

Peter Just

Big Data und Künstliche Intelligenz im Bestandsmanagement

Ein erster Blick auf das Potenzial medienbezogener Leistungskennzahlen

Öffentliche Bibliotheken richten sich bei der Auswahl ihrer Medien stärker als Wissenschaftliche Bibliotheken an der Nachfrage der Nutzer aus. Eine genaue Messung der Nutzung ist daher wichtig, um die Leistungsfähigkeit eines Bestandes einzuschätzen. Existierende Leistungskennzahlen (Umsatz, Verfügbarkeit, Absenz, Aktivierungsgrad, Hitliste und Nullerliste) beziehen sich immer auf eine Anzahl mehrerer Medien (Bestand oder Bestandssegment). In diesem Aufsatz möchte ich anhand des Umsatzes aufzeigen, welche neuen Anwendungsmöglichkeiten entstehen, wenn man Leistungskennzahlen nicht nur für einen Bestand, sondern auch für jedes einzelne Medium berechnet.

Bisherige Berechnung des Umsatzes

Im »Handbuch Bestandsmanagement in Öffentlichen Bibliotheken«¹ wird der Umsatz wie folgt berechnet:

$$\text{Umsatz} = \frac{\text{Summe der Ausleihen im Bestand des Jahres } X}{\text{Bestand zum Ende des Jahres } X}$$

Titel	im Bestand seit	Ausleihen 2018	Bestandsumsatz 2018
Sinnlose Wettbewerbe	21.12.2012	6	6,5
Speed Cleaning	30.05.2018	7	6,5
Naturheilverfahren bei Borreliose	28.10.2016	9	6,5
Schwindsucht	04.09.2018	4	6,5

Tabelle 1: Berechnung des Umsatzes eines Bestands nach der herkömmlichen Methode

Der Umsatz drückt damit aus, »wie viele Ausleihen eine Medieneinheit im Bestand durchschnittlich reflektiert«. Als Ergebnis erhalten wir einen Durchschnittswert für ein Jahr X, der für alle Medieneinheiten im Bestand gilt (siehe Tabelle 1). Der Umsatz für den Bestand ist in diesem Beispiel 6,5 Ausleihen pro Medium. Der Umsatz des einzelnen Mediums wird in der Regel nicht bestimmt, entspricht aber der Anzahl der Ausleihen des Mediums. Wenn wir nun die Medien nach ihrem Umsatz sortieren, um ihre Leistung untereinander zu vergleichen, dann ist das Buch von »Naturheilverfahren bei Borreliose« mit neun Ausleihen das leistungsstärkste, gefolgt von dem Titel »Speed Cleaning« mit sieben Ausleihen, »Sinnlose Wettbewerbe« mit sechs, und »Schwindsucht« mit vier Ausleihen.

Meiner Einschätzung nach ist eine solche Sortierung der Liste so ungenau, dass sie nicht nutzbar ist. Das Problem ist, dass die Zeiteinheit, in der sie misst (Kalenderjahr), zu groß ist. Titel, die im Laufe des Jahres erworben werden, haben weniger Chancen entliehen zu werden als Titel, die das ganze Jahr im Bestand sind. Das in dem obigen Beispiel genannte Buch »Speed Cleaning« hatte in 2018 zum Beispiel circa 40 Prozent weniger Chancen entliehen zu werden, als die Medien die das ganze Jahr über im Bestand waren. Eine Berechnung einer Leistungskennzahl für einzelne Medien, mit der diese anschließend untereinander verglichen werden können, ist mit der bisherigen Berechnungsmethode damit nicht möglich.

Neue Berechnungsmethode für den Umsatz

Neue Berechnungsmethode für den Umsatz

Daher soll hier eine Berechnung des Umsatzes vorgeschlagen werden, die

- a) mit der tatsächlichen Anzahl an Tagen rechnet, die ein Medium im Bestand ist und

Titel	Ausleihen 2018	im Bestand seit	Tage im Bestand in 2018	Medienumsatz 2018 pro 30 Tage	Bestandsumsatz 2018 pro 30 Tage
Sinnlose Wettbewerbe	6	21.12.2012	365	0,493150685	0,81
Speed Cleaning	7	30.05.2018	215	0,976744186	0,81
Naturheilverfahren bei Borreliose	9	28.10.2016	365	0,739726027	0,81
Schwindsucht	4	04.09.2018	118	1,016949153	0,81

Tabelle 2: Berechnung des Bestandsumsatzes nach neuer Berechnungsmethode (Erhebungsdatum der Daten: 31.Dezember 2018)

b) mit einer Zeiteinheit rechnet, die diese tatsächlichen Tage im Bestand auch abbilden kann und so die Darstellung des Umsatzes für ein einzelnes Medium erlaubt. Die vorgeschlagene Gleichung soll sein:

$$\text{Medienumsatz im Kalenderjahr pro 30 Tage} = (\text{Ausleihen des Mediums im Kalenderjahr} / \text{Tage des Mediums im Bestand im Kalenderjahr}) * 30 \text{ Tage}$$

Mit »Tage des Mediums im Bestand im Kalenderjahr« ist die Anzahl der Tage gemeint, die das Medium in dem Kalenderjahr im Bestand war. Die Zahl ergibt sich aus der Differenz zwischen dem Freischaltungsdatum (Datum ab wann ein Medium zum ersten Mal ausleihbar war) und dem Erhebungsdatum.²

Die Zeiteinheit »30 Tage« dient vor allem der besseren Darstellbarkeit. Wir könnten auch den Medienumsatz pro Tag berechnen. Die Aussage »Das Medium ›Speed Cleaning‹ wurde in 2018 durchschnittlich pro Tag 0,0325581 Mal ausgeliehen« ist zwar korrekt, aber schwer handhabbar.

Aus den einzelnen Medienumsätzen lässt sich auch der Umsatz eines jeweils festzulegenden Bestandes berechnen. Die Gleichung hierfür lautet:

$$\text{Summe der einzelnen Medienumsätze} / \text{Anzahl der Medienumsätze}$$

Auf die Medien in Tabelle 1 angewendet ergibt das Ergebnisse, die in Tabelle 2 dargestellt sind.

Welche Vorteile hat nun die Neuberechnung?

a) Sie ist genauer, da sie den Zeitraum zugrunde legt, in denen das Medium tatsächlich ausgeliehen werden kann.³ So weist zum Beispiel das in Tabelle 1 als letztes gerankte Medium »Schwindsucht« mit vier Ausleihen in Tabelle 2 den besten Medienumsatz auf.

b) Durch diese zusätzliche Genauigkeit lassen sich die einzelnen Medienumsätze überhaupt erst aussagekräftig in Beziehung setzen.

c) Es kann weiterhin für einen Bestand oder ein Bestandssegment ein Umsatz berechnet werden.⁴ Die Zusammensetzung dieses Bestandsumsatzes kann allerdings jederzeit in Einzelwerte aufgeschlüsselt und so wesentlich aussagekräftiger analysiert werden.

d) Der Umsatz eines Bestandes lässt sich genauer analysieren, da sich die Verteilung der Einzelwerte mit den Werkzeugen der quantitativen Datenanalyse betrachten lassen (weitere Lageparameter, Streuung et cetera).

1 Frauke Schade, Konrad Umlauf: *Handbuch Bestandsmanagement in Öffentlichen Bibliotheken*. Berlin: de Gruyter Saur, 2012 (Bibliothek- und Informationspraxis; 46), S. 248

2 Beispiel: Medium »Speed Cleaning«: Im Bestand seit 30. Mai 2018, Erhebungsdatum 31. Dezember 2018 = 215 Tage im Bestand.

3 Bei dem Medienumsatz bewerten wir jede Ausleihe gleich – egal wie lange die Nutzer*in das Medium tatsächlich entliehen hatte. Eine noch genauere Leistungskennzahl (Medienabsenz) erhalten wir, wenn wir zusätzlich noch die tatsächliche Anzahl an Ausleihtagen erheben würden. Wenn diese Zahlen erhoben werden können, ist die Medienabsenz dem hier vorgestellten Medienumsatz vorzuziehen.

4 Aufgrund der unterschiedlichen Berechnungsmethode ist der Umsatz aber auch nicht vergleichbar mit dem nach der alten Methode berechneten Umsatz.

5 Eine Hitliste »ist ein Ausdruck derjenigen Exemplare aus dem Bibliothekssystem, die im Zeitraum z häufiger als x-mal ausgeliehen wurden« (Schade (Anm.1) S. 249).

6 Bei Notationen mit schlechtem Medienumsatz kann die Zahl der Medien verringert werden. Für Notationen mit hoher Performance können zusätzliche Titel erworben werden. Bei Notationen mit hoher Anzahl an Titeln und hohem Umsatz kann über die Aufteilung der Notation nachgedacht werden.

Einige Anwendungsmöglichkeiten im klassischen Bestandsmanagement

Mit solchen für das jeweilige Medium berechneten Leistungskennzahlen lassen sich Bestände genauer, detaillierter und flexibler analysieren. Folgende Vorteile für klassische Arbeiten im Bestandsmanagement gibt es:

1. Eine nach Leistungskennzahlen geordnete Liste kann zur Stafflung genutzt werden. Klassischerweise wird dazu eine Hitliste⁵ herangezogen. Da diese nicht beachtet, wie lange ein Medium im Bestand ist (siehe oben), ist eine Liste sortiert nach Medienumsätzen genauer.

2. Eine nach Medienumsätzen geordnete Liste kann zur Aussonderung genutzt werden. Im Gegensatz zu der üblicherweise genutzten Nullerliste können so auch Titel mit geringer Performance erfasst werden.

3. Auf Medieneinheiten bezogenen Leistungskennzahlen können neu kombiniert werden (ohne neue Datenerhebung, ohne neue Datenbereinigung). Hier lassen sich zahlreiche Anwendungsmöglichkeiten vorstellen, zum Beispiel kann eine Berechnung von Medienumsätzen nach Notation für die Freihand- und Systematik-Revision genutzt werden.⁶ Denkbar wäre auch eine Berechnung nach Erwerbungsarten, die die Leistung verschiedener Erwerbungsarten wie Standing Order, Approval Plan, Stafflung et cetera miteinander vergleicht.

Welche neuen Anwendungsmöglichkeiten im Bestandsmanagement gibt es?

Darüber hinaus gibt es aber auch gänzlich neue Möglichkeiten mit medienbezogenen Leistungskennzahlen. Speichert man die Leistungskennzahlen eines Mediums im Bibliothekssystem (quasi als »Leistungsmetadaten«), ließen sich diese Leistungskennzahlen auf andere Informationen des Mediums übertragen, zum Beispiel

- auf die Autor*in: Es ließe sich ein durchschnittlicher Medienumsatz der Medien einer Autor*in errechnen.
- auf die Schlagworte: Es ließe sich ein durchschnittlicher Medienumsatz von Medien errechnen, die alle ein bestimmtes Schlagwort zugewiesen bekommen haben,
- auf die Notation: Es ließe sich ein durchschnittlicher Medienumsatz pro Notation errechnen.

Die so entstandenen Metadaten können für einen Empfehlungsdienst (Recommender-Systeme) genutzt werden. Bei der Beurteilung von Neuerscheinungen kann die Lektor*in sich nun informieren:

Hat die Autor*in bereits Medien publiziert und welchen durchschnittlichen Medienumsatz haben diese erzielt? Welche Schlagworte sind dem Medium zugewiesen und welchen durchschnittlichen Medienumsatz erreichen diese bisher in meiner Bibliothek? Welchen durchschnittlichen Medienumsatz haben die Medien mit der der Neuerscheinung zugewiesenen Notation?

Darüber hinaus kann ein solches System zum lernenden System werden, wenn die Leistungskennzahlen von Medien, die aufgrund der Empfehlung des Recommender-Systems erworben wurden, wiederum in die Leistungskennzahlen einfließen.

Peter Just hat Bibliothekswissenschaft, Politikwissenschaft und Soziologie an der HU-Berlin (M.A.) studiert und arbeitet zurzeit an der Zentral- und Landesbibliothek Berlin. Arbeitsschwerpunkte sind die Analyse, Implementierung und Sichtbarmachung digitaler Angebote und die Entwicklung, Erhebung und Anwendung von Leistungskennzahlen im Bestandsmanagement. – Kontakt: peter.just@zlb.de



Fazit

Leistungskennzahlen für das einzelne Medium lassen sich wesentlich genauer berechnen, als zurzeit üblich. Wenn wir die uns vertrauten Berechnungsmethoden präzisieren, können wir sowohl klassische Aufgaben im Bestandsmanagement verbessern als auch neue Anwendungen etablieren. Daher sollten Bibliotheken die bisherigen Leistungsindikatoren überarbeiten und den neuen Möglichkeiten der Datenverarbeitung anpassen.

ANZEIGE

Ideal für Ihre
XXL-Medien!

Erleben Sie den Platzgewinn: NORIS Media-Boxen

- Ideale Verpackung für Hörbuch-Kollektionen oder Serien von bis zu 30 Discs
- Flexibel dank Varianten für 8, 14, 24 oder 30 Discs und einzeln einheftbaren Doppeltaschen
- Kinderleicht eingearbeitet – so geht's:
<http://bit.ly/NORIS-Media-Boxen-Video>

Infos und Bestellung:
<http://bit.ly/NORIS-Media-Boxen>