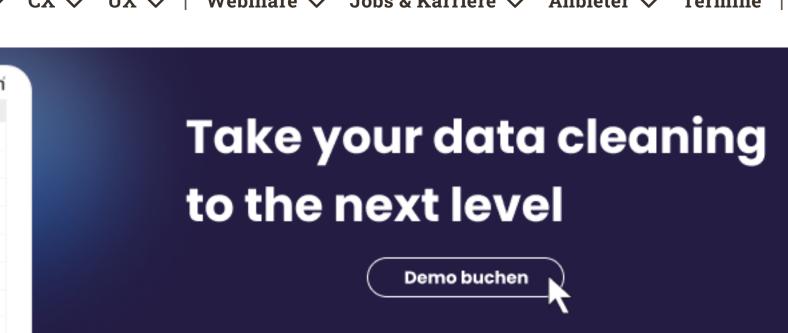
CONSULTING.de

PERSONALintern.de

Login

marktforschung.de - Das Portal für Marktforschung, Data Analytics und Insights.



Top-Themen: dossier.PLUS: UX Research Days 2024 ♦ Watchlist 2025 TOCii - Top Of Class

marktforschung.de > Marktforschung > Marktforschung als Data Pipeline: Datenflüsse kontinuierlich Analysieren

PAUL SIMMERING, Q | AGENTUR Marktforschung als Data Pipeline: Datenflüsse kontinuierlich Analysieren

iii veröffentlicht am 11.05.2021

j'ai lancé une malédiction

Gibt es eine Möglichkeit, Marktforschungsergebnisse stetig aktuell zu halten und deren Altern entgegenzuwirken? Paul Simmering,



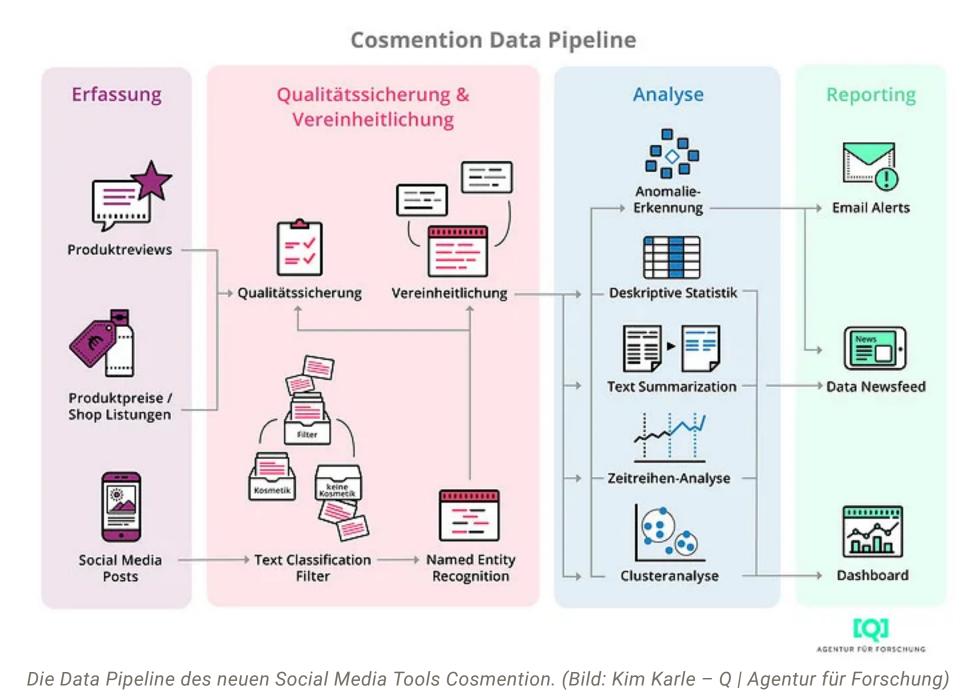
Von Jahr zu Jahr wird das Datenangebot größer: Es werden intern bei mehr Unternehmen systematisch Daten erfasst, die Tools für Umfragen werden besser und die Zahl der externen Datenanbieter steigt. Quantitative Marktforschungs-Projekte greifen sich typischerweise einen Datensatz mit einem festen Start- und Enddatum und produzieren einen Report, der für diesen Zeitraum Insights liefert. Damit beginnt der Report zu altern: in FMCG Märkten wie Mode oder Kosmetik sind die Ergebnisse bald nicht mehr aktuell. Eine Alternative zu diesem Vorgehen ist eine Data Pipeline, d.h. ein kontinuierlicher Marktforschungsprozess, bei dem Ergebnisse immer aktuell gehalten werden. Die Q | Agentur für Forschung hat mit diesem Ansatz das KI-gestützte Social Media Tool Cosmention entwickelt. Cosmention findet neue Kosmetik Trends und ermöglicht es, Produkte, Brands, Zutaten, Texturen und Hashtags zu beobachten. Dieser Artikel erklärt den Pipeline-Ansatz, zeigt die Usecases sowie die Vor- und Nachteile.

Kontinuierliche Datenerfassung

Anstatt Daten als einen festen Datensatz zu sehen, versteht man die Daten als einen Fluss. Sie strömen kontinuierlich in die Analyse, werden verarbeitet und kommuniziert. Mögliche Datenquellen sind Programmierschnittstellen (APIs), z.B. von Social Media Plattformen, automatisiert eingehende Fragebögen, Produkt-Reviews und andere Feedback-Systeme. Marktdaten wie Einkaufspreise, Produktionsmengen und Absatzzahlen können darin auch eingehen. In einer Data Pipeline werden Daten automatisch geprüft, bereinigt und angeglichen, sodass die Datenquellen kombiniert werden können. Datenquellen können unterschiedliche Namen für dieselbe Sache verwenden, z.B. Produktnamen können ausgeschrieben werden oder als GTIN (Global Trade Item Number), oder ASIN (Amazon Standard Identification Number) benannt sein. Für die Analyse muss eine eindeutige Bezeichnung verwendet werden, um Duplikate und Fehlzuweisungen zu vermeiden.

Data Pipelines konstruieren

Data Pipelines werden iterativ weiterentwickelt. Mit neuen Datenquellen, Analysebausteinen und Reporting-Outputs können sie erweitert werden, um mehr Perspektiven und Use Cases abzudecken. Eine Möglichkeit der Umsetzung bieten Microservices in der Cloud, also kleine Apps die regelmäßig ausgeführt werden und von einem Workflow-Management Programm wie Apache Airflow gesteuert werden.



Die Frequenz einer Data Pipeline drückt aus, wie lange es dauert, bis ein neuer Datenpunkt im

Reporting landet. Typische Optionen sind wöchentlich, täglich oder direkt in Echtzeit. Die

technischen Anforderungen für Echtzeit-Analysen sind bedeutend höher als für tägliche Updates. Wenn manuelle Arbeitsschritte nötig sind, werden hohe Frequenzen arbeitsintensiv. Die Grundidee ist daher: alles automatisieren, was sich automatisieren lässt. Machine Learning in der Data Pipeline

Aufgaben wie Inhalte kategorisieren und codieren können von neuronalen Netzen, welche Texte, Audio und Bilder verarbeiten, erledigt werden. Das Trainingsmaterial für Supervised Learning muss aber zunächst von Menschen annotiert werden. Besonders lohnt sich der Einsatz von Unsupervised Learning Methoden, etwa zur Cluster-Analyse. Diese bekommen durch die Pipeline durchgehend neues Trainingsmaterial eingegeben und lernen automatisch mit.

Beim Betrieb einer Machine Learning Pipeline über einen längeren Zeitraum muss Model Drift vermieden werden. Es tritt auf, wenn sich die eingehenden Rohdaten verändern, während das Modell, welches mit älteren Daten trainiert wurde, gleichbleibt. Dadurch verringert sich die Genauigkeit der Analyse. Um das zu verhindern ist regelmäßiges Testen und Updaten der Modelle mit neuen Trainingsdaten notwendig.

Anomalieerkennung

Die Flut von eingehenden Daten kann nicht händisch geprüft werden und ein Dashboard wird meist nicht ständig beobachtet. Stattdessen können Anomalien, d.h. ungewöhnliche Entwicklungen, automatisch erkannt werden. Beispiele für Anomalien bei Cosmention sind neue Hashtags, Produkte, die kurzfristig häufiger diskutiert werden oder Produkte, die besonders positive oder negative Reviews erhalten. Für die Erkennung von Anomalien gibt es zahlreiche Machine Learning Modelle. Die Bedingungen für ein Alerting können aber auch mit einfachem SQL definiert werden. Dabei ist Finetuning erforderlich, um nicht zu oft Alerts zu senden (Spam), aber auch um wichtige Entwicklungen nicht zu verpassen.

Automatisches Reporting

Der letzte Baustein in der Data Pipeline ist das Reporting. Dashboards bieten Kunden eine immer aktuelle Übersicht und Möglichkeiten zu selbstständiger Analyse mit Filtern. Details und Anomalien können darin aber untergehen. Ein Dashboard kann mit einem Newsfeed ergänzt werden, in dem die Anomalien dargestellt werden. Bei Cosmention orientierte man sich am Format des Twitter Newsfeeds. Dieser wurde personalisierbar gemacht, indem Kunden einstellen können, welche Produktkategorien für sie interessant sind. Die News-Stories aus dem Feed können zusätzlich als E-Mail verschickt werden und von dort auf das Dashboard verweisen.

Unterschiede zu klassischem Reporting

Eine Data Pipeline bleibt stets aktuell und liefert Insights auf direktem Weg, was Kunden einen Timing-Vorteil im Markt schaffen kann. Im kontinuierlichen Betrieb wird die Data Pipeline immer weiter verbessert, sodass sie eine höhere Genauigkeit erreichen wird als eine ad hoc Analyse. Zusätzlich wird dadurch ein Datensatz mit historischen Daten aufgebaut, welcher für übergeordnete, strategische Fragen ausgewertet werden kann.

Der Preis für diese Vorteile sind Komplexität und Investment beim Einrichten der Pipeline. Das Thema Operations gewinnt an Bedeutung, denn für eine zuverlässige Pipeline ist das Monitoring, Testen und Updaten der Module entscheidend. Bis alles ineinander greift, müssen viele Hürden überwunden werden und Prozesse mit internen und externen Datenanbietern, sowie mit den Nutzenden der Outputs abgestimmt werden. Und auch wenn alles steht, kommt es durch Ausfälle oder Änderungen bei Datenquellen sowie auch im Anforderungsprofil der Kunden zu neuem Arbeitsaufwand.

Marktforschenden nicht typisch ist. Daher ist ein interdisziplinäres Team von Programmierenden und Marktforschenden wichtig. Das Projekt darf sich nicht zu einer Technokratie entwickeln, sondern Kunden müssen weiter im Mittelpunkt stehen. Gleichzeitig ist die Entwicklung einer Engineering Culture und einem hohen Software-Qualitätsverständnis wichtig. Die Anlage einer Data Pipeline lohnt sich vor allem in der betrieblichen Marktforschung sowie bei langfristiger Betrachtung eines Marktes für mehrere Kunden. Der Nutzen aus kontinuierlichen Updates ist branchenspezifisch. Besonders in FMCG mit kurzen Trendzyklen zahlt sich das Investment aus.

Eine Data Pipeline ist ein Softwareprojekt und sie erfordert ein neues Skillset, dass bei

Über den Autor



Paul Simmering ist Data Scientist bei Q | Agentur für Forschung. Er leitet die Entwicklung von Cosmention, dem KI-gestützten Social Media Monitoring Tool für Kosmetik. Daneben ist er an Projekten in FMCG und Pharma beteiligt und beschäftigt sich mit Natural Language Processing, Big Data, Netzwerkanalysen und interaktiver Datenvisualisierung.



Weitere Informationen zum Unternehmen auf marktforschung.de:

Q | Agentur für Forschung GmbH In Research We Trust Q | Agentur für Forschung GmbH ist eine mehrfach ausgezeichnete Insights- und Marktforschungs-Agentur. Wir finden die Antworten, die unsere Kunden suchen - und schaffen mit ihnen Grundlagen für Lösungen, die wirklich...

Kommentare:

Noch keine Kommentare zu diesem Artikel. Machen Sie gerne den Anfang!

Bitte loggen Sie sich ein, um zu kommentieren: Jetzt kostenfrei Registrieren 2+ 🙎 Login

Mehr zum Dossier: KI in Marktanalyse und Kundenmanagement **Artikel dieser Ausgabe:**

KI in Marktanalyse und Kundenmanagement

EDITORIAL & MEHR

Marktforschung als Data Pipeline: Datenflüsse kontinuierlich Analysieren

PAUL SIMMERING, Q | AGENTUR

lyse aus User Generated Content

CHRISTIAN WINKLER

PROF. DR. STEFANIE SCHOLZ, DR.

KI-gestützte Persona-Ana-

Künstliche Intelligenz beim Ad-Test: Revolution im Anmarsch

INTERVIEW MIT MARC-DAVID

ROMPF, DLA DIGITAL LEADERS AD-

PROF. DR. MICHAEL FRETSCHNER,

SMART IMPACT

DR. MICHAELA MENKEN, BOLD&BLUNT GMBH Bias in / bias out - Warum wir mehr Datenethik brauchen

ANDRÉ LANG, INSIUS UG Black Box KI: Worauf Sie beim Einsatz achten sollten

Was KI-basierte Research-**Assistenz heute leisten** kann und was (noch) nicht

PAVOL SIKULA, ASKBRIAN

VISORY "Wer in diesem Wettbewerb bestehen möchte, muss intern AI-Skills aufbauen"

MANTO

ten sein"

HANS-PETER RIPPER VON SY-ANDERA GADEIB, MARKTFOR-SCHERIN Psychographische Seg-KI in der Marktforschung – mentierung: Das "Warum" Die richtige Balance aus

ELEMENTS

CHRISTOPHER HARMS, SKOPOS

Mit Data Thinking von der

Idee zum Datenprodukt: KI

zielgerichtet einsetzen

Arche Neo: Über den Einsatz von Archetypen in der Marktforschung

DR. KLAUS HOLTHAUSEN

hinter Kaufentscheidungen entdecken INTERVIEW MIT TJARK DÖNNI, TALIOX GMBH "Eine große Herausforde-

rung kann die Beschaffung

qualitativ hochwertiger Da-

DR. HANS-JÜRGEN FRIESS, IPSOS & DR. NORBERT ARNOLD, KAS Künstliche Intelligenz muss für die Menschen erlebbar werden

Automation und Intuition

Mehr Praxis, weniger Hype: KI zündet die nächste Stufe

BENEDIKT BONNMANN, ADESSO

die Marktforschung verändern

KI IN DER MARKTFORSCHUNG

Wie KI-basierte Methoden



N RD LIGHT BRANCHENNEWS

Weitere Highlights auf marktforschung.de



ins Management-Team, Bonsai Strategy...

und dem ADM

Neues von September, Bon-

sai Strategy, Ipsos, YouGov

September freut sich über den

Aufstieg von Eva Daniela Bock



von marktforschung.de.

marktforschung de pesche

Die tägliche Online-Zeitung von marktforschung de seit 2007

äußerst knappen Voting sicherte sich Telius. Ai den...

7. START-UP ONLINE PITCHES

Telius.Ai triumphiert! -

Präsentationen und einem

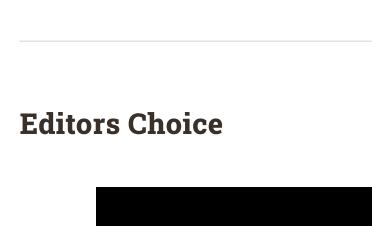
den Finale

Nach fesselnden

Nachbericht zum packen-

branchenrelevanten Insights, die Sie in der Produktentwicklung, Vermarktung und Vertrieb

voranbringen.





::: alle anzeigen

Smart News Fachverlag GmbH Max-Ernst-Str. 4 D-50354 Hürth (bei Köln) Tel.: +49 (0)2233 - 460 78 60

 □ redaktion(at)marktforschung.de für alle anderen Anliegen: ■ service(at)marktforschung.de

Redaktionelle Anliegen bitte an:

 Autoren Archiv Lexikon Studien

Quicklinks

 Kontakt / Anfahrt Über uns Impressum / Datenschutz

Podcasts

© 2007 bis 2024 Alle Rechte bei marktforschung.de