

Das DGVFM-Datenbankprojekt – eine zukunftsweisende Kooperation zwischen Versicherungsindustrie und Wissenschaft

von Marco Degginger, Axel Grätz, Marc Hilber, Kai Hoff, Verena Reiter und Matthias Scherer



■ *Statistische Analysen und die Modellierung von Risiken sind seit Jahrzehnten Schlüsseltechnologien im Versicherungswesen und der Umgang mit interessanten Datensätzen gehört für die meisten Aktuarinnen und Aktuarinnen zum Tagesgeschäft. Es ist zu erwarten, dass der Stellenwert datenbezogener Forschung und Entwicklung weiterhin rasant anwachsen wird, getrieben nicht zuletzt durch neue Techniken im Bereich Data Science.*

Als Datengrundlage zur (Weiter-)Entwicklung von unternehmensinternen Methoden und Modellen ist der eigene Versicherungsbestand für Aktuarinnen und Aktuarinnen ganz selbstverständlich vorhanden. In der akademischen Forschung und Lehre hingegen, verfügen Forschende häufig über keinen Zugang zu realen Datensätzen aus der Versicherungswirtschaft. Als Ersatz werden dann veraltete Daten, Datensätze aus anderen Ländern oder auch fiktive Datensätze benutzt – oder die statistische Forschung wird auf andere Gebiete (mit leichterem Zugang zu Daten) fokussiert. In Konsequenz findet wenig empirische Forschung im Versicherungswesen statt – eine Situation, die für alle Beteiligten Verbesserungspotential bietet. Darüber hinaus wird auch in der Lehre die Chance verpasst, Studierenden durch die Arbeit mit spannenden Datensätzen aus der Versicherungswirtschaft einen Einblick in das aktuarielle Berufsbild zu bieten, was wiederum die Sichtbarkeit und Attraktivität der Versicherungsbranche als Arbeitgeber für mathematisch interessierte Studierende erhöhen würde.

Vision: das DGVFM-Datenbankprojekt

Die DGVFM hat es sich angesichts obiger Ausgangslage zum Ziel gesetzt, Daten aus mehreren Gebieten der Versicherungswirtschaft zu sammeln und für die Forschung nutzbar zu machen. Dies ist Motivation und Ziel des Datenbankprojektes des DGVFM-Ausschusses „Forschung und Transfer“. Konkret sollen Datensätze von Unternehmen aus verschiedenen Gebieten der Finanz- und Versicherungsmathematik gesammelt und aggregiert werden, um sie in anonymisierter Form in Forschung und Lehre zu verwenden. Auf dieser Grundlage können innovative statistische Modelle und Methoden entwickelt werden. Für die Forschung ist die Möglichkeit der Validierung und Reproduktion der veröffentlichten Erkenntnisse immens wichtig, was durch eine klar definierte Datengrundlage geboten wäre. Im Zeitalter von Data Science und modernen statistischen Verfahren erscheint solch eine Datengrundlage längst überfällig. Notwendige Voraussetzung für die Erstellung einer Datenbank mit versicherungsspezifischen Daten ist die Kooperation zwischen Wissenschaft und Versicherungsindustrie, denn letztere muss die benötig-

ten Datensätze zur Verfügung stellen. Das Statut und die Kompetenzen der DGVFM bieten eine ideale Plattform, um diese Kooperation erfolgreich zu organisieren. Eine solche Datenbank würde allen Mitgliedern von DAV und DGVFM sowie anderen interessierten Personen zu wissenschaftlichen Zwecken zur Verfügung stehen.

Vorteile für beteiligte Unternehmen und die Versicherungsbranche

Für beteiligte Unternehmen ergeben sich zahlreiche mittelbare und unmittelbare Vorteile. In der Entwicklung können beispielsweise eigene Ansätze und Modelle auf einem weiteren Datensatz validiert werden. Von verschiedenen Versicherern befüllte Datenpools sind insbesondere von großem Nutzen für neue Versicherungsprodukte mit kurzer Schadenshistorie; Cyber stellt ein gutes Beispiel hierfür dar. In der Literatur veröffentlichte Ergebnisse und Modelle

Der Ausbau der Zusammenarbeit zwischen Wirtschaft und Wissenschaft und damit einer Kooperation zwischen Versicherungsunternehmen und akademischen Einrichtungen macht Anwendungsbereiche sichtbar, verbessert den Wissenstransfer und fördert innovative Lösungen.

müssten lediglich auf das eigene Unternehmen angepasst werden, was wiederum den Aufwand für eigene Entwicklungen reduzieren könnte. In einigen Ländern, z. B. der Schweiz und den USA, ist es bereits üblich, dass Unternehmen der Wissenschaft Daten zur Verfügung stellen. Dies hat zur Folge, dass für diese Länder mehr empirische Forschungsergebnisse erzielt und für deren Unternehmen passgenauere Modelle entwickelt werden. Die Barriere, welche in Deutschland durch den fehlenden Datenaustausch von Forschung und Praxis besteht, erweist sich mittelfristig als Nachteil im internationalen Vergleich.

Eine Beteiligung an einer solchen versicherungsmathematischen Datenbank fördert auch die Sichtbarkeit als innovatives Unternehmen, das sich aktiv im Wissenstransfer einsetzt. Damit einher geht eine Sichtbarkeit bei Studierenden der Finanz- und Versicherungsmathematik, welche die Datensätze im Studium nutzen. So könnten Studierende, über die Arbeit mit realen Daten, einen guten Einblick in das aktuarielle Berufsfeld erlangen und idealerweise für eine Tätigkeit in der Versicherungsbranche begeistert werden.

Nicht zuletzt können die Datensätze auch der Weiterbildung der unternehmensinternen Aktuare und Aktuarinnen dienen, indem beispielsweise die DAV-Fachgruppe „Actuarial Data Science“ durch ihre Nutzung realitätsnahe Weiterbildungsveranstaltungen anbieten kann. Wenn Use-Cases aus der Praxis verfügbar wären, könnte auch die DAV-Ausbildung besser auf die konkreten Anforderungen in den Unternehmen angepasst werden.

Vorteile für den Wissenschaftsstandort

Das Fehlen realistischer Daten ist derzeit eines der größten Hindernisse in der angewandten versicherungsmathematischen Forschung. Als Konsequenz fehlender Datensätze findet oft rein theoretische Forschung statt und aus Sicht der Industrie ist diese nicht immer anwendungsrelevant. Alternativ werden immer wieder die gleichen, weil frei verfügbaren, Datensätze als Beispiel herangezogen. Ein Klassiker sind die „Danish Fire Insurance Claims“. Dieser und ähnliche frei verfügbare Datensätze sind oft veraltet, besitzen wenige Kovariablen und sind nicht repräsentativ. Sie eignen sich daher meist nur bedingt für das jeweilige Forschungsziel und führen ferner nicht dazu, dass die versicherungsmathematische Forschung an Attraktivität gewinnt. Forschende, die an Finanz- und Versicherungsmathematik Interesse haben, entscheiden sich oft nicht zuletzt aufgrund der leichten Datenverfügbarkeit im Finanzwesen für diesen Bereich und nicht für das Versicherungswesen.

Einwenden könnte man nun, dass es bereits aktive Kooperationen zwischen einzelnen Unternehmen und Universitäten gibt, welche für eine spezielle Forschungsfrage Zugriff auf firmeninterne Daten erlauben. Solche Kooperationen sind sicherlich ein begrüßenswerter Start, lösen jedoch nicht ein wichtiges Problem. Gute empirische Forschung muss gemäß den FAIR-Prinzipien „Findable, Accessable, Interoperable and Reusable“ verifizierbar und reproduzierbar sein. Dies ist im Rahmen von Verträgen mit einzelnen Forschenden nicht möglich. Forschungsergebnisse müssen auch in Peer-Review-Verfahren verifizierbar sein und andere Wissenschaftler und Wissenschaftlerinnen sollten die Möglichkeit haben, die entwickelten Methoden, Modelle

und Hypothesen zu reproduzieren und zu verbessern. Die Anwendbarkeit der FAIR-Prinzipien ist ein großer Vorteil einer der Wissenschaft allgemein zugänglichen versicherungsmathematischen Datenbank. Damit steigt nicht nur die Qualität der Forschung, sondern auch die internationale Anerkennung der erzielten Ergebnisse. Ohne garantierte FAIR-Prinzipien ist die Veröffentlichung von Forschungs-

Ein gemeinsamer Datenpool sorgt somit nicht nur für eine bessere Datenbasis, sondern verstärkt auch die Attraktivität des Versicherungssektors für empirisch Forschende, was letztendlich die Konkurrenzfähigkeit des Wissenschaftsstandortes Deutschland verbessert.

ergebnissen in renommierten Journalen zunehmend nicht mehr gestattet. Ein gemeinsamer Datenpool sorgt somit nicht nur für eine bessere Datenbasis, sondern verstärkt auch die Attraktivität des Versicherungssektors für empirisch Forschende, was letztendlich die Konkurrenzfähigkeit des Wissenschaftsstandortes Deutschland verbessert.

In Zeiten rückläufiger Studierendenzahlen in der Mathematik, und in Konkurrenz mit anderen MINT-Angeboten (z.B. den aktuell boomenden Data-Science Studiengängen), können Statistik- und IT-affine Studierende nur für ein Fachgebiet bzw. eine Branche begeistert werden, wenn der „Rohstoff“ Daten verfügbar ist. Vielen Firmen aus dem Software- und Internet-Umfeld ist dies längst bewusst und sie versorgen Lehrende und Studierende aktiv mit guten Datensätzen, um so Werbung für ihre Branche zu machen. Durch konkrete Beispiele aus der Praxis steigt auch die Qualität der Ausbildung maßgeblich und Studierende können besser auf eine aktuarielle Tätigkeit vorbereitet werden.

Juristische Klärung der Machbarkeit

Eine Herausforderung des Datenbankprojekts besteht darin, dass bei der Erstellung einer versicherungsmathematischen Datenbank verschiedene rechtliche Anforderungen einzuhalten sind. Für Forschende und Versicherer gleichermaßen ist es essentiell zu klären, welche Art von Daten ein Versicherungsunternehmen (VU) zu wissenschaftlichen Zwecken teilen darf. Zu prüfen ist in diesem Zusammenhang die Datenschutz-Grundverordnung (DSGVO). Die DSGVO findet auf die Verarbeitung personenbezogener, nicht dagegen anonymer Daten, Anwendung. Aus Sicht der Forschung bleiben die Daten nutzbar, auch

wenn der Personenbezug der Datensätze aufgelöst wird. Damit ist entscheidend, wann Daten nach der DSGVO als anonymisiert gelten und infolgedessen ohne datenschutzrechtliche Einschränkungen genutzt und mit der Forschungsöffentlichkeit geteilt werden dürfen.

Dr. Marc Hilber, Marco Degginger und Dr. Axel Grätz von der Kanzlei Oppenhoff & Partner beraten die DGVFM juristisch bei der Gestaltung des Datenbankprojekts und haben die rechtliche Zulässigkeit im Hinblick auf das Datenschutzrecht, aber auch dem strafrechtlichen Geheimnisschutz gemäß § 203 StGB, dem Kartellrecht sowie dem Immaterialgüterrecht, insbesondere nach dem Urheberrechtsgesetz (UrhG) geprüft. Oppenhoff kam zu dem Ergebnis, dass das Datenbankprojekt rechtskonform gestaltet werden kann. Im Einzelnen:

(1) Datenschutzrechtlich ist das Datenbankprojekt nach folgender Maßgabe zulässig:

(i) Die Erstellung der versicherungsmathematischen Datenbank aus den personenbezogenen Originaldaten der VU („Forschungsdaten“) stellt gegenüber der ursprünglichen Erhebung der Versichertendaten eine gemäß Art. 5 Abs. 1 lit. b), 6 Abs. 4 DSGVO zulässige Zweckänderung dar, da bereits dieser Verarbeitungsschritt darauf abzielt, die wissenschaftliche Forschung und eine Verarbeitung der Daten zu statistischen Zwecken zu ermöglichen. Darin liegt gleichzeitig ein berechtigtes Interesse der VU und DGVFM (Art. 6 Abs. 1 lit. f) DSGVO in Verbindung mit dem Forschungsprivileg gemäß Art. 5 Abs. 1 lit. b) und Art. 89 DSGVO).

(ii) Wird der Personenbezug bei der Erzeugung der Forschungsdaten nicht wieder herstellbar aufgehoben, sind diese als anonym anzusehen und unterfallen damit nicht mehr dem Anwendungsbereich der DSGVO. Selbst wenn die Forschungsdaten lediglich pseudonym und damit noch personenbezogen sind, können diese nach Art. 6 Abs. 1 lit. f) DSGVO für statistische und wissenschaftliche Zwecke weiterverarbeitet werden. Zulässig ist auf dieser Grundlage (i) die Übermittlung der Forschungsdaten an die DGVFM als wissenschaftliche Vertretung, (ii) die Bereitstellung der Datenbank samt Forschungsdaten durch die DGVFM an die teilnehmenden Versicherungsunternehmen (VU) oder akademische Forschungseinrichtungen und (iii) deren Nutzung, jeweils im Rahmen des Datenbankprojekts, d. h. zu statistischen und wissenschaftlichen Auswertungszwecken.

(iii) Die teilnehmenden VU trifft gem. Art. 13 DSGVO die Pflicht, die Versicherten über die Datenverarbeitungen zur

Erzeugung der Forschungsdaten zu informieren. Die VU haben die Versicherten in ihren Datenschutzerklärungen über die Verarbeitung ihrer Daten zu statistischen Zwecken zu informieren und müssen ihre Datenschutzerklärung erforderlichenfalls aktualisieren. Es lässt sich gut vertreten, dass eine individuelle Ansprache der Versicherten entbehrlich ist. Weder die VU noch die DGVFM müssen die Betroffenen dagegen über die Datenverarbeitungen im Rahmen der Nutzung der Datenbank informieren, da insofern die Ausnahmeregelung des Art. 14 Abs. 5 lit. b) DSGVO greift.

(iv) Betrachtet man die Forschungsdaten als pseudonym und damit personenbezogenen, müssten die DGVFM und die teilnehmenden VU einerseits einen Auftragsverarbeitungsvertrag (Art. 28 DSGVO) schließen, falls die Erzeugung der Forschungsdaten durch die DGVFM im Auftrag der jeweiligen VU erfolgen soll. Diese Verpflichtung entfällt, wenn die Forschungsdaten anonym sind.

(2) Geheimnisschutz

Daten aus der Kranken-, Lebens- und Unfallversicherung, die Rückschlüsse auf den Versicherten zulassen, sind als Geheimnisse gemäß § 203 Abs. 1 Nr. 7 StGB geschützt und dürfen nicht Dritten gegenüber offengelegt werden. In der ersten Phase des Datenbankprojekts werden daher ausschließlich Daten aus dem Bereich Schaden genutzt. Wege zur Nutzung von Daten aus der Kranken-, Lebens- und Unfallversicherung werden in der zweiten Phase intensiv juristisch geprüft. Um strafrechtliche Risiken bei der Umsetzung des Datenbankprojekts auszuschließen, müssen diese Daten gesichert anonym sein.

(3) Kartellrecht

Die Anforderungen an einen kartellrechtlich zulässigen Informationsaustausch können erfüllt werden. Die Daten werden von der DGVFM als neutrale Stelle gesammelt und verwaltet. Außerdem wird die Datenbank über Daten von mindestens fünf Versicherungsunternehmen verfügen, die zudem nur so eingestellt werden sollen, dass eine Zuordnung von Datensätzen zu einzelnen Versicherungsunternehmen unmöglich ist. Sollten über die Datenbank allerdings auch sensible Preisdaten wie aktuelle und zu-



Mehr zum Datenbankprojekt erfahren Sie auch auf der Jahrestagung von DAV/DGVFM, unter anderem im Programm von ASTIN. Lernen Sie uns und unsere Partner – die Sozietät Oppenhoff & Partner und den Stifterverband – kennen und überzeugen Sie sich selbst von der Wichtigkeit und der guten Umsetzbarkeit unserer Initiative.

künftige Prämienbestandteile, Provisionshöhen, Rabatte oder Zuschläge zugänglich gemacht werden, empfehlen Hilber, Degginger und Grätz zusätzlich zu einer Anonymisierung auch eine Aggregation jedenfalls dieser Art von Daten.

(4) Immaterialgüterrecht

An den jeweiligen einzelnen Forschungsdatenbeständen der VU und an der von der DGVFM gesammelten Gesamtmenge der Forschungsdaten bestehen wahrscheinlich

Das positive Ergebnis, dass das Datenbankprojekt rechtskonform ausgestaltet werden kann, ist ein entscheidender Schritt auf dem Weg zur Umsetzung.

gem. § 87a UrhG (Datenbankrecht) Ausschließlichkeitsrechte, da die Strukturierung und Aufbereitung der Daten eine wesentliche Investition erfordert. Die VU sollten der DGVFM daher (vorsorglich) das nicht-ausschließliche Recht einräumen, die jeweils bereitgestellten Forschungsdaten im Rahmen der Datenbank vereinbarungsgemäß zu nutzen, d. h. für die Zwecke statistischer und wissenschaftlicher Auswertungen teilnehmender VU und Forschungseinrichtungen bereitzustellen.

Das positive Ergebnis, dass das Datenbankprojekt rechtskonform ausgestaltet werden kann, ist ein entscheidender Schritt auf dem Weg zur Umsetzung. Auf Basis des Rechtsvermerks und der gewonnenen Erkenntnisse können sowohl die Initiatoren der Datenbank als auch die Versicherungsunternehmen auf einer rechtssicheren Basis weitere Schritte klären.

Stiffterverband und Forschungsdatengesetz

Mit der Erstellung einer wissenschaftsöffentlichen Datenbank aus Unternehmensdaten leistet die DGVFM Pionierarbeit in vielen Bereichen, wie auch bereits die Notwendigkeit für die aufwändige juristische Klärung verdeutlicht. Für Projekte dieser Art gibt es bisher kaum etablierte Beispiele, besonders im Hinblick auf rechtliche Vorgänge, die durch ihre Transparenz und Zugänglichkeit als Orientierungshilfe dienen könnten.

Die DGVFM engagiert sich seit 2023 in der Initiative „Datagroup Business 2 Science“ des Stiffterverbandes für die Deutsche Wissenschaft, um durch deren Netzwerk Partner und Unterstützung für ihr Datenbankprojekt zu

erhalten. Die Initiative des Stiffterverbandes zielt darauf ab, den Datenaustausch zwischen Wirtschaft und Wissenschaft zu verbessern. Durch die Vernetzung von Akteuren aus verschiedenen Sektoren fördert sie das Verständnis der aktuellen Situation, macht Herausforderungen sichtbar und bietet anhand praktischer Beispiele Lösungsansätze an. Das Projekt der DGVFM, eine versicherungs- und finanzmathematische Datenbank aufzubauen, dient als Pilotprojekt innerhalb dieser Initiative und kann als Orientierung für Fachgesellschaften, wissenschaftliche Institutionen und Unternehmen dienen. Zudem leistet die Datagroup einen Beitrag zu datenpolitischen Diskussionen, wie der Debatte um das Forschungsdatengesetz, unterstreicht dabei die Wichtigkeit der Datenverfügbarkeit für Forschung und Innovation und unterstützt mit konkreten Handreichungen die Entwicklung politischer Maßnahmen in diesem Bereich.

Chancen und Möglichkeiten

Der europäische Gesetzgeber verfolgt mit seinen aktuellen Regulierungsbestrebungen eine breitere Verfügbarkeit von Daten. Mit der Verordnung über harmonisierte Vorschriften für einen fairen Datenzugang und eine faire Datennutzung (kurz: Data Act) sollen Datensilos durch Datenzugangs- und Datenweitergabeansprüche aufgebrochen werden. Zusätzlich soll eine breitere Akzeptanz für mithilfe von KI durchgeführte Datenanalysen geschaffen werden, indem die KI-Verordnung den Unternehmen ein konkretes Pflichtenprogramm auferlegt, das an das vom eingesetzten KI-System ausgehende Risiko gebunden ist. Das für Deutschland derzeit in Erarbeitung befindliche Forschungsdatengesetz wäre hier ein guter erster Schritt. Im Koalitionsvertrag der Bundesregierung wurde ein solches Gesetz vereinbart, um Forschungsdaten effektiver nutzen zu können. Ein bereits stattgefundenes Konsultationsverfahren hat laut dem Bundesministerium für Bildung und Forschung den gesetzgeberischen Handlungsbedarf in den Bereichen Datenschutz, Zugang zu öffentlichen Daten und Forschungsk Kooperationen untermauert.

Eine solche gesetzliche Grundlage, aber vor allem auch konkrete Initiativen wie die versicherungsmathematische Datenbank, ermöglichen eine stärkere Vernetzung und Unterstützung der wissenschaftlichen Forschung im Bereich der Versicherungs- und Finanzmathematik durch Bereitstellung von qualitativ hochwertigen, relevanten Daten und einer Förderung der Zusammenarbeit von Forschung und Praxis. Der Ausbau der Zusammenarbeit zwischen Wirtschaft und Wissenschaft und damit einer Kooperation zwischen Versicherungsunternehmen und akademischen Einrichtungen macht Anwendungsbereiche sichtbar, verbessert den Wissenstransfer und fördert innovative

Lösungen. In der Konsequenz einer optimierten Datenlage und -verfügbarkeit würden die oben bereits genannten FAIR-Prinzipien durch deren konsequente Anwendung eine Stärkung erfahren. Damit steigt in der Folge auch die Reputation der am Wissenschaftsstandort Deutschland erbrachten versicherungsmathematischen Forschung.

Bei einer erfolgreichen Umsetzung der angestrebten versicherungsmathematischen Datenbank ermöglicht dies auch die Schaffung von Synergien mit anderen nationalen und internationalen Datenprojekten und -initiativen, um die Wirkung und Effektivität des Datenbankprojekts von DGVFM und DAV zu erhöhen. Ferner erlaubt ein frühes Engagement von DAV und DGVFM in diesem Bereich einen Beitrag zu datenpolitischen Entwicklungen, also einer aktiven Mitgestaltung datenpolitischer Rahmenbe-

dingungen und Diskussionen, insbesondere in Bezug auf Datenschutz, Datenethik und die Förderung offener Datenkulturen. Dies würde nicht nur die DAV, sondern vor allem auch die DGVFM als Vertretung der finanz- und versicherungsmathematischen Forschung in Deutschland in eine prominente und wirkmächtige Position versetzen.

Nun braucht es als nächsten Schritt das Commitment und Engagement der Versicherungsunternehmen, um das gestartete Projekt erfolgreich zu etablieren, denn ohne diese Zusammenarbeit wird die Umsetzung der angestrebten versicherungsmathematischen Datenbank nicht möglich sein. Sowohl Anfragen als auch Zusagen werden von Prof. Dr. Matthias Scherer (scherer@tum.de) und Dr. Verena Reiter (verena.reiter@aktuar.de) gerne aufgenommen. ▀



Über die Autoren



→ **Marco Degginger**, Rechtsanwalt Marco Degginger ist Junior Partner in der Sozietät Oppenhoff. Er berät und vertritt Unternehmen zu Fragen des IT- und Datenschutzrechts.



→ **Dr. Axel Grätz** Rechtsanwalt Axel Grätz ist Associate bei Sozietät Oppenhoff. Er berät und vertritt Unternehmen zu sämtlichen Fragen des IT- und Datenschutzrechts sowie des E-Commerce.



→ **Dr. Marc Hilber LL.M.** Rechtsanwalt Marc Hilber ist Partner in der Sozietät Oppenhoff in den Bereichen IT-Recht und Datenschutz sowie Commercial. Er berät und vertritt Unternehmen bei technologiebezogenen Transaktionen.



→ **Kai Hoff** Kai Hoff ist Projektmanager und Data Strategist im Bereich "Programm und Förderung" des Stifterverbandes für die Deutsche Wissenschaft e.V. Er betreut federführend die Initiative „Datagroup. Business 2 Science“.



→ **Dr. Verena Reiter** Verena Reiter ist Referentin der Deutschen Gesellschaft für Finanz- und Versicherungsmathematik e. V. und koordiniert das Datenbankprojekt der DGVFM.



→ **Prof. Dr. Matthias Scherer** Matthias Scherer ist Professor für Risk and Insurance an der Technischen Universität München. Er ist Mitglied im Vorstand der DGVFM und verantwortlich für das DGVFM Datenbankprojekt.