

## Content Syndication am Beispiel der Kompetenznetze in der Medizin

Löbe M

Institut für Medizinische Informatik, Statistik und Epidemiologie (IMISE), Universität Leipzig  
matthias.loebe@imise.uni-leipzig.de

**Einleitung und Fragestellung** Medizinische Forschungsverbände in Deutschland nutzen die Möglichkeiten des Internets im internationalen wie interdisziplinären Vergleich zu wenig aus. Eine sinnvolle Anwendung moderner Internetstandards ist das Verbreiten und Zusammenführen von Inhalten aus unterschiedlichen Quellen (Content Syndication). Mittels RSS-Nachrichtenkanälen (*RSS-Feeds*) lassen sich Inhalte verbreiten, von Anwendern abonnieren und maschinell verarbeiten. Im vorliegenden Beitrag wird die prototypische Umsetzung eines Nachrichtenportals am Beispiel der Kompetenznetze in der Medizin aufgezeigt. Das Portal arbeitet autark, es bezieht seine Inhalte ausschließlich aus den Webpräsenzen der Netze. Die Schaffung eines zentralen Informationsdienstes soll die Dissemination der Projektergebnisse fördern.

**Material und Methoden** Die Kompetenznetze in der Medizin [1] publizieren umfangreiche und aktuelle Informationen auf ihren Internetportalen. Das Schreiben und Veröffentlichen selbst erstellter Artikel ist zeitaufwändig; diese Web-Dokumente stellen somit einen echten geschaffenen Wert dar. Trotzdem werden die Informationen häufig nicht oder nur von wenigen Nutzern gelesen. Dafür können sowohl der geringe Bekanntheitsgrad der Website als auch Fehler bei der Informationsstruktur oder der Navigation verantwortlich sein. Eine Möglichkeit, mehr Nutzer zu erreichen, bestünde im Publizieren einer Information an mehreren Orten gleichzeitig. Allerdings zeigt die Erfahrung, dass es sehr aufwendig ist, Daten an mehreren Stellen konsistent und aktuell zu halten.

Eine andere Möglichkeit ist der Austausch von Dokumenten zwischen den Betreibern von Websites, das so genannte Content Syndication. Hier werden Internetdokumente oder zumindest Zusammenfassungen der Inhalte ausgetauscht und auf externen Seiten verlinkt. Vorteile sind hierbei die größere Verbreitung eigener Inhalte und der geringe Aufwand, da der Austausch (zumindest theoretisch) automatisch erfolgt. Als Standard für Content Syndication hat sich Rich Site Summary (RSS) herausgebildet. In diesem einfach strukturierten XML-Datenformat werden Nachrichtenkanäle mit einzelnen Artikeln definiert, die z.B. Metadaten wie Überschrift, Autor, Erstellungszeit, Kurzzusammenfassung und einen Link auf die Vollversion enthalten. Idealerweise werden solche RSS-Feeds von einer Webverwaltungssoftware automatisch bereitgestellt, indem sie vorhandene Inhalte auf Wunsch in RSS transformieren.

RSS-Feeds können auf verschiedene Arten konsumiert werden. Entweder benutzt der Anwender ein Extra-Programm ähnlich einem Webbrowser. Der Vorteil gegenüber einer normalen HTML-Seite besteht in einer aktuellen, zentralen Liste aller Änderungen, welche sich automatisch aktualisiert. So kann ein Journalist bspw. viele Websites überwachen, ohne alle per Hand nach Änderungen absuchen zu müssen. Eine weitere Verwertungsart von RSS-Feeds ist die Syndizierung oder Aggregation, das heißt das Sammeln und Darstellen von Neuigkeiten aus verschiedenen Quellen, z.B. aktuelle Nachrichten über den Informationsdienst Wissenschaft, Kompetenznetze.de usw. Im Ergebnis entsteht eine aggregierte HTML-Seite oder ein neuer RSS-Feed. Da es sich bei RSS-Feeds um maschinenlesbare Daten in einem offenen XML-Format handelt, ist die Weiterverwertung nicht eingeschränkt. Auch Suchmaschinen können diese Dokumente auswerten

Zur Umsetzung eines Content-Syndication-Prototyps bietet sich die zentrale Seite der Kompetenznetze in der Medizin an. Diese Seite verfügt derzeit nur über rudimentäre Inhalte (Kurzbeschreibungen der Netze), könnte aber eine wichtige Anlaufstelle für Interessierte über die Arbeit der Kompetenznetze werden, z.B. für Journalisten (Abonnement des aggregierten RSS-Feeds). Es ist nicht anzunehmen, dass alle Kompetenznetze RSS-Datenquellen bereitstellen. Es muss im Gegenteil davon ausgegangen werden, dass viele dieser an sich grundlegenden Informationen auf den bisherigen Webseiten der Verbände fehlen. Daher müssen die RSS-Feeds durch externes Scannen und Interpretieren der Webseiten erzeugt werden. Dieser Prozess muss sich ferner automatisieren lassen, da die Feeds täglich aktualisiert werden sollen und eine menschliche Qualitätssicherung des Imports auf Dauer zu aufwendig wäre. Mögliche Quelldaten für die RSS-Syndizierung sind:

- Termine und Veranstaltungsdaten, aktuelle Mitteilungen aus dem Netz (Zielgruppe Verbundpartner)
- Newsletter, Pressemitteilungen, Mediadaten (Zielgruppe Presse und Industrie)
- Patientenseminare, Informationen zu laufende klinische Studien (Zielgruppe Patienten und Angehörige)
- Forschungsberichte, Publikationen, Stellenangebote (Zielgruppe Wissenschaftler)

Obwohl RSS kein neuer Standard ist, hat er noch keine durchdringende Verbreitung gefunden. Nur jeder achte Internetnutzer kann den Begriff zuordnen [7]. Der Grund hierfür liegt in der Inkompatibilität verschiedener Versionen von RSS [2-4], die im Hinblick auf eine universelle Nutzbarkeit alle unterstützt werden müssen.

**Ergebnisse** Im Rahmen des Projekts „Web Qualitätsmanagement“ der Telematikplattform für Medizinische Forschungsnetze (TMF) [6] wurde ein Katalog syndizierbarer Daten erstellt und in der AG Öffentlichkeitsarbeit diskutiert. Es wurde ferner eine Umfrage in den Verbänden durchgeführt und mögliche Datenquellen eruiert. Erste Versuche der Integration verschiedener Datenquellen zeigen die prinzipielle Durchführbarkeit des Konzepts. Als besonderes Problem lässt sich schon hier die schlechte Qualität der Primärdaten ausmachen, besonders der Verstoß gegen die formalen Strukturgrammatiken und die geringe Persistenz der Ressourcenreferenzen.

**Diskussion** Spätestens mit der im Laufe dieses Jahres erwarteten Version 7 des Internet Explorers verfügen alle bedeutenden Browser über eine integrierte RSS-Funktionalität. Damit entfällt das wesentlichste Hemmnis einer umfassenden Durchdringung des RSS-Standards: der Bedarf, eine Software lokal zu installieren. Die benutzerfreundliche Integration in den Webbrowser wird die Verbreitung RSS-basierter Dienste stark beschleunigen.

### Literatur

- [1] <http://www.kompetenznetze-medizin.de/>
- [2] RDF Site Summary (RSS) 1.0 <http://web.resource.org/rss/1.0/>
- [3] Really Simple Syndication (RSS) 2.0 <http://blogs.law.harvard.edu/tech/rss>
- [4] ATOM 1.0 <http://www.atomenabled.org/>
- [5] Hammersley B. Developing Feeds with RSS and Atom. O'Reilly Media, 2005.
- [6] [http://www.tmf-ev.de/site/DE/int/AG/QM/projekte/Web-QS/c\\_Web-QS.php](http://www.tmf-ev.de/site/DE/int/AG/QM/projekte/Web-QS/c_Web-QS.php)
- [7] Grossnickle J, Board T et al. RSS—Crossing into the Mainstream, Oktober 2005, Studie der Marktforschungsagentur Ipsos Insight im Auftrag von Yahoo!