# Privatsphäre im Internet – es gibt sie wirklich!

## WIE NEUESTE TECHNOLOGIEN DER INFORMATIONSVERARBEITUNG FÜR SICHERHEIT SORGEN KÖNNEN

**Die Popularität des Internets wächst mit jeder Minute und ein Ende ist nicht in Sicht. Neue Dienste und Soziale Netzwerke wie Facebook, Ebay, YouTube oder Skype bieten besonders jungen Menschen eine Plattform, um Freundschaften zu schließen und Informationen auszutauschen. Viele Nutzer sind sich dabei nicht bewusst, dass genau wie im realen Leben im Internet Gefahren lauern und schon bei kleinen Unachtsamkeiten schnell erhebliche persönliche Schäden entstehen können. Eine Wissenschaftlerin und zwei Wissenschaftler vom Forschungszentrum L3S erläutern, wie Anwender im sicheren Umgang mit privaten Informationen unterstützt werden können.**

Obwohl es kaum denkbar ist, dass jemand private Informationen wie das Geburtsdatum oder die Kreditkartennummer einer unbekannten Person auf der Straße mitteilt, werden dennoch im Internet solche Daten tagtäglich von verschiedenen Webseiten abgefragt und oft öffentlich zugänglich gemacht. Nach der Bekanntgabe solcher Informationen hat der Benutzer weder Einfluss darauf, wo und wie lange sie gespeichert werden, noch Informationen darüber, wer Zugang zu den Daten hat. Hinzu kommt, dass es für die meisten Dienste tatsächlich gar nicht nötig ist, wirklich alle Informationen zu erfragen – ein Teil der Benutzerdaten genügt meist, um eine bestimmte Transaktion ausführen zu können.

Das Forschungszentrum L3S entwickelt in verschiedenen Projekten Konzepte und Softwaremodule, die es dem Benutzer ermöglichen, Informationen gezielt und kontrolliert im Internet auszutauschen. Gleichzeitig wird dabei die Vertrauenswürdigkeit des jeweiligen Dienstes analysiert und überprüft. So wird es dem Anwender leicht gemacht zu entscheiden, ob bestimmte Daten zur Verfügung gestellt werden sollen oder nicht.

### Techniken der Künstlichen Intelligenz für die Sicherheit

Die Idee, dass Computer selbständig und intelligent agieren, entstand bereits in den 50er Jahren und führte zu der Begründung des Forschungsgebietes »Künstliche Intelligenz«. Über die Jahre wurden aus den ursprünglich sehr ambitionierten Visionen konkrete Techniken, die Informationssystemen helfen, automatisch Entscheidungen zu treffen. Heutzutage unterstützen uns solche Entscheidungen in diversen Bereichen des alltäglichen Lebens, sei es die Digitalkamera, die entscheidet, ob eine abgebildete Person lächelt oder nicht, oder

ein Navigationssystem, das – basierend auf Verkehrssituation, Staumeldungen und persönliche Präferenzen des Fahrers – eine Routenempfehlung gibt. Solche Mechanismen sind natürlich auch für sicherheitsrelevante Entscheidungen interessant: Sind die Zertifikate, die ein Online-Shop liefert, so vertrauenswürdig, dass ich meine Kreditkartendaten ohne Bedenken übermitteln kann? Dürfen beliebige Benutzer auf ein bestimmtes Foto in meinem Facebook-Account zugreifen? Werden meine Daten vertrauenswürdig gespeichert? Im Forschungszentrum L3S wird unter anderem nach den Antworten auf solche Fragen geforscht.

### Policy-Sprachen und automatisierte Verhandlungen über Vertrauen

Zum Beispiel werden hier formale Regelsprachen (sogenannte Policy-Sprachen) entwickelt, mit deren Hilfe Sicherheitsregeln für Systeme definiert und relevante und gesicherte Informationen für Entscheidungen zusammengestellt werden können. Diese Regeln werden von bestimmten Systemmodulen, sogenannten Softwareagenten, interpretiert, die letztendlich das Verhalten des ganzen Systems beeinflussen.


Hochsensible Daten können nach diesen Regeln zwischen mehreren Softwareagenten schrittweise ausgetauscht werden und die Softwareagenten lernen sich dabei langsam kennen. Der Kennenlernprozess findet dabei, ähnlich wie im wirklichen Leben, zwischen zwei sich unbekannten Personen statt; mit jedem Datenaustausch wird das Vertrauen stufenweise aufgebaut. So kann der Benutzer sicher im Netz stöbern, einkaufen und darin andere Leute treffen. Die Maschine alarmiert, falls zu viele persönliche Daten vom Internetsurfer verlangt werden, und übernimmt hierfür die technischen Einzelheiten wie zum Beispiel die Zertifikatsüberprüfung.

### Webdienste: Identifikation der Sicherheitsrisiken für den Benutzer

Doch auch die beste Technik ist wenig hilfreich, wenn mögliche Gefahren von Benutzern einfach ignoriert werden. Insbesondere der jüngeren Generation scheint die Wichtigkeit der Privatsphäre im Internet nicht bewusst zu sein. Für eine Untersuchung der Verbraucherzentrale Schleswig-Holstein im März dieses Jahres wurden 5.564 Schülerinnen und Schüler an 26 Schulen des Landes zu ihrem Internetverhalten befragt. 90 Prozent der befragten Jugendlichen sind in sozialen Netzwerken vertreten, 33 Prozent haben öffentliche Profile. Persönliche Daten und Bilder würden dabei häufig bedenkenlos ins Netz gestellt. 80 Prozent der Befragten gaben an, Fotos ins Netz gestellt zu haben, die sie ihren Lehrern oder Eltern nicht zeigen würden. Dabei wird in unserem Zeitalter das Internet häufig als eine Art Datenbank für die Personensuche benutzt und das nicht nur von Privatpersonen. Zum Beispiel beauftragen immer mehr Arbeitgeber spezielle Suchagenturen, die Informationen über die Bewerber sammeln. Damit kann der erste Eindruck vom Bewerber noch vor dem Bewerbungsgespräch entstehen, falls es, nach der Informationsauswertung, überhaupt noch dazu kommen sollte.

Die Forschungsarbeit im L3S lässt heute schon erkennen, dass eine automatische Identifikation privater Bilder möglich ist und der Bedarf an Systemen, die den Benutzer bei der Wahl der Sicherheitsbeschränkung für seine Daten unterstützen können, gedeckt werden kann. Die Entscheidungen korrelieren häufig mit spezifischen Bildmerkmalen und können somit automatisiert werden. Solche Merkmale sind zum Beispiel die im Internet populären Textdaten wie Titel, Beschreibung und die Schlüsselwörter, die an die Bilder »angehängt« werden. Die geeignete Vorgehensweise für die Klassifikation sowie die Auswahl der Merkmale sind dabei der Schlüssel zum Erfolg. Der Benutzer kann damit rechtzeitig auf mögliche Sicherheitsrisiken aufmerksam gemacht werden und entsprechend reagieren. Die vom L3S durchgeführte Studie zeigt, dass Wörter wie **Lächeln**, **Kind**, **Familie**, **Porträt** weitaus häufiger mit privaten Bildern assoziiert werden als **Himmel**, **Weg**, **Fluss** und **Natur**. Auch Gesichter und Menschenmengen können automatisch erkannt werden. Es spielt eine Rolle, ob das Bild im Haus oder außerhalb gemacht worden ist, ob ein Blitz benutzt wurde, wie hell das Bild ist und welche Farben dominieren. Die automatische Erkennung der privaten Ressourcen erreicht bereits eine Genauigkeit über 80 Prozent und weitere Verbesserungsschritte folgen.

Abbildung 1
*Das Forschungszentrum L3S steht für grundlagen- und anwendungsorientierte Forschung im Bereich Web Science.*
*Quelle: www.L3S.de*

Abbildung 2
*Das Vertrauen zwischen zwei Softwareagenten wird stufenweise aufgebaut, ähnlich wie es im wirklichen Leben zwischen zwei sich unbekannten Personen passiert.*
*Quelle: www.fotolia.de*

Abbildung 3
*Besonders für junge Menschen ist der Umgang mit sozialen Netzwerken zu einer Selbstverständlichkeit geworden.*
*Quelle: www.L3S.de*

### Vertrauenswürdige Datenspeicherung

Wenn man im Bereich der digitalen Privatsphäre arbeitet, stellt sich heraus, dass es nur wenige Daten gibt, die wirklich privat sind und nur von einem Benutzer verwendet werden.

Abbildung 4
*Der Benutzer möchte die Kontrolle über seine privaten Daten jederzeit behalten.*
*Quelle: www.fotolia.de*

**Dipl.-Volksw. Gabriele Herrmann-Krotz**
Jahrgang 1960, ist seit 2009 Geschäftsführerin des Forschungszentrums L3S der Leibniz Universität Hannover.
Kontakt: herrmann@L3S.de

**M. Sc. Dipl.-Inf. (FH) Sergej Zerr**
Jahrgang 1977, arbeitet seit 2008 als wissenschaftlicher Mitarbeiter am Forschungszentrum L3S der Leibniz Universität Hannover.
Kontakt: zerr@L3S.de

**Dipl.-Inf. Philipp Kärger**
Jahrgang 1981, ist seit 2006 als wissenschaftlicher Mitarbeiter am Forschungszentrum L3S der Leibniz Universität Hannover tätig.
Kontakt: kaerger@L3S.de

Zum Briefverkehr gehören mindestens zwei Personen, und von dem Wunsch, private Fotos und Videos der Welt zu zeigen, wissen wir spätestens seitdem Flickr und YouTube gegründet worden sind. Das Tagebuch würde hier ein Beispiel für Privatinformation darstellen, wären da nicht die Blogs … Auch private Information möchte demnach kontrolliert geteilt werden. Um dies zu ermöglichen, müssen die Daten in irgendeiner Form veröffentlicht werden. Der sogenannte »Index« ist eine kompakte Datenstruktur, die zum schnellen und präzisen Informationsgewinn in einer großen Datenmenge bei den Suchmaschinen verwendet wird. Es ist eine Art digitale Übersichtskarte, mit deren Hilfe die Dokumente zu den Suchbegriffen gefunden werden können. Leider kann man auch dieser Karte eine große Menge privater Informationen entziehen. Der Index befindet sich auf einem Server der Suchmaschine und der Benutzer hat keinerlei Kontrolle über die Daten. Für das Forschungsteam des L3S ist dies eine weitere Herausforderung.

Ein wahrscheinlichkeitsbasiertes Verfahren wurde entwickelt, um die Daten so zu speichern, dass der rechtmäßige Gebrauch möglich ist. Im Falle einer Entwendung sind die Daten für den Angreifer dann nutzlos. Um dies zu ermöglichen, werden die Begriffe, die in den Dokumenten vorkommen, im Index zusammengefasst. So wird nur eine Zeichenfolge sowohl für den Namen »Günter Mustermann« als auch für Begriffe wie »Milch«, oder »Sonne« gespeichert. Auf diese Weise ist es für den Angreifer nicht mehr möglich, ohne weiteres festzustellen, ob sich die Dokumente zu einem bestimmten Suchbegriff in der Datenmenge befinden und falls ja, welcher Person genau sie zugeordnet sind. Eine berechtigte Person jedoch kann mit Hilfe ihres Schlüssels die richtige Information wiedergewinnen. So ein Index ist sowohl effizient als auch kontrollierbar sicher.

## Fazit

Es gibt eine große Menge von Anwendungen, die von der Forschung im Bereich Datensicherheit profitieren können. Immer mehr Menschen entdecken für sich tagtäglich soziale Internetdienste und die Möglichkeiten, immer mehr Daten zu veröffentlichen. Diese Entwicklung ist so rasant, dass die Sicherheitsforschung aufrüsten muss, um den Bedarf des Marktes zu decken. Auch erfahrene Benutzer brauchen eine technische Unterstützung sowohl beim Austausch als auch bei der Veröffentlichung und Speicherung ihrer Daten. Das L3S Expertenteam stellt sich dieser Herausforderung und leistet einen signifikanten Forschungsbeitrag für eine sichere Internetzukunft.

## Weiterführende Informationen

- Abel, F.; Marenzi, I.; Nejdl, W.; Zerr, S.; Sharing Distributed Resources in Learn-Web2.0 – European Conference on Technology Enhanced Learning (EC-TEL 2009)
- Statistische Erfassung zum Internetverhalten Jugendlicher und Heranwachsender; http://www.verbraucherzentrale-sh.de/mediabig/109191A.pdf
- Kärger, P.; Olmedilla, D.; Passant, A; Polleres, A; Proceedings of the Second Workshop on Trust and Privacy on the Social and Semantic Web – (CEUR 2010)
- Zerr, S.; Demidova, E.; Chernov, S.; deskweb2.0: Combining Desktop and Social Search – Desktop2010 Workshop at 33d Special Interest Group on Information Retrieval (SIGIR 2010)
- Zerr, S.; Nejdl, W; Olmedilla, D.; Siberski,W.; Zerber+R: Top-k Retrieval from a Confidential Index – 12th International Conference on Extending Database Technology (EDBT 2009)