndatensatz einen Hash-Wert hat:

```
404 <kdq xmlns="http://www.brz.gv.at/eproc/kdq/20180626">
405   <header>
406     <publisher>BRZ GmbH</publisher>
407     <contact-person>Philip Helger</contact-person>
408     <contact-email>philip.helger@brz.gv.at</contact-email>
409   </header>
410   <item id="12345677" lastmod="2017-12-18T14:13:00Z">
411     <url>https://test.evs.at/kerndaten/613c0b76-e3f5-11e7-80c1-9a214cf093ae</url>
412     <hash-value algorithm="SHA-
413 256">c0535e4be2b79ffd93291305436bf889314e4a3faec05ecffcb7df31ad9e51a</hash-value>
414   </item>
415   <item id="12345678" lastmod="2017-12-18T14:11:00Z">
416     <url>https://test.evs.at/kerndaten/613c0e14-e3f5-11e7-80c1-9a214cf093ae</url>
417   </item>
418 </kdq>
```

419