

Einsatz von Social Software und Semantic Web-Technologien im Patentbereich

Dr. Jasmin Schmitz

Institut für Forschungsinformation und Qualitätssicherung (iFQ), Berlin

schmitz@forschungsinfo.de

Abstract: Der Beitrag stellt überblicksartig Beispiele für den Einsatz von Social Software und Semantic Web-Technologien bei der Patentprüfung sowie der Erschließung von und Recherche nach Patenten vor.

Aufbau eines Semantic Web zur Retrievalunterstützung

Stichwortrecherchen sind gerade in solchen Technologiebereichen unzureichend, deren Vokabular u.a. über viele Synonyme verfügt wie z.B. die Biomedizin. Zudem eignen sich diese nicht, um semantische Assoziationen zwischen Konzepten, Patenten, Erfindern und Unternehmen aufzuzeigen. Die Anwendung BioPatentMiner [Mukherjea & Bamba 2004] generiert daher aus dem Inhalt von Patenten und Ontologien ein Semantic Web. Das System besteht aus folgenden Komponenten:

- Parser zum Herunterladen von Patenten von Patentämtern,
- *BioAnnotator* identifiziert und klassifiziert die Terme aus Patenten. Dazu wird die hier Unified Medical Language System (UMLS) genutzt, die medizinische Begriffe und Beziehungen aus unterschiedlichen Sprachen und Disziplinen enthält,
- *SemWebBuilder* generiert das Semantic Web. Folgende Relationen werden abgebildet:

```
<patentA refers_to patentB>  
<inventorC invented patentD>  
<assigneeE assigned patentF>  
<patentG has_term bioTerminH>
```

- *SemWebRetriever* ermöglicht eine stichwortbasierte Recherche.

Während eine einfache Stichwortsuche bei einer Beispielrecherche beim USPTO lediglich zu 29 Treffern führte, ergab die Recherche mit dem BioPatentMiner 196 Treffer, da zusätzliche Verbindungen gefunden wurden.

Zusätzlich können die Ergebnisse auch gerant oder visualisiert werden, um Recherche und Auswertung intuitiver zu gestalten und explorative Suchen zu ermöglichen.

Die derzeitigen Entwicklungen sind rein prototypisch und beschäftigen sich vielfach mit dem Gebiet „Biomedizin“ [siehe auch die Entwicklung von Spyropoulos & Botsivaly 2009]. Eine umfassende Rechercheunterstützung mittels Semantic Web für alle Technikgebiete steht derzeit nicht in Aussicht. Gerade bei Technologien, deren Anwendung sich über unterschiedliche Gebiete erstreckt, müssen mehrere Ontologien integriert (Merging) und aktuell gehalten werden [Mukherjea 2005].

Kollaborative Inhaltserschließung

Hsieh et al. [2009] haben ein Desktop-Tagging System entwickelt, welches die Verständigung über den Inhalt und Bewertung von Patenten erleichtern soll. Die Nutzer können Tags für den Nutzungskontext und die technischen Eigenschaften vergeben und diese werden in einer Matrix abgelegt.

Hierüber ist einsehbar, welche Tags einzelne Nutzer verwendet haben und welche gemeinsam vergeben wurden. Daneben können Tags auch geclustert und visualisiert werden. So können Verbindungen identifiziert und neue Themen aufgespürt werden.

Ein weiteres Beispiel ist ein von Giereth & Ertl [2008] vorgeschlagenes Wiki, welches über folgende Funktionen verfügt:

- Ablagemöglichkeit und Annotierfunktion für Suchanfrage zur erneuten Verwendung (zusätzlich: Möglichkeit zum automatischen Update),
- automatische Generierung von Wiki-Seiten,
- für angemeldete Nutzer: Kommentier-, Bewertungs- und Klassifizierungsfunktion für ganze Patente oder einzelne Passagen,
- Visualisierungsmöglichkeiten.

Bei diesen Anwendungen stellt sich die Frage nach dem Zusatznutzen gegenüber der herkömmlichen Patentverwaltungssoftware. Zudem müssen Aspekte der Datensicherheit und Fragen des Zugriffs berücksichtigt werden [siehe hierzu auch Schmitz 2010, 308-9].

Einsatz von Social Software bei der Patentprüfung

Die Patentämter sind zunehmend überlastet: Allein in den USA hat sich die Anzahl der Anmeldungen seit 2008 verdoppelt. Patentprüfer sind mit einem Rückstand von etwa einer Million Anmeldungen konfrontiert, dem einzelnen Prüfer bleiben lediglich 20 Stunden zur Prüfung einer Anmeldung [Center für Patent Innovations 2008, 3]. Auf Initiative der New York Law School [Noveck 2009, 7ff.] hat das US Patent- und Markenamt USPTO 2007 einen Pilotversuch mit Fokus auf die Bereiche Software und Computer Hardware initiiert, um die fachlich interessierte Öffentlichkeit an der Patentprüfung zu beteiligen. Den überlasteten Prüfern soll damit eine Hilfestellung gegeben und die Qualität der Patentprüfung verbessert werden. Um Patentanmelder für diesen Pilotversuch zu gewinnen, wurde ihnen eine bevorzugte Bearbeitung zugesichert. Zur Kommunikation zwischen Patentprüfern und Öffentlichkeit wurde die Plattform www.peertopatent.org [Peer to Patent 2007] entwickelt. Sie dient in erster Linie dazu, dass angemeldete Nutzer relevante Dokumente zum Stand der Technik einreichen können, die von den Patentprüfern zur Bewertung herangezogen werden. Die Entscheidung, ob ein Patent erteilt werden soll, verbleibt nach wie vor beim Patentprüfer. Der Ablauf gestaltet sich folgendermaßen:

- Die Anmeldungen werden von den Nutzern durchgesehen und kommentiert,
- Nutzer recherchieren Dokumente zum Stand der Technik und laden diese hoch,
- Nutzer kommentieren und bewerten die Einreichungen,
- die zehn besten Dokumente werden dem Patentprüfer übermittelt.

Nutzer können zudem weitere Experten einladen. Diejenigen Nutzer, die einen maßgeblichen Beitrag zur Stand der Technik-Recherche geleistet haben, werden namentlich genannt. Eine Befragung der beteiligten Patentprüfer am Ende des ersten Pilotjahres ergab u.a. [Center für Patent Innovations 2008, 6]:

- 92% würden gerne weiterhin mit Peer Review-Unterstützung arbeiten,
- 24% gaben an, dass die eingereichte Literatur Informationen enthielt, die durch eigene Recherchen nicht gefunden wurden,
- 21% stellten fest, dass die eingereichte Literatur teilweise aus Recherchequellen stammt, die vom USPTO nicht lizenziert wurden.

2008 wurde der Pilotversuch um ein weiteres Jahr verlängert und auf Geschäftsmodelle ausgeweitet [Noveck 2009, 13]. Neben anderen Patentämtern hat Großbritannien 2011 ebenfalls einen Pilotversuch gestartet [Peer to Patent UK 2011].

Generell wird das Rekrutieren von Experten als Schwierigkeit angesehen: Da die Anmeldungen sehr unterschiedlich sind, kann ein Teil nur von einzelnen Experten bewertet werden, was eine Ausweitung des Projekts auf andere Technikgebiete erschwert [Noveck 2009, 101]. Zudem ist der Unterhalt der Plattform mit zusätzlichen Kosten verbunden [Ghafele et al. 2011].

Literatur

- Center für Patent Innovations (2008). Peer to Patent: First Anniversary Report. URL: <http://dotank.nyls.edu/community/patent/P2PAnniversaryreport.pdf> [zuletzt abgerufen am 13.10.2011].
- Ghafele, R., Gibert, B., & DiGiammarino, P. (2011). Driving innovation through patent application review: The power of crowdsourcing prior art search. *Journal of Intellectual Property Rights*, 16(4), 303-308.
- Giereth, M., & Ertl, T. (2008). Visualization enhanced semantic wikis for patent information. *Proceedings of the 12th International Conference on Information Visualisation*, Los Alamitos, California, 185-190.
- Hsieh, W., Situ, J., Chen, Y., & Chou, S. (2009). A collaborative desktop tagging system for group knowledge management based on concept space. *Expert Systems with Application*, 36(5), 9513-9523.
- Mukherjea, S., & Bamba, B. (2004). BioPatentMiner: An information retrieval system for biomedical patents. *Proceedings of the 30th VLDB Conference*, Toronto, Canada, 1066-1077.
- Mukherjea, S. (2005). Information retrieval and knowledge discovery utilizing a biomedical semantic web. *Briefings in Bioinformatics* 6(3), 252-262.
- Noveck, B.S. (2009). Wiki Government. How technology can make government better, democracy stronger, and citizens more powerful. Washington D.C.
- Peer to Patent (2007). <http://peertopatent.org/> [zuletzt abgerufen am 13.10.2011].
- Peer to Patent (UK) (2011). <http://peertopatent.org.uk/> [zuletzt abgerufen am 13.10.2011].
- Schmitz, J. (2010). Patentinformetrie: Analyse und Verdichtung von technischen Schutzrechtsinformationen. Frankfurt a. M.
- Spyropoulos, B., & Botsivaly, M. (2009). Improving the interpretation of patent data by the employment of semantic-web formalism in the specific context of biomedical technology. *IFMBE Proceedings* 25(12), 257-260.