Kein Nachdruck, keine Veröffentlichung im Internet oder einem Intranet ohne Genehmigung des Verlags und der GDNÄ

Warum das digitale Forschungslabor IT-Sicherheitsforschung braucht

Kreative Angriffe und wie man sie abwehrt

Die Digitalisierung hält auch in Forschungsinstituten und -laboren Einzug. Angefangen von computergesteuerten Gerätschaften zum Durchführen von Experimenten und zum Austausch und der Auswertung der gesammelten Daten, bis hin zu Cloud-Lösungen, sind IT-Systeme inzwischen in alle Bereiche der Forschung involviert. Neben klassischen Angriffen auf diese IT-Systeme spielen dabei Privatsphäre und subtilere Angriffe eine immer größere Rolle. In seinem Vortrag thematisierte Professor Michael Backes unter anderem, wie die Privatsphäre in medizinischen Studien unter Erhaltung der Nutzbarkeit der Daten gewährleistet werden kann und welche Möglichkeiten das Maschinelle Lernen bringt. Zum Abschluss ging er auf einige der ausgefallensten, hochkreativen Angriffe auf Labore und medizinische Geräte ein, wie etwa Angriffe auf DNA-Sequencer durch spezielle DNA-Proben.

Prof. Dr. Dr. h. c. **Michael Backes** (Jahrgang 1978) studierte nach Ableisten des Zivildienstes Informatik an der Universität des Saarlandes und schloss sein Studium 2001 mit dem Diplom ab. Das Mathematikdiplom und die Promotion folgten 2002. Nach Forschungen am IBM Research Forschungslabor Zürich wurde er 2005 auf den Lehrstuhl für Informationssicherheit und Kryptographie an der Universität des Saarlandes berufen.
Er ist Gründungsdirektor und Vorsitzender der Geschäftsführung des CISPA Helmholtz Center for Information Security. Seine aktuellen Forschungsschwerpunkte sind die Entwicklung von Methoden zum Schutz persönlicher Daten, insbesondere in der Medizin; die Erforschung vertrauenswürdiger Methoden zum Maschinellen Lernen sowie universelle Lösungen in der Software- und Systemsicherheit. Herr Backes wurde wiederholt ausgezeichnet, u.a. mit dem ERC Synergy Grant, dem höchstdotierten EU-Forschungspreis. Er ist Mitglied in der Akademie der Technikwissenschaften acatech und ist Ehrendoktor der Université de Lorraine, die ihn im Oktober 2018 für seine Leistungen im Bereich der Cybersicherheit und für die hervorragende deutsch-französische Zusammenarbeit zwischen CISPA und der Université Lorraine (Nancy, Metz) würdigte.

CISPA Helmholtz Center for Information Security, Stuhlsatzenhaus 5, 66123 Saarbrücken

Backes zeigte in seinem Vortrag einige Kehrseiten der Digitalisierung, denn Hackerangriffe auf Autos oder Krankenhäuser sind schon längst Realität. Er nahm das Publikum zunächst auf eine Zeitreise in die Entwicklung des Internets und der sich parallel dazu entwickelnden Cybersicherheit mit: Während es 1980 noch kein Internet gab, wurde das langsame Internet 1990 fast nur für E-Mails genutzt. Datenschutz war in dieser Phase kein großes Problem, da kaum personenbezogene Daten übermittelt wurden. Nutzer gingen davon aus, die volle Kontrolle über ihre Daten zu haben. Das hat sich grundlegend geändert. Heute bietet das nutzerzentrierte Internet ein riesiges Ausmaß an veröffentlichten personenbezogenen Informationen. Das Internet ist zu einer Plattform für die Kommunikation zwischen Millionen von Nutzern geworden. Es ist ein globales Phänomen, das alle Lebensbereiche erfasst (Abb. 1).

„Wir bieten heute als Nutzer eine viel größere Angriffsfläche für Datenmissbrauch als noch vor 15 Jahren. Denn wer weiß schon, was eine installierte App von Drittanbietern wirklich im Hintergrund mit unseren Daten macht? Das Thema Vertrauen ist heute anders gelagert“, betonte Prof. Backes. *„Der Datenschutz ist heute ebenfalls nicht mehr in dem Maße wie früher gegeben. Zu viele persönliche Daten kursieren im Internet, von denen wir nicht wissen, wie sie genutzt werden, was mit ihnen passiert. Dafür haben wir den Überblick und Durchblick verloren in unserer hochgradig vernetzten Welt.“*

Das ginge selbst Fachleuten so, räumte er ein, und darum stelle sich die Frage, welche Sicherheitsgarantien die IT-Forscher heute überhaupt geben können. Er schätzt die Lage kritisch ein:

„Uns fehlen Grundlagen. Es fehlt uns das fundamentale Verständnis von vielen Dingen – wir verstehen noch nicht fundamental, wie man Datenschutz in einer hochgradig vernetzten

©Wissenschaftliche Verlagsgesellschaft Stuttgart

Kein Nachdruck, keine Veröffentlichung im Internet oder einem Intranet ohne Genehmigung des Verlags und der GDNÄ

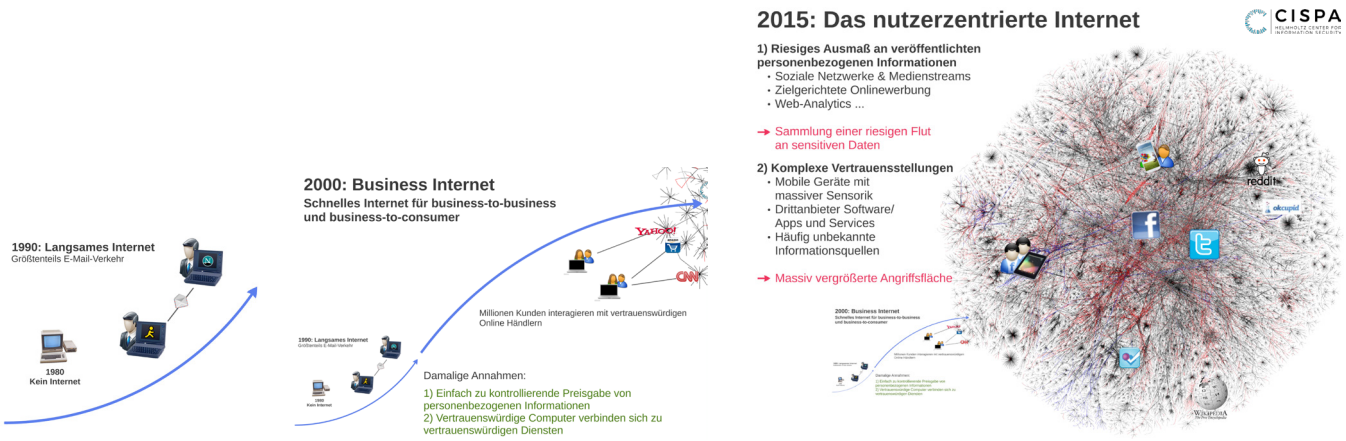

Abb. 1. Vom langsamen Internet zur heutigen Situation: Wir hinterlassen immer mehr Spuren im Internet und werden dadurch immer leichter angreifbar.

Welt wirklich sicherstellen kann. Wir verstehen noch nicht, wie man Lösungen baut, die von Menschen sicher genutzt werden können.“ Bestes Beispiel seien Passwörter: *„Ein Passwort soll so komplex sein, dass es keiner erraten kann, aber man soll es sich merken – das ist nicht wirklich benutzerfreundlich.“*

Immerhin, so beruhigte er, haben wir ein tiefes Verständnis von der Kryptographie – sie sei in der Praxis nie das schwächste Glied – ein solcher Stand sei auch für die anderen Bereiche anzustreben. Einen Grund für die unbefriedigende Situation sieht Backes in einem inhärenten Wettrüsten. Forscher und Gesetzgeber reagieren auf Fallbasis, sind getrieben von Erfordernissen und Risiken aktueller Technologien. Sie bieten zwar Vorschläge zum Beseitigen von Symptomen der Sicherheitslücken, reagieren aber immer nur auf das, was unmittelbar bevorsteht. Aber die Grundursachen werden oftmals nicht verstanden. *„Wir brauchen disruptive Veränderungen, wir brauchen Durchbrüche auf einer ganz tiefen Ebene“*, mahnte Backes an. *„Cybersicherheitsforschung ist eine umfassende, höchst anspruchsvolle Querschnittstechnologie ähnlich wie die KI-Forschung. Es gibt unglaublich viele Arten, wie wir angegriffen werden können.“* Backes beschränkte sich im Folgenden auf drei Bereiche, die für die Sicherheitsforschung besonders relevant sind:

Schutz der Privatsphäre im Internet

Selbst eine Anonymisierung von Daten führe nicht dazu, im Netz unerkannt zu bleiben. Durch Datenabgleich und Profilbildung in großen Datensätzen lassen sich sehr wohl Rückschlüsse ziehen, die am Ende eine Identifizierung von anonymisierten oder pseudonymisierten Personen ermöglichen. Backes führte hierzu einige interessante Beispiele an: Etwa bei als belanglos eingeschätzten Bewertungen von Filmen auf verschiedenen Filmportalen. Bei der Datenbank *imdb* sind diese mit Klarnamen verbunden, bei der Datenbank *Netflix* anonymisiert. Dennoch konnten Rückschlüsse bei *Netflix* über die Bewertungen bei *imdb* gezeigt und die Personen hinter den Pseudonymen eindeutig identifiziert werden. Was hier harmlos ist, bekommt eine andere Bedeutung, wenn sich dieselbe Person etwa wegen gesundheitlicher Probleme im Internet austauscht und dabei Spuren hinterlässt, die eine Verknüpfung mit personalisierten Daten erlauben.

Selbst bei aller Vorsicht bei Besuchen im Internet können intelligente Suchmaschinen und Algorithmen vieles über anonymisierte Personen herausfinden, etwa durch den Abgleich von direkten Profildaten, indirekten Profilen von Freunden, mit denen man vernetzt ist, semantischen Daten (welche Suchbegriffe gebe ich bei Suchmaschinen wie Google ein?) und statistischen Daten (wer sucht nach einer speziellen Krankheit und hat gleichzeitig Konzertkarten für eine extravagante Künstlerin im Internet bestellt?). *„Im Internet ist man immer gläsern, unsere Suchhistorie im Internet macht uns auffindbar und trotz Anonymisierung individualisierbar. Alles, was Sie machen, kann in einem bestimmten Kontext personenbezogen sein“*, warnte Backes eindringlich. Da kein Mensch die Verknüpfungsmöglichkeiten voraussehen könne, arbeitet man am CISPA an Systemen, die alle möglichen Szenarien durchspielen, um den Benutzer vor möglichen Gefahren zu warnen, ehe er eine Entscheidung per Mausklick fällt, die er bereuen könnte.

Datenschutz in Gesundheitsdatenbanken

Dass Datenschutz im medizinischen Bereich besonders relevant ist, steht außer Frage. Entsprechend vorsichtig geht man mit Genomdaten um. Schwieriger ist aber der Umgang mit epigenetischen Daten. Während das Genom anzeigt, was ein Organismus *prinzipiell kann*, verraten epigenetische Daten etwas über die *aktuelle Konstitution*. Es handelt sich eigentlich um noch interessantere Daten, deren Brisanz aber von den Medizinern sehr unterschiedlich eingeschätzt wurde, wie Backes am Beispiel der mikro-RNA-Daten und der Daten zur DNA-Methylierung ausführte.

Die mikro-RNA ist an der aktuellen Umsetzung der in der DNA gespeicherten Information beteiligt und ihr Gehalt bzw. ihre Spezifität ist starken Schwankungen unterworfen. Man glaubte daher, die hohe Variabilität mache es unmöglich, eine Profilbildung durchzuführen, da sich schon nach wenigen Wochen ein ganz anderes Bild ergäbe. Dies erwies sich als grundlegend falsch: Man kann über ein Jahr hinweg Profilbildung betreiben. Wer im Besitz dieser Daten ist, hat damit Wissen über den Gesundheitszustand einer Person. Zum Glück gibt es einen Ansatz (genannt Differential Privacy), bei dem man sich davon löst, dass alle Anfragen an die Daten hundertprozen-

©Wissenschaftliche Verlagsgesellschaft Stuttgart

Kein Nachdruck, keine Veröffentlichung im Internet oder einem Intranet ohne Genehmigung des Verlags und der GDNÄ

tig korrekt sein müssen. Man legt quasi einen Rauschfilter über die Daten. Damit ist die Antwort nur „fast korrekt", aber für Biomediziner benutzbar. Sie ist allerdings nicht ausreichend, um ein Individuum herauszukristallisieren. Dies ist nach Meinung von Backes ein Beispiel dafür, wie man durch eine fundamental andere Vorgehensweise für Datensicherheit sorgen kann.

Bei der DNA-Methylierung, die zu einer länger anhaltenden Modifizierung der Genaktivität führt, war man von Anfang an vorsichtig. Daten über DNA-Methylierungen erlauben wertvolle Rückschlüsse auf die DNA und den Gesundheitszustand und sind für die personalisierte Medizin wichtig. Um aber eine Diagnose stellen zu können, ist man auf hochgenaue Daten angewiesen, die man daher nicht einfach „verrauschen" lassen kann. Eine Lösung besteht darin, die Daten so aufzubereiten, dass sie nur die behandlungsrelevanten Informationen liefern, nach denen die Algorithmen fragen. Auch hier ist dann eine Profilbildung nicht möglich.

„Sie geben also nur das preis, was die andere Seite braucht, um Sie perfekt behandeln zu können, alles andere behalten Sie für sich. Das ist kompliziert, aber es ist möglich, und in diese Richtung gehen wir bei der personalisierten Medizin und beim Datenschutz", sagte Backes.

Autonome Systeme

Während uns die Gefahr, über das Internet ausgespäht oder manipuliert zu werden bewusst ist, aber meist wenig berührt, fühlen wir uns unwohl, wenn wir unser Schicksal autonomen Systemen anvertrauen sollen.

Maschinelles Lernen erfolgt ähnlich wie bei einem Kind, das lernt, mit der Umgebung klar zu kommen. Es lernt Dinge, vergleicht diese, verfeinert das Erlernte durch Erfahrung und formt auf diese Weise Wissen.

Zwar sind die Systeme gut, aber wir sollten uns von der Intuition verabschieden, dass sie genau das lernen, was sie lernen sollen. Ein Problem sei auch, wie beispielsweise ein System bei einem selbstfahrenden Auto auf eine neue Situation reagiert, die es vorher nicht gelernt hat, z. B. ein Geisterfahrer oder ein Kind, das plötzlich auf die Straße läuft. Auch Cyberangriffe auf autonomes Fahren werden durch „bösartige" Apps immer wahrscheinlicher; diese könnten entweder die Kontrolle über das Fahrzeug übernehmen oder ganze Systeme ausfallen lassen.

Backes machte deutlich, dass das Thema Cybersicherheitsforschung für Deutschland höchste Bedeutung hat, denn die deutsche Industrie ist durch den Trend zu autonomen Systemen im Kern betroffen, sowohl hinsichtlich der Chancen, aber auch der Risiken. Anlass zur Panik sieht er allerdings nicht. Es komme gerade in diesem Bereich nicht darauf an, der erste auf dem Markt zu sein, sondern der erste, der das Versprechen einlöst, sichere und verlässliche datenschutzfreundliche autonome Systeme anzubieten.

Zur Reduzierung der Risiken in diesem Feld hat sich das CISPA Helmholtz Center in Saarbrücken laut Backes eine Mission gegeben: *„Unser langfristiges und grundlegendes Ziel ist, substantielle Fortschritte im Wettrüsten zu erzielen, indem ganze Klassen von Angriffen ausgeschlossen werden. Wir wollen die Kernursachen aktueller und zukünftiger Probleme verstehen, präzise Grundlagen für IT-Sicherheit- und Privacy-Forschung entwickeln und innovative anwendungsorientierte Forschung mit disruptiver Grundlagenforschung verzahnen."*

Seitenkanalangriffe

Welche außergewöhnlichen Angriffe auf Systeme und Geräte beim Datenklau heute schon möglich sind, illustrierte Backes abschließend an einigen Beispielen. James Bond kann dabei sogar noch etwas lernen. So sei es möglich, aus den Tischvibrationen oder auch aus den Geräuschen, die beim Tippen auf Tastaturen entstehen, mit einem empfindlichen Smartphone, das danebenliegt, herauszufiltern, was gerade mit der Tastatur geschrieben wurde. Anhand der von feinen Luftbewegungen hervorgerufenen Blattbewegungen von Zimmerpflanzen gelinge es sogar, Gespräche in einem Zimmer zu „belauschen". Ebenso können Spiegelungen von Bildschirmen auf danebenstehenden Tassen, Kaffeekannen oder Trinkflaschen aus bis zu 20 Metern Entfernung mit guten Teleskopen erkannt werden. Man könne also quasi um die Ecke schauen, ohne den Bildschirm direkt zusehen. Geradezu abenteuerlich mutete es an, dass es mit synthetischer DNA sogar möglich sei, in DNA-Sequenzierautomaten einzudringen und über diese Schadsoftware auf Rechnern zu installieren. – Dem neuen Helmholtz-Zentrum, dem 19. seiner Art, wird die Arbeit mit Sicherheit nicht ausgehen.

Abb. 2. Das neue CISPA Helmholtz Center for Information Security. Als Großforschungseinrichtung des Bundes soll es mit bis zu 1000 Mitarbeitern die Voraussetzungen für einen umfassenden Datenschutz und die Sicherheit cybertechnischer Systeme schaffen. CISPA steht für Center for IT Security, Privacy and Accountability und war der Name eines 2011 gegründeten BMBF-Kompetenzzentrums, aus dem das neue Helmholtz-Zentrum hervorging. [Photo CISPA]

©Wissenschaftliche Verlagsgesellschaft Stuttgart