

TopicZoomWeb-Tags



Die TopicZoom WebTags sind automatisierte Metadaten, die das Thema eines Textes spezifizieren. Sie können sowohl zur Verschlagwortung von Texten (z.B. in Form von Tag-Clouds), als auch zur Strukturierung und Navigationserleichterung auf Archiven verwendet werden. Die TopicZoom Web-Tags sind sowohl als HTML als auch als XML Service zugreifbar.

HTML: Die TopicZoom WebTags können mit einem Browser über folgenden Link erreicht werden:

<http://twittopic.topiczoom.de/start.de.html>

Hier kann ein beliebiger Text in UTF8 eingegeben werden. Mittels klick auf den "Index"-Button wird dieser Text mittels der TopicZoom Map automatisiert indexiert werden. Die Themen werden auf der folgenden Seite nach dem jeweiligen thematischen Gewicht geordnet, angezeigt.

XML: Automatisiert sind die TopicZoom WebTags mit einer XML Schnittstelle abrufbar. Diese kann über den folgenden link mittels eines HTTP-POST-Requests angesprochen werden:

<http://twittopic.topiczoom.de/quickindex.xml>

Im Content-Block des HTTP-POST-Requests muss der zu annotierende Text in UTF-8 enthalten sein.

Beispiel eines HTTP-POST-Requests:

```
-----  
POST /quickindex.xml HTTP/1.1  
Host: twittopic.topiczoom.de  
Content-length: 28
```

```
Chardonnay ist ein Weißwein.  
-----
```

Die Antwort erfolgt in XML in UTF-8. Das Root-Tag des XML Tokens ist ein **TZTopicSet**. Die beinhaltet für jedes erkannte Thema/Schlagwort ein **TZTopic**-Tag. Jedes TZTopic-Tag kann folgende Attribute haben:

1. txt-Attribut: in diesem Attribut steht der Name bzw. die Bezeichnung des jeweiligen Tags.
2. weight-Attribut: In diesem Attribut ist das absolute Gewicht des Themas angegeben. Je höher diese Zahl umso sicherer kann dem Text dieses Thema zugeordnet werden.
3. DoG-Attribut: DoG steht für „Degree of Generality“. Umso niedriger die der DoG-Wert ist, umso allgemeiner ist das Thema, umso höher umso spezifischer.
4. Sig-Attribut: Sig steht für Significance. Hierbei wird bewertet, wie ungewöhnlich ein Thema im Vergleich zu einem großen Nachrichten- und Artikelarchiv ist.
5. direct-Attribut: Gibt an ob und wie oft das Schlagwort direkt im Text erwähnt wird. 0 bedeutet, dass nur spezifischere Themen erwähnt werden, aus denen eine Indikation des gegebenen Themas abgeleitet wurde.
6. RDFID-Attribut: Ein eindeutiger Identifikator des Themas.
7. Diversity-Attribut: Ein Wert, der angibt in wie weit spezifischere verwandte Themen im Text gefunden wurden. Ein hoher Wert deutet eine hohe Relevanz des Themas im Text und gleichzeitig eine gute Zusammenfassende Wirkung des Schlagworts an. 0,5 bedeutet, dass das Schlagwort nur über mehrere Abstraktionsebenen den Text beschreibt. Werte größer als 1 stellen sehr hohe Relevanzen da.
8. TSCCS-Attribut: Dieses Attribut stellt ein Mapping zum „Thema Subject Classification Code System“ von EDITEUR.org bereit
9. TZTYPE-Attribut: Eine Einsortierung von Einträgen in verschiedene Entitätsklaaen wie z.B. „geo“, „Person“ oder „time“;



Beispiel des Content-Blocks einer (möglichen) zugehörigen und gültigen Antwort:

```
<soap:Envelope xmlns:soap="http://www.w3.org/2003/05/soap-envelope">
<soap:Header>
</soap:Header>
<soap:Body>
<TZTopicSet xmlns:tzns="http://twittopic.topiczoom.de/TZNS" date="Sat Sep 5 17:42:21 2015">
<TopicZoomWebTagsInfo info="Die aktuellen Nutzungsbedingungen der TopicZoomWeb-Tags finden sie
unter http://www.topiczoom.de/wp-content/uploads/2012/02/TopicZoomWebTag\_V2.0.pdf/>

<TZTopic txt="Speisen und Getränke" weight="2" DoG="5" Sig="6.18236702368784" direct="0"
RDFID="20287964" Diversity="1.5" TSCCS="0" TZTYPE="0" />
<TZTopic txt="Chardonnay" weight="1" DoG="11" Sig="6.3529646649187" direct="1"
RDFID="20273837" Diversity="1" TSCCS="0" TZTYPE="0" />
<TZTopic txt="Getränke" weight="2" DoG="6" Sig="6.40148052741537" direct="0" RDFID="20765188"
Diversity="1.5" TSCCS="WBX" TZTYPE="0" />
<TZTopic txt="Alkoholische Getränke" weight="2" DoG="8" Sig="7.17650024232593" direct="0"
RDFID="20900916" Diversity="1.5" TSCCS="WBXD" TZTYPE="0" />
<TZTopic txt="Wein" weight="2" DoG="9" Sig="7.6625042129181" direct="0" RDFID="20900899"
Diversity="1.5" TSCCS="WBZ" TZTYPE="0" />
<TZTopic txt="Lebensmittel" weight="2" DoG="4" Sig="5.39270894274705" direct="0"
RDFID="21884553" Diversity="1.5" TSCCS="0" TZTYPE="0" />
</TZTopicSet>
</soap:Body>
</soap:Envelope>
```

Die Darstellung ist hierbei analog zur HTML-Version, jedoch um die etliche Attribute erweitert, die ein Mapping in unsere Ontologiestruktur ermöglicht. Unsere Ontologiestruktur liegt in SKOS-RDF-XML vor und ermöglicht eine Navigation auf den Themengebieten (z.B. eine Verallgemeinerung, eine Präzisierung etc.) sowie ein Mapping in weitere Klassifikationssystem (IPTC etc.). Bei weiteren Fragen können Sie uns gerne kontaktieren.

Datenschutzerklärung und Nutzungsbedingungen der TopicZoom GmbH (nachfolgend TopicZoom genannt) im Internet zur Verfügung gestellten TopicZoom Web-Tags

Einführung

TopicZoom hat auf Basis einer eigener Technologie ein Produkt entwickelt, um Anwender bei der Verschlagwortung von unstrukturierten, natürlichsprachigen Texten zu unterstützen. Der Benutzer der TopicZoom Web-Tags verwendet den Service mittels der oben beschriebenen Schnittstellen. Dabei werden die Textdaten an TopicZoom übermittelt. Nach Berechnung der Themenkategorien wird das Ergebnis in einem Fenster unterhalb des Texteingabefeldes dargestellt. Nachfolgend erklärt Topiczoom die Übermittlung, Verarbeitung und Nutzungsbedingungen dieser elektronischen Daten.

Datenschutzerklärung

TopicZoom erhebt und verwendet die durch die TopicZoom Web-Tags an TopicZoom übertragenen Daten ausschließlich im Rahmen der Bestimmungen des Datenschutzrechts der Bundesrepublik Deutschland. Nachfolgend unterrichten wir Sie über Art, Umfang und Zwecke der Erhebung und Verwendung dieser Daten.

Die Daten zur Auswertung und Berechnung der Themenkategorien werden über das Internet an einen TopicZoom Server übermittelt. Danach wird der Text tokenisiert und indiziert. Die Berechnung der Themenkategorien erfolgt auf Basis des Index mit einer durch TopicZoom entwickelten semantischen Ontologie. Dabei werden von TopicZoom ausschließlich die von TopicZoom erzeugten Indexdaten erfasst und gespeichert. Die Rechte des ursprünglichen Textes bleiben dabei unberührt.

Nach Berechnung der Themenkategorien auf dem TopicZoom Server, werden diese Informationen über das Internet zurück an den anfragenden Client geschickt.

Die in der Datenbank gespeicherten Indexinformationen über das Dokument sind von jeglichen personenbezogenen Daten getrennt. Eine Zuordnung dieser Daten zu einer bestimmten Person ist der TopicZoom GmbH nicht möglich.

Auf Anordnung der zuständigen öffentlichen Stellen muss TopicZoom im Einzelfall Auskunft über die Daten der Dokumente erteilen, soweit dies für Zwecke der Strafverfolgung, zur Gefahrenabwehr, zur Erfüllung der gesetzlichen Aufgaben der Verfassungsschutzbehörden oder des Militärischen Abschirmdienstes oder zur Durchsetzung der Rechte am geistigen Eigentum erforderlich ist.



Nutzungsbedingungen

Haftungsausschluss

TopicZoom übernimmt keine Gewährleistung dafür, dass die bereitgestellten Themenkategorien zu dem Dokument vollständig, richtig und in jedem Fall aktuell sind. Die TopicZoom GmbH behält sich das Recht vor, ohne vorherige Ankündigung, Änderungen oder Ergänzungen der Themenkategorien vorzunehmen oder die Bereitstellung der TopicZoom Web-Tags gänzlich zu beenden. TopicZoom übernimmt keine Gewährleistung dafür, dass die Funktionen unterbrechungs- und fehlerfrei ablaufen, dass Fehler behoben werden oder dass der Server, der sie zur Verfügung stellt, frei von Viren oder sonstigen schädlichen Komponenten ist. TopicZoom kann unter keinen Umständen und unter keiner rechtlichen Handhabe für spezifische oder mittelbare Schäden, die sich aus dem Gebrauch oder der nicht gegebenen Möglichkeit zum Gebrauch der bereitgestellten Themenkategorien ergeben, haftbar gemacht werden. Dies gilt auch dann, wenn die TopicZoom oder deren Erfüllungsgehilfen auf die Möglichkeit eines solchen Schadeneintritts aufmerksam gemacht wurden. Diese Haftungsausschlüsse kommen nicht zur Anwendung, wenn ein Schaden aus grober Fahrlässigkeit oder Vorsatz der TopicZoom herrührt, oder grundsätzlich, wenn es sich um Körperverletzung handelt.

Kommerzielle Nutzung

Einer kommerziellen Nutzung der durch TopicZoom berechneten Web-Tags wird bis auf Widerruf ausdrücklich zugestimmt. Sollten Sie eine weitergehende kommerzielle Nutzung der TopicZoom Technologie, der zugehörigen Themenkategorien und der daraus entstehenden Informationen planen oder Interesse an weiteren diesbezüglichen Informationen und Features haben, setzen Sie sich bitte mit uns in Verbindung: (<mailto:info@topiczoom.com>).

Sonstige Bestimmungen

Die in digitaler Form überlassenen Web-Tags dürfen nur mit schriftlicher Zustimmung von TopicZoom vervielfältigt oder Dritten zugänglich gemacht werden. Ein Rechtsanspruch auf Gestattung der Nutzung durch Dritte besteht nicht. Der Nutzer verpflichtet sich, die in den Web-Tags enthaltenen Schutzvermerke, wie Copyright Vermerke und andere Rechtsvorbehalte unverändert beizubehalten.

Allgemeine Bestimmung

TopicZoom behält es sich vor, die vorgenannten Klauseln jederzeit zu ändern, zu ergänzen oder abschnittsweise zu kürzen.

Copyright TopicZoom GmbH – Februar 2016

<http://www.topiczoom.de/impressum.html>