

Anleitung zur Verwendung des Webservice des Katalogs der historischen Fotografien (Interreg-Projekt "Lichtbild. Kulturschatz Historische Fotografie")

Inhaltsverzeichnis

Vorwort	1
1. Aufruf des Webservice und dessen Parametrisierung.....	2
2. Suchanfragen	3
2.1. Abfragen der Bilder.....	4
3. Datenschema	5
3.1. XML- und JSON-Response-Format des Webservice.....	7

Vorwort

Im Rahmen des Interreg-Projektes „Kulturschatz Lichtbild“ wurde ein Webservice mit einem Suchmotor erstellt, der eine Suche über den gesamten erfassten Bestand der historischen Bilder erlaubt.

Der Webservice erlaubt das Auffinden und das Herunterladen der von den folgenden Projektpartnern zur Verfügung gestellten Bildern: der Verein Tiroler Archiv für photographische Dokumentation und Kunst (TAP) in Lienz, die Stadtgemeinde Bruneck, das Südtiroler Landesamt für Film und Medien und die Südtiroler Landesabteilung Museen.

Dieser Webservice kann im Sinne der offenen Daten von jedem Unternehmen, jeder Einzelperson, aber auch von einzelnen Museen verwendet werden, um die Daten herunterzuladen oder einfach im Internet darzustellen.

Ein Export, also die Erstellung einer Kopie der Daten aus der zentralen Datenbank, ist nicht notwendig und zielführend, da der Webservice den Zugriff auf alle aktuellen Daten der zentralen Datenbank erlaubt, die kontinuierlich gepflegt und erweitert wird.

Der Webservice ist als RESTful konzipiert, d.h. die Abfragen können über einfache http-Befehle durchgeführt werden und einfach im Browser ausprobiert werden.

Dieses Dokument beschreibt, wie der Webservice aufgerufen werden kann.

1. Aufruf des Webservice und dessen Parametrisierung

Der RESTful-Webservice kann über eine HTTP-Schnittstelle (so auch über einen einfachen WebBrowser) aufgerufen werden.

Die Adresse des Webservice (wird ab hier als *ENDPOINT* bezeichnet) lautet <http://daten.buerger-netz.bz.it/services/kksSearch/collect/lichtbild>

Der Webservice ist parametrisierbar, sodass er die vom Anwender (oder von der Anwendungsapplikation) geforderten Daten präzise ausgeben kann.

Da es sich im Interreg-Projekt „Lichtbild“ um offene Daten handelt, ist keine Authentifizierung des Anwenders (Benutzername oder Passwort) erforderlich.

Die Daten werden verschlüsselt übertragen (SSL).

Die Parameter, die dem RESTful-Webservice zur Wiedergabe der erforderlichen Daten hinzugefügt werden können, sind in der folgenden Tabelle wiedergegeben.

Die Parameter zur Definition der Abfragen sind:

Parameter		Beispiel
q	Suche	q=Rafenstein (sucht im default-Suchfeld nach dem Begriff „Rafenstein“) q=*: * liefert alle Ergebnisse zurück q=CL_de:Waldmüller (sucht im Feld CL_de nach „Waldmüller“)
rows	Maximal-Anzahl der zurückzugebenden Records, default = 10	rows=100
start	Gibt an, ab welcher Position die Rückgabe der Daten starten soll, default=0	start=0
fl	Beistrich-getrennte Liste der Felder, die in der Antwort zurückgegeben werden sollen. Falls der Parameter fehlt, werden vorkonfigurierte Felder verwendet. fl=* liefert alle Felder zurück	fl=TI_de,BE_de,OB_de,IN

wt	Response-Writer: damit kann das Rückgabeformat des Service angegeben werden. Mögliche Werte sind: <ul style="list-style-type: none"> • xml (default) • json • csv 	wt=json
fq	Filter-Query: für häufig verwendete queries, z.B. bei Facettensuche (cached die Ergebnisse)	fq=MUS:MAL
indent	Parameter für die Formatierung des Outputs	indent=on
q.op	Art der Verknüpfung bei mehreren Suchkriterien: <ul style="list-style-type: none"> • AND • OR (default) 	q.op=OR
size	Parameter für die Größenangabe der downzuladenden Bilder	size=s m o]

2. Suchanfragen

Die Suche kann in einzelnen, spezifizierbaren Feldern erfolgen, die in der Suchanfrage angegeben werden können. Wird kein Feld angegeben, wird standardmäßig die deutsche Beschreibung (Feld BE_de) verwendet.

Die Suchanfrage wird über Parameter in der URL gesteuert.

Die Suchkriterien können mit OR oder AND verknüpft werden. Falls nicht anders angegeben, wird OR verwendet. Dies kann für die gesamte Suche mit dem Parameter q.op geändert werden (q.op=OR).

Gesucht werden kann in allen Feldern, die in der Tabelle in Kapitel [3. Datenschema](#) angeführt wird.

Die Suchanfrage `ENDPOINT?q=TI_de:Rafenstein` liefert z.B. die ersten Treffer, die im deutschen Titel das Wort „Rafenstein“ enthalten (welche Felder ausgegeben werden, ist ebenfalls im Kapitel [3. Datenschema](#) dokumentiert).

Die Suchanfrage kann beliebig viele Parameter enthalten, die Reihenfolge der Parameter spielt dabei keine Rolle.

Die einzelnen Parameter werden durch das kaufmännische Und (&) getrennt.

Hier einige Beispiele:

Die ersten 20 Resultate zurückliefern:

`ENDPOINT?q=TI_de:Rafenstein&start=0&rows=20`

Suche im Materialfeld nach ‚Fotopapier‘ und Hersteller ‚Fränzl‘:

`ENDPOINT/select?q=MA_de:Fotopapier AND VV_de:Fränzl`

oder

`ENDPOINT/select?q=MA_de:Fotopapier VV_de:Fränzl&q.op=AND`

Suche der neuesten 20 Datensätze:

ENDPOINT/select?q=*&sort=creation desc&rows=20

Wenn mehrere Wörter in einem Feld gesucht werden sollen, müssen diese in Klammern () eingeschlossen oder das Suchfeld für jedes Wort wiederholt werden:

q=CP_de:(Bozen Freiheitsstraße)

bzw.

q=CP_de:Bozen CP_de: Freiheitsstraße

Diese Suche

q=CP_de:Bozen Freiheitsstraße

wird hingegen folgendermaßen interpretiert

q= CP_de:Bozen BE_de:Freiheitsstraße

Wichtig: es werden nur GET-Anfragen unterstützt, andere http-Verben sind nicht erlaubt.

2.1. Abfragen der Bilder

Die Bilder werden im Feld B1p_url angegeben. Dieses Feld beinhaltet die URL, unter der sie abgerufen werden können.

Das Format ist folgendes:

https://cert.provinz.bz.it/services/kksSearch/image?file=<filename>&mus=<MUSEUMSKÜRZEL[&size=s|m|l|o]>

wobei die Werte in den Spitzklammern (Dateiname und die Institution) durch den entsprechenden Wert aus der Antwort des Webservice ersetzt wird, während der Parameter „size“ optional vom Anwender oder von der Anwendungsapplikation nach Abfrage des Webservice hinzugefügt werden kann.

Mögliche Werte für „size“ sind: s für small, m für medium, l für large und o für Originalbild (siehe folgende Tabelle):

Parameter		Bildbreite (px)
s	small (default)	180
m	medium	360
l	large	720
o	Original (nur, wenn auch das Original freigegeben wurde)	Originalgröße (oder „l“ wenn das Original nicht freigegeben wurde)

Falls der Parameter „size“ nicht explizit angegeben wird, wird als Standard-Voreinstellung „s“ genommen (180px).

Falls ein Bild nicht existiert oder dieses nicht öffentlich ist, wird Http-Status 404 – Not Found zurückgegeben.

Folgende Formate werden momentan unterstützt: jpg, bmp, gif.

3. Datenschema

Das Datenschema, also die Struktur der vom RESTful-Webservice ausgegebenen Daten, ist in der folgenden Tabelle wiedergegeben.

In der ersten Spalte der Tabelle ist der Parameter angegeben. In der zweiten Spalte sind die Feldkürzel (tags) wiedergegeben; diese tags sind die eigentlichen Feldnamen, die von der Schnittstelle ausgegeben werden.

Bei mehrsprachigen Feldern wird an das Feldkürzel _de für Deutsch oder _it für Italienisch angehängt (z.B: BE_de für die deutsche Beschreibung).

Bei einigen Thesaurustermini werden auch die Synonyme und Oberbegriffe eingeschlossen, diese Felder enden mit „_syn“. So enthält das Feld MA_de_syn die deutsche Materialbezeichnung samt Synonymen und Oberbegriffen, während MA_de genau das ausgewählte Material des Objekts enthält.

Abhängig von den verwendeten Parametern kann das Resultat folgende Felder aufweisen:

Feldbeschreibung	Tag	Pflichtfeld	wiederholbar	Bemerkungen
Priref	Priref	X		Alphanumerischer Primärschlüssel
Inventarnummer	IN	X		
Objektbezeichnung DE	OB_de		X	
Objektbezeichnung IT	OB_it		X	
Synonym Objektbezeichnung DE	OB_de_syn		X	enthält auch Synonyme und Oberbegriffe
Synonym Objektbezeichnung IT	OB_it_syn		X	
Beschreibung DE	BE_de		X	
Beschreibung DE	BE_it		X	
Hist.-krit. Angaben DE	B3_de		X	Historisch kritische Angaben in deutscher Sprache
Hist.-krit. Angaben IT	B3_it		X	Historisch kritische Angaben in italienischer Sprache
Hersteller DE	VV_de		X	
Hersteller IT	VV_it		X	
Sachgruppe DE	OC_de		X	
Sachgruppe IT	OC_it		X	

Synonym Sachgruppe DE	OC_de_syn		X	Synonyme und Oberbegriffe eingeschlossen
Synonym Sachgruppe IT	OC_it_syn		X	
Datierung von	DS			Datum im Format yyyy.mmdd. Beispiel 1995.1231
Datierung bis	DE			
Dargestellter Ort DE	CP_de		X	
Dargestellter Ort IT	CP_it		X	
Synonym Dargestellter Ort DE	CP_de_syn		X	Synonyme und Oberbegriffe eingeschlossen
Synonym Dargestellter Ort IT	CP_it_syn		X	
Material DE	MA_de		X	
Material IT	MA_it		X	
Synonym Material DE	MA_de_syn		X	Synonyme und Oberbegriffe eingeschlossen
Synonym Material IT	MA_it_syn		X	
Technik DE	TK_de		X	
Technik IT	TK_it		X	
Synonym Technik DE	TK_de_syn		X	Synonyme und Oberbegriffe eingeschlossen
Synonym Technik IT	TK_it_syn		X	
Schlagwort DE	ip_de		X	
Schlagwort IT	ip_it		X	
Maße DE	dim_de		X	
Maße IT	dim_it		X	
Bild	B1p		X	Name der Bilddatei
Bild-URL	B1p_url		X	URL der Bilddatei
Institutionskürzel	MUS			Dieses Feld beinhaltet ein Kürzel aus 3 Buchstaben, das die Institution (z.B. Museum) angibt, die das Objekt besitzt.
Erstellungsdatum	creation			

Änderungsdatum (letztes)	modification			ISO-Datums-Format in UTC-Zeit. Millisekunden sind optional.
Timestamp	timestamp			Zeitstempel wann das Objekt im Index aktualisiert wurde. Format wie für die Felder creation/modification
Geografische Koordinaten	CP_geo		X	X-Y-Koordinaten des Ortes in dem das Bild aufgenommen wurde im Bezugssystem WGS 84 (EPSG:4326) z.B.: 46.498267, 11.354752 für Waltherplatz Bozen

3.1. XML- und JSON-Response-Format des Webservice

Der Webservice kann die Daten in den drei Formaten XML (default), JSON und CSV ausgeben.

Die Formate XML und JSON besitzen gegenüber dem rein tabellarischen CSV-Format zusätzliche Elemente:

"responseHeader", das einfache Metadaten (Informationen) zum Response selber zusammenfasst:

- status: immer 0 außer wenn ein Fehler aufgetreten ist
- QTime: die Anzahl an Millisekunden, die der Server für die Anfrage benötigt hat
- params: Parameter der Suche

Die Ergebnisse, also die Daten der gefundenen Bilder, werden unter dem Element <result> (XML) bzw. "response" (JSON) zusammengefasst. Einige der Attribute sind:

- numFound: Gesamtanzahl an Treffern.
- start: identisch mit dem übergebenen „start“-Parameter

Der Inhalt des Resultats ist eine Liste von <doc>-Elementen. Jedes einzelne Element entspricht einer entsprechenden Fotografie. Die Kind-Elemente des <doc>-Elementes sind die Attribute des Dokuments.

Folgende einfache Elemente sind möglich:

- str: ein Text-Feld
- int: ein Integer im Bereich -2^{31} bis $2^{31} - 1$
- float: eine floating-point Zahl im Bereich $1.4e-45$ bis $3.4e38$
- long: ein Integer im Bereich -2^{63} bis $2^{63} - 1$
- double: ein Integer im Bereich $4.9e-324$ bis $1.8e308$
- date: ein Datum im ISO-8601-Format wie z.B. 1965-11-30T05:00:00Z (Z steht für die GMT-Zeitzone)
- bool: ein boolscher Wert: true oder false

Folgende zusammengesetzten Elemente sind möglich:

- `lst`: eine benannte Liste. Jedes Kind-Element hat auch ein `name`-Attribut
- `arr`: ein Array von Werten. Jedes Kind-Element ist Bestandteil dieses Arrays