Perspektiven vertrauenswürdiger Aussagen im Semantic Web

Lutz Suhrbier

Freie Universität Berlin, Institut für Informatik

Netzbasierte Informationssysteme

Takustraße 9, D-14195 Berlin, http://nbi.inf.fu-berlin.de

suhrbier@inf.fu-berlin.de

Zusammenfassung

Das Semantic Web ergänzt das World Wide Web durch maschineninterpretierbare Meta-Informationen, welche die Semantik von Inhalten anhand von Aussagen beschreiben. Die Ausarbeitung einer Lösung für das Problem der automatisierten Überprüfbarkeit solcher Aussagen in Bezug auf die Vertrauenswürdigkeit des daraus generierbaren Wissens, ist das Ziel des in diesem Beitrag entwickelten Modells.

1 Einleitung

Das Semantic Web - nach den Ideen von Tim Berners-Lee - ergänzt das World Wide Web durch Hinzufügen maschineninterpretierbarer Informationen. Diese Meta-Informationen beschreiben die Semantik bereitgestellter Inhalte in Form von Aussagen. Die Vision dieser Idee besteht in der Entwicklung autonom agierender Dienste, die aus vernetzten semantischen Aussagen durch Ausführung effizienter Inferenzmechanismen neue, qualitativ hochwertigere Information ableiten [BeHL01].

Von der Semantic Web Initiative des W3C [Mill06] erarbeitete Standards wie das Resource Description Framework (RDF) [MaMi04], dem Resource Description Framework Schema (RDFS) [BrGM04] oder die Web Ontology Language (OWL) [McHa04], sowie darauf aufbauende Werkzeuge stehen zum Erzeugen aussagenbasierter Wissensräume bereit. An deren Umsetzung wird aktuell in verschiedenen Themengebieten mit unterschiedlichen Aspekten gearbeitet. Eine gerade im Hinblick auf potentielle legislative, administrative oder kommerzielle Einsetzbarkeit von Semantic Web Technologien grundlegende Frage blieb bislang jedoch fast unbeachtet: Wie kann die Vertrauenswürdigkeit von Aussagen gewährleistet und, daraus folgend, ein möglichst hoher Korrektheitsgrad der hieraus abgeleiteten Schlussfolgerungen erreicht werden?

Bisherige Ansätze stellten zur Ermittlung der Vertrauensstellung von Aussagen vornehmlich die Anwendung von Semantic Web Technologien in den Vordergrund. Einige Arbeiten setzen Semantic Web Technologien ein, um durch Kombination von Informationen aus sozialen

Netzwerken mit verfügbaren Hintergrundinformationen und unter Berücksichtigung unterschiedlicher Bewertungskriterien die Korrektheit von Informationen zu überprüfen. In [BiOl04] wird z.B. eine Vertrauensarchitektur vorgeschlagen, die auf der Definition vertrauensbasierter Politiken aus einer Kombination aus reputations-, kontext- und inhaltsbasierten Vertrauensmechanismen beruht. Im Rahmen des Projekts Rei wurde eine in diese Richtung wirkende Ontologie zur Definition entsprechender Politiken vorgeschlagen und auch umgesetzt [FiJo02][Kaga05].

Insbesondere bei Reputationssystemen beruht die Bewertung der Glaubwürdigkeit von Aussagen jedoch auf den subjektiven Einschätzungen des Anwenders. Die Vernachlässigung einer unabhängigen, zweifelsfreien Identifikation der Urheber von Aussagen und Informationen beeinträchtigt aber unter Umständen die Korrektheit der hergeleiteten Information. Im Hinblick auf die Vision autonom agierender Dienste, die im Namen ihres Auftraggebers z.B. in der Lage wären, Termine zu vereinbaren, Geräte zu steuern oder gar verbindliche Verträge abzuschließen, ist eine solche Form der Vertrauensbildung jedoch als unzureichend anzusehen.

Mit Techniken wie der Digitalen Signatur und Public Key Infrastrukturen stehen aus dem Gebiet der IT-Sicherheit auch vom Gesetzgeber anerkannte Lösungsansätze bereit. In Berners-Lee's grundlegenden Ideen zum Semantic Web [BeHL01] spielen Signaturen zur Verifikation von Informationen eine Rolle. Ebenso gehören Signaturen als Grundlage der Vertrauensschicht zum Schichtenmodell der W3C Semantic Web Activity [KoMi02].

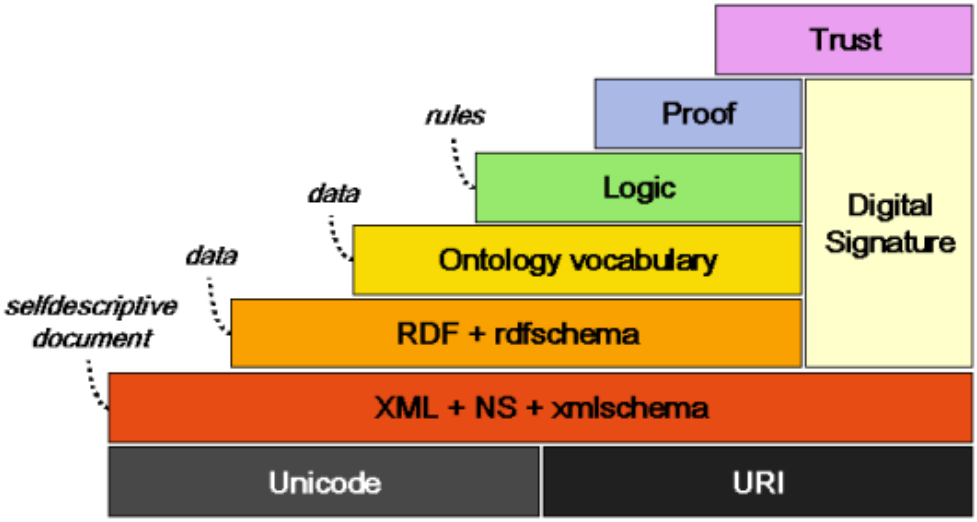

Abbildung 1 Semantic Web Schichtenmodell [KoMi02]

Andere Arbeiten integrieren Digitale Signaturen in ihre Entwürfe, z.B. um die traditionelle Autorisation anhand von Semantic Web Technologien umzusetzen. Hierbei werden z.B. Authentisierungsprotokolle zur Integration externer Agenten in semantische Plattformen [ZhKa01] oder Ontologien zur Definition und Durchsetzung von Zugriffspolitiken.entwickelt [UBJ+03][Denk02] entwickelt. Das in [CBHS05]vorgestellte Semantic Web Publishing (SWP) Vokabular verlagert den Anwendungsbereich der Vertrauensschicht von einzelnen Diensten hin zu einem global verknüpfbaren Netz von verteilten Aussagen. SWP unterstützt

den Einsatz traditioneller Sicherheitsmechanismen, wie Signaturen und X.509-Zertifikate, um diese mit Zusicherungen über die Glaubwürdigkeit von Aussagen zu verknüpfen.

Ein Nachteil der Integration klassischer Sicherheitsmechanismen ins Semantic Web besteht jedoch im Mangel einer nahtlosen semantischen Einbettung in ein entsprechend vorbereitetes semantisches Umfeld. Ziel dieser Einbettung ist es, die Verifikation der Urheber von Aussagen so mit dem Kern des Semantic Web zu verbinden, dass diese bei Bedarf von jeder sicherheits- oder vertrauenssensitiven Anwendung des Semantic Web integriert werden kann. Letzteres soll nach Möglichkeit erfolgen, ohne das semantische Umfeld zu verlassen und ohne verschiedene nicht semantische Sicherheitsmechanismen adaptieren zu müssen.

Daher schlägt dieser Beitrag die Entwicklung einer semantischen Sicherheitsarchitektur zur Bereitstellung vertrauenswürdiger Aussagen im Semantic Web vor. Der Beitrag beschreibt zunächst grundlegende Anforderungen und Konzepte an vertrauenswürdige Aussagen. Anschließend wird deren Abbildung im Semantic Web erläutert. Danach wird der Entwurf einer Sicherheitsinfrastruktur vorgestellt. Schließlich werden Perspektiven für den Einsatz vertrauenswürdiger Aussagen anhand von Beispielszenarien aufgezeigt.

2 Anforderungen und Konzepte

Vor einer Beschreibung der Abbildung vertrauenswürdiger Aussagen im Semantic Web bedarf es zunächst einmal einer Definition der Eigenschaften dieser Aussagen. Jede vertrauenswürdige Aussage muss folgende sicherheitstechnisch relevanten Anforderungen erfüllen:

- Die Aussage besitzt einen Autor
- Autoren sind zweifelsfrei identifizierbar
- Autoren und Aussagen sind eindeutig einander zuzuordnen
- Die Integrität (Fälschungssicherheit) dieser Zuordnung ist sichergestellt
- Die Nichtabstreitbarkeit von Aussagen durch Autoren ist sichergestellt

Ausgehend von der Grundannahme, dass jede Aussage einen menschlichen oder maschinellen Autor besitzt, ist als Voraussetzung für die Vertrauenswürdigkeit von Aussagen die eindeutige Identifizierbarkeit dieses Autors erforderlich. Es sind demnach Bedingungen herzustellen, welche es erlauben, die Authentizität eines Autors jederzeit zu verifizieren. Zusätzlich ist es unabdingbar, Autoren und Aussagen so eindeutig miteinander zu verknüpfen, dass Aussagen eines Autors weder durch Dritte verfälscht, noch vom Autor geleugnet werden können. Selbstverständlich müssen diese Eigenschaften jederzeit überprüfbar sein.

Durch die Verifizierbarkeit der Verknüpfungen zwischen Autoren und Aussagen erlangen Letztere eine höhere Wertigkeit als solche, deren Urheber nicht zweifelsfrei festzustellen ist. Schon aus soziologischen Erwägungen heraus ist zu erwarten, dass die Bereitschaft vorsätzlich falsche Informationen zu veröffentlichen abnimmt, wenn man als Urheber dieser Information leicht erkennbar ist. Zudem können Autoren mit einer hohen Frequenz an Falschaussagen schon mit relativ einfachen Methoden wie Blacklists von der Verarbeitung z.B. in einem Inferenzprozess ausgeschlossen werden. Von der damit einhergehenden Verkleinerung des Suchraums ist außerdem eine signifikante Effizienzsteigerung bezüglich der Informationsauswertung zu erwarten.

Mit der Definition folgender Qualifizierungskonzepte, inklusive der Qualifizierung als Autor, lassen sich vertrauenswürdige Aussagen als sichere Basis für Rating- und Reputationssysteme einsetzen:

- „Aussage X hat Autor Y“ bzw. „Y ist Autor von Aussage X“
- „Aussage X wird (nicht) bestätigt durch Y“ bzw. „Y bestätigt Aussage X (nicht)“
- „Aussage X wird (amtlich) beglaubigt durch Y“ bzw. „Y beglaubigt (amtlich) Aussage X“

Mit dem Qualifizierungskonzept „hat Autor“ bestätigt ein Autor, dass er der Urheber einer bestimmten Aussage ist.

Das Konzept der Bestätigung erlaubt es Dritten, die Korrektheit beliebiger Aussagen zu bestätigen. Der Bestätigende wird dadurch ebenfalls zum Autor, nämlich zum Autor einer positiven oder negativen Bestätigung. Die Anzahl beziehungsweise das Verhältnis von positiven zu negativen Bestätigungen gestattet nun eine quantifizierbare Bewertung der Korrektheit von Aussagen. Diese Bewertung kann als Entscheidungsgrundlage für die Berücksichtigung von Aussagen z.B. in Such- und Inferenzprozessen herangezogen werden (z.B. akzeptiere alle Aussagen mit mehr als 90% positiven Bestätigungen). Durch die Entwicklung darauf aufbauender Algorithmen ist von einer weiteren Effizienz- und Qualitätssteigerung von Bewertungsmechanismen auszugehen.

Über das Konzept der Beglaubigung können besonders anerkannte und entsprechend gekennzeichnete Entitäten unantastbare, „atomare“ Aussagen treffen. Die Integration von Beglaubigungen in der Informationsgewinnung kann beim Aufbau semantischer Wissensnetze z.B. in Firmen, Organisationen oder im wissenschaftlichen Bereich zu wertvollen Performancesteigerungen führen. Bezüglich der amtlichen Beglaubigung ist hier in erster Linie an den Gesetzgeber gedacht, der z.B. Gesetze oder Richtlinien in Kraft setzen sowie in bestimmten Bereichen Befugnisse an öffentliche Einrichtungen delegieren kann. So könnte, z.B. über entsprechende Beglaubigungen im Hochschulrahmengesetz, der Gesetzgeber an Universitäten das Recht delegieren, beglaubigte Aussagen über ihren Lehrbetrieb (z.B. Fachbereiche, Professoren, Lehrveranstaltungen) zu veröffentlichen. Aussagen eines, durch verifizierbare Beglaubigungen ausgewiesenen, Universitätsprofessors bezüglich seines Fachgebiets könnten somit zur Bewertung fachspezifischer Aussagen eine besondere Gewichtung erhalten. Insbesondere dienen Beglaubigungen aber als Instrument um „Massenbestätigungen“ in grundlegenden oder sensiblen Bereichen zu verhindern. Somit würde z.B. verhindert, dass die „Masse“ der Steuerzahler durch millionenfache Bestätigungen die staatlich festgelegten Steuergesetze „überstimmt“ und z.B. Steuersätze selber festlegt.

3 Abbildung im Semantic Web

Das Semantic Web ist eng mit dem Resource Description Framework (RDF) verknüpft [MaMi04], welches das Modell zur Präsentation von Metadaten spezifiziert. Metadaten beinhalten in diesem Modell zusätzliche semantische Informationen über beliebige Ressourcen. Diese Informationen werden in Form von Aussagen abgelegt, die in so genannten „Statements“ in Form von Tripel bestehend aus Subjekt, Prädikat und Objekt dargestellt werden. Der Autor eines Dokuments kann beispielsweise folgendermaßen dargestellt werden: Das Dokument „http://www.example.org/index.html (Subjekt) hat einen Autor (Prädikat) mit dem Namen Lutz Suhrbier (Objekt)”.

Die Bestandteile einer Aussage werden durch Uniform Resource Identifier (URI) gekennzeichnet. Aussagenteile können somit gleichsam als Subjekt weiterer Aussagen dienen. Die durch Hinzufügen weiterer Aussagen entstehenden Aussagegruppen werden als Graphen abgebildet. Diese Graphen bestehen aus Knoten (Subjekte und Objekte) und Kanten (Prädikate) und sind ebenfalls miteinander verknüpfbar. Vokabulare hingegen bestehen aus einer Menge von Prädikaten und Subjekten, die mit Hilfe von z.B. RDF Schema (RDFS) oder der Web Ontology Language (OWL) Taxonomien oder Ontologien bestimmter Themengebiete formal beschreiben [BrGM04][McHa04].

Um vertrauenswürdige Aussagen möglichst eng an den Semantic Web Kern zu binden, werden die oben beschriebenen autorenbezogenen Konzepte im RDF-Format abgebildet. Hierbei ist die Entwicklung eines entsprechenden RDF-Vokabulars in RDF Schema erforderlich, um vertrauenswürdige Aussagen als optionale Erweiterung beliebiger Aussagen(-gruppen) auf einer grundlegenden Ebene des Semantic Web Schichtenmodells anzusiedeln (siehe Abbildung 1).

Für die Spezifikation eines geeigneten RDF-Vokabulars für vertrauenswürdige Aussagen stehen insbesondere folgende Probleme zur Lösung an:

- Eindeutige, persistente Referenz von Aussagen über URIs
- Zweifelsfreie Authentisierung von Autoren (RDF-Zertifikate)

RDF empfiehlt Reifikation als Mechanismus zur Referenz von Aussagen in RDF. Dieser Mechanismus ist jedoch relativ kompliziert in der Anwendung und birgt zudem weitere bekannte Probleme [CaSt04]. Mit Named Graphs [CBHS05] existiert jedoch ein einfach anzuwendender Mechanismus zur URI-basierten Kennzeichnung von RDF-Graphen beliebiger Größe in RDF. Zudem werden Named Graphs von SPARQL, der aktuellen W3C Query Language, unterstützt [PrSe05]. Das W3C führt Named Graphs als Teil der Semantic Web Interest Group [Carr04].

Die Eindeutigkeit von Aussagen und damit auch ihrer Referenz kann durch die Integration des Fingerabdrucks (Digest) des jeweiligen Aussagentripels in den URI-Bezeichner des Named Graph Mechanismus gewährleistet werden. Algorithmen zur Berechnung von Digests von RDF Graphen beschreiben die Arbeiten von Sayers und Carroll [SaKa04][Carr03]. Die Idee ist nun, diese Digests von RDF-Graphen in eine adäquate, URI-basierte Form zu bringen, so dass sie Aussagen nicht nur eindeutig beschreiben, sondern diese gleichzeitig eindeutig kennzeichnen und somit als global eindeutiger Bezeichner für Suchanfragen dienen. Im Bereich von Peer-to-peer-Netzwerken werden ähnliche Mechanismen bereits eingesetzt und sind unter dem Begriff „Content Hash“ bekannt [LiKR04].

Darüber hinaus legt die enge Verknüpfung zwischen RDF und URIs die Definition eines Schemas für Uniform Resource Names (URN) [Moat97] oder Persistent Uniform Resource Locators (PURL) [SWJF06] nahe. Die Verfügbarkeit persistenter und kompatibler Bezeichner für Named Graphs zusammen mit der Integration entsprechender Verzeichnisdienste in bestehende Such- und Inferenzmechanismen des Semantic Web eröffnet nicht nur für die langfristige Referenz vertrauenswürdiger Aussagen vielfältige Perspektiven.

Im Bereich der „traditionellen“ IT-Sicherheit gewährleisten üblicherweise X.509-Zertifkate und Public Key Infrastrukturen die Authentizität von Entitäten [IETF05]. Übertragen auf die oben definierten autorenbezogene Konzepte für vertrauenswürdige Aussagen, ist die Spezifi-

kation von RDF-Zertifikaten zunächst einmal gleichbedeutend mit der Anwendung des autorenbezogenen Konzepts der (amtlichen) Beglaubigung auf

- Metadaten einer Entität
- Öffentlichen Schlüssel dieser Entität
- Attributen zur Kennzeichnung der Eigenschaft (amtlicher) Beglaubigungen

Hinsichtlich der Gleichwertigkeit gegenüber bestehenden Infrastrukturen sollten RDF-Zertifikate weitere Informationen beinhalten, die den Attributen der etablierten X.509-Zertifikate entsprechen. Das Ausstellen dieser Beglaubigung erfolgt durch das Erzeugen einer digitalen RDF-Signatur über entitätsbezogene Metadaten. Aus der Verallgemeinerung dieses Spezialfalls ergibt sich, dass alle autorenbezogenen Konzepte RDF-Signaturen über die von Ihnen bescheinigten Metadaten in Form von Aussagen(-gruppen) erzeugen müssen, um deren Echtheit zu dokumentieren.

Analog zur Erzeugung „traditioneller“ digitaler Signaturen, umfasst eine RDF-Signatur mindestens folgende Informationen:

- Referenz (Fingerabdruck) über
  - die signierten Aussagentripel
  - die Identität des Signierenden und
  - den verwendeten Signaturalgorithmus
- Signaturdaten

Methoden zum Signieren von RDF-Graphen beschreiben die Arbeiten von Carroll [Carr03], Sayers [SaKa04] und Tummarello [TMPP05]. Zur Darstellung der Identität des Signierenden dienen seine oben erwähnten Metadaten. Aus dem Gebiet der „klassischen“ IT-Sicherheit stehen zur Erzeugung der Signaturdaten hinreichend bewährte Verfahren zur Verfügung (siehe [IETF05]). Die Aufgabe besteht daher in der Formulierung eines geeigneten Vokabulars zur Abbildung von RDF-Signaturen.

Die Verfügbarkeit einer Spezifikation von RDF-Signaturen stellt wiederum die Basis zur Formulierung eines Vokabulars für RDF-Zertifikate dar. Dies folgt aus der obigen Erläuterung, nach der RDF-Zertifikate einen Spezialfall der Anwendung von RDF-Signaturen darstellen.

Eine sicherheitstechnisch einwandfreie Umsetzung der Vokabulare RDF-Zertifikat und RDF-Signatur vorausgesetzt, genügen diese den Anforderungen an Vokabulare für vertrauensbasierte Aussagen bezüglich der zweifelsfreien Identifikation von Autoren bzw. der eindeutigen Zuordnung von Aussagen und Autoren. Sie sind demnach als Grundlage zur Spezifikation eines Vokabulars für vertrauenswürdige Aussagen anwendbar. Zusammen mit der eindeutigen, persistenten und URI-basierten Referenz von Aussagen stehen diese als grundlegende Konzepte zur Spezifikation einer Sicherheitsarchitektur für das Semantic Web bereit.

4 Entwurf einer Sicherheitsinfrastruktur

Im Bereich der „traditionellen“ IT-Sicherheit erfordern kryptographische Techniken wie Zertifikate oder Signaturen im Allgemeinen den Aufbau einer entsprechenden Public-Key-Infrastruktur (PKI). Eine solche Infrastruktur stellt Dienste bereit, die die Ausstellung, Veröffentlichung und Prüfung digitaler Zertifikate durchführen. Grundlegende Dienste, wie sie z.B.

Certification Authorities (CAs), Registration Authorities (RAs) und Verzeichnisdienste in Bezug auf Zertifikate erbringen, sind auch in einer Semantic Web Sicherheitsinfrastruktur zum Ausstellen von RDF-Zertifikaten unentbehrlich.

Weitgehenden Veränderungen unterliegen jedoch Dienste zur Verteilung und Prüfung von RDF-Zertifikaten. Hinsichtlich der Verteilung von RDF-Zertifikaten würde es zunächst genügen, wenn CAs bzw. RAs RDF-Graphen mit allen von ihnen ausgestellten Zertifikaten veröffentlichten. Liegt als Referenz bereits ein eindeutiger, persistenter URI (siehe Abschnitt 3) des gefragten RDF-Zertifikats vor, so könnten z.B. auf Named Graphs basierende Verzeichnisdienste schnell den zugehörigen RDF-Graphen oder einen Verweis auf diesen liefern. Ein solcher Verzeichnisdienst könnte z.B. als Semantic Web Service implementiert sein, der Methoden zur Ermittlung von RDF-Graphen auf Basis eindeutiger, persistenter URIs anbietet.

Durch die Konzentration solcher Suchanfragen auf entsprechend spezialisierte und performante Verzeichnisdienste ist eine bedeutende Effizienzsteigerung für den Anwender zu erwarten. Im Gegenzug wird der allgemeine Datenverkehr von einer Vielzahl unnötiger Suchanfragen entlastet, mit denen die benötigten Informationen andernfalls zusammengetragen werden müssten. Weiterhin tragen effektive Suchdienste dazu bei, die zu erwartende Vielzahl redundanter lokaler Speicherungen von RDF-Graphen deutlich zu reduzieren. Dies wäre dann z.B. nur für besonders zeitkritische Anwendungen oder Caches erforderlich. Schließlich ließe sich ein solcher Dienst nahtlos in bestehende Semantic Web Anwendungen integrieren und könnte generell im Semantic Web zur Effizienzsteigerung genutzt werden.

Dennoch erscheint es allein aus Performancegründen sinnvoll, die Veröffentlichung und Prüfung von RDF-Zertifikaten in einem spezialisierten Verzeichnisdienst für RDF-Zertifikate zu kanalisieren. Dieser bietet dafür zusätzliche Methoden, z.B. für Suchanfragen nach bestimmten Zertifikatsinhabern, den Status ausgestellter RDF-Zertifikate oder zur Verifikation autorenbasierter Konzepte anhand signierter RDF-Aussagen. Ein solcher Dienst könnte ebenfalls als Semantic Web Service etabliert werden und Suchanfragen, z.B. unter Verwendung der aktuellen Suchsprache des Semantic Web, der SPARQL RDF Query Language [PrSe05], entgegennehmen. Die Beantwortung der Anfragen erfolgt in Form von RDF-Graphen aus dem internen RDF-Zertifikatsspeicher. Dieses Verfahren hat den Vorteil, dass es nahtlos in bestehende semantische Umgebungen integrierbar ist. Die Entwicklung spezieller Adapter, z.B. für externe LDAP-, OCSP- oder XKMS-Anfragen [WaHk97] [MAM+99] [HaMy05], und damit aufwändige Transformationen vom traditionellen ins semantische Umfeld könnten entfallen. Außerdem erlauben solche semantischen Suchanfragen kombinierte Anfragen, die sich automatisch aus dem internen RDF-Zertifikatsgraphen des Dienstanbieters beantworten lassen. Mit einer Suchanfrage könnte man z.B. nur gültige, aber auch abgelaufene Zertifikate bestimmter Entitäten ermitteln. Weitere Anwendungsmöglichkeiten bestehen in der Verifikation vertrauenswürdiger Aussagen, die in Kooperation mit dem oben skizzierten (einfachen) Verzeichnisdienst auf Basis eindeutiger, persistenter URIs implementierbar sind.

Zusammenfassend besteht der Entwurf einer Semantic Web Sicherheitsinfrastruktur aus folgenden Komponenten:

- CA bzw. RA zur Erstellung von Zertifikaten
- Verzeichnisdienst für Named Graphs auf Basis eindeutiger, persistenter URIs
- Verzeichnisdienst für RDF-Zertifikate
- Verifikationsdienst für RDF-Signaturen (autorenbasierte Konzepte)

Abbildung 2 stellt die Zusammenarbeit und die Abhängigkeiten der Komponenten einer Semantic Web Sicherheitsarchitektur im Zusammenhang mit der fiktiven Bewertung vertrauenswürdiger Aussagen durch eine Inferenzmaschine dar.

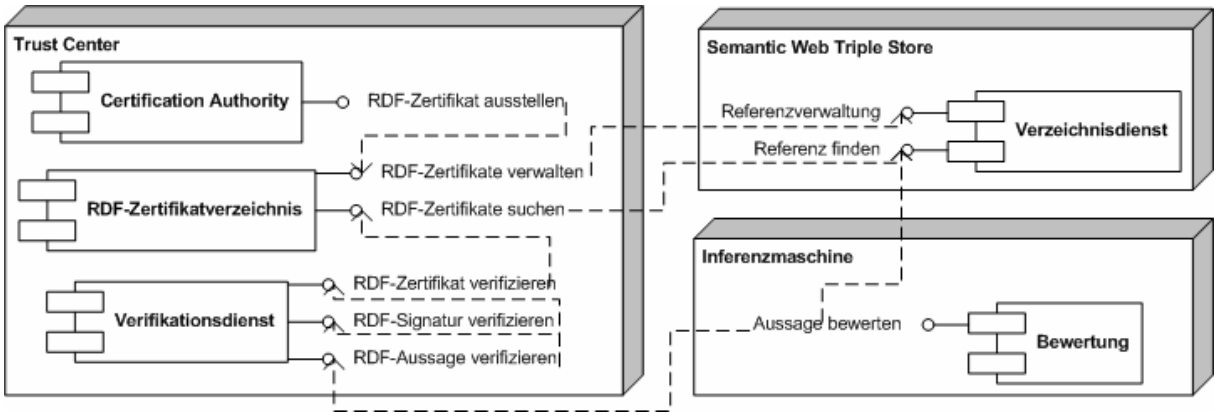

Abbildung 2 Komponenten der Semantic Web Sicherheitsarchitektur

5 Perspektiven

Vertrauenswürdige Aussagen beeinflussen das Semantic Web in mehrfacher Hinsicht. Zunächst einmal ergibt sich durch die Absicherung der Quelle per RDF-Signatur eine neue Qualität bezüglich der Glaubwürdigkeit bzw. Korrektheit von Aussagen. Die Herkunft einer Aussage manifestiert sich nun nicht mehr ausschließlich an der (leicht fälschbaren) IP-Adresse oder URL eines semantischen Dienstes, sondern direkt an der digitalen Signatur ihres Urhebers. Diese Eigenschaft ermöglicht es z.B. Reputationssystemen sensible, vertrauensbasierte Bewertungen auf einer deutlich solideren Grundlage vorzunehmen.

Weiterhin gestatten autorenbezogene Konzepte eine Qualitäts- bzw. Performancesteigerung semantischer Such- und Inferenzmechanismen im Allgemeinen. Durch die Bevorzugung vertrauenswürdiger Aussagen ist nicht nur eine deutliche Verkleinerung des Suchraums zu erwarten, sondern auch ein deutlicher Qualitätsgewinn. Dies gründet sich insbesondere auf der Erwartung, dass die Bereitschaft von Autoren (bewusst) ungenaue oder falsche Informationen zu veröffentlichen deutlich abnimmt, wenn die Quelle zweifelsfrei zurückzuverfolgen ist. Gleichzeitig steigt natürlich auch das Risiko für Autoren für die Bereitstellung von Meta-Informationen zur Rechenschaft gezogen zu werden. Dies birgt letztlich die Gefahr, dass sich deutlich weniger Autoren bereit finden könnten, Meta-Information bereitzustellen. Eine seriöse Abschätzung kann zum jetzigen Zeitpunkt allerdings nicht gegeben werden.

Zudem erlaubt die Berücksichtigung autorenbezogener Konzepte, wie Bestätigung oder (amtlicher) Beglaubigung, eine weitere Einschränkung des Suchraums. Je nach Anwendungsgebiet kann die verwendete Bewertungspolitik z.B. durch variable Gewichtung der verschiedenen autorenbezogenen Konzepte für die jeweiligen Anforderungen optimiert werden. Während für einfache Suchanfragen die Existenz einer ausreichenden Anzahl von Bestätigungen für eine Aussage genügen würde, erfordern finanzielle Transaktionen zumindest beglaubigte Informationen über Kontodaten und Kontoinhaber.

Ein interessantes Szenario stellt z.B. ein semantisches Wiki dar. Die Bevorzugung vertrauenswürdiger Beiträge trägt zu einer weitgehenden Automatisierung der Moderation bei. Sie hilft den Arbeitsaufwand des Moderators zu reduzieren, indem das System eine Vorauswahl von Beiträgen nach Qualitätskriterien anbietet oder bei Vorliegen bestimmter Merkmale sogar selbständig hinzufügt.

Nicht zuletzt entsteht durch das Konzept der amtlichen Beglaubigung die Möglichkeit, eine Art „atomarer“ Aussagen in einem unüberschaubaren Angebot an Informationen zu strukturieren. Hiermit können z.B. staatliche oder öffentliche Einrichtungen ermächtigt werden, Aussagen mit hoheitlichem Charakter zu kennzeichnen. Als Beispiele solcher Aussagen kommen Gesetze, öffentliche Bekanntmachungen, Gerichtsentscheide oder die besondere Auszeichnung öffentlicher Einrichtungen in Betracht. Dieses Verfahren könnte ebenfalls auf die internen Strukturen privater Organisationen (z.B. Firmen, internationale Konzerne) übertragen werden. Auf wissenschaftlichem Gebiet wäre sicherlich der Aufbau abgesicherter, fachspezifischer Informationssysteme (z.B. durch semantische WiKis) interessant, die analog zu Fachkonferenzen Beiträge in Form von Aussagen akzeptieren und die sich anhand autorenbezogener Konzepte wie Bestätigung und Beglaubigung weitgehend selbständig organisieren.

6 Fazit und Ausblick

Es ist nicht zu leugnen, dass die Entwicklung des Semantic Web sich noch im Anfangsstadium befindet. Aber genau dieser Umstand bietet die Gelegenheit, wirkungsvolle Sicherheitsmechanismen von Beginn an als wesentlichen Bestandteil im Kern des Semantic Web zu verankern. Begreift man frei nach Bruce Schneier „Sicherheit als Prozess“, so erreicht man damit, dass Sicherheitsaspekte den weiteren Entwicklungsprozess des Semantic Web stärker beeinflussen.

Vertrauenswürdige Aussagen eröffnen im Kontext des Semantic Web eine neue Qualität der Wissenserschließung und Wissensrepräsentation. Die in Abschnitt 5 erläuterten Perspektiven zeigen sicher nur einen kleinen Teil möglicher Anwendungen auf.

Die Problemstellungen dieses Beitrags sind gleichzeitig auch Gegenstand der Untersuchungen im Rahmen meiner Dissertationsarbeit auf diesem Gebiet. In der nächsten Zukunft sollen zunächst die Grundlagen für die Anwendung der in diesem Beitrag skizzierten Konzepte anhand der Entwicklung und Implementierung der Komponenten geschaffen werden. Der Nachweis der Tragfähigkeit soll nach Möglichkeit durch Integration dieser Architektur in ein geeignetes, semantisches Projektumfeld erfolgen.

Literatur

[BeHL01] Tim Berners-Lee, James Hendler, Ora Lassila: The Semantic Web. In: Scientific American (2001)

[BiOl04] Christian Bizer, Radoslaw Oldakowski: Using Context- and Content-Based Trust Policies on the Semantic Web. In: Proceedings of the Thirteenth International World Wide Web Conference (WWW2004) (2004) 228-229

[BrGM04] Dan Brickley, R.V. Guha, Brian McBride: RDF Vocabulary Description Language 1.0: RDF Schema - W3C Recommendation 10 February 2004. In:

http://www.w3.org/TR/rdf-schema/, gesehen am 10.Januar 2006.

[CBHS05] Jeremy Carroll, Christian Bizer, Patrick Hayes, Patrick Stickler: Named Graphs, Provenance and Trust. In: Proceedings of the 14th international conference on World Wide Web (2005) 613-622.

[CaSt04] Jeremy J. Carroll, Patrick Stickler: TriX : RDF Triples in XML. In: HPLabs Technical Reports, HPL-2004-56 (2004)

[Carr03] Jeremy J. Carroll: Signing RDF Graphs. In: Lecture notes in computer science, 2870, Springer (2003) 369-384.

[Carr04] Jeremy Carroll: Named Graphs. In: http://www.w3.org/2004/03/trix/, gesehen am 07.Januar 2006.

[Denk02] Grit Denker: Towards Security in DAML. In: SRI International Internal Report (2002)

[FiJo02] Tim Finin: Anupam Joshi: Agents, Trust, and Information Access on the Semantic Web. In: ACM SIGMOD, 31 (4) (2002) 30-35.

[Hamy05] Phillip Hallam-Baker, Shivaram H. Mysore: XML Key Management Specification XKMS 2.0 - W3C Recommendation 28 June 2005. In: http://www.w3.org/TR/xkms2/, gesehen am 10 January 2006.

[IETF05] IETF Secretariat: Public-Key Infrastructure (X.509) (pkix). In: http://www.ietf.org/html.charters/pkix-charter.html, gesehen am 09.September 2005.

[Kaga05] Lalana Kagal: Rei Ontology Specifications, Ver 2.0. In: http://www.csee.umbc.edu/%7Elkagal1/rei/, gesehen am 04.Januar 2006.

[KoMi02] Marja-Ritta Koivunen, Eric Miller: W3C Semantic Web Activity. In: Helsinki Institute for Information Technology (HIIT): Semantic Web Kick-Off in Finland - Vision, Technologies, Research, and Applications, HIIT Publications (2002) 27-43.

[LiKR04] Jian Liang, Rakesh Kumar, Keith W. Ross: Understanding Kazaa. In: Polytechnic University, Department of Computer and Information Science (2004)

[MAM+99] M. Myers, R. Ankney, A. Malpani, S. Galperin, C. Adams: X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP. In: IETF: Request for Comments 2560 (1999)

[MaMi04] Frank Manola, Eric Miller: RDF Primer - W3C Recommendation 10 February 2004. In: http://www.w3.org/TR/rdf-primer/, gesehen am 10.Januar 2006.

[McHa04] Deborah L. McGuinness, Frank van Harmelen: OWL Web Ontology Language Overview - W3C Recommendation 10 February 2004. In: <http://www.w3.org/TR/2004/REC-owl-features-20040210/>, gesehen am 10.Januar 2006.

[Mill06] Eric Miller: Semantic Web Activity Statement. In: http://www.w3.org/2001/sw/Activity, gesehen am 10.Januar 2006.

[Moat97] R. Moats: URN Syntax. In: IETF Network Working Group: Request for Com-

ments (2141) (1997)

[PrSe05] Eric Prud'hommeaux, Andy Seaborne: SPARQL Query Language for RDF - W3C Working Draft 23 November 2005. In: http://www.w3.org/TR/rdf-sparql-query/, gesehen am 07.Januar 2006.

[SWJF06] Keith Shafer, Stuart Weibel, Erik Jul, Jon Fausey: Introduction to Persistent Uniform Resource Locators. In: http://purl.oclc.org/docs/inet96.html, gesehen am 09.Januar 2006.

[SaKa04] Craig Sayers, Alan H. Karp: RDF Graph Digest Techniques and Potential Applications. In: HPLabs Technical Reports, HPL-2004-95 (2004)

[TMPP05] Giovanni Tummarello, Christian Morbidoni, Paolo Puliti, Francesco Piazza: Signing individual fragments of an RDF graph. In: Proceedings of the Fourteenth International World Wide Web Conference (WWW2005) (2005)

[UBJ+03] A. Uszok, J. Bradshaw, R. Jeffers, N. Suri, P. Hayes, M. Breedy, L. Bunch, M. Johnson, S. Kulkarni, J. Lott: KAoS Policy and Domain Services: Toward a Description-Logic Approach to Policy Representation, Deconfliction and Enforcement. In: Proceedings of the IEEE Workshop on Policy 2003 (2003)

[WaHk97] M Wahl, T. Howes, S. Kille: Lightweight Directory Access Protocol (v3). In: IETF: Request for Comments 2251 (1997)

[ZhKa01] Min ZHANG, Ahmed KARMOUCH: Adding Security Features to FIPA Agent Platforms. In: School of Information Technology & Engineering (SITE), Multimedia & Mobile Agent Research Laboratory (2001)