

DIGITAL.CORNER

Mit KI heterogene Daten verknüpfen und unentdeckte Muster erkennen

DZD: Intelligente Datenbanken sollen präzisere Diagnosen und individualisierte Therapien ermöglichen



Professor Martin Hrabě de Angelis
Vorstand des Deutschen Zentrums für Diabetesforschung; Helmholtz Zentrum, München sowie TU München.

Foto: Privat



Dr. Alexander Jarasch
Leiter Bioinformatik und Datenmanagement im Deutschen Zentrum für Diabetesforschung

Foto: Privat

Bei der Erforschung, Diagnostik und Therapie von Diabetes werden viele Daten aus unterschiedlichsten Quellen gesammelt, die sich bislang nur schlecht zusammenführen und nutzen lassen. Im Deutschen Zentrum für Diabetesforschung (DZD) will man die verborgenen Datensätze heben und setzt dabei auf Mustererkennung durch maschinelles Lernen.

Was mangelnde Interoperabilität bedeutet, ist den Teams in Diabetespraxen vertraut: Die Glukosewerte aus Messgerät X lassen sich nicht in Software Y einlesen, ganz zu schweigen von den Aktivitätsprotokollen des Fitnesstrackers oder den Kurven des Langzeit-EKG. Dabei würden alle Daten miteinander verknüpft ein sehr aufschlussreiches Bild ergeben, das bei der Optimierung der Diabetes-therapie unterstützen kann. Immerhin handelt es sich bei Diabetes um eine komplexe Erkrankung, bei der neben genetischen und epigenetischen Faktoren auch der Lebensstil und die Umwelt eine Rolle spielen.

Es gibt also eine Vielzahl von Datenquellen, die potenziell für die Diabetes-therapie genutzt werden können – sofern es denn gelingt, die Datensätze zu heben.

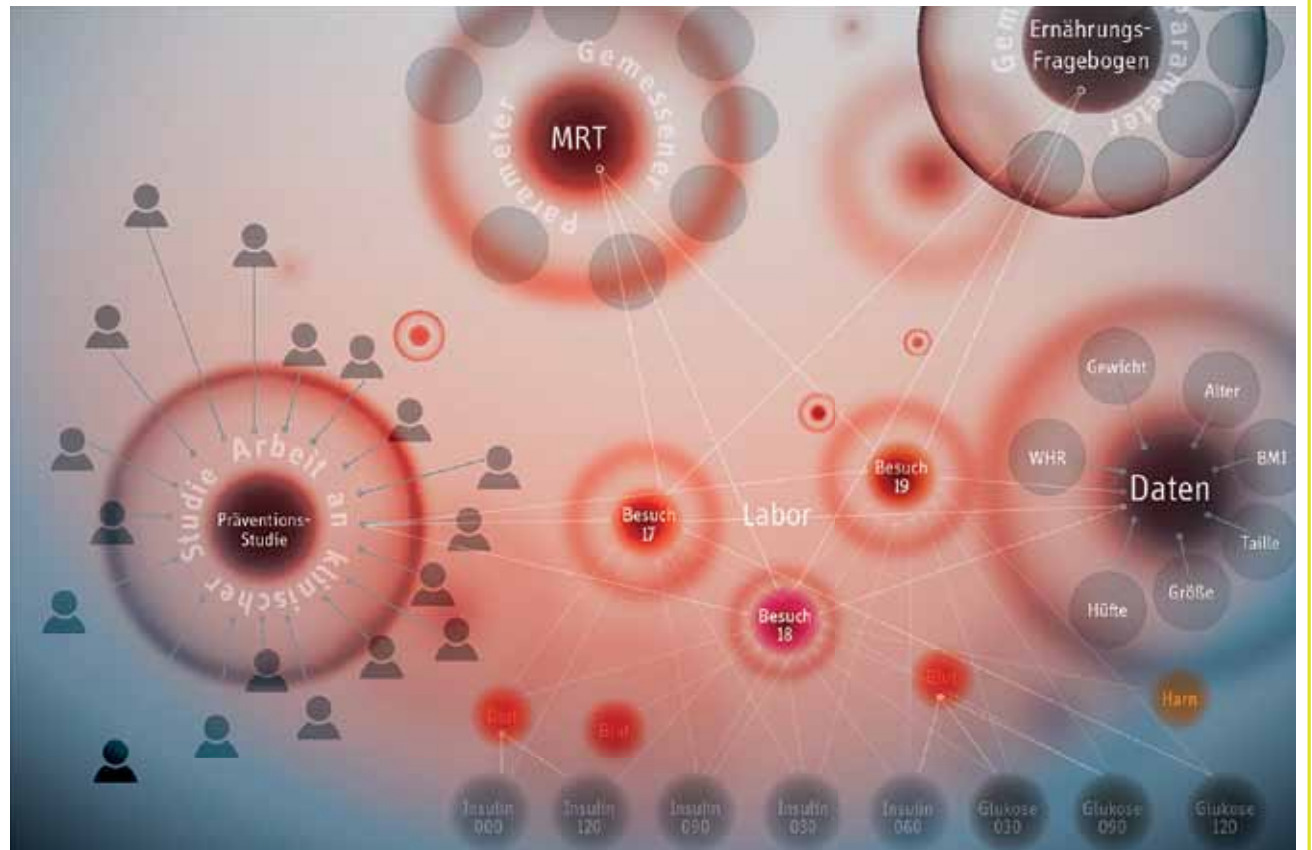
Vor ähnlichen Herausforderungen stehen auch Einrichtungen wie das Deutsche Zentrum für Diabetesforschung (DZD): Hier fallen Daten aus unterschiedlichen wissenschaftlichen Zweigen und zu verschiedenen Aspekten der Erkrankung an – etwa aus Kohorten, Studien, Bioproben, präklinischen Modellen, Geno- und Phänotypisierung oder Genanalysen. Um diese Daten besser nutzen zu können, wurden das Datenmanagement und die Anwendung von KI im DZD in den vergangenen Jahren kontinuierlich ausgebaut: „Wir arbeiten daran, alle diese Daten in eine Form zu überführen, in der sie vergleichbar und analysierbar sind“, erklärt Dr. Alexander Jarasch, Leiter Bioinformatik und Datenmanagement im DZD.¹

Preisgekröntes Projekt „Graphs to Fight Diabetes“

Durchaus mit Erfolg: So wurde das von ihm geleitete Projekt „Graphs to Fight Diabetes“ Ende Januar in Berlin im Rahmen des DiaTec-Kongresses mit dem bytes4diabetes-Award ausgezeichnet. Der Preis wird von der BERLIN-CHEMIE AG gestiftet. Über die Preisvergabe entscheidet eine Jury, u.a. bestehend aus Mitgliedern des Zukunftsboards Digitalisierung (siehe Kasten). Bei „Graphs to Fight Diabetes“ handelt es sich um eine Graphdatenbank, die Daten aus unterschiedlichen Quellen zugänglich, wiederverwertbar und übergreifend nutzbar machen soll. Mittels Graphentechnologie führt sie die Daten zusammen und bildet ihre Beziehungen untereinander in sogenannten Knoten und Kanten ab. „Die Graphdatenbank umfasst bis dato 2,8 Milliarden Knoten und 4,9 Milliarden Kanten aus Forschungsdaten der vergangenen zehn Jahre – und sie wächst ständig weiter. Mit herkömmlichen Methoden kann man diese Datenmengen unmöglich auswerten“, berichtet Dr. Jarasch.

Diabetesrelevante Daten für die COVID-19-Forschung

Diese algorithmusgesteuerte Form der Datenverarbeitung erleichtert das Auffinden von Einzelinformationen, aber auch die Visualisierung, weil



Die Graphdatenbank des DZD umfasst aktuell 2,8 Milliarden „Knoten“ und 4,9 Milliarden „Kanten“ aus Forschungsdaten der vergangenen zehn Jahre. Foto: DZD

sie Querverbindungen, Zusammenhänge und Muster sichtbar macht. „Wir entwickeln damit eine neue Herangehensweise in der Diabetesforschung“, beschreibt DZD-Vorstand Professor Martin Hrabě de Angelis das Projekt. Aktuell nutzt das DZD seine Daten u.a. im Rahmen verschiedener Forschungsinitiativen und Studien zu COVID-19. In dem Projekt CovidGraph² etwa werden öffentlich zugängliche Literaturquellen, Patentschriften zu COVID-19 sowie Datensätze aus Genom- und molekularbiologischen Datenbanken verknüpft und in einem Graphen anschaulich dargestellt. Dies soll Forschenden dabei helfen, sich schnell und effizient in der wachsenden Wissensbasis zu COVID-19 zurechtzufinden. Derzeit wird der CovidGraph noch weiter ausgebaut. So steuert das DZD u.a. diabetesrelevante Daten bei. Künftig könnte der CovidGraph dann helfen, neue Erkenntnisse zu gewinnen, welche Gefahr speziell bei Diabetes vom Coronavirus SARS-CoV-2 ausgeht. „Die Graphdatenbank funktioniert dabei ähnlich wie etwa der Algorithmus von Google. Man kann Suchbegriffe eingeben – etwa „ACE2-Rezeptor“ und „Pankreas“ – und erhält dann sämtliche Verknüpfungen graphisch anschaulich angezeigt“, so Dr. Jarasch.

„Neue Technologien können bei der Diagnostik unterstützen“

Das DZD arbeitet in sieben Forschungsschwerpunkten (Academies) an definierten Schlüsselthemen und Fragestellungen – wie zum Beispiel, welche Rolle bestimmte Gene bei der Diabetesentstehung und beim Auftreten von Diabetes und Folgeerkrankungen spielen.

Die Graphentechnologie kann helfen, Antworten auf Fragestellungen aus den Academies zu finden. „Es gibt beim Typ-2-Diabetes und auch beim Prädiabetes große Unterschiede zwischen den verschiedenen Subgruppen, z.B. was die Wahrscheinlichkeit einer späteren Niereninsuffizienz angeht“, erklärt Prof. Hrabě de Angelis. „Mit Graphentechnologie und maschinellem Lernen könnte es künftig gelingen, Risikogruppen präziser zu identifizieren.“

Künstliche Intelligenz kann Ärzte nicht ersetzen

Von besonderem Interesse sind dabei die Übergangspunkte (transition points), an denen sich eine Erkrankung bzw. Folgeerkrankung manifestiert: „Woran lässt sich das Risiko von Menschen mit Prädiabetes genau festmachen? Was passiert auf molekularer Ebene, wenn sich aus einem Prädiabetes ein manifester Typ-2-Diabetes entwickelt? Ab wann genau kann die Betazelle das Ungleichgewicht im Stoffwechsel nicht länger kompensieren? Oder wodurch zeichnet sich der Punkt aus, an dem Einlagerungen von Fett in der Leber zur nicht-alkoholischen Hepatosteatose führen?“, beschreibt er mögliche

Fragestellungen, bei deren Beantwortung maschinelles Lernen unterstützen kann.

Bei anderen Projekten ist bereits ein klinischer Nutzen greifbar. So gelang es Computerbiologie-Experten in einem Pilotprojekt aus München,³ mit Techniken des maschinellen Lernens diabetische Retinopathien besser zu klassifizieren. Sie trainierten den Algorithmus darauf, anhand von Netzhaut-Bilddaten kranke von gesunden Augen zu unterscheiden und die Patienten gemäß der Schwere und des Fortschreitens der Erkrankung einzuordnen. In einem weiteren Projekt⁴ konnten Wissenschaftler aus München und Dresden – ebenfalls mithilfe maschinellen Lernens – den Risiko-Score für Typ-1-Diabetes deutlich verbessern. Der neue Risikotest bezieht bis zu 41 Risikogen-Regionen für Typ-1-Diabetes ein und könnte helfen, Typ-1-Diabetes bei Risikokindern früher zu diagnostizieren. „Künstliche Intelligenz und Graphdatenbanken werden Ärztinnen und Ärzte niemals ersetzen können“, betont Prof. Hrabě de Angelis. „Doch die neuen Technologien können sie entlasten und bei Diagnostik und Therapiefindung unterstützen, sodass sie letztlich mehr Zeit haben, sich den Patientinnen und Patienten zu widmen“, ergänzte Dr. Jarasch.

1. www.dzd-ev.de
2. https://covidgraph.org
3. Eulenberg P et al. Nature Communications 2017; 8: 463
4. Bonifacio E. et al. PLoS Medicine 2018; doi:10.1371/journal.pmed.1002548



Das Zukunftsboard Digitalisierung

Mit dem Zukunftsboard Digitalisierung (zd) möchte die BERLIN-CHEMIE AG dazu beitragen, den Digitalisierungsprozess in der Diabetologie aktiv voranzutreiben. Zurzeit gehören dem zd zehn feste Experten an. Darunter niedergelassene und klinisch tätige Diabetologen, Experten für Diabetestechnologie, Vertreter von Krankenkassen und Patienten. Das zd wird geleitet von Professor Dr. Bernhard Kulzer, Bad Mergentheim. Das zd möchte aufzeigen, welche Chancen und Nutzen die Digitalisierung bietet, aber auch Risiken identifizieren. Mehr Informationen unter www.zukunftsboard-digitalisierung.de und www.medical-tribune.de/digital-corner.