

Big Data in der Pharmaindustrie

Sammeln, speichern – und dann?

Alle reden von Big Data. Auch für die Pharmaindustrie ist das ein wichtiges Zukunftsthema. Aber um welche Daten geht es hier eigentlich? Welche sind besonders wichtig, und wie müssen sie erfasst werden, um künftig sinnvoll genutzt werden zu können? Um Licht ins Datendunkel zu bringen, haben wir mit Experten gesprochen, die sich auf die Verarbeitung großer Datenmengen spezialisiert haben und der Pharmaindustrie Tipps für den Umgang mit Big Data geben können.

» Kann man Pharmadaten kategorisieren? Oliver Bracht, Geschäftsführer und Chief Data Scientist bei eoda, teilt die Daten in drei Klassen ein – je nach Stand der Produktentwicklung: „Im ersten Schritt geht es um Daten, die im Zuge der statistischen Versuchsplanung bei der Produktentwicklung, ‚Design of Experiments‘, relevant sind.“ Mithilfe von Analysen soll dabei die beste Kombination von Inhaltsstoffen mit möglichst wenigen Versuchsreihen gefunden und die aufwendige Entwicklungsphase verkürzt werden.

Im zweiten Schritt, der Produktionsphase werden laut Bracht sämtliche Sensordaten der Fertigungsmaschinen erhoben, um eine möglichst genaue Ausfallwahrscheinlichkeit ebendieser vorherzusagen. „Hier sprechen wir von einem typischen Predictive-Maintenance-Anwendungsfall, bei dem es darum geht, die Downtime der Produktionsanlagen zu minimieren“, erklärt der eoda-Chef.

Im dritten Schritt würden die Daten aus Einkauf, Vertrieb und Logistik verwendet und bestimmte Fragen geklärt, wie etwa: Unterliegen bestimmte zur Produktion benötigte Substanzen preislichen Schwankungen am Markt? Oder: Wie können Lieferengpässe vermieden werden? Wie lässt sich die Produktion auf den voraussichtlichen Absatz im nächsten Monat abstimmen oder welche Vertriebskampag-

nen haben die größten Erfolgsaussichten? „Hier können der strategische Einkauf, der Vertrieb und die Logistik von Optimierungspotenzialen profitieren, indem verschiedene Datenstränge aus diesen Bereichen verknüpft werden“, erklärt Bracht.

Kirsten Weiss – Chief Digital Officer bei ciSio – ordnet hingegen die verschiedenen Datentypen in zwei Gruppen ein (s. Kasten S. 35): Sie unterscheidet zwischen internen und externen Business-Daten. Zu der ersten Gruppe zählen Unternehmensdaten, die bei der Produktentwicklung anfallen. Die zweite Gruppe bilden unter anderem Patientendaten aus mobilen Anwendungen oder Pressespiegel zu Produkten und dem Image des Unternehmens.

Nicht zuletzt spielen die Daten auch im Pharmamarketing eine große Rolle. Darauf weist der Geschäftsführer von TWT Digital Health, Klaus Mueller, hin. Wenn nämlich Pharmaunternehmen Multi-Channel-Konzepte entwickeln, bei denen die Zielgruppen über verschiedene Kanäle wie Website, Apotheke oder Außendienst erreicht und die Ansprachen in den Kanälen aufeinander abgestimmt werden sollen.

Ob nun eine chronologische Klassifizierung der Daten (nach Produktzyklus), oder eine unternehmenszentrierte (intern oder extern) – letztendlich muss jedes Unternehmen für sich selbst entscheiden, wie es einen Überblick über die relevanten

Daten bekommt. Wichtig ist nur, dass früh genug entsprechende Ressourcen geschaffen werden. Und das bedeutet: Ausreichend Manpower mit entsprechender Sachkompetenz und ein Team, das sich jetzt Gedanken über die künftige Digitalisierungsstrategie des Unternehmens und die Bedeutung von Big Data macht.

Welche Fähigkeiten nötig?

Wie wichtig eine rechtzeitige Rekrutierung von kompetenten Mitarbeitern ist, betont auch Chief Digital Officer von ciSio Kirsten Weiss. Sie weiß auch, worauf es dabei ankommt: „In erster Linie müssen Fachkräfte eingestellt werden, die in der Lage sind, aus den vielfältigen Daten die Geschäftsmöglichkeiten für das eigene Unternehmen herauslesen zu können.“

Die Digital-Verantwortliche nennt in diesem Zusammenhang zwei Kompetenzprofile und entsprechende Aufgaben. Demnach sei der Data Engineer im Wesentlichen für das Zusammenführen von Daten verantwortlich. Er kreierte aus den zur Verfügung stehenden Daten und Technologien eine Landschaft, in der sich der Data Scientist ausleben kann. Sein Wissen beschränke sich nicht nur auf die im Unternehmen vorhandenen Daten und ihre Speicherorte, so Weiss. Dieser wisse auch, wie man diese Daten am besten in eine zentrale Analyseinfrastruktur integriert, welche Technologien sich dazu eignen und welche zusätzlichen externen Daten zur Anreicherung genutzt werden können.

Als zweites Profil führt die Expertin den Management Scientist an, der zwischen der Technik und dem Fachbereich vermittelt. Die Leistung des Management Scientist besteht darin, die Sprache des technisch-



Oliver Bracht (Geschäftsführer und Chief Data Scientist bei eoda), Klaus Mueller (Geschäftsführer TWT Digital Health), Kirsten Weiss (Chief Digital Officer von ciSio)

und datenunversierten Mitarbeiters in die des Data Scientists zu übersetzen. Dies umfasse die Spezifizierung der eigentlichen fachlichen Problemdefinition, die Übersetzung und das Schärfen der dahinterliegenden analytischen Fragestellung, und gehe über die Identifikation der benötigten Daten, das Management der operativen Analyse bis hin zur Kommunikation von analytischen Ergebnissen und Handlungsempfehlungen. Für diese anspruchsvolle Aufgabe benötige der Management Scientist ein sehr gutes Verständnis von sowohl analytischen Methoden und Vorgängen als auch fachlichen Prozessen und Auswirkungen. Im Projektalltag kommen dem Data Scientist laut Weiss folgende Aufgaben zu:

- die Datensuche und die Klärung der Frage: Welche Daten stehen zur Verfügung, beziehungsweise welche lassen sich zusätzlich besorgen?
- die Datenbereinigung – Aufbereitung der Daten für die anschließende Analyse,
- die Offline-Datenanalyse: Wie lassen sich aus den vorliegenden Daten die gewünschten Informationen extrahieren?
- das Überführen der Ergebnisse in ein produktives System zur Online-Analyse.

Die Expertin von ciSio rät, die Daten so früh wie möglich zu sammeln und zu speichern und damit so genannte Data Lakes zu schaffen. Sie empfiehlt außerdem eine Kombination von „low cost“-Systemen, die unstrukturierte und strukturierte Daten speichern können, und „high cost“-In-Memory-Systemen, mit denen in Echtzeit schnell sämtliche KPIs gerechnet werden können. Für die Ableitung zukünftiger Trends schlägt sie Analysewerkzeuge wie etwa „Predictive Analysis“ vor.

Unternehmen, denen es an Datenexperten mangelt, können auch auf externe Unterstützung zurückgreifen. Bracht empfiehlt Unternehmen, die Data Science für sich nutzen wollen, ein eigenes Analytik-Team aufzubauen, das auch von externen Data Science Spezialisten unterstützt wird. „Dieses Team ist im besten Falle interdisziplinär aufgestellt und wissensdurstig.“ Außerdem bricht er eine Lanze für die Open-Source-Technologie: „Mit ihr erlangt man bessere Einblicke und transparente Ergebnisse.“ Die Lösungen seien dadurch nachvollziehbar.

Zusätzlich setze man auf ein breites Methodenset, verrät der Experte, und nutze verschiedene Methoden für individuelle Fragestellungen. „Das raten wir unseren Kunden auch“, so Bracht. Das Team sollte außerdem aus seiner Sicht bereit sein, sich ständig weiterzuentwickeln. „Data

Science ist eine Entdeckungstour – man muss offen sein für Abzweigungen.“ Oft ließen sich die spannendsten Erkenntnisse dort erlangen, wo sie im Vorfeld nicht vermutet würden.

Die Erkenntnisse

„Unter anderem können Data Lakes in Verbindung mit Prognosen auch bestens im Rechnungswesen eingesetzt werden“, erklärt Weiss, „um diverse Schätzgrößen, die im Rahmen von IFRS-Buchungen in Zukunft mehr und mehr benötigt werden, zu berechnen“. Es sei daher durchaus denkbar, etwa die Gewährleistungsanteile, die nicht mehr als Umsatz gelegt und bestmöglich abgeschätzt werden müssen, zu berechnen. „Dies kann entweder auf Basis der Historie bei vergleichbaren Produkten gemacht werden oder sogar in Kombination mit Informationen zu den betroffenen Produkten aus dem Internet.“ Somit werde eine genauere Einschätzung erreicht, welches Gewährleistungsrisiko vor allem bei neueren Produktlinien besteht.

Bei der Nutzung großer Datenmengen aus neuen Technologien oder bestehenden Datenquellen stehe die Pharmabranche allerdings noch ganz am Anfang, so die Einschätzung des TWT Digital Health-Geschäftsführers Mueller. Er ist jedoch fest davon überzeugt, dass Big Data auch die Pharmaindustrie revolutionieren wird.

Doch welche Erkenntnisse können noch aus den Data Lakes gewonnen werden? Und sind alle Ansammlungen von Daten gleich wichtig für die Zukunft? Mueller: „Der eigentliche Schatz sind die klinischen Daten und deren Verknüpfung mit pathologischen Parametern und patientenspezifischen Modellen.“ Vor allem in der Personalisierten Medizin verspricht sich der Geschäftsführer wichtige Erkenntnisse auf Basis von Big Data: „Individuelle Therapiepläne wie in der Onkologie werden wir perspektivisch auch bei vielen anderen Krankheiten sehen. Und sie werden auch dort zu einer breiten Stratifizierung der therapeutischen Optionen führen. Und

Überblick über die Pharma-Daten

Kirsten Weiss, Chief Digital Officer bei ciSio, unterscheidet zwischen internen und externen Business Daten in der Pharmaindustrie. Zu den **internen Business-Daten** zählt sie:

- Produktionsdaten,
- Chargen,
- Finanzdaten,
- Qualitätsdaten,
- Daten aus Forschung und Entwicklung (Wirkstoffentwicklung, klinische Prüfung, Prüfsubstanzen),
- Daten des Produkt-Launchs, Zulassungsdaten,
- Patentverwaltung,
- Vertriebsdaten,
- Auftragseingang,
- Umsatz,
- Margen.

Zu den **externen Business-Daten** gehören:

- Mobile Dienste/-Daten, wie gesundheitsrelevante Eckdaten, z.B. Bewegung, Schlaf, Blutdruck, Zuckerwerte, Puls etc.,
- Benchmarks,
- Patentabläufe,
- Ablauf von Produkten mit Unterlagenschutz (durch Abläufe des Unterlagenschutzes verringert sich das Volumen des geschützten Marktes sukzessive),
- Alert für Änderungen bzw. neue gesetzliche Vorschriften,
- Meta-Search-Erkenntnisse aus Clippings, Social-Media-Plattformen etc. zu Fragen wie: Was wird über mein Unternehmen, ein bestimmtes Medikament und/oder Produkt geschrieben?

das nicht nur bei Neuzulassungen. Für das gesamte Arsenal bewährter Arzneimittel, die heute eher nach dem Gießkannenprinzip eingesetzt werden, werden dann individuelle Behandlungspläne erstellt.“

Gerade bei den Volkskrankheiten, wo teilweise sehr unterschiedliche Behandlungskonzepte zur Verfügung stünden, könne das ein Ende der „Trial and error“-Strategie bedeuten, so die Vision des Fachmanns. Er stellt sich außerdem einen Datenkreislauf vor, bei dem die Informationen selbstverständlich nicht nur in eine Richtung fließen: „Natürlich liegt es im ureigensten Interesse der Pharmahersteller, die Wirkung der eigenen Präparate bei den unterschiedlichsten Bevölkerungsgruppen eingehend zu untersuchen und zu dokumentieren, auch lange nach der Markteinführung. Und diese Daten fließen dann wieder in die Produktentwicklung zurück.“ ☞