

Redaktionsgespräch mit Tarek Al-Wazir

„Der Erhalt der Datensouveränität hat oberste Priorität“

◀ Daten werden mehr und mehr zu einer Schlüsselressource für die Wirtschaft, sind aber nicht unbedingt immer und überall vorhanden. Das Land Hessen engagiert sich bei dem Thema: Was genau hat es mit dem Financial Big Data Cluster auf sich?

Ihrer Aussage stimme ich voll und ganz zu. Gerade in der Finanzwirtschaft sind oftmals große Datenmengen bei Instituten und anderen Marktakteuren vorhanden, allerdings verfügt jeder Marktakteur nur über seine eigenen Daten. Die Gesamtdatenmenge ist somit sehr frag-

Finanzplatz initiiert. Wesentliche Mitinitiatoren und Treiber sind das Tech-Quartier und die Deutsche Börse AG.

◀ Sie haben kürzlich eine Machbarkeitsstudie zum Aufbau des Financial Big Data Clusters vorgestellt. Wie weit sind Sie hier und wie lange wird es dauern, bis diese Plattform eine Lösung darstellen kann?

Ein wichtiger Baustein beim Aufbau des FBDC ist das Förderprojekt „Untersuchung der Eignung eines Financial Big Data Clusters (FBDC) zur Absicherung

Erste kurzfristige Erkenntnisse erwarten wir uns bereits im Laufe dieses Jahres. Parallel zu den verschiedenen Arbeitspaketen testen wir mit sogenannten Demonstratoren anhand vorbestimmter Daten die praktische Machbarkeit des Projekts. Die Tests laufen mit Blick auf die Frage, ob gesetzliche Nachjustierungen erforderlich sind, um ein erfolgreiches Projekt zu garantieren.

◀ Das Financial Big Data Cluster ist Teil des Gaia-X-Projekts. Wie kommt dieses voran?

Ziel von Gaia-X ist es, gemeinsame Anforderungen an eine europäische Dateninfrastruktur zu entwickeln. Gaia-X ist in verschiedene sogenannte Hubs unterteilt, welche wiederum in einzelne thematische Domänen, etwa Finanzen, Gesundheit und Mobilität eingeteilt sind. Es gibt nationale und länderübergreifende sowie domänenspezifische und domänenübergreifende Arbeitsgruppen. Die verschiedenen Domänen werden jeweils von einem Paten geleitet und koordiniert. Das hessische Wirtschaftsministerium bringt sich in das Projekt Gaia-X durch die Übernahme der Patenrolle für die deutsche Domäne „Finanzwesen“ ein.

Das FBDC ist dabei der zentrale Anwendungsfall der Finanzdomäne im deutschen Gaia-X-Hub. Die Eingliederung des FBDC in das Gaia-X-Projekt hat zum Vorteil, dass zum einen so eine Vernetzung mit anderen Domänen im deutschen Hub, aber auch mit den Finanzdomänen in anderen Ländern möglich ist. Zum anderen ermöglichen themenübergreifende Arbeitsgruppen beim Aufbau des FBDC europäische Standards zu implementieren.

„Ziel von Gaia-X ist es, gemeinsame Anforderungen an eine europäische Dateninfrastruktur zu entwickeln.“

mentiert, sodass sich die Frage stellt, wie diese Daten sinnvoll genutzt und ein Mehrwert daraus generiert werden kann. Es fehlt der deutschen Finanzwirtschaft bisher schlicht die Möglichkeit, Analysen und KI-Entwicklungen auf Basis von Daten durchzuführen, die über öffentlich verfügbare Daten sowie Daten aus dem jeweils eigenen Haus hinausgehen – dieses Problem wollen wir mit dem Financial Big Data Cluster (FBDC) lösen.

Das FBDC soll die Entwicklung von KI-Anwendungen insbesondere im Finanzsektor ermöglichen, indem es als Basis eine geeignete digitale Dateninfrastruktur zur Verfügung stellt, die das Zusammenführen von Daten verschiedener Inhaber ermöglicht. Das Gesamtprojekt FBDC wurde 2018 vom hessischen Wirtschaftsministerium und weiteren Partnern am

der Datensouveränität im Finanzsektor“, kurz mit dem Schlagwort „Safe-FBDC“ benannt. Das Förderprojekt ist Teil des Innovationswettbewerbs „Künstliche Intelligenz als Treiber für volkswirtschaftlich relevante Ökosysteme“ des Bundeswirtschaftsministeriums.

Der Arbeitsplan des Gesamtvorhabens sieht eine 36-monatige Projektlaufzeit von Januar 2021 bis Dezember 2023 vor. Im Rahmen dieses Projektes werden mehrere Arbeitspakete bearbeitet, die sich unter anderem mit der Zusammenführung von Daten im Bereich der Sustainable Finance, der Geldwäscheprävention und der Verbesserung der Marktintegrität befassen. Im Rahmen des Projektes sollen, unter anderem, neue KI-basierte Methoden auf den Anwendungsfällen erforscht, entwickelt und prototypisch validiert werden.

Bislang sind im deutschen Gaia-X-Hub eine Vielzahl von Unternehmen und Experten vertreten. Alle Domänen aus den verschiedenen Wirtschaftsbereichen haben die Arbeiten an ihren „use cases“ aufgenommen und tauschen sich regelmäßig über den Fortschritt aus.

Das FBDC basiert auf freiwilliger Basis und baut auf Kollaboration. Inwieweit machen Sie die Erfahrung, dass die verschiedenen Akteure aus Politik, Verwaltung, Universitäten sowie Unternehmen des Finanzplatzes Frankfurt bereit sind, ihre Daten miteinander zu teilen?

Der Bedarf nach einer Verknüpfung der am Markt vorhandenen, aber bislang fragmentierten Daten bei den verschiedenen Marktakteuren ist groß. Nach unserer Erfahrung ist das größte bestehende Hindernis und der Grund für die bisherige Zurückhaltung der Marktakteure, ihre Daten zu teilen, die Befürchtung, gegen gesetzliche Vorgaben zu verstoßen.

Das FBDC wird genau diese Lücke schließen. Es wird sicherstellen, dass ein Höchstmaß an Datensicherheit durch die Nutzung des neuesten Standes der Technik sichergestellt wird – sowohl infrastrukturell als auch durch vertragliche Vereinbarungen mit den Nutzern. Das übergeordnete Ziel des FBDC ist es, Daten unter strikter Einhaltung aller einschlägigen gesetzlichen Regelungen zusammenzuführen,

men kann, ist die rege Beteiligung von Marktakteuren aus allen relevanten Bereichen zu sehen: Verschiedene Banken und Finanzdienstleister, Universitäten, Start-ups, Politik und Aufsichtsbehörden haben sich dem Projekt angeschlossen.

Das FBDC soll hierbei kein Data Hub sein, sondern vielmehr als Intermediär zur Sammlung von rechtssicheren Daten dienen. Wie genau wird diese Sicherheit gewährleistet und wie sieht es mit der Überprüfung von Verfügbarkeit sowie Qualität dieser Daten aus?

Das initiale FBDC-Konzept sah einen zentralen Daten-Hub vor. Dieses wurde in den letzten Monaten zu einem Konzept eines dezentralen Daten- und Infrastrukturökosystems weiterentwickelt, das im Einklang mit den Gaia-X-Architekturprinzipien steht und die im Rahmen der Machbarkeitsstudie offengelegten rechtlichen Hürden und geäußerten Bedenken bei der Datennutzung berücksichtigt. Das dezentrale System wird ermöglichen, dass Datenbestände institutionsübergreifend aufgefunden sowie Dateninhaber und Datennutzer über einen Datenkatalog zusammengeführt werden können: Daten und ihre Inhaber können auf der Basis von Metadaten-Beschreibungen von interessierten Datennutzern aufgefunden werden, ohne dass die Daten den Dateninhaber verlassen müssen und ohne dass sensible Inhalte sichtbar sind.

„Grund für die Zurückhaltung, Daten zu teilen, ist die Befürchtung, gegen gesetzliche Vorgaben zu verstoßen.“

ren, um einen Mehrwert zu generieren. Dies sind insbesondere rechtliche Vorgaben im Bereich des Datenschutzes und bestehender Verschwiegenheitspflichten. Der Erhalt der Datensouveränität hat hierbei oberste Priorität. Dass ein solches Vorhaben grundsätzlich möglich ist, hat die im Februar finalisierte Machbarkeitsstudie von Deloitte zum FBDC gezeigt.

Als erster Beleg dafür, dass das FBDC Befürchtungen der Marktakteure ausräu-

Außerdem wird das FBDC einen Katalog der vorhandenen Technologien und Dienstleister führen. Diese werden durch das FBDC identifiziert, zertifiziert und ermöglichen so deren effiziente Einbindung durch Dateninhaber beziehungsweise Datennutzer bei der Umsetzung von Anwendungsfällen. Und nicht zuletzt: Das FBDC wird fachliche Beratung bei der Umsetzung konkreter Anwendungsfälle anbieten, indem es seine domänenspezifische Expertise zur Verfü-



Tarek Al-Wazir



Hessischer Minister für Wirtschaft, Energie, Verkehr und Wohnen, Wiesbaden

Das Projekt Gaia-X hat sich zum Ziel gesetzt, eine leistungs- und wettbewerbsfähige Dateninfrastruktur für Europa aufzubauen. Teil des Projektes ist das Financial Big Data Cluster (FBDC), das Finanzmarktteilnehmern ermöglichen soll, Daten sicher untereinander auszutauschen, um so KI- und datengestützte Geschäftsmodelle vorantreiben zu können. Tarek Al-Wazir, dessen hessisches Wirtschaftsministerium die Patenrolle für den deutschen Bereich „Finanzwesen“ im Gesamtprojekt übernimmt, erläutert im Gespräch mit der Redaktion, wie neben der Verbesserung der Wettbewerbsfähigkeit auch die Prävention und Bekämpfung von Geldwäsche durch das FBDC gestützt werden können. Dabei geht er vor allem auf die Vorteile von Machine Learning zur Erkennung von Auffälligkeiten ein. Das System würde von der Bereitstellung von verschiedensten Datensätzen aus unterschiedlichen Finanzinstituten wesentlich effizienter und präziser als herkömmliche Modelle, welche oftmals auf limitierten Datensätzen und festen Regeln fußen. Er betont auch, dass bei allem Austausch der Datenschutz und die Datensouveränität der Teilnehmer am System oberste Priorität haben. (Red.)

gung stellt und Dateninhaber und Datennutzer bei der Auswahl geeigneter Technologien und bei der fachlich und rechtlichen Technologieauswahl berät.

Ich möchte betonen, dass in jedem Fall die maßgebliche Letztentscheidung über das „Ob“ und das „Wie“ der Datennutzung dem oder den Dateninhaber(n) vorbehalten bleibt. Verträge über die Bereit-

stellung und Nutzung von Daten werden allein zwischen Inhaber und Nutzer abgeschlossen.

„Mit dem Financial Big Data Cluster möchten wir datengesteuerte Geschäftsmodelle im Finanzsektor in ganz Europa fördern“, sagten Sie bei der Vorstellung. Droht dadurch unter Umständen eine Benachteiligung etablierter Banken?

Nein, in keinsten Weise! Datengesteuerte Geschäftsmodelle sind auch für etablierte Banken unerlässlich, um im Wettbewerb

Die heutigen Methoden basieren vor allem auf vordefinierten Regeln. Diese erkennen Anomalien oder Auffälligkeiten unabhängig vom Normalverhalten der einzelnen Teilnehmer. Unter Einsatz von Machine Learning jedoch ist eine wesentlich individuellere Betrachtung des Verhaltens einzelner Akteure oder Netzwerke möglich, die das Auffinden von Anomalien oder Auffälligkeiten verbessert.

Des Weiteren beschränken sich die bestehenden Regeln weitestgehend auf bekannte Betrugsmuster. Mit Ansätzen aus

dem können. Glauben Sie, dass die Plattform hier zukünftig entlasten kann?

Absolut! Eine Zielsetzung der Plattform ist, dass diese auch den Sicherheitsstandards für eine Nutzung durch die entsprechenden Behörden soweit entspricht, dass hier Daten analysiert werden können. Viel wichtiger als die Analyse ist aber, dass dann auch eine Rückmeldung vonseiten der zuständigen Behörden zur Qualität der Meldungen möglich wird. Das ist heute zum Teil nur sehr eingeschränkt möglich, wäre aber für die Entwicklung besserer Modelle zur Erkennung von echten Geldwäschefällen elementar wichtig.

„Eine Vernetzung der Daten ist erforderlich, damit Geldwäscheprävention effektiv betrieben werden kann.“

der Zukunft bestehen zu können. Vielmehr würde nach meiner Auffassung ein realer Nachteil für die deutsche und europäische Finanzwirtschaft gerade dann entstehen, wenn sie bei dieser Entwicklung zurückfällt. Insbesondere angloamerikanische und asiatische Dienstleister sowie die Hyperscaler sind nicht zuletzt aufgrund abweichender Datenschutzregularien deutlich aktiver und entwickeln zunehmend innovative Lösungen, die von Kunden nachgefragt werden.

Banken und andere Finanzdienstleister, nicht nur Fintechs, lagern bereits jetzt viele Daten in Clouds aus und nehmen dabei die Dienstleistungen diverser Serviceanbieter, insbesondere auch der Hyperscaler in Anspruch. Die bei solchen Datenweitergaben bestehenden Bauchschmerzen vieler Banken im Hinblick auf Datenschutz und andere rechtliche Vorgaben, auch die aufsichtsrechtlichen Vorgaben, können durch die Dateninfrastruktur von Gaia-X und des Financial Big Data Cluster in der Zukunft beseitigt werden. Beide haben sich der vollumfänglichen Einhaltung der bestehenden rechtlichen Vorgaben zum Schutz von Daten verschrieben.

Wo liegt der wesentliche Mehrwert beim Einsatz vom sogenannten Machine Learning bei der Geldwäschebekämpfung?

dem Machine Learning können in Zukunft auch neue Muster früher entdeckt und untersucht werden. Abschließend kann Machine Learning auch eingesetzt werden, um durch Hinzunahme von anderen externen Informationsquellen die Rate der fehlerhaften Meldungen – sogenannte False Positives – zu reduzieren.

Wie viel Evidenz muss in den Daten vorliegen, um nicht nur eine Anomalie, sondern einen Trend oder ein Netzwerk feststellen zu können?

Die genaue Datenmenge ist immer auch von der Gestaltung und Komplexität des Modells abhängig. Wichtig ist vor allem, dass ausreichend und repräsentative Da-

Was ist der nächste Schritt, nachdem verdächtige Transaktionen identifiziert wurden? An welcher Stelle werden die Regulatoren/Behörden eingeschaltet?

Die im Rahmen des FBDC entwickelten Modelle sollen parallel oder anstelle der heute existierenden Systeme zur Geldwäschebekämpfung eingesetzt werden – natürlich mit dem Anspruch, besser zu funktionieren. An den Berichtsketten zur Meldung von Verdachtsfällen ändert sich dadurch nichts.

Was passiert, wenn das System eine Falschmeldung abgibt? Wie können bei der Früherkennung Falschmeldungen ausgeschlossen werden?

Falschmeldungen sind bei modellbasierter Erkennung immer möglich, sowohl mit den aktuell verwendeten Modellen

„Datengesteuerte Geschäftsmodelle sind für Banken unerlässlich, um im Wettbewerb bestehen zu können.“

ten zum Trainieren der Modelle vorliegen. Genau diese Grundlage soll durch das Financial Big Data Cluster geschaffen werden.

Bei den Behörden geht eine enorme Flut von Verdachtsmeldungen ein, die auch zum Teil überfor-

als auch mit solchen, die neu entwickelt werden sollen. Falschmeldungen können jedoch gegebenenfalls durch Machine Learning reduziert werden. Ehe es zu einer Verdachtsmeldung kommt, werden zudem die vom Modell generierten Meldungen von Experten in den Häusern eingehend geprüft.



Ist das System vollautomatisiert oder kommen stellenweise noch Menschen zum Einsatz? Beispielsweise bei der Bestätigung einer Anfrage nach Daten oder der Weitergabe einer Verdachtsmeldung an die Behörden.

Die Modelle haben in erster Linie zum Ziel, den Experten, die sich in der Finanzindustrie mit Geldwäschebekämpfung befassen, bei der Bewältigung der extrem großen Datenmengen zu helfen. Durch automatisierte und verbesserte Erkennung von Mustern kann der Einsatz dieser Experten fokussierter erfolgen und somit die Bekämpfung der Geldwäsche effektiver gestaltet werden.

Was sind die konkreten Pläne für erste Umsetzungen? Wie genau sollen die sogenannten Demonstranten entwickelt werden und welche Akteure sind hier konkret beteiligt?

Im Rahmen des Projekts Safe-FBDC werden erste Datenpools unter Teilnahme einer Vielzahl von Banken aufgebaut – dieses Projekt ist aber mit einer Laufzeit von drei Jahren eher langfristig orientiert. Parallel entwickelt die Deutsche Börse mit einzelnen Aufsichtsbehörden wie der Deutschen Bundesbank kleinere, sogenannte Demonstratoren, um den Mehrwert der Vernetzung von verschiedenen sensitiven Datensätzen bei der Modellierung zu demonstrieren. Hier sollten erste Ergebnisse noch in diesem Jahr vorliegen.

Wäre es unter Umständen denkbar, dass das Projekt langfristig betrachtet in die Richtung eines Data-Hubs gelenkt wird, vielleicht unter der Voraussetzung, dass die regulatorischen Leitplanken ein Stück weit hierfür geöffnet werden würden?

Hiervon sind wir aus verschiedenen Gründen noch weit entfernt. Zum einen stellt sich grundsätzlich die Frage, ob der mögliche Mehrwert die erheblich größeren Risiken, etwa in Bezug auf Datensicherheit oder rechtskonforme Nutzung der

gehen, um herauszufinden, wie weit wir kommen. Sollten sich auf unserem Weg Hürden auftun, werden wir prüfen, ob es unter Abwägungsaspekten sinnvoll ist, entsprechende Änderungen anzustoßen. Doch davon unabhängig bin ich sehr zuversichtlich, dass mit dem FBDC bereits unter dem geltenden Rechtsrahmen große Fortschritte erreicht werden können.

Was wären Ihrer Meinung nach sinnvolle Instrumente, um die Geldwäscheprävention weiter zu unterstützen?

„Das FBDC als Data-Hub ist nicht ohne Gesetzesänderungen, insbesondere beim Datenschutz, möglich.“

Daten, überhaupt überwiegt. Zudem wird es das FBDC schon in der aktuell geplanten Ausgestaltung ermöglichen, Daten aus einer Vielzahl von Quellen bei entsprechender Zustimmung der Dateneinhaber für Analysen einzubeziehen, ohne dass die Daten innerhalb des FBDC langfristig gespeichert werden müssen.

Das FBDC als Data-Hub, in dem Daten gespeichert und dauerhaft verfügbar sind, ist nicht ohne weitreichende gesetzliche Änderungen, insbesondere hinsichtlich der datenschutzrechtlichen Vorgaben und gesetzlicher Verschwiegenheitspflichten, möglich. Ich denke daher, wir sind gut beraten, einen Schritt nach dem anderen zu

Der Einsatz von modernen Methoden aus dem Feld des Machine Learning ist mit Sicherheit vielversprechend. Vor allem die Ansätze des Federated Learning, also des Trainierens von Modellen, ohne dass ein Teilen der Daten und somit eine Beeinträchtigung der Datensouveränität erforderlich ist, ist vielversprechend und hoch spannend, da so datenschutzrechtliche und andere Anforderungen eingehalten werden. Eine zumindest logische Vernetzung der Daten ist aber dennoch erforderlich, damit diese Ansätze effektiv eingesetzt werden können. Mehr Standardisierung in der Erhebung und Speicherung von relevanten Daten wäre somit auch ein wichtiger Faktor.

Bleiben Sie immer auf dem neuesten Stand!

Ihre Kreditwesen-Redaktion informiert täglich in der Rubrik „Tagesmeldungen“.

Folgen Sie uns auf



oder besuchen Sie uns unter

www.kreditwesen.de/tagesmeldungen