Datenschutzrecht und Big Data – ein Anwendungsfall aus der privaten Krankenversicherung

In vielen Wertschöpfungsketten und Prozessen haben Big-Data-Anwendungen in den vergangenen Jahren Einzug gehalten und gewinnen auch im Versicherungswesen immer stärker an Bedeutung. Damit rücken sie zunehmend in den Mittelpunkt öffentlicher Diskussionen über ihre ethischen und datenschutzrechtlichen Grenzen, wie der nachfolgende Text illustriert.

Gerade im Gesundheitsbereich sind die Potenziale von Big-Data-Anwendungen für jeden ersichtlich. Es ist kaum vorstellbar, dass sie technisch nicht genutzt werden. Und es erscheint auf den ersten Blick auch kaum vorstellbar, dass die Menschen mit Hinweis auf den Datenschutz darauf verzichten, die aktuellsten medizinischen und technischen Möglichkeiten zur Behandlung zu nutzen. Doch die Realität sieht anders aus: Sowohl die EU-Datenschutzgrundverordnung (DSGVO) als auch das Bundesdatenschutzgesetz reglementieren strikt, in welchem Umfang Daten verwendet werden dürfen.

Ist Hepatitis-C-Erkrankung vorhersagbar?

Dieses Spannungsfeld zwischen potenziellem Nutzen einer Big-Data-Anwendung und den Datenschutzherausforderungen für einen privaten Krankenversicherer soll das folgende Beispiel veranschaulichen: Am 5. April 2019 hat ein großer deutscher Krankenversicherer die Pressemitteilung „Deutsche Leberstiftung: Mit künstlicher Intelligenz der Hepatitis C auf der Spur“ herausgegeben. Ihr zufolge sind nach Schätzung der Weltgesundheitsorganisation weltweit mindestens 70 Millionen Menschen an einer chronischen Hepatitis C erkrankt. In den letzten Jahren wurden große Fortschritte in der Therapie gemacht, die jedoch nur durchgeführt werden kann, wenn die Krankheit auch erkannt wird. In Deutschland wird die Zahl der mit Hepatitis C Infizierten auf 250.000 geschätzt, viele wissen aber nichts von ihrer Erkrankung.

Das Projekt DETECT hat anonymisierte Gesundheitsdaten von mehr als 1,8 Millionen Versicherten von zwei Krankenversicherungen aus dem Zeitraum von 2009 bis 2014 ausgewertet. Die Krankenversicherungsdaten wurden mithilfe eines KI-Systems, genauer eines neuronalen Netzwerkes, auf Auffälligkeiten für das Vorliegen einer möglichen Hepatitis C analysiert. Dafür wurden in einem ersten Schritt die Datensätze der bekannten Hepatitis-C-Erkrankungen extrahiert und auf versteckte Hinweise auf eine Infektion untersucht. Solche Hinweise können Müdigkeit, Gelenkschmerzen, Schilddrüsenerkrankungen, Depressionen oder Diabetes sein.

Mithilfe des neuronalen Netzes wurde dann folgenden zentralen Fragestellungen nachgegangen: Können die bekannten Hepatitis-C-Diagnosen wiedergefunden werden? Und wie viele Versicherte zeigen auffällige Ähnlichkeiten ihrer Gesundheitsdaten im Vergleich zu den Hepatitis-C-Erkrankten? Das Ergebnis: Mit dem neuronalem Netz wurden alle bekannten Hepatitis-C-Erkrankungen wiedergefunden. Darüber hinaus wurden 2.217 Versicherte aufgrund ähnlicher „sozio-medizinischer“ Daten dem Cluster zugeordnet und hatten damit eine potenziell höhere Wahrscheinlichkeit, an einer noch nicht erkannten Hepatitis C erkrankt zu sein.

Versicherte haben Recht auf Nichtwissen

Für einen Krankenversicherer stellt sich die Frage, was er mit dieser Erkenntnis anfangen kann. Soll er die poten-

ziell betroffenen Versicherten gezielt auf die neuartige Therapie mit guten Heilungschancen ansprechen? Erwarten die Versicherten, dass er aus dem Datenbestand Erkenntnisse herauszieht, die gezielte Diagnosen und Therapien unterstützen? In dem Beispiel ist beides nicht möglich, da die Daten entsprechend der Vorgaben der Ethikkommission anonymisiert werden mussten und keine Rückschlüsse auf konkrete Personen möglich waren. Denn der deutsche Datenschutz sieht das Recht auf Nichtwissen vor, das Ausfluss des allgemeinen Persönlichkeitsrechts ist. Die Datenschutzgesetze konkretisieren dieses Grundrecht, sie sind die rechtlichen Vorgaben für die Versicherer: Ihre Handlungsgrundlage für die Verarbeitung von Daten ist meist die gesetzliche Erlaubnis, Daten zur Erfüllung des Vertrags zu verarbeiten oder eine Einwilligung, die sich zum Beispiel auf die Datenerhebung von Gesundheitsdaten für den Erstattungsprozess bezieht.

Ist auch die Big-Data-Anwendung zur Erkennung von potenziellen Hepatitis-C-Infizierten eine von diesen Erlaubnissen erfasste Datenverarbeitung? Der Versicherte hat im Zweifel seinen Wohnort hinterlegt, damit er vom Versicherer angeschrieben werden kann und die noch in der Bisex-Welt Versicherten haben ihr Geschlecht angegeben, um in einen richtigen Tarif eingestuft zu werden. Die Einwilligung in die Verarbeitung von Gesundheitsdaten umfasst gegebenenfalls nur die Erhebung von Diagnosen zum Zweck der Erstattung. Wenn diese Stammdaten nun mit den Gesundheitsdaten in einer Anwendung verarbeitet werden, um mögliche Erkrankungen zu entdecken, würde eine Zweckänderung vorliegen. Mit jedem neuen Verarbeitungszweck müsste die Einwilligungserklärung angepasst werden. Die neue Einwilligungserklärung müsste auch von allen Bestandsversicherten eingeholt werden, deren Einwilligungserklärung den Zweck der Big-Data-Anwendung noch nicht enthielt. Dies ist ein erheblicher Aufwand, mit wahrscheinlich überschaubarem Erfolg. Denn die Antwortquoten der Versicherten sind in der Regel recht gering. Dabei benötigen aber Big-Data-Anwendungen große Datenmengen, um zu belastbaren Ergebnissen zu kommen.

Ausweg Anonymisierung?

Eine Lösungsmöglichkeit ist die Anonymisierung der Daten. Dadurch entfallen der Personenbezug und die oben genannten Probleme. Der Versicherer kann dann aber nur noch die Versichertengemeinschaft pauschal darauf hinweisen, dass ein Teil von ihr mit einer höheren Wahrscheinlichkeit mit Hepatitis C infiziert ist.

Eine weitere Möglichkeit wäre, auf den ursprünglich verfolgten Zweck zur Datenverarbeitung aufzusetzen und für die anschließende Big-Data-Anwendung eine Zweckänderung nach Maßgabe der Datenschutzgrundverordnung anzunehmen. Die Datenverarbeitung muss dann eng im Datenverarbeitungsprozess an den ursprünglichen Zweck anknüpfen. Diese Betrachtung sollte sehr sorgfältig angestellt werden. Je sensibler die Daten sind, umso enger muss der ursprüngliche Zweck mit dem neuen zusammenliegen. Eine solche Einschätzung ist sicherlich nicht einfach vorzunehmen, zumal die Sanktionen bei einem Verstoß gegen das Datenschutzrecht erheblich sind. Solange es keine Rechtsprechung oder Literatur zu vergleichbaren Sachverhalten gibt, sprechen sich Juristen dafür aus, für die Zweckänderung eine externe Expertise einzuholen.

Dies trifft auch auf eine Datenschutzfolgenabschätzung zu. Laut DSGVO muss vor der Aufnahme des Betriebs eine Abschätzung der Folgen der Big-Data-Anwendung vorgenommen und dokumentiert werden. Den Leitlinien zur Datenschutzfolgenabschätzung zufolge muss dies unter anderem bei der Verarbeitung sensibler Daten oder von großen Datenmengen, der Zusammenführung beziehungsweise der Kombination von Daten erfolgen. Auch der Einsatz neuer Technologien oder biometrische Verfahren erfordern dieses Vorgehen.

Fazit

Gesellschaftliche Diskussion über Datennutzung unerlässlich

Die Chancen für einzelne Versicherte, durch eine gezielte Information des Versicherers von einer schwerwiegenden Erkrankung geheilt zu werden, werden geschmälert, weil die von den Versicherern erhobenen Daten einem besonderen Schutz unterliegen. Der Versicherer darf diese nur zu dem Zweck verwenden, zu dem er sie erhoben hat. Es gibt zwar die Möglichkeit der Zweckänderung, die jedoch sorgfältig angewendet werden muss. Zudem bringt die notwendige Datenschutzfolgenabschätzung einigen Aufwand mit sich.
Schlussendlich belegt dieses Beispiel: In den kommenden Jahren wird es umfassender gesellschaftlicher Diskussionen über den Nutzen und die Risiken von Big-Data-Anwendungen im Gesundheitswesen bedürfen. Diese Debatten werden der Grundstein für die weitere Ausgestaltung unserer Datenschutzgesetze sein, die letztendlich der Ausdruck von Regeln sind, die sich die Gesellschaft gegeben hat und die sie akzeptiert. Die deutschen Aktuare werden diese Diskussionen mit ihrem jahrzehntelangen Fachwissen im Umgang mit sensiblen Versichertendaten aktiv begleiten, um im Interesse der Kunden bestmögliche Lösungsansätze im Spannungsfeld zwischen Datenschutz und Datennutzung zu entwickeln.