



Verwendung von Daten und Modellen in datengetriebenen Geschäftsmodellen

Jonathan Misler

Datengetriebene Geschäftsmodelle – seit jeher existent und derzeit lediglich „en vogue“?

Das Thema datengetriebene Geschäftsmodelle ist nicht neu, dennoch erfährt es aktuell eine hohe Aufmerksamkeit. Unternehmen wie Amazon, Alipay oder Google befeuern mit ihrem Erfolg in der Datennutzung die Diskussion. Dieser Artikel greift die aktuellen Diskussionen um dieses Thema auf und erörtert die Wichtigkeit von Daten und Auswertungsmodellen für den Erfolg eines datengetriebenen Geschäftsmodells.

Datengetriebene Geschäftsmodelle sind keine Erfindung der heutigen Zeit. Insbesondere Bankgeschäfte waren schon immer von Daten getrieben. Es galt ausreichend Liquidität vorzuhalten, um benötigtes Guthaben auszahlen zu können und zugleich Kredite vergeben zu können. Auch sollten die Kredite gewinnbringend sein, weswegen eine Risikoeinwertung vorzunehmen war. Diese Tätigkeiten erforderten eine abstrakte Bewertung der Faktenlage und konnten nur gewinnbringend durchgeführt werden, wenn das Geschäft oder die Gesamtheit der Geschäfte die reale Situation gut abbildete. Dafür galt es auch den Überblick über Guthaben, Zinsen und Tilgungen zu bewahren.

Diese abstrakte Bewertung von Cash-Flow und Risiken erfolgte durch Modelle unter Verwendung der Daten aus der eigenen Buchhaltung oder anderen Quellen.

Die Bedeutung des Zusammenspiels von Daten und Modellen

Wie bereits festgestellt sind datengetriebene Geschäftsmodelle nicht neu. Im Grunde basiert die gesamte Wirtschaft auf Entscheidungen, welche in der Regel besser ausfallen, je besser die Datengrundlage und deren Antizipation in Modellen für Ergebnisse und die damit verfolgten Ziele sind.

Es gibt demnach auch nicht den einen Schlüssel, der sofort die richtige Lösung liefert. Gerade in der aktuellen Corona-Krise sehen wir, wie Wissenschaftler täglich versuchen die Daten durch ihre Modelle auszuwerten, um so Hebel zu finden, mit denen die Krise abgemildert oder beendet werden kann. Es werden Daten erhoben, qualitätsgesichert und in den Modellen verwendet. Die Modelle werden angepasst und die Ergebnisse ausgewertet, um dem Ziel, dem Verstehen der Zusammenhänge der Krise näher zu kommen. So verhält es sich auch bei einem datengetriebenen Geschäftsmodell. Der Schlüssel dabei liegt in dem Aufzeigen des Nutzens, kundenseitig oder effizienzbasierend, durch Abbildung der Kausalität aus Korrelationen.

Das Aufzeigen dieser Kausalität durch Korrelationen ist die Grundlage nicht nur für erfolgreiche Maßnahmen in einer Krise, sondern auch für den Erfolg von datengetriebenen Geschäftsmodellen. Die im letzten Jahrhundert entwickelten Technologien zur elektronischen Verarbeitung von Daten in Verbindung mit den bekannten statistischen Methoden (bspw. der Einsatz der verschiedenen Verteilungen) geben die Möglichkeit diesen Erfolg zu vervielfältigen. Dieses ist auch der Grund, warum datengetriebene Geschäftsmodelle heute wesentlich gewinnbringender eingesetzt werden können als noch vor einigen Jahrzehnten.

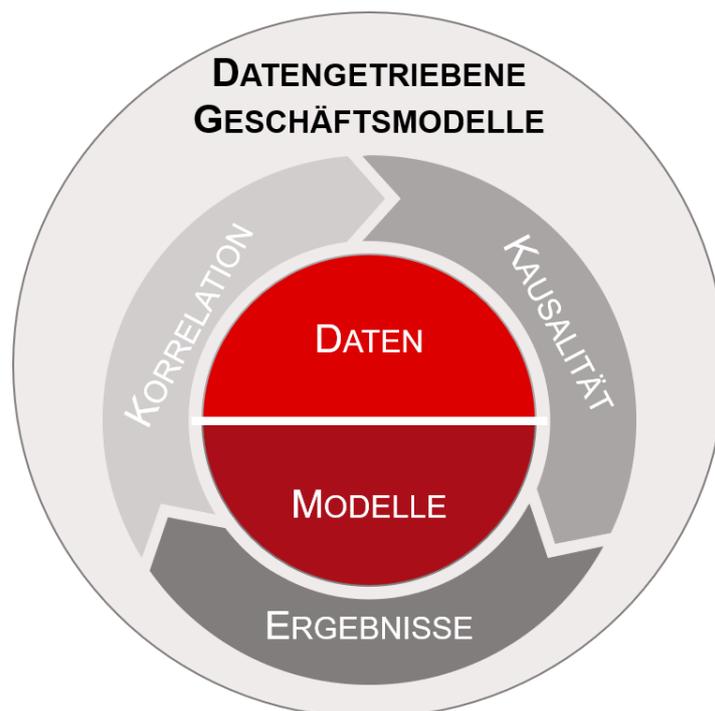


Abbildung 1 - Zusammenspiel von Daten und Modell

Die statistischen Methoden ermöglichen den Aufbau von Modellen zur Auswertung von Daten und so die Ermittlung von Korrelationsgraden. Generell wird davon ausgegangen, dass ein sehr hoher Grad der Korrelationen zwischen zwei Variablen eine Kausalität dieser Variablen darstellt. Zugleich ist die Annahme, dass eine hohe Korrelation automatisch Kausalität bedeutet, mit Vorsicht zu betrachten. Ein Beispiel: Statistisch korreliert die Storchpopulation mit der Anzahl der neugeborenen Kinder in einem Gebiet. Man könnte also statistisch davon ausgehen, dass der Storch die Kinder bringt. Der tatsächliche

gemeinsame Faktor ist jedoch ein anderer. Im ländlichen Raum werden mehr Kinder geboren, und dort leben auch mehr Störche. Der kausale Zusammenhang basiert hier auf den ländlichen Bedingungen, von denen beide Zahlen profitieren. Somit würde es vermutlich nicht helfen, mehr Störche in die Stadt zu bringen, um die Geburtenrate anzuheben. Die statistischen Modelle sollten demnach als Qualitätssicherung auf andere mögliche gemeinsame Faktoren geprüft werden, bevor auf den Ergebnissen eines solchen Modells Entscheidungen basiert werden. Die technologischen Entwicklungen erlauben hier eine starke Beschleunigung.

Mit den Möglichkeiten der heute zur Verfügung stehenden Technologie können die Storchenpopulationen und Geburtsraten relativ schnell verglichen werden. Auch ist es dadurch möglich, ein solches Modell mit anderen Parametern, wie bspw. der Bevölkerungsdichte des Ortes der Geburtsrate und Storchenpopulation, abzugleichen und eine Variable zu finden, die ebenfalls stark korreliert. Dadurch würde die direkte Kausalität der Geburtsrate und der Storchenpopulation in Frage gestellt werden.

Die Grundlage für den Erfolg eines datengetriebenen Geschäftsmodell ist demnach die Fähigkeit das eigene Geschäftsmodell und die dahinterstehenden Modelle und Daten schneller und qualitativer zur Verfügung zu haben, hinterfragen und weiterentwickeln zu können als die Wettbewerber. Entscheidend ist demnach die dynamische Effizienz.

Datengetriebene Geschäftsmodelle sind nicht per se erfolgreich

Um die Folgen des Erdbebens im Jahr 1923 und der generellen Abkühlung des weltweiten Wirtschaftsklimas entgegen zu wirken, gab die Bank of Japan sogenannte „Erdbebenbonds“ aus. Diese waren vor allem für Banken gedacht, die zuvor stark die Kapazitätsausweitungen der Unternehmen unterstützt hatten, die jetzt wirtschaftliche Probleme hatten. Im Januar 1927 nahm die Bank of Japan diese Bonds zurück. Durch Gerüchte unter den Kunden dieser Banken gab es Bank-Runs und 37 Banken mussten schließen.

In dieser Situation gab es zwei Auslöser, die dafür sorgten, dass die datengetriebenen Geschäftsmodelle dieser Banken nicht erfolgreich waren.

Der ausschlaggebende Faktor für die Insolvenz dieser Banken war der Bank-Run der Kunden. Die Banken hatten nicht ausreichend Bargeld eingelagert, um die Bedürfnisse der Kunden decken zu können. Die Modelle zur Vorhaltung der Liquidität hatten den Bedarf dieser Situation unterschätzt.

Die Erstellung solcher Modelle basiert in der Regel auf Annahmen, wie bspw. vergangenen Werten oder begründeten Schätzungen. Es könnte z.B. davon ausgegangen werden, dass Banken zum Wochenende hin mehr Bargeld vorhalten müssen. Die Kunden bereiten sich auf das Wochenende vor, wenn die Banken geschlossen haben und sie somit nicht an Bargeld herankommen. Andere Faktoren könnten saisonaler Natur oder auch situativ sein. In diesem Beispiel hätten die Modelle situativ durch die Gerüchte einen wesentlich höheren Bestand vorsehen müssen. Man hätte sich bei der Berechnung dieses Szenarios an der Auszahlungskapazität orientieren können. Die Fragestellung für das Modell wäre gewesen,

wie viel Geld kann praktisch innerhalb eines Tages ausgezahlt werden? Auf dieser Basis hätte die Entscheidung getroffen werden können, mehr Bargeld vorrätig zu halten.

In der Regel ist ein solches Modell daher komplexer. Es hat verschiedene Effekte und Situationen abzubilden, um jeweils richtig entscheiden zu können. Menschen tun dies meist aus dem Gefühl bzw. basierend auf der gesammelten Erfahrung. Statistische Modelle können dieses aber präziser, da hier Thesen direkt geprüft werden können.

Modelle können auf verschiedenen Arten ein datengetriebenes Geschäftsmodell unterstützen. Eine Möglichkeit ist direkt mit diesem (mathematischen) Modell Gewinne zu erzielen. Das Modell gibt als Ergebnis Chancen für Investitionen aus, ggf. verbunden mit einer Eintrittswahrscheinlichkeit, nach der gehandelt werden sollte. Eine andere Möglichkeit ist die indirekte Unterstützung, wie Google dieses mit auf Algorithmen basierenden Werbeanzeigen tut oder in dem oben genannten Beispiel der Liquiditätsvorhaltung. Amazon fährt eine gemischte Strategie, indem die Algorithmen Nutzerverhalten auswerten und auf der Basis sowohl den Kunden Kaufvorschläge unterbreiten als auch die Logistik steuern. So steigern die Algorithmen die Einnahmen im klassischen Onlinehandel und reduzieren parallel die Kosten durch effizienten Einsatz der Ressourcen.

Die Grundlage für Entscheidungen kann auf diese Weise durch die Modelle gestärkt werden. Die zu Grunde liegenden (statistischen/mathematischen) Modelle sind für datengetriebene Geschäftsmodelle somit essenziell.

Daten justieren das Modell

Der chronologisch erste Faktor für die Schließung der Banken war das Erdbeben gepaart mit dem generellen Abkühlen der Weltwirtschaft durch den vorangegangenen 1. Weltkrieg. Die Banken hatten auf extensive Kredite für die Ausweitungen von Produktionen gesetzt. Das Risiko des Ausfalls wurde in den Modellen dabei rückblickend unterschätzt. Auch könnten die Daten zur Validierung dieser Modelle gefehlt haben. Produktionsbetriebe im industriellen Maßstab waren damals noch neu. Das Resultat war das Versagen des Geschäftsmodells, weswegen die Stützung durch die Notenbank notwendig wurde.

Der Kern der Kreditvergabe ist die Prognose des Ausfalls der Kredite, ausgedrückt durch die Ausfallwahrscheinlichkeit. Die Ausfallwahrscheinlichkeit eines Kredits wird durch die Daten des einzelnen Kunden über das zuvor bestimmte Ausfallmodell bestimmt. Es werden also die Kundendaten erfasst, die für das Modell benötigt werden (bspw. Einkommenshöhe, Familienstand, Sicherheit des Einkommens).

Die benötigten Daten sind dann für das Modell aufzubereiten (bspw. Ermittlung des durchschnittlichen monatlichen Einkommens). In der Regel sind die Modelle statistische Modelle, welche Korrelationen ausrechnen und dadurch eine Fehlerrate beinhalten. Der Einsatz qualitativ höherwertiger Daten verringert die Fehlerrate des Modells und kann somit den Erfolg des datengetriebenen Geschäftsmodells steigern.

Qualitativ höherwertige Daten unterstützen das Modell, indem die darin getroffenen Annahmen in den Daten wiederzufinden sind. So sind Einkommen aus einem abhängigen Arbeitsverhältnis z.B. anders zu berücksichtigen als das volatilere Einkommen aus einer unabhängigen Beschäftigung. Im Mittel können beide gleich sein, jedoch ist das Risiko einen Kredit zumindest stunden zu müssen höher bei einem volatileren Einkommen. Entsprechend der im Modell getroffenen Annahmen wird es eher auf die eine oder andere Variante ausgerichtet sein.

Die Daten sind dementsprechend angepasst für das Modell bereitzustellen. Sind diese nicht verfügbar, muss das Modell für die andersartigen Daten angepasst werden.

Ein verlässliches Ergebnis benötigt daher Daten und Modelle, die einander unterstützen. Andernfalls kann die Verwendung nicht passender Daten zu einer ungewollten Justierung des Modells und einer Verzerrung der Ergebnisse führen.

Fazit und Ausblick

Ein erfolgreiches datengetriebenes Geschäftsmodell nutzt die Synergien aus den Modellen und der zur Verfügung stehenden Daten. Dabei wird die Qualität der Daten beachtet und im Modell berücksichtigt. Zugleich werden die Daten so erhoben, dass sie in dem Modell verwendet werden können, ohne zu Verzerrungen zu führen.

In der heutigen Zeit stehen in der Regel ausreichend viele Daten bereit oder können relativ kurzfristig erhoben werden. Häufig werden diese jedoch nicht in dem Maße genutzt, wie sie genutzt werden können, da die Anpassungen der Systeme und Berechnungen zu kompliziert sind. Gerade eine Automatisierung dieser Prozesse auch durch Nutzung von künstlicher Intelligenz findet dadurch eher selten Anwendung.

Um die aktuellen datengetriebenen Geschäftsmodelle im Sinn einer dynamischen Effizienz anzupassen, wären daher die Strukturen zur Datenhaltung, -aufbereitung und -verarbeitung in diesem Sinne anzupassen



Jonathan Mislér ist Associate Manager bei Eurogroup Consulting und Experte für Datenmanagement und Umsetzung von IT-Projekten in der Finanzbranche. Seine Beratungsprojekte beschäftigen sich mit der Einführung von Datenmanagement mit dem Fokus auf Datenqualitätsmanagement, Transformationen von Unternehmen sowie der Einführung eines Kernbanksystems. Jonathan Mislér hat nach seinem dualen Studium der BWL einen Master of Science in Economics, Finance and Philosophy abgeschlossen.

Über den DataNavigator

Unterschiedliche Autoren beleuchten in der Reihe DataNavigator die vielfältigen Facetten des Datenmanagements und zeigen Ihnen interessante Aspekte auf. Der DataNavigator wird von Eurogroup Consulting regelmäßig zweimonatlich herausgegeben und möchte Ihnen interessante Denkanstöße liefern. Wir freuen uns, Ihnen mit dem DataNavigator unsere Erfahrungen weiterzugeben und stehen Ihnen für die Diskussion gerne zur Verfügung! Sie erreichen Autoren und Herausgeber bei Anregungen und Feedback unter der E-Mail-Adresse datanavigator@eurogroupconsulting.de.

Bisher erschienen:

- #1 Dr. Marco Weiß – *Welt und Wesen des Datenmanagements*; September 2020
- #2 Mario Kubik – „DLT“ – *Potenzial zur Zukunftstechnologie?*; November 2020
- #3 Jonathan Misler – *Verwendung von Daten und Modellen in datengetriebenen Geschäftsmodellen*; Januar 2021