

## Nutzung von RSS-Feeds im deutschen Verbandswesen

Eine Untersuchung von 1.769 Websites

**Autor:**

*Dipl.-Pol. Ralf-Thomas Hillebrand*

*politik & internet Kommunikationsberatung*

*Internet: [www.politik-und-internet.de](http://www.politik-und-internet.de)*

*E-Mail: [rth@politik-und-internet.de](mailto:rth@politik-und-internet.de)*

*Blogbeitrag zu dieser Untersuchung: <http://polkomm.net/rss-studie>*

## **1. Summary**

### **1.1. Gegenstand der Untersuchung**

Die Technologie von RSS-Feeds ist eine der wesentlichen Grundlagen des sogenannten Echtzeit-Webs. Betrachtet man das World Wide Web als eine Art dezentral organisiertes Speichermedium für Publikationen aller Art, also gewissermaßen als globale Bibliothek der Menschheit, stellt die RSS-Technologie quasi einen Nachrichtenstream über Neuanschaffungen dieser Bibliothek dar. RSS erlaubt den Betreibern von Websites ihre Veröffentlichungen in einen chronologischen Kontext zu stellen. Erst ein solcher chronologischer Kontext verleiht einer Botschaft in der Logik des maschinell gesteuerten Echtzeitwebs Aktualität – Voraussetzung für Nachrichtenwert. Insofern kommt der RSS-Technologie eine massive Bedeutung in der politischen Kommunikation zu, deren geringer Aufwand zudem in einem minimalen Verhältnis zum hohen Nutzen steht<sup>1</sup>.

Grundlegende Fragestellung dieser Untersuchung war, inwieweit in Deutschland aktive Verbände im Rahmen ihrer Online-Kommunikation RSS-Feeds (der Formate RSS, RDF und Atom<sup>2</sup>) einsetzen, um auf ihre aktuellen Internet-Publikationen hinzuweisen.

Die „Öffentliche Liste über die beim Bundestag registrierten Verbände und deren Vertreter“ weist in ihrer aktuellen amtlichen Fassung (Stand 30. April 2010) insgesamt 2.136 Verbände aus. Das gemeinhin als „Lobbyliste“ bekannte Verzeichnis gibt also einen weit reichenden Überblick über das deutsche Verbandswesen.

Auf der Basis der Lobbyliste wurde untersucht,

- a) welche Verbände eine eigene Website unterhalten,
- b) welche Websites im HTML-Header ihrer Homepage RSS-Feeds technisch korrekt angeben,
- c) welche angegebenen RSS-Feeds vorhanden und technisch einlesbar sind,
- d) welche Feeds aktuelle Inhalte anzeigen sowie
- e) welche Feeds im Sinne von W3C<sup>3</sup> technisch valide sind.

### **1.2. Vorgehensweise**

Im ersten Schritt wurden aus der Lobbyliste alle von den Verbänden angegebenen Website-Adressen (URLs) extrahiert.

Im zweiten Schritt wurden die Homepages der Websites von einem Robot eingelese und daraufhin untersucht, ob sie im HTML-Header etwaige RSS-Feeds korrekt deklarieren.

Im dritten Schritt wurde untersucht, ob sich die Feeds einlesen lassen.

---

<sup>1</sup> Zu Möglichkeiten und zur faktischen Nutzung mehr auf den folgenden Seiten dieses Berichts.

<sup>2</sup> Auch wenn es technisch nicht ganz exakt ist, wurden im Rahmen dieser Untersuchung die verschiedenen Formate RSS, RDF und Atom unter dem Oberbegriff RSS zusammengefasst. Dies ist insofern inkorrekt, als Atom technisch betrachtet nicht ein Unterformat von RSS ist, sondern ein Nachfolger des RSS-Formats.

<sup>3</sup> Das World Wide Web Consortium (W3C) ist das führende internationale Konsortium zur Entwicklung von Web-Standards. Sein Ziel ist, dass Protokolle und Richtlinien entwickelt werden, die ein langfristiges Wachstum des World Wide Web sichern. Siehe dazu auch: <http://www.w3c.de/about/overview.html>

Im vierten Schritt wurden die eingelesenen Feeds daraufhin untersucht, ob sie aktuelle Nachrichten beinhalten.

In einem fünften Schritt wurden jene Feeds, die alle zuvor genannten Voraussetzungen erfüllten, einem technischen Validierungs-Test nach den Spezifikationen von W3C unterzogen.

### 1.3. Ergebnisse

Von den 2.136 beim Bundestag gemeldeten Verbänden haben 1.769 eine eigene Webpräsenz angegeben (82,8 Prozent).

Von den 1.769 angegebenen Websites weisen nur 332 im Header ihrer Homepage, also der unter dem angegebenen URL als erstes sichtbaren HTML-Seite, mindestens einen RSS-Feed korrekt aus (18,8 Prozent der Verbände mit eigenen Websites, 15,5 Prozent aller Verbände).

Bei den 332 Verbandswebsites mit korrekt deklarierten Feeds ließen sich die Feeds von 19 Websites nicht einlesen, da sie nicht vorhanden, technisch fehlerhaft oder passwortgeschützt waren (5,7 Prozent der Verbände mit korrekt deklarierten Feeds). Nur 313 Verbände hatten einen technisch einlesbaren RSS-Feed (17,7 Prozent der Verbände mit eigenen Websites, 14,7 Prozent aller Verbände).

Zum Zeitpunkt der Untersuchung (Abruf via Internet am 18. und 19. Januar 2011) beinhalteten 39 Feeds der 313 Verbände mit einem technisch einwandfreien RSS-Feed keinerlei aktuelle Inhalte (12,5 Prozent). Bei 25 Verbänden stammt die letzte Meldung aus 2009 oder den Jahren davor, bei 14 weiteren Verbänden zeigte der RSS-Feed überhaupt keine Inhalte.

Nur 274 Verbände stellen demzufolge einlesbare RSS-Feeds bereit und verbreiten darüber auch Inhalte (15,5 Prozent der Verbände mit eigenen Websites, 12,8 Prozent aller Verbände).

Die RSS-Feeds von 108 der 274 Verbände (39,4 Prozent) wurde als technisch nicht valide identifiziert, da sie Syntax-Fehler enthielten. Nur 166 Verbände setzen demzufolge RSS in ihrer Verbandskommunikation fehlerfrei ein (7,8 Prozent aller Verbände, 9,4 Prozent der Verbände mit eigenen Websites).

### 1.4. Methodische Grenzen der Untersuchung

Die vorliegende Untersuchung berücksichtigt gezielt nur jene RSS-Feeds, die im Header der Homepage technisch korrekt deklariert wurden. Denn nur so ist eine sogenannte *Auto-Detection* des Feeds möglich, das heißt:

Nur bei technisch korrekter Deklaration des RSS-Feeds

- sind handelsübliche Webbrowser in der Lage ein oranges RSS-Symbol anzuzeigen und den Internetbenutzer dadurch auf die Existenz eines Feeds hinzuweisen und

- können Webrobots (wie etwa Suchmaschinencrawler oder Newsaggregatoren) RSS-Feeds einlesen und zu deren automatisierter Weiterverbreitung beitragen.

Das Untersuchungsergebnis lässt insofern keine exakten Aussagen darüber zu wie viele Verbände RSS-Feeds eventuell irgendwo im Inhalt ihrer Website, möglicherweise auch erst auf Unterseiten, anbieten.

Nach Stichproben grob geschätzt fallen daher existierende RSS-Feeds von 1 bis 2 Prozent der Verbände durch das Untersuchungsraaster. Aus grundsätzlichen methodologischen Erwägungen zählen hierzu auch 15 Verbände, die mittels einer von W3C missbilligten und aus diversen Gründen technisch obsoleten Browserweiterleitung von einer Vorschaltseite auf die eigentliche Homepage weiterleiten, wo dann RSS-Feeds deklariert sind.

## **2. Zur Methodik**

### **2.1. Bedeutung von RSS**

RSS, einst als *Rich Site Summary* entwickelt, wird in seiner aktuellen technischen Spezifikation der Version 2.0 nunmehr mit *Really Simple Syndication* übersetzt.

Die ursprüngliche Intention von *Rich Site Summary* war: Da eine Website in ihrer Struktur prinzipiell Wichtiges prominent (also etwa auf der Homepage) platziert, nicht aber unbedingt Neues, muss der Besucher zeitaufwändig nach Neuem suchen, wenn er einen Webauftritt kontinuierlich beobachten möchte. Denn beispielsweise eine aktuelle Personalie oder ein bevorstehender Termin sind unter Umständen nur tief verschachtelt auf Unterseiten der Webpräsenz zu finden, während die Homepage möglicherweise von langfristig gleichbleibenden Inhalten der Selbstdarstellung des Websitebetreibers bestimmt wird. Hier hilft RSS dem Besucher, indem es Hinweise über jegliche neue Informationen auf der Website als Nachrichtenfeed anbietet. Der Besucher kann solche Nachrichten mit einem üblichen Browser oder einer FeedReader-Software abonnieren. Er wird danach automatisch darauf hingewiesen, wenn die Website neue Inhalte bietet. Dazu muss er die Website nicht erneut besuchen.

Die hinter *Really Simple Syndication* stehende Idee geht noch weiter: Hier steht der Austausch von Inhalten, die Syndication, im Vordergrund – und zwar der maschinell-automatisierte Austausch. RSS erlaubt nämlich nicht nur Internetnutzer auf neu im Internet veröffentlichte Information hinzuweisen, sondern RSS erlaubt auch die Inhalte selbst der globalen Nachrichtenmaschinerie des Internets zur unbegrenzten Weiterverbreitung zur Verfügung zu stellen. Denn die Inhalte von RSS-Feeds sind technisch so exakt strukturiert, dass sie jederzeit von entsprechender Software oder entsprechend programmierten Websites in neuem Kontext dargestellt werden können. Dabei kann der Inhalt auch maschinell auf das Vorhandensein bestimmter Suchworte untersucht und nur dann angezeigt werden, wenn die Suchworte darin enthalten sind – wohlgemerkt automatisiert. Nachrichten oder zusammenfassende Hinweise auf Nachrichten können sich so im Extremfall über verschiedene Medien hinweg frei im Internet ausbreiten<sup>4</sup>.

Im Sinne von *Rich Site Summary* stagniert die Nutzung von RSS – so wird jedenfalls unter Experten diskutiert<sup>5</sup>. Ursache für derartige Befunde ist die geringere Nutzung von RSS-Readern aufgrund neuer Echtzeit-Technologien und -Plattformen wie etwa Twitter.

Gleichzeitig nimmt jedoch die Zahl von web- wie auch gerätebasierten Diensten zu, die auf RSS aufbauen:

- Mit *Yahoo Pipes* offeriert Yahoo einen mächtigen RSS-Aggregator, mit dem sich RSS-Feeds zu individuellen Nachrichtenfeeds kombinieren lassen<sup>6</sup>. Hierbei lässt sich vielerlei Suchmaschinen-

---

<sup>4</sup> Ein einfaches, in der Praxis gängiges Beispiel sind von einem Content-Management-System erzeugte RSS-Feeds mit Pressemitteilungen, die automatisch auf den Facebook-Account des Herausgebers gepostet werden und dort innerhalb der Community durch „Teilen“ von User zu User weitergeleitet werden.

<sup>5</sup> Siehe dazu: (1) Steve Gillmore „Rest in Peace, RSS“, TechCrunchIT vom 05.05.2009 <http://www.techcrunchit.com/2009/05/05/rest-in-peace-rss/> und (2) Chelsi Nakano „Ask.com kills Bloglines, is RSS Dead?“, CMS Wire vom 13.09.2010 <http://www.cmswire.com/cms/web-content/askcom-kills-bloglines-is-rss-dead-008580.php>.

<sup>6</sup> Siehe dazu: Daniel Schlicht „Ab durch die Röhre!“, Zeit online vom 11.09.2009 <http://www.zeit.de/digital/internet/2009-09/yahoo-pipes>

technologie – bis hin zu semantischen Suchmethoden – auf die RSS-Feeds anwenden, so dass man über entsprechende selbst zu bestimmende Filter sehr präzise themenspezifische Feeds erzeugen kann.

- Mit dem *Google Reader* bietet der führende Suchmaschinenkonzern ein Tool, das beliebige RSS-Feeds aggregiert. Deren Inhalte können nicht nur nach Suchbegriffen gefiltert werden, der Google Reader erlaubt über kostenlose Erweiterungen auch bequemes Offline-Lesen von RSS-Feeds und dies auch speziell im Hinblick auf Smartphones. Der Google Reader ermöglicht auch das Empfehlen von RSS-Inhalten gegenüber anderen Google-Benutzern (sogenanntes „Teilen“).
- Sowohl Medien als auch Recherchedienste bieten zunehmend entsprechend aufbereitete Aggregate von RSS-Feeds als eigenen Nachrichtendienst an. Als Beispiel sei das *Handelsblatt* genannt, das zu Jahresbeginn 2011 eine eigene Website<sup>7</sup> startete, die von Unternehmen per RSS bereitgestellte börsenrelevante Nachrichten nahezu in Echtzeit sammelt und darstellt.
- Mit zunehmender Verbreitung von Smartphones, die sich zur Darstellung von ganzen Websites nur bedingt eignen, erlangen RSS-Feeds eine besondere Bedeutung. So werden für alle Handy-Betriebssysteme derzeit ständig neue RSS-Reader entwickelt<sup>8</sup>. Dieser Trend setzt sich auch beim iPad von Apple fort, zu dem es bereits Software gibt, die aus RSS-Feeds eine individuell gestaltete „Zeitung“ generiert<sup>9</sup>.
- Grafische Benutzeroberflächen von PC-Betriebssystemen und Webdiensten (wie zum Beispiel iGoogle) nutzen zunehmend Widgets, kleine Unterprogramme, die im Leerlaufzustand des Geräts oder Webdienstes aktuelle Informationen anzeigen<sup>10</sup>. Nahezu alle Widgets, die Neuigkeiten anzeigen, basieren auf RSS-Feeds. Widgets gibt es mittlerweile beispielsweise für fast alle PC-Betriebssysteme, aber auch für die persönliche Gestaltung von Websites von Google und Yahoo und nicht zuletzt für Smartphones.
- Soziale Netzwerke ersetzen RSS nicht, wie gelegentlich behauptet wird, sondern sie nutzen die Technologie vielmehr im Hintergrund als Schnittstelle. Für das automatisierte Veröffentlichen von Inhalten in Sozialen Netzwerken gibt es eine wachsende Zahl von Diensten, die einen RSS-Feed einlesen und dessen Inhalte auf Facebook, Twitter und andere Netzwerke posten. Ein Beispiel ist Twitterfeed.com, ein Dienst, der beliebige RSS-Feeds auf Twitter und gleichzeitig auf Facebook veröffentlicht. Nach eigenen Angaben verbreitet Twitterfeed.com über 2,5 Millionen RSS-Feeds in Sozialen Netzwerken weiter. Twitter selbst stellt seine Inhalte übrigens wiederum als RSS-Feed zur Verfügung.

Die Nutzung von RSS verlagert sich also vom persönlichen Gebrauch durch User in RSS-Readern hin zu einem Standard im von Maschinen gesteuerten Content-Tausch zwischen verschiedenen Medien – angetrieben insbesondere durch die wachsende Bedeutung Sozialer Netzwerke und mobiler Endgeräte.

---

<sup>7</sup> <http://www.financial-informer.de/>

<sup>8</sup> Siehe dazu auch: Jack Wallen „The five best news aggregator Android apps“, TechRepublic vom 05.11.2010

<http://www.techrepublic.com/blog/smartphones/the-five-best-news-aggregator-android-apps/1858>

<sup>9</sup> Siehe dazu: Martin Weigert „Applikation macht aus RSS-Feeds eine iPad-Zeitung“, Netzwertig.com vom 26.05.2010

<http://netzwertig.com/2010/05/26/paper-pile-applikation-macht-aus-rss-feeds-eine-ipad-zeitung/>

<sup>10</sup> Einige Beispiele für Windows 7: <http://www.windows7download.com/win7-widgets/news/index304-311-o-d.html>

## 2.2. Primäres Kriterium: Technisch einwandfreier Pfad zur Auto-Detection der Feeds

Die im vorigen Abschnitt genannten Funktionen von RSS bauen in vielen Fällen darauf auf, dass Websites ihre RSS-Feeds korrekt im Header deklarieren, sodass Software-Robots den Feed finden und automatisch einlesen können (*auto detection*). Insofern ist eine technisch einwandfreie Deklaration in einer zunehmenden Zahl von Fällen existenziell für die Nutzung des RSS-Feed einer Website.

Darüber hinaus nutzt eine überwiegende Mehrheit der am Markt befindlichen Webbrowser *auto detection*, um dem User das Vorhandensein von Feeds anzuzeigen. Ein irgendwo im Inhalt einer Website versteckter Feed wird hingegen aller Wahrscheinlichkeit nach oft von Besuchern nicht entdeckt und daher auch nicht genutzt.

Die kommunikative Wirkung eines RSS-Feeds ist daher direkt von seiner korrekten Deklaration im HTML-Header abhängig – ein nicht aufzufindender RSS-Feed erreicht die Adressaten nicht oder nur selten. Anfängliche Überlegungen, ob es im Zuge der vorliegenden Untersuchung sinnvoll und realisierbar wäre, die Websites der 2.163 einbezogenen Verbände komplett auch nach Hyperlinks zu nicht korrekt deklarierten RSS-Feeds abzusuchen, wurden daher verworfen. Als primäres Untersuchungskriterium wurde die technisch korrekte Deklaration im HTML-Header der Homepage ausgewählt.

Die Deklaration von RSS-Feeds im HTML-Header einer Website erfolgt durch Einfügen einer Zeile HTML-Code wie der folgenden:

```
<link rel="alternate" type="application/rss+xml" title="[pol:komm:net]-RSS"
href="http://www.polkomm.net/blog/?feed=rss2" />
```

Die Internet Engineering Task Force, eine Schwesterorganisation von W3C, hat die Syntax im Jahr 2006 in einem Draft beschrieben<sup>11</sup>. Sie hat sich als Standard durchgesetzt und wird beispielsweise von allen gängigen, RSS-fähigen Browsern verstanden.

Um im Hinblick auf das Kriterium einer korrekten Deklaration konsequent zu sein, wurde darüber hinaus bei der Untersuchung jenes HTML-Dokument als Homepage einer Website definiert, das nach den Regeln des Hypertext Transfer Protocols (HTTP) als erstes nach Aufruf der Domain als erfolgreich übertragen eingestuft wurde (HTTP-Status-Code 200).

Websites, die von einer Vorschaltseite ohne RSS-Deklaration mittels des obsoleten Meta-Refresh-Tags<sup>12</sup> im HTML-Header auf die eigentliche, nachgeschaltete Homepage verweisen, wurden daher als Websites mit nicht korrekter RSS-Deklaration gewertet, selbst wenn diese auf der zweiten Seite korrekt war. Denn W3C hat die Verwendung des Meta-Refresh-Tags bereits im Jahre 2000 missbilligt<sup>13</sup> und stattdessen HTTP-konforme Weiterleitungstechniken empfohlen<sup>14</sup>. Es kann daher nicht davon ausgegangen werden, dass Robots oder RSS-Software einer solchen Weiterleitung folgen und ein *auto*

<sup>11</sup> Siehe: <http://www.rssboard.org/rss-mime-type-application.txt>

<sup>12</sup> Siehe: [http://en.wikipedia.org/wiki/Meta\\_refresh](http://en.wikipedia.org/wiki/Meta_refresh)

<sup>13</sup> Ausführlicher begründet hat W3C das in diesem Dokument von 2007: <http://www.w3.org/QA/Tips/reback>

<sup>14</sup> Das Hypertext Transfer Protocol sieht vor, dass Dokumente, die nicht unter dem aufgerufenen, sondern einem anderen URL vorliegen, mit den HTTP Status Codes 301 (Moved Permanently), 302 (Found) oder 307 (Temporary Redirect) und dem entsprechenden neuen URL beantwortet werden. Browser laden dann üblicherweise automatisch das Dokument mit dem neu angegebenen URL, ohne dass der Benutzer dies wahrnimmt oder gar aktiv werden muss.

*detect* erfolgreich ist. Dies auch deshalb, weil Dokumente mit einer Meta-Refresh-Anweisung den HTTP-Statuscode 200 („OK“) mitliefern – und damit einen abgeschlossenen Download anzeigen.

### **2.3. Eigene Website: Prinzip der Selbstauskunft**

Die Datenbasis für die vorliegende Untersuchung bildet die „Öffentliche Liste über die Registrierung von Verbänden und deren Vertretern“ des Deutschen Bundestages in der Fassung vom 30. April 2010. Durch Veröffentlichung im Bundesanzeiger vom 3. Mai 2010 ist sie die gegenwärtig gültige „Amtliche Fassung“.

Die Registrierung von Verbänden wird von der Geschäftsordnung des Deutschen Bundestages geregelt. Dabei wird die Angabe von E-Mail- und Internetadresse explizit als „erforderlich“ bezeichnet. Die Selbstauskunft der Verbände zu ihrer Webpräsenz hat insofern annähernd normativen Charakter<sup>15</sup>.

Es kann zwar nicht ausgeschlossen werden, dass einzelne Verbände trotz Bestehen einer eigenen Website diese dem Bundestag nicht bekannt gegeben haben. Methodisch ist jedoch die Selbstauskunft gegenüber einem Verfassungsorgan die eindeutigste Datenbasis. Denn sie entscheidet jegliche Zweifelsfälle im Sinne des jeweiligen Verbands. Zweifelsfälle bestehen beispielsweise in folgenden Fällen:

- In Verbandskonglomeraten – bestehend etwa aus Wirtschafts- und Arbeitgeber-Verband derselben Branche – werden Werbauftritte gelegentlich gemeinsam betrieben. Inwieweit der einzelne Verband in solche Fällen „einen eigenen Werbauftritt“ betreibt, lässt sich von außen nur schwer entscheiden.
- Insbesondere kleinere Verbänden werden bisweilen maßgeblich von einem einzigen Unternehmen getragen. Die Online-Kommunikation des Verbands findet in solchen Konstellationen nicht selten über eine Website statt, die nur schwer eindeutig entweder dem Unternehmen oder aber dem Verband zuzuordnen ist.

Da knapp 83 Prozent der Verbände eine Angabe zu ihrer Website gemacht haben, gewährleistet dies in jedem Falle die statistische Signifikanz der in diesem Bericht dargestellten Untersuchungsergebnisse.

### **2.4. Syndication: Automatisierte Verbreitung der Verbandsbotschaften im Internet**

Die Intention der vorliegenden Untersuchung ist herauszufinden, inwieweit RSS genutzt wird, um die Botschaften der untersuchten Verbandswebsites unverzüglich auffindbar zu machen (*Rich Site Summary*) oder gar in den globalen Nachrichtenstream einzubringen (*Really Simple Syndication*).

Sofern deklarierte RSS-Feeds passwortgeschützt waren und somit nur einem geschlossenen Benutzerkreis zur Verfügung standen, wurden sie in dieser Untersuchung als nicht einlesbar gewertet. Dies war in zwei Fällen gegeben.

---

<sup>15</sup> <http://www.bundestag.de/dokumente/parlamentsarchiv/sachgeb/lobbyliste/index.html>

### **3. Weitere Ergebnisse im Detail**

#### **3.1. Ermittlung der Webpräsenzen**

Die Ermittlung der Adressen (URLs<sup>16</sup>) der Verbandswebsites erfolgte durch elektronische Suche in einem PDF-Exemplar der „Lobbyliste“. Gelegentlich vorkommende fehlerhafte Schreibweisen (wie etwa [www.bdi.ev](http://www.bdi.ev) statt [www.bdi.eu](http://www.bdi.eu)) wurden nachträglich korrigiert. Doppelnennungen<sup>17</sup> wurden aus dem Resultat entfernt. Bei Angabe mehrerer Websites wurde jeweils nur jene im Resultat belassen, und zwar jene, auf der die kontinuierliche verbandspolitische Arbeit in Deutschland überwiegend dokumentiert war<sup>18</sup>.

Es wurden unter den 2.136 beim Bundestag gemeldeten Verbänden 1.769 Verbände ermittelt, die eine Internetadresse für ihre Webpräsenz angegeben hatten.

#### **3.2. Download der Homepages und Untersuchung des Headers auf RSS-Deklaration**

Es wurde ein Robot programmiert, der vom 18. bis zum 19. Januar 2011 den HTML-Code der Homepages abrief und den `<head>`-Bereich maschinell nach `<link>`-Tags der folgenden MIME Media-Types durchsuchte:

- *application/rss+xml*
- *application/rdf+xml*
- *application/atom+xml*
- *application/xml*

Der vierte genannte Media-Type ist zwar streng genommen keine syntaktisch ganz korrekte Deklaration für einen RSS-Feed, aber da das weitverbreitete, in erster Linie von Bloggern verwendete Content-Management-System *WordPress* seine RSS-Feeds fälschlicherweise lange Zeit so deklarierte, wurde diese Deklaration (die gleichwohl nur sehr selten vorgefunden wurde) im Verlauf der Untersuchung als korrekt toleriert.

Der Robot war so programmiert, dass er Redirects über die HTTP-Status-Codes 301, 302 und 307 verfolgte, das technisch obsolete Tag „Meta Refresh“ jedoch entsprechend der Missbilligung durch W3C ignorierte<sup>19</sup>. In 78 Fällen ließ sich der Robot daher zunächst nicht auf die eigentliche Homepage weiterleiten. Unter diesen 78 Websites befanden sich auch jene von prominenten Verbänden wie der *Bundesvereinigung der Deutschen Arbeitgeberverbände*, die gegenüber dem Bundestag nicht mehr die früher verwendete Webadresse *bda-online.de* angegeben hatte, sondern *arbeitgeber.de*, sowie der *Verbraucherzentrale Bundesverband*, die *Gewerkschaft Erziehung und Wissenschaften* und der *deutsche Hotel- und Gaststättenverband DEHOGA*.

---

<sup>16</sup> Unified Resource Locator, Internetadresse der Syntax *[Protocol/Schema]:([Subdomain].)[Domain].[Top-Level-Domain]*

<sup>17</sup> Beispielsweise führt die Liste die Website des BKK Bundesverbands unter dem Verband selbst, aber zudem unter seinem Bonner Büro erneut auf.

<sup>18</sup> So hat der Bundesverband der Deutschen Entsorgungs-, Wasser- und Rohstoffwirtschaft e.V. mit [www.bde-berlin.de](http://www.bde-berlin.de) und [www.bde-bruessel.be](http://www.bde-bruessel.be) zwei Websites angegeben, von denen nur die erstere im Weiteren berücksichtigt wurde.

<sup>19</sup> Siehe dazu auch: Abschnitt 2.2. „Primäres Kriterium: Technisch einwandfreier Pfad zur Auto-Detection der Feeds“

Eine Auswertung dieser erheblichen Anzahl von Websites mit obsoletter Weiterleitung zur eigentlichen Homepage ergab, dass der Grund dafür sehr häufig der Einsatz eines Content-Management-Systems (CMS) ist und offensichtlich Unkenntnis darüber herrscht, wie man eine Homepage, für die das CMS eine komplexere Webadresse mit Pfaden und Parametern erwartet, trotzdem unter der bloßen Internetdomain angezeigt bekommt<sup>20</sup>.

Schließlich wurden die 78 Homepages gemäß der im HTTP-Refresh-Tag angegebenen Adresse eingelesen und auf die Deklaration von RSS-Feeds hin untersucht. In zwei Fällen musste dabei sogar einem zweifachen Redirect mittels HTTP-Refresh-Tag gefolgt werden. Es wurden auf 15 der Websites korrekt deklarierte RSS-Feeds vorgefunden, die aus weiter oben genannten methodischen Erwägungen aber nicht in das Endergebnis dieser Untersuchung einbezogen wurden, darunter auch auf *arbeitgeber.de*.

Stichprobenartig wurde – hauptsächlich im Fall größerer Verbände – auch untersucht, ob die Websites RSS-Feeds anbieten, jedoch keine Deklaration im HTML-Header vorgenommen wurde. Dies war beispielsweise beim *Bundesverband der Deutschen Industrie* und beim *Verbraucherzentrale Bundesverband* der Fall. Ein Kuriosum bot die Website des *Netzwerk Europäische Bewegung Deutschland (EBD)*: Die syntaktisch korrekte Deklaration des RSS-Feeds im Header wurde auskommentiert<sup>21</sup>, so dass sie nicht funktioniert.

Die Stichproben lassen nur eine Schätzung darüber zu, wie viele Verbände RSS anbieten, ohne eine Deklaration im Header vornehmen – und ihren Feed damit vom maschinengesteuerten Zugriff definitiv ausnehmen. Die Zahl dürfte im niedrigen zweistelligen Bereich liegen.

Insgesamt 332 Verbandswebsites lieferten schließlich eine W3C- und HTTP-konforme Deklaration von RSS-Feeds im Header ihrer Homepage.

### 3.3. Einlesen der Feeds

Zum Download der 332 RSS-Feeds wurde der frei verfügbare *Feedreader*<sup>22</sup> verwendet.

Bei 19 Websites ließ sich der deklarierte RSS-Feed nicht herunterladen, wofür es eine ganze Reihe von Ursachen gab:

- So existierte in einigen Fällen das entsprechende RSS-Dokument nicht (HTTP-Error 404 – Not found), wie etwa beim *Republikanischen Anwältinnen - und Anwälteverein*.

---

<sup>20</sup> Gängige Webserversoftware, wie etwa der Marktführer Apache, ermöglicht dies beispielsweise mittels eines einfachen Eintrages in einer Konfigurationsdatei. Für komplexere Weiterleitungen steht darüber hinaus mit *Mod\_Rewrite* ein Softwaremodul zur Verfügung, über das sich Adressen sämtlicher Dokumente eines Webauftritts nahezu beliebig manipulieren lassen. Und schließlich lassen sich über PHP- oder Perl-Scripting auch mit geringem Aufwand W3-konforme HTTP-Redirects mit den Status-Codes 301, 302 und 307 realisieren.

<sup>21</sup> Bei der Programmierung von HTML oder Software haben Programmierer in nahezu allen Programmiersprachen die Möglichkeit Kommentare einzufügen, die mittels einer bestimmten Zeichenkombination gekennzeichnet sind, damit sie sich vom auszuführenden Code unterscheiden. Mit solchen Zeichenkombinationen kann aber auch funktionaler Programmcode gegen die Ausführung gesperrt werden. In HTML haben Kommentare die folgende Syntax: `<!-- Kommentartext -->`

<sup>22</sup> Herunterzuladen unter: [www.feedreader.com](http://www.feedreader.com)

- In anderen Fällen konnte der Webserver das RSS-Dokument aufgrund von Mängeln in der Programmierung nicht erstellen (HTTP-Error 500 – Internal Server Error) wie beim *Bundesverband IT-Mittelstand*.
- Und in einer Reihe von Fällen enthielt das ausgelieferte Dokument kein valides, vom FeedReader lesbares RSS, so etwa beim *Bundesverband mittelständische Wirtschaft* und beim *Bundesverband Niere*.

Auffällig war im Falle dieser nicht vorhandenen oder technisch fehlerhaften RSS-Feeds der hohe Anteil von Webpräsenzen, die mit dem freien CMS *Typo 3* betrieben wurden (wie etwa beim *Republikanischen Anwältinnen - und Anwälteverein* und beim *Bundesverband mittelständische Wirtschaft*). Als kostenloses und quelltextoffenes CMS erfreut sich *Typo 3* wachsender Beliebtheit, allerdings wird der Aufwand einer Anpassung des universell konzipierten Systems an die Erfordernisse verbandspolitischer Kommunikation sehr oft unterschätzt.

Ein kleiner Anteil der im Rahmen des Untersuchungszeitraums am 18. und 19. Januar mangels Validität nicht einlesbaren RSS-Feeds wies offensichtlich nur temporäre Fehler auf und war einige Tage später wieder fehlerfrei. Dies weist darauf hin, dass die in den RSS-Feed übernommenen Inhalte die Validität des RSS-Codes zu korrumpieren in der Lage waren. Auch in solchen Fällen liegt in aller Regel eine mangelhafte Programmierung des Feeds vor, die sich aber immer nur dann auswirkt, wenn beispielsweise Zeichen wie „&“, „<“ oder „>“ in einem der Inhalte des RSS-Feeds auftauchen und nicht entsprechend maskiert sind<sup>23</sup>.

313 jener Verbände, die in Ihrem Homepage-Header RSS-Feeds korrekt auswiesen, verfügten über einen einlesbaren Feed.

### 3.4. Weitere Technische Mängel in RSS-Feeds

Bei den 313 Feeds, die der *FeedReader* darzustellen in der Lage war und die somit augenscheinlich ihren kommunikativen Zweck erfüllen konnten, tauchten dennoch technische Schwächen auf.

Beispielsweise ist der RSS-Feed der *Bundespressekonferenz* eine Vermischung von verschiedenen Zeichensätzen: Während der Code des RSS-Files laut Deklaration ISO-8859-1 sein soll (Zeichensatz für Mitteleuropa mit deutschen Umlauten) importiert das CMS Inhalte, die in UTF-8 notiert sind. Die Folge ist, dass überall wo der Feed Umlaute zeigen sollte, kryptische Kombinationen aus jeweils zwei Zeichen zu sehen sind<sup>24</sup>.

Der *Bundesverband Deutscher Milchviehhalter* bietet auf seiner Website *bdm-verband.org* zwei RSS-Feeds an, die – folgenreich! – offenbar versuchen Probleme mit dem Euro-Zeichen „€“ zu umgehen. Das Euro-Zeichen wurde erst vor rund zehn Jahren eingeführt und war zu diesem Zeitpunkt in den

<sup>23</sup> Die beiden häufigsten Fehlerquellen, wegen derer RSS-Feeds infolge aufgenommenen Inhalte ihre Validität einbüßen, sind: (1) die Verwendung von reservierten Zeichen (wie „&“, „<“ und „>“) im Inhalt und (2) die Vermengung unterschiedlicher Zeichensätze (wie UTF und ISO-8859-1). So kann beispielsweise die Erwähnung der Textilmärke „H&M“ in einer Meldung des RSS-Feeds nur wegen des Kaufmanns-Umlauts den Feed korrumpieren; genauso kann dies geschehen, wenn in einem als UTF ausgewiesenen RSS-Feed ein Umlaut vorkommt.

<sup>24</sup> <http://www.bundespressekonferenz.de/documents/appointments.xml>

verbreiteten Zeichensätzen nicht enthalten. Damals schuf man deshalb den neuen, zusätzlichen Zeichensatz ISO-8859-15. Dessen Verbreitung vollzieht sich aber, insbesondere außerhalb des Euro-Raumes, langsamer als erhofft. RSS-Software, seien es RSS-Reader oder Robots, sind daher häufig nicht in der Lage in ISO-8859-15 notiertes RSS darzustellen. So scheiterte auch der FeedReader an der Darstellung des Feeds des *Bundesverbandes Deutscher Milchviehhalter*. Da es prinzipiell syntaktisch korrekt ist einen RSS-Feed auch in ISO-8859-15 bereitzustellen, wurden die Feeds im Gesamtergebnis dieser Untersuchung dennoch als lesbar gewertet.

Der *Automobilclub von Deutschland* stellt seinen RSS-Feed via *HyperText Transfer Protocol Secure* (HTTPS) bereit. HTTPS ist eine Technologie, mit der Daten im Internet verschlüsselt übertragen werden, damit sie auf dem Übertragungsweg nicht mitgelesen werden können. HTTPS wird beispielsweise bei Bezahlvorgängen im Onlinehandel oder beim Internetbanking genutzt, außerdem schreibt das deutsche Datenschutzrecht HTTPS bei der Übertragung schützenswerter persönlicher Daten via Internet vor. Im Falle von Nachrichten, die sich möglichst frei im Internet verbreiten sollen, ist ein solcher Schutz jedoch nicht erforderlich. Vielmehr kann die Verschlüsselung in Einzelfällen sogar kontraproduktiv, denn – auch wenn übliche Browser HTTPS beherrschen – ist nicht davon auszugehen, dass auch alle mobilen Endgeräte oder Webrobots damit problemlos umgehen können.

### 3.5. Inhaltliche Auswertung der Feeds

Die 313 einlesbaren Feeds wurden in Bezug auf das Vorhandensein von Inhalte und das Datum der letzten Meldung untersucht.

Bei 14 Verbänden zeigte der RSS-Feed keinerlei Inhalte an.

Bei 25 Verbänden stammt die letzte Meldung aus 2009 oder den Jahren davor.

Unter den letztgenannten befindet sich auch der *Verband Deutscher Privatschulverbände*, dessen technisch einwandfreier RSS-Feed unter dem Datum vom 7. Juli 2006 einen einzigen Inhalt hat, nämlich eine Begrüßung für Besucher der Website. Die Website des VDP, die ebenfalls mit einem kostenlosen und quelltextoffenen CMS aufgesetzt ist, nämlich mit *Joomla*, ist ein weiteres Beispiel für eine mangelhafte Anpassung eines freien Universal-CMS an die Erfordernisse der verbandspolitischen Arbeit: Denn nur durch Zufall stößt man unter Umständen darauf, dass der Verband, der mehrmals monatlich Pressemitteilungen herausgibt, auch einen RSS-Feed zu diesen Pressemitteilungen anbietet – allerdings ist der Feed ausschließlich im HTML-Header jener Seite deklariert, auf der sich die Übersicht über alle Pressemitteilungen befindet – und eben nicht auf der Homepage.

Nur 274 Verbände stellen demzufolge einlesbare RSS-Feeds bereit und verbreiten darüber auch Inhalte. Das sind 15,5 Prozent der Verbände mit eigenen Websites und 12,8 Prozent aller beim Bundestag gemeldeten Verbände.

### 3.6. Syntaktische Validierung

Alle RSS-Formate sind Dialekte der als „Auszeichnungssprache“ bezeichneten *Extensible Markup Language*, kurz XML. Eine der wesentlichen Aufgaben von XML ist die exakte und zweifelsfreie Strukturierung von Information, sodass die Information maschinenlesbar und mithin von Computern automatisiert verarbeitbar ist. XML hat deshalb eine sehr strikte Spezifikation – und damit auch RSS. So sind beispielsweise Datumsangaben innerhalb von RSS-Feeds immer im MIME-Format<sup>25</sup> zu machen (Beispiel: „Sun, 06 Feb 2011 12:56:38 GMT“).

W3C bietet daher einen *Feed Validation Service* als Internetdienst an, mit dem man RSS-Feeds auf Einhaltung der Spezifikation testen kann<sup>26</sup>. Mit diesem Tool wurden die 274 einlesbaren und mit Inhalten bestückten Feeds abschließend auf Validität untersucht. Die Feeds von 108 Verbänden bestanden diesen Test nicht.

Es muss zunächst noch einmal angemerkt werden, dass diese nicht-validen RSS-Feeds vom FeedReader dargestellt werden konnten und daher ihren kommunikativen Zweck im Wesentlichen zu erfüllen in der Lage sein dürften. Doch kann eine fehlerhafte Syntax von Fall zu Fall die maschinelle Lesbar- und Verarbeitbarkeit beeinträchtigen. In einer ganzen Reihe von Fällen wurden beispielsweise die Veröffentlichungszeitpunkte der einzelnen RSS-Nachrichten nicht im erforderlichen MIME-Format angegeben. Folge war, dass zwar die Inhalte dargestellt wurden, aber der FeedReader keinen sinnvollen Veröffentlichungszeitpunkt angeben konnte. Man darf davon ausgehen, dass es nicht jeder Software, die mit Nachrichten umgeht, egal ist, von wann eine Information stammt. Die Gefahr besteht deshalb zweifellos, dass sich solche RSS-Feeds – nur aufgrund der Verwendung des falschen Datumsformats – nicht in erwünschtem Maße weiterverbreiten.

Es erstaunt gleichwohl, dass 39,4 Prozent jener 274 Verbände, die RSS einsetzen und technisch korrekt deklarieren, Feeds mit fehlerhafter Syntax publizieren. Darunter waren im Untersuchungszeitraum IT-spezifische Verbände, wie *Linux Verband* und *Verband der Elektrotechnik Elektronik Informationstechnik*, große Gewerkschaften wie *Deutscher Gewerkschaftsbund IG Metall* und der mitgliederreiche *Zentralverband des Deutschen Handwerks*.

Nur 166 Verbände setzen demzufolge RSS in ihrer Verbandskommunikation fehlerfrei ein. Das sind 7,8 Prozent aller Verbände und 9,4 Prozent der Verbände mit eigenen Websites.

---

<sup>25</sup> Dieses Datumformat wurde bereits 1982 in der RFC 822 festgelegt und gilt heute eher als kompliziert und unpraktisch. Siehe: <http://www.ietf.org/rfc/rfc0822.txt>

<sup>26</sup> <http://validator.w3.org/feed/>

**Autor:**

*Dipl.-Pol. Ralf-Thomas Hillebrand*

*politik & internet Kommunikationsberatung*

*Internet: [www.politik-und-internet.de](http://www.politik-und-internet.de)*

*E-Mail: [rth@politik-und-internet.de](mailto:rth@politik-und-internet.de)*

*Blogbeitrag zu dieser Untersuchung: <http://polkomm.net/rss-studie>*