

Günter Jürgensmeier

Die elektronische Bargfelder Ausgabe (eBA) 2.0

<https://arno-schmidt-stiftung.de/eba/search>

Noch kaputter by Design

Vor einem Jahr, am 10.12.2021, machte ich darauf aufmerksam, dass die seit Januar 2019 im Netz verfügbare eBA (elektronische Bargfelder Ausgabe der Werke Arno Schmidts) kaputt ist, und zwar »kaputt by design«, also nicht reparierbar. Das Hauptproblem besteht darin, dass die als Grundlage verwendete Open-Source-Software (Apache Solr) für Literatur ungeeignet ist, da sie auf Business-Webseiten mit kurzen Informationseinheiten ausgerichtet ist, wie sie bei Katalogen, Produktlisten, Videosammlungen und ähnlichen Datenbeständen anzutreffen sind. Mein Fazit lautete, dass die eBA wegen des ungeeigneten Unterbaus auch durch Überarbeitungen nie »gut« werden könne, sondern aufgrund der unüberwindbaren Probleme immer »ungenügend« und für wissenschaftliches Arbeiten ungeeignet bleiben würde.

»Niemand hat die Absicht,
eine Mauer zu errichten!«

Bei der ersten Ausgabe der eBA von 2019 war der Seitenwechsel die Mauer, über die hinweg Zusammengehörendes nicht gefunden werden konnte. Da scheint man sich bei der Pagina GmbH gedacht zu haben: »Halt mal mein Bier, das kann ich *noch* schlechter!« Nun ist *jeder Absatz*

eine durch ein solides Mäuerchen vollständig isolierte Sucheinheit!

»Äu-ßerst-merkwürdich!!« Denn das bedeutet, dass sich der Umfang der zwar logisch zueinander gehörigen, aber nicht gemeinsam durchsuchbaren Passagen grob geschätzt verachtfacht hat! Jeder Schmidt-Leser wird sich erinnern, wie Gesprächspartner sich in den Funkdialogen unterbrechen, wie in den Typoskripten ein Protagonist den Satz eines andern fortsetzt oder wie Schmidt Teile einer Aussage dadurch hervorhebt, dass er sie auf eine separate Zeile stellt. Diese Teilabsätze stehen nun für sich und keine Suchmacht der Welt bringt mit diesem System wieder zusammen, was zusammen gehört.

Ein Beispiel muss reichen: Ich erinnere mich, dass Schmidt irgendwo sagt, dass *Finnegans Wake* verschlüsselt ist. Wenn ich das nun bei der eBA eingabe findet sie – nichts. Wenn man die Stelle anderweitig ausfindig macht, sieht man auch sofort warum:

2

Seit 20 Jahren nun befindet sich 1 Buch in unserer Welt, über das bereits eine ganze Literatur zu entstehen im Begriff ist :

James Joyce, «FINNEGANS WAKE».

Ein Buch, so stark verschlüsselt, daß selbst im Herkunftsland, im angelsächsischen Sprachraum, ein «Skeleton Key» notwendig wurde, um wenigstens 40% davon annähernd in Einfach=Englisch umzuschreiben. Solch kommentierendes Umschreiben ist nötig, weil Joyce mit Sprache und Orthografie derart frei geschaltet hat – zuweilen schön und sinnreich; zuweilen abstrus – daß schon das bloße «Erkennen» der Basis=Worte Schwierig-

Schmidt hat Autor und Titel durch zentriertes Anordnen in einer separaten Zeile hervorgehoben, »verschlüsselt« steht erst im folgenden Absatz. So etwas kann nun an einigen 1000 Stellen passieren.

»Äu-ßerst-merkwürdich!!«


Was hat sich sonst getan?

AND ist nun der Standardoperator, es werden also Such-
einheiten gefunden, in denen *alle* angegebenen Wörter
enthalten sind.

Die Stoppwörter sind raus, man kann nun also auch
nach Phrasen suchen, die nur aus häufigen Wörtern be-
stehen wie »nur Alles in Allem« oder »bis ich NICHT mehr
bin«.

Man kann mit »phrase:« nach Phrasen inkl. Wildcards
und Fuzzy-Operator suchen, beispielsweise mit **phrase:**
"Kaff auch Mare~ Cris*". Wenn man allerdings nach einem
Einzelwort suchen möchte, z.B. nur nach "Cris*", dann
geht das nur, wenn man die Anführungen weglässt:
phrase:Cris*. (Und wenn man bei der Suche in der Nase
popelt, muss man vermutlich *wieder was anderes* ange-
ben!) Immerhin gibt die Software auch Fundstellen für
phrase:"ni* ni* ni* ni*" aus, nämlich das leitmotivische
»Nichts Niemand Nirgends Nie!« aus *Kaff*. Die Markierung
der Fundstellen und die Ausgabe des Kontexts gelingen
der Software dann allerdings nicht mehr:

*Nichts Niemand Nirgends Nie ! : Nichts Niemand Nirgends Nie ! : (die
Dreschmaschine rüttelte schtändig dazwischen, wir konnten
sagen & denken was wir wollten. Also lieber bloß zukukken.)*

*»'dollaus. -«; und ihr Fuß zeigte liederlich eben=dort hinüber : ein
Knecht hob 1 Arm; (vorn=dran also  ich 1 Faust) : sofort*

#	Fundkontext	Fundstelle
1		Kaff auch Mare Crisium, Abt. 1, Bd. 3, S. 11
2		Kaff auch Mare Crisium, Abt. 1, Bd. 3, S. 12
3		Kaff auch Mare Crisium, Abt. 1, Bd. 3, S. 76
4		Kaff auch Mare Crisium, Abt. 1, Bd. 3, S. 171
5		Kaff auch Mare Crisium,

In der BA kommt zweimal der Spruch »más tiempo para el amor« vor. Wenn ich danach mit "**más tiempo para el amor**" suche, wird nur eine Stelle gefunden, da die Standardsuche die zweite Stelle nur dann findet, wenn man die dort verwendeten Anführungen mit angibt: "**»más tiempo para el amor«**". Das passiert nicht nur bei diesen Anführungen, sondern oft auch bei =, : oder nach Satzzeichen, die auf « folgen und alle mit angegeben werden müssen!

»Äu-Berst-merkwürdich!!«

Wenn man beide Stellen gleichzeitig und ohne Vorkennntnis der verwendeten Anführungen und Satzzeichen, also mit *weniger exakten* Angaben finden möchte, muss man »exakt« verwenden (!), also **exakt:"más tiempo para el amor"**. Die Programmierer sind hier wohl schon bei der Planung, spätestens aber beim Programmieren irgendwo falsch abgebogen.

»Äu-Berst-merkwürdich!!«

Mit doppeltem Bindestrich zusammengesetzte Wörter werden nun in vielen Fällen sowohl bei der Suche nach einzelnen Bestandteilen als auch bei der Eingabe ohne das = gefunden, also z.B.

la bimmel la bammel

oder

la=bimmel la=bammel

oder

labimmel labammel

Leider bin ich auch hier wieder auf Ausnahmen gestoßen, wo der doppelte Bindestrich eingegeben werden musste (seufz); z.B. funktioniert

"Nobel Mittelmäßigkeiten"

nur mit =.

Wo hat sich nichts getan?

»Um die Trefferquote zu erhöhen, werden« noch immer »für einzelne Suchbegriffe standardmäßig auch Varianten zu Umlauten, Numeri und Flexionsformen berücksichtigt.« Nun sind Computer bekanntlich doof und schon bei notdürftig vom Englischen auf einfachstes Hochdeutsch angepassten Sprachroutinen oft noch doofer; es ist also viel Mist unter den Funden zu finden, der unnötig viel Arbeit macht. Andererseits kommt das Progrämmle schon mit einfachen Wörtern wie »hörn« & »sehn« und »gehn« & »stehn« nicht klar, erkennt also keine Flexionsformen mehr; über die Eignung für Schmidts Verschreibkunst brauchen wir da gar nicht erst diskutieren.

Weiterhin hält diese Sprachfeh/funktion den Namen »Schaeffer« für jemanden, der schäfft & schafft und zeigt entsprechend viele »Funde« an (231 statt 9). Möchte man diesen Flexionskrepel vermeiden, und vermeint das mit »exakt« erreichen zu können, wird man enttäuscht, Formen wie »SCHAEFFER's« werden dann nicht gefunden. Erst mit **phrase:Schaeffer*** werden alle Neune entdeckt. Da Schmidt »Schaeffer« aber auch gelegentlich mit »ä« schreibt (z.B. im Briefwechsel mit Bernd Rauschenbach), wäre es schön gewesen, wenn auch die Vorkommen von »Schäffer*« mit ausgegeben worden wären, aber die muss man mit einer zweiten Suche herauszubekommen versuchen.

Ich kann hier unmöglich alles Fehlerhafte aufführen, das immer noch vorhanden ist. Ich empfehle, die in der Rezension der ersten Version monierten Probleme durchzuprobieren und dabei z.B. mit der Suche nach **"Zettel's Traum"** / **"Zettels Traum"** oder **Detâ'j** / **Detâ''j** zu beginnen. Es hat sich etwas geändert, besser wurd's dadurch nicht.

Hier nur noch drei Hinweise:

Bereits in der ersten Version der eBA fehlende Teile der BA, etwa Schmidts Nachbemerkung zu ZT und Auszüge aus Alice und Arno Schmidts Tagebüchern, wurden in dieser *nicht* ergänzt.

Bei Briefen, die auf der vorherigen Seite beginnen, ist der Briefschreiber nicht erkennbar und muss umständlich in der gedruckten Ausgabe nachgeschlagen werden.

Die Informationen zu den Suchmöglichkeiten sind weiterhin zu knapp gehalten. So werden wichtige Konzepte wie Platzhalter/Wildcards (*, ?; zuerst fälschlich als »Trunkierung« bezeichnet) und Reguläre Ausdrücke auch diesmal nicht ausreichend erläutert.

Welche *neuen* Probleme gibt es?

Es ist immer wieder vorgekommen, dass in der Ganzseitenanzeige falsche Passagen markiert waren, die keinerlei Bezug zur Abfrage aufwiesen.

Das Suchergebnis kann man nach Relevanz sortieren. Was hier jedoch die »Relevanz« ausmacht, ist nicht zu erkennen. Die Fundstellen scheinen völlig zufällig durcheinander geschüttelt zu werden. Wahrscheinlich ist hier eine ähnlich unwissende Sprachfunktion am Werk wie beim Flexionskrepel.

Bei der großen Unsicherheit der Germanisten in puncto Entstehungszeiten von Schmidts Werken wäre es von sehr großer Relevanz, wenn man die Fundstellen in chronologischer Reihenfolge anordnen könnte. (Die Daten dafür sind in den Textdateien enthalten.)

Fazit

Der Geschäftszweig »Publikationstechnologien« der Firma Pagina hat ein wenig an der eBA herumgedoktert und sie dabei noch kaputter gemacht.

Sehen Arno Schmidts Werke aus wie Kataloge, Produktlisten, Aufstellungen, Verzeichnisse, Mediensammlungen oder Websites? Nein! Warum versucht man bei Pagina dann, die Werke in die für solche Datenbestände gedachte Struktur von Apache Solr zu quetschen? Ich versteh es nicht! Es ist »äu-Berst-merkwürdich!!«

Warum wird dem Nutzer eine so intrikate Abfrageform zugemutet? Warum kann er nicht einfach wie bei Google Wörter und/oder Phrasen in "" (zusätzlich mit beliebigen Wildcards und Fuzzy-Operator) eingeben und die Software kümmert sich um den Rest? Warum muss der Nutzer umständlich **"»äu-Berst-merkwürdich!!«"** oder **exakt: "äu-Berst-merkwürdich"** eingeben, um den Spruch zu finden? Ich versteh es nicht! Es ist »äu-Berst-merkwürdich!!«

Wie kommt man auf solche Abfrageformen? Warum wird ein solches System vor der Veröffentlichung nicht getestet? Fragen über Fragen.

In dieser zweiten Version der eBA ist die Zahl der nicht findbaren Zusammenhänge verachtfacht worden, die Zahl der schlecht findbaren Wörter verhundertfacht. Die eBA ist damit kein Find- sondern ein Such-dich-tot-Mittel! Sie bleibt »a kludge: An ill-assorted collection of poorly-matching parts, forming a distressing whole.« Lesen Sie daher bitte das Fazit meiner furiosen Besprechung der ersten Version.

»Wir Auguren helfen uns schließlich noch aus dem Wirrwarr heraus, aber, wie es im ›Schimmelreiter‹ sehr richtig heißt: ›Gott gnad de Annern !«. –