

Günter Jürgensmeier

Die elektronische Bargfelder Ausgabe (eBA)
<https://arno-schmidt-stiftung.de/eba/search>

Kaputt by Design

Nachdem ich 1996/97 als Übergangslösung die *Bargfelder Ausgabe auf CD-ROM* erstellt hatte, wurde ich von der Arno Schmidt Stiftung nach dem Satz von *Zettel's Traum* auch mit der Programmierung einer elektronischen Bargfelder Ausgabe beauftragt. Leider war es mir aus privat-gesundheitlichen und geschäftlich-organisatorischen Gründen nicht möglich, dieses Projekt zu beenden. Ich stellte jedoch eine »E-BA-Engine« (die Serverseite) und eine einfache Benutzeroberfläche dafür fertig:

Abends unversehens als glücklichen Besitzer eines Paares derber Plattfüße wiederfindet? <Life begins at forty>; und bitter nicken.¶

[1/4:0111:019]

Freilich, wenn man Geld hätte Ich wüßte es jetzt schon richtig anzuwenden : ein winziges Häuschen in der Heide (achttausend höchstens; nicht wie diese Bausparkassen, die mit Zwanzigtausend um sich werfen, als wär's ein bloßer Silbenfall); im Ställchen eine Isetta; Eintausend erlesene Bücher : einmal in aller Ruhe die <Insel Felsenburg> durchgehen können, den <Nachsommer>, oder Lessing von **A bis Z;** zur Nacht ein richtiges Bett zum Drinniederlegen (nicht mehr dieses dürre indianerrote Gestelle von Schlafcouch!); nichts mehr ums liebe Brot schreiben zu brauchen, keine <experimentelle Prosa> mehr, keine feinsinnigen <Essays>, keine <Nachtprogramme>; an Uhren werden nur die lautlosen geduldet, die mit Sand und Sonne, oder höchstens im Korridor eine alte Standuhr, die alle Ewigkeiten, nachdem man vieles und vielfältiges gedacht hat, vor sich hin <Mnja> sagen. Den Mond untergehen sehen, über Wiesen einsamkeiten, ganz rot würde das silberne Meer zuenden sein, wenn es einack in Dornstaud und

	Vor	Kontext	Kurz	Bd	S	Z
0001	den <Nachsommer>, oder Lessing von	A bis Z; zur Nacht ein richtiges Bett zum Drinniederlegen (n	kSAu	1/4	0111	023
0002	g; die Summe ihrer Liebe galt, von	A bis Z, : Tieren.« / Allerdings auch das wieder in einer am	dMG	2/2	0410	039
0003	: KREUZ- UND QUERZÜGE DES RITTERS	A BIS Z. / Und : LEBENSÄUFE IN AUFSTIEGENDER LINIE¶) Wñh	bES	8/3	0122	017
0004	r, Fallada, L. Frank, uswusw., von	A bis Z, der ganze <Kürschner> durch, verwebten Reminiszenze	eGB	3/4	0055	017
0005	d verworrne Kreuz- & Querzüge von	A bis Z; <Die Schlacht am Berge Ripplah>; kleine & große Städ	eBPM	3/4	0074	016
0006	n=Abbildungen : MAY's Werk ist von	A bis Z anal stigmatisiert! (Vgl. W. LA BARRE's und H. DICKS	Sit	3/2	0089	037
0007	gt, = ich in meinem Fall also dies	<A bis Z> - mag der theilnehmende Leser den II. Band, ganz h	Sit	3/2	0259	014
0008	Unsinn; meinetwegen ist Unsinn von	A bis Z. Aber es giebt in dem Unsinn Grade: verzeihlichen; e	Jul	4/4	0020	133

8 Fundstellen in 0.140 / 0.006 Sek.

Die Fortsetzung, d.h. die Erstellung der Clientseite (des Webinterfaces), wurde der Pagina GmbH Publikationstechnologien übertragen. Aus mir unbekanntem Gründen ignorierte die Pagina GmbH jedoch alle gelieferten Vorarbeiten, das ganze Schmidt-Know-How, alle Tests & Tunings und programmierte eine eigene Lösung auf der Basis der Open-Source-Software Apache Solr.

Damit war das Unheil in der Schmidt-Welt.

Als im Januar 2019 die eBA freigeschaltet wurde, habe ich sie natürlich gleich ausprobiert, mußte sie jedoch schon nach kurzer Zeit als Werkzeug verwerfen: die Suchergebnisse waren überflüssig umfangreich, Wörter mit Sonderzeichen oder Apostroph machten Schwierigkeiten und das schönste daran, die Ganzseitenanzeige mit Hervorhebung der Funde, machte das Sichten der Fundstellen zu einer so zähen Angelegenheit, daß dies bei den zig Suchen, die ich manchmal innerhalb kürzester Zeit machen muß, nicht zu bewältigen war.

Als mein Chefredakteur Arno Tator Anfang September zu tweeten begann, sollten die Tweets bei Schmidt-Zitaten Links auf die Stellen in der eBA enthalten. Bei der Formulierung der Abfragen stieß er immer wieder auf kaum zu überwindende Schwierigkeiten. Er hat die Fälle an mich weitergereicht, und ich habe mir angesehen, woran es liegt.

Die eBA ist kaputt, und zwar »kaputt by design«, also nicht reparierbar.

Das Hauptproblem ist, daß die als Grundlage verwendete Open-Source-Software (Apache Solr) auf Webseiten ausgerichtet ist und für Literatur ungeeignet. Webseiten sind meist relativ kompakte Texteinheiten, bei denen die Suche nach Begriffen irgendwo auf der Seite sinnvoll ist.

Bei längeren Texteinheiten, wie etwa Kapitel, Szenen und Bilder bei Schmidt, ist es selten sinnvoll, Fundstellen zu melden, wo die gesuchten Begriffe weit auseinander stehen. Um das zu vermeiden ist bei der eBA »die Einheit, in welcher nach Treffern gesucht wird, die Druckseite der Bargfelder Ausgabe.«

Es ist (geline gesagt) eine wunderliche Entscheidung, den Text in zufällig beim Satz entstandene Einheiten aufzuteilen! Das führt dazu, daß z.B. das Spätwerk nun in rund 2200 vollständig voneinander isolierte Texteinheiten zerstückelt ist. Jeder Seitenwechsel ist eine unüberwindliche Grenze, die den eventuellen Zusammenhang mit dem folgenden bzw. vorhergehenden Text rigoros abschneidet. Die Folgen:

Ich erinnere mich, daß in AmG ein »Rundfunksender Paphos« erwähnt wird, die Suche nach **Rundfunksender AND Paphos** liefert jedoch kein Ergebnis, da die Wörter auf verschiedenen Seiten stehen (AmG 89f.). Ebenso bei **Promiscuität AND Perversionen** (AmG 168f.) oder **Eule AND Fleder*** (AmG 183f.). Die Abfrage **fußlang* AND finger** zeigt eine Seite in Poc, wo »fußlanger« und »Finger« weit auseinander stehen; eine Zeile, in der »fing« steht, wird als Zugabe noch mit markiert (»fing, finger, am fingersten«?). Die exakte Abfrage **"fußlangen Finger"** liefert auch kein Ergebnis, denn mitten durch die »fußlangen¶Finger des Bastard Marwenne« geht AmG 123f. ein Seitenwechsel.

Einige wenige weitere Beispiele für mit der eBA nicht findbare Stellen sollen das Problem deutlich machen:

"duftnde gestrip" (AmG 24f.)

markieren: !). Am EinhornLoch, im Grauholz: er schaut ungemein zeugerisch drein. Und drückt dann, lieb und muthig, das Gesicht ins **duftnde**

Gestrip: ! / Auch sie fährt mit Händen lustich auf und nieder an ihm; - ; auf Gliedsuche: - , - das sie vergeblich zu entblöß'n strebt. / Der Präser klemmt ihn sichtlich ... -)

"Erdachtn Gespräche" AND Gründgens (SdA 23f.)

GUIDE (angeregt. Ach!): »Der will an Deine Suse ran? (Ach, daher auch dies Greta=Grin'sn schon von weitem. – Am Hang dort; wo die >3 tanzndn Meerkatzen<; ah=ja.). Etwas kurz, wie?; (aber stemMich! Und Suse auf Dessn Schoß?; das heißt freilich, eine PulverTonne auf ein KohlenFeuer setzn!). – (?): ach Dér iss das?!«; (Der im >Grünen Boten<, und in der >Eidora<, die Sonette verzapft. Und diese knollijn **»Erdachtn Gespräche<:»**)

>Darwin und die Venus von Milo<; >Themistokles und Emilia Galotti<; >Simson mit einem Superintendentn<; >Magnus Hirschfeld und Gustav **»Gründgens<:»**; >Talma und Melanchthon<): »Wahnsinnich=komisch! Allein dies Pseudonym: >Halicacabum<!«

Makart* AND Hackländer (AmG 44f.)

BASTARD MARWENNE: »Ich will mein'n eignen Bart fressen, wenn ich solchn **»MakartKitsch ...«;** (er beendet seinen Satz mit dumpfem Aufstoß'n).»

GRETE: »Also für mich iss der Name **»HACKLÄNDER»** etwas, vor dem ich (vergebens!) bis ans äußerste Meer fliehen möchte. (Nicht, daß ich deswegn das Schwatzwerk Eurer modernen übertrainiertn Pegasusse lieber hätte.)«;

»Es ist irr=ratio=nal.«

Ein weiteres Hauptproblem ist, daß die eBA den Abstand von Wörtern nicht berücksichtigt. Google z.B. läßt in seinen Algorithmus die Nähe von Wörtern einfließen, wodurch die Google-Suche so gute Ergebnisse liefert. Die eBA kennt die Nähe von Wörtern nur bei der Phrasensuche (aufeinanderfolgende Wörter) oder mit expliziter Angabe des Abstands "**Wort1 Wort2**"~4, die nicht universell verwendet werden kann und auch jedesmal erneut angegeben werden muß. Einige Beispiele für die Nutzlosigkeit vieler Ergebnisse:

Bei ***silb* AND *weiß*** liefert die E-BA-Engine 42 Fundstellen, die eBA wegen des beliebigen Abstands zwischen den Wörtern 425 Ergebnisse!

Die Phrase »Messieurs, wir erheben uns von den Plätzen!« kommt bei Schmidt öfter vor, die E-BA-Engine findet mit der Abfrage **erh?b* AND plätz*** 17 Vorkommen dieser Muster nah beieinander:

Messieurs : wir	erheben uns von den Plätzen! – Wenn man das eine Schock wahr	wNeu
Messieurs : wir	erheben uns von den Plätzen! ...) Und merkwürdigste Menschen	wZie
: Messieurs, wir	erheben uns von den Plätzen!) Und nun zu dem vorliegenden Un	eRKB
: Messieurs, wir	erheben uns von den Plätzen! Wie kann sich ein Volk bloß ein	Poc
: Messieurs, wir	erheben uns von den Plätzen! – Karl Philipp Moritz war am 15	eAAR
- Messieurs, wir	erheben uns von den Plätzen : was wird dereinst die Nachwelt	eSMm
Messieurs : wir	erheben uns von den Plätzen : es ist viel geschehen in diese	eGTr
nsern Plätzen zu	erheben haben. – Ich verbleibe, mit dem Ausdruck der Bewunde	bKol
Messieurs : wir	erheben uns von den Plätzen! – nie, daß sich Ethik sehr wohl	dKlp
: Messieurs, wir	erheben uns von den Plätzen! – Moritz war am 15. September 1	eAAR
: Messieurs, wir	erheben uns von den Plätzen! – am 4. Februar 1794 dekretiert	eFoF
Messieurs : wir	erheben uns von den Plätzen! In Österreich reformiert Joseph	Fou
; Messieurs, wir	erheben uns wieder einmal von den Plätzen! – hatten den Mut,	Fou
ler : Messieurs,	erheben wir uns von den Plätzen! – nicht Einer katholisch wa	eWdt
von den Plätzen	erhebt, die Biografie eines Barons schreibt, und ihn nun dau	bWM
von ihren Plätzn	erhobm; und auch Tanndte ließ noch einmal die Hände ruhen, u	Kaf
: Messieurs, wir	erheben uns von den Plätzen) Es ist schon erstaunlich, wie d	bWM

Die eBA liefert mit der Abfrage 91 Ergebnisse. Schon beim ersten kommt der Abfrage die Normalisierung »a« = »ä« ins Gehege, es stehen zufällig »Platz« und »erhaben« auf einer Seite. Hier, wie auch bei der nächsten Fundstelle, ist der Abstand zwischen den Wörtern zu groß, um relevant zu sein. Das wiederholt sich nun mit großer Regelmäßigkeit und macht einen großen Teil der »Funde« aus. Die eBA berücksichtigt den Abstand zwischen gesuchten Wörtern nicht und wertet als Fund jedes zufällige Vorkommen von Entsprechungen der Wildcards oder Flexionsformen in einer Texteinheit, d.i. eine zufällig beim Satz entstandene Seite.

Die Funde werden mit den größeren Seiten der Typoskripte noch absurder, die Flexionsformen ebenso: »Platz, platzten, platzte« und alle Wörter, die mit Platz... beginnen, z.B. »Platzkonzerte« ... werden »gefunden«, allerdings nur bei **plätz***, bei **erh?b*** entsprechen die Funde dem Muster.

»Es ist irr=ratio=na!.«

Diese Fehler sind nicht behebbar, da bei einer Vergrößerung der Texteinheiten nur noch mehr irrelevante Stellen gefunden würden und eine Sichtung der Ergebnisse immer aufwendiger wäre.

Darüber hinaus gibt es aber noch ein ganze Reihe weiterer Fehler, die jedoch behebbar wären, wenn es sich bei der verpfuschten Grundlage denn lohnen würde:

1. »Der OR-Operator ist der Standard-Operator.«

Alle populären Suchmaschinen verwenden als Standard-Operator **AND**. Es mag zwar sein, daß jemand, der **Schnabel Felsenburg** eingibt, alle Seiten sehen möchte, auf denen »Schnabel« *oder* »Felsenburg« steht, aber das könnte er auch durch zwei Suchanfragen nach den einzelnen Wörtern erreichen. Wer nach **August Lafontaine** sucht, findet mit der eBA 313 Seiten, da neben »Lafontaine« jedes Vorkommen des Namens »August« (und »Auguste!) sowie des Monats August geliefert wird.

»Es ist irr=ratio=nal.«

Erst **August AND Lafontaine** findet die höchstwahrscheinlich gewünschten 10 Seiten mit »August« *und* »Lafontaine«.

Die Abfrage (**II OR Zweit***) **AND Programm*** liefert außer dem Gesuchten viele zufällige Vorkommen der Wörter auf einer Seite.

Auch die Suche nach Phrasen und Zitaten macht nur mit der Suche nach *allen* Wörtern Sinn, **AND** sollte daher der Standard-Operator sein, die »Treffer«-Listen werden sonst regelmäßig zu umfangreich.

2. »Ignorieren sehr alltäglicher Wörter (>Stopwords)<«

Die Software verwendet Stoppwörter! STOPPWÖRTER bei literarischen Texten!! Bei Texten von ARNO SCHMIDT!!! Bei einem Werkumfang von für moderne Computer geradezu läppischen 4 Mio. Wörtern! Mit solchen Folgen:

Ein häufiges Shakespeare-Zitat bei Schmidt ist »nimmt nur Alles in Allem« (*Hamlet*, I/2). Meine E-BA-Engine findet bei der Suche nach **"nur Alles in Allem"**:

nimmt nur Alles in Allem, das Wichtigste! / Gruß!¶Arno Schmi	bKol	B/5	0111	:	003
nimmt nur Alles in Allem - ausreicht für Deine urinatorisch	Kaf	1/3	0237	:	016
nimmt nur Alles in Allem, so ziemlich quitt. Sei's drum. Als	bKol	B/5	0149	:	002
nimmt nur Alles in Allem, unschwer zu erreichen gewesen sein	eNCC	3/4	0139	:	015
nimmt nur Alles in Allem, nichts als das 1, freilich unschät	eNCC	3/4	0149	:	024
nimm nur Alles in Allem); und endlich 3.) aufgrund des Ortes	bES	B/3	0243	:	001
nimmt nur Alles in Allem, auch eines Vereinigenden! - akzept	eADt	3/4	0410	:	020
nimmt nur Alles in Allem, eben doch mehr sei, <Pickwick> ges	dTaT	2/3	0261	:	020
nimmt nur Alles=in=Allem, : >Feinde ringsum!<. Gebe es doch	ZT3	4/1	0474	:	041

Die eBA findet mit der Abfrage **"nur Alles in Allem"** nichts; **"nimmt nur Alles in Allem"** findet 46 Seiten mit »nimmt«. (Die Flexionen werden diesmal erfreulicherweise nicht gefunden!) Alle anderen Wörter stehen offensichtlich in der Stoppwörterliste.

Die Suche nach **"bis ich nicht mehr bin"** (ein Feuerbach-Zitat) findet 4346 Seiten mit »mehr«, weil die anderen Wörter ignoriert werden!

Die Suche nach dem **"Ding an sich"** findet 969 Seiten mit Vorkommen von »Ding, Dinge, Dingen, Dinger, Dingern, Dinges, Dings« anstelle des einen Vorkommens der Phrase in *Sitara*.

Bedürfnis gedeckt haben mag. <Feilesufluk> scheidet von vornherein aus; Bücher der Art nahm er, womit er im SUREHAND I, 408 gar noch prahlt, so gut wie nie in die Hand, (und unter Wendungen wie dem ¶<Ding an sich>¶ hätte er ja auch vermutlich etwas ganz falsches verstanden). Ich kann die wahren Verhältnisse wiederum nur kurz skizzieren, da es auch in dieser Hinsicht an Vorarbeiten noch gänzlich mangelt. -¶						
					[3/2:0237:026]	
MAY selbst hat eine Schilderung aus dem Jahre 1880 (83?), also aus relativ früher Zeit, gegeben, wie ihn, den Jungverheirateten, seine Frau Emma mit einer ihr befreundeten Familie Metzner bekannt machte, »deren Glieder alle angeordnete Spiritisten waren, und seit dem Tode des alten Ballen //Eau						
	Vor	Kontext	Kurz	Bd	S	Z
¶ 0001	Hand, (und unter Wendungen wie dem	<Ding an sich> hätte er ja auch vermutlich etwas ganz falsch	Sit	3/2	0237	: 026

»Es ist irr=ratio=nal.«

3. »Sollten Sie Phrasen mit sehr alltäglichen Wörtern suchen, verwenden Sie besser die exakte Suche.«

In der *Anleitung zur Suche* wird für Fälle mit Wörtern aus der Stoppwörterliste auf die exakte Suche verwiesen, die jedoch unbenutzbar ist, weil man damit *vorher* wissen muß, ob und welche Satzzeichen eventuell um die gefundene Stelle drumherum stehen! (s.a. den nächsten Punkt)

Zuerst habe ich geglaubt das Konstrukt **exakt:/dabei** [.,!%&V-;:;><„"...'()„»«]*denkt Jhering[.,!%&V-;:;><„"...'()„»«]*/ in der Anleitung sei ein Scherz, aber nein, es ist den Programmierern wohl Ernst damit. Dabei bräuchte man nur die Stoppwörterliste herauszunehmen, dann wäre das **exakt** überhaupt nicht nötig!

»Es ist irr=ratio=nal.«

(Bei der Suche nach der Phrase "**dabei denkt Jhering**" wird die Stelle gefunden, weil »denkt Jhering« nur einmal vorkommt.)

4. »Um die Trefferquote zu erhöhen, werden für einzelne Suchbegriffe standardmäßig auch Varianzen zu Umlauten, Numeri und Flexionsformen berücksichtigt.«

Mit »Varianzen« (ein statistischer Begriff) sind offensichtlich die Flexionsformen &c. gemeint, das wären eher Varianten.

Computer sind doof und Datenbanken verstehen *nichts* von Sprache, *gar nichts*. Hier nun solche elektronische Banausen in die Suche hineinfuschen zu lassen, ist schon für gewöhnliche Texte unvernünftig, für literarische aber geradezu aberwitzig!

Bei Schmidt, dessen Werk über weite Bereiche durch seine »Verschreibkunst« nicht mehr den üblichen Schreib-

weisen entspricht, kann das überhaupt nur für die einigermaßen standardsprachlichen Bereiche funktionieren; d.h. gelegentlich funktioniert's so lala, meist erkennt die Software aber schon die Wörter nicht.

Einige Beispiele dafür, wie lachhaft das System ist:

Die Suche nach (Marieluise) **Fleißer** listet 103 Seiten mit Wörtern wie »Fleiß, Fleißes, fleiß«, die Software scheint also von einem Adjektiv »fleiß, fleißer, am fleißesten« auszugehen. Der Versuch, das durch eine Suche nach **Fleißer*** zu verbessern, liefert denselben Quatsch. (Tatsächlich kommt die Frau Fleißer 21-mal in der BA vor.)

Die Suche nach **Dickens** listet 916 Seiten mit den Wörtern »dick, Dick (Name), dicke, die Dicke, dickem, dicken, den Dicken, Dickens (Hurra!), dicker, ein Dicker, dickeren, dickes, Dicks (Name), dickst, dickste, dicksten«. So ein Pfusch! Tatsächlich kommt »Dickens« 271 Mal in der BA vor. Die Abfrage "**Dickens**" liefert denselben Unsinn. Der Versuch, das durch eine Suche nach **Dickens*** zu verbessern, führt zu lediglich 7 Seiten mit »Dickens'sch...« und »Dickensmadamm«, aber »Dickens« wird nicht gefunden. Die Abfrage **exakt:"Dickens"** findet 95 Seiten. Nur die Abfrage **exakt:/*Dickens*/** findet nach langem Grübeln ein einigermaßen plausibles Ergebnis. Das kann man doch niemandem zumuten!

Ähnlich sieht es bei der Suche nach **Platen** aus, es werden auch »plate« und »plates« gefunden (»der Plate, die Platen, des Plates«?) Die Suche nach **Platen*** verbessert nichts.

Das Zitat aus Shakespeares *Heinrich IV.*, 1. Teil:
»Je, Heinz! 's ist mein Beruf, Heinz; 's ist einem Menschen nicht zu verargen, daß er in seinem Berufe arbeitet.«
ist mit der E-BA-Engine leicht zu finden mit einer Suche nach "**mein Beruf**":

mein Beruf, und ein Mensch muß in seinem Beruf arbeiten))).	bES	8/3	0023	:	004
mein Beruf, und ein jeder muß doch in seinem Beruf bleiben«	bES	8/3	0038	:	001
mein Beruf bald sei, 9mein wunderholdes Lieb als Gattin zu be	Fou	3/1	0513	:	007
mein Beruf, Heinz; 1 Mensch muß in seinem Beruf arbeitn'.) /	Kaf	1/3	0182	:	005
mein Beruf, Heinz. 'S ist einem Menschen nicht zu verargen,	Sit	3/2	0144	:	026
Mein Beruf, Heinz!<)9(: da schwic das Händel schtille : ?) /	ZT3	4/1	0419	:	026
Mein Beruf erfordert'S schlicht!,) - den Mann auf der Ställe	ZT4	4/1	0589	:	037
mein Beruf. - (?): Nein.«; (zu den Mädchin=runter will's nic	SdA	4/2	0181	:	186
mein Beruf, Heinz). Und erhält dafür einen wild=belohnenden	Jul	4/4	0078	:	170

Die eBA liefert alle 285 Vorkommen von »Beruf«, dessen Flexionen und »berufen(e/r/n)«; die Abfrage "**Beruf Heinz**" funktioniert zwar, liefert jedoch auch nur die dazu passenden 4 Ergebnisse.

Bei der Suche nach **Tropfen~ AND Schnaps~** sollte idealerweise die eine Stelle »klein'n Trapfn Schnops« (AmG 160) gefunden werden, da jedoch für »Schnaps« auch die ›Varianzen‹ »schnappt(e)(n)«, »schlapp«, »Schapp«, »schnäller«, »Schnabl«, »schnuppenn« und »Schnalle« für ähnlich gehalten werden, gibt es 13 Ergebnisse, bei denen diese falschen Wörter auch meistens noch weit auseinander stehen.

Nazar findet (ohne *) auch »Nazarener«! Auch berücksichtigt diese Flexionsfunktion nur Endungen: **Mondsichel** findet nicht »Mondessichel«!

Das sollte reichen, die gänzliche Unkenntnis der Software in puncto (deutsche) Sprache klar zu machen. Das Verfahren vervielfacht die Ergebnisse lediglich mit völlig unnützen Funden.

»Es ist irr=ratio=nal.«

5. »Unschärfe Suche ... Beispiel: Nobelpreis~ findet Nobel=Preis.«

Nein, **Nobelpreis~** findet »Nobel=Preis« *nicht*, nur die Abfrage "**Nobel Preis**" findet diese 3 Vorkommen.

In *Kaff* wird bei der Suche nach **Spritze~** »S=pritze« gefunden, »s=prechn« bei der Suche nach **sprechn~** und »S=tügg« bei der Suche nach **Stügg~**.

Andererseits werden bei der einfachen Suche nach »Spritze, sprechn, Stügg« (ohne ~) die Vorkommen mit doppeltem Bindestrich *nicht* gefunden! Man muß also wissen, in welchen Varianten Wörter vorkommen, um sie suchen und finden zu können!

Die Suche nach **verschmidtst~** sollte auch folgende Schreibweisen finden:

ver=schmidt=stesten Stücke des bekannten & beliebten Autors.	bWM	B/2	0174	:	013
ver=schmidt=sten Stücken ist weder Zeit noch Energie. / Was	bAA	B/1	0219	:	009
ver=schmidt=ste Stücke sind nicht mehr möglich : aus ehrlich	ePls	3/4	0031	:	001
ver=schmidt=stesten Stücke des bekannten Autors.)¶Ansonsten	bES	B/3	0148	:	006
ver=schmidt=stesten Unbehülflichkeit, mehrere Bedeutungen in	eS&B	3/4	0255	:	001

Tut sie aber nicht.

Die Suche nach **Nep*ot*** sollte folgende Schreibweisen finden: »Nepotismus«, »Nepp=o=tismus« und »Nepp=Potissmuß«. Tut sie aber wieder nicht. Wenn man also nicht auf Buchstaben und Bindestrich genau weiß, in welchen Schreibweisen ein Wort vorkommt, hat man keine Chance, alle Vorkommen zu finden.

Als Beispiel wird in der Anleitung noch genannt: **incommodieren~** findet »incommodir'n« oder »incommodierten«. Es findet allerdings *nicht* die Stelle »vielleicht incommodiere ich Sie bloß« (Kurt Marti an AS, 23.8.1957).

Kurze Wörter führen zu *viel* zu umfangreichen Ergebnissen, **Humor~** findet z.B. 678 Seiten mit den Wörtern »Amor, amore, Aurore, Auroren, Autor, Autoren, Autores, DuMír, Dämmren, Eudore, Furore, furores, hamar, Hamor,

hemon, hör, hören, Hörern, huître, hombre, homogene, Homogenes, Homos, hore, Horen, HörenS, Hugo, Hugar, Huhmohr, HuhMors, hum, human, humanen, Humanste, humble, humbler, humid, Humma, Humml, Humore, humored, humorigen, humour, Humnur, Humpl, Humpm, Humpn, Humus, Huren, Huronen, Husar, Husaren, husir, kümre, luxor, more, pudor, rumohren, rumoren, rumour, sumir, uxor, Zuhörer, zuvor«.

Die Abfrage **Øveræssjøen~** findet dagegen nur die beiden exakten Übereinstimmungen in Anmerkungen zu Briefen; die eine Stelle in der WK2-Schreibweise »Øve-raasjøen«, die Schmidt verwendete, wird nicht gefunden.

Øveraasjøen und Umgebung ansehen 1:4000 mit Höhenschichtlini	f8rs	S/1	0018	:	006
Øveræssjøen am Romsdalsfjord - abkommandiert gewesen. Landve	bKol	B/5	0228	:	003
Øveræssjøen am Romsdalsfjord, wo Schmidt im Zweiten Weltkrie	bKol	B/5	0024	:	007

Das Problem besteht offensichtlich darin, daß der Operator **~** immer pauschal 3 Unterschiede erlaubt. Das ist bei kurzen Wörtern zu viel, es entstehen unüberschaubare Trefferlisten; bei langen Wörtern ist es oft zu wenig und die Software hält durchaus ähnliche Wörter für nicht ähnlich genug. »Es ist irr=ratio=nal.«

Apropos »irr=ratio=nal«: dieses Wort sollte man bei der eBA-Suche nach **irrational~** eigentlich *auch* in *Großer Kain* finden, aber wieder Fehlanzeige, es werden nur die gewöhnlichen Schreibweisen gefunden, von der anderen muß ich 1.) wissen, daß es sie gibt und sie 2.) mit z.B. "**irr ratio**" suchen.

Für "**pudendum~ muliebre~**" gibt es 7 Fundstellen in der E-BA-Engine:

pudendum muliebre; und, last not least eben auch »Brüste« sc	ZT1	4/1	0046	:	020
pudendum muliebre) : wobei Er die trüb=wildn Sorten präferi	ZT2	4/1	0205	:	007
pudendum muliebre / & voyieren : ?¶(ab hier nickte P regelmä	ZT2	4/1	0206	:	031
Puh'den'dumm molli'ebre!)¶(Fr kuschilte, (nicht unselig), a	ZT2	4/1	0257	:	005
pudendum muliebre : >touch bun for luck = a practice observe	ZT6	4/1	0999	:	020
puDÄNdumm mülliebre<¶(21.x.66, ESCHEDÉ, abms 18 h)¶(: wie et	ZT7	4/1	1136	:	015
pudendum muliebre >eingeführt<.¶iv. Schreibweise : Abgesehen	ZTN	4/1	1501	:	029

Die eBA liefert nur die 5 exakten Übereinstimmungen, in Phrasen funktioniert die Fuzzy-Suche also offensichtlich nicht. Es ist daher eine Abfrage wie **pudendum~ AND muliebre~** erforderlich, was auch »puDÄNdumm mülliebre« liefert und noch eine weitere Seite, wo »pudend« und »Muliebria« weit auseinander vorkommen, »Puh'den'dumm molli'ebre« wird jedoch nicht gefunden.

Die Abfrage **Phta*tateeta** sollte auch die beiden »Phtah=ta=te=eta« finden, tut sie aber nicht.

Die Fuzzy-Suche **verschiedenen~ AND Gegenden~** sollte eigentlich auch »ferschidenen gehenden« finden, das sollte sie können, tut sie aber bei der eBA auch nicht.

In der E-BA-Engine ist pro 6 angefangene Buchstaben jeweils eine Änderung erlaubt, also für Wörter mit 1 bis 6 Buchstaben 1 Änderung, für Wörter mit 7 bis 12 Buchstaben 2 Änderungen, für Wörter mit 13 bis 18 Buchstaben 3 Änderungen, usw. Die Abhängigkeit von der Wortlänge erwies sich nach langen Testreihen als vernünftigste Lösung.

Man kann sowieso nicht erwarten, daß phonetische Schreibweisen wie »aabmz« oder gar »aaabmz« gefunden werden, wenn man nach **abends~** sucht; erlaubte man das, gäbe es immer zu viele Treffer. Hier muß man einen Kompromiß eingehen, und bei der Suche nach **humor~** das Ergebnis »Huhmohr« opfern, um die Fundstellenliste überschaubar zu halten. Man hat ja auch noch die Wildcards und kann mit ***hu*mo*** nach »des Wortes Kernen« suchen, das liefert eh das bessere Ergebnis.

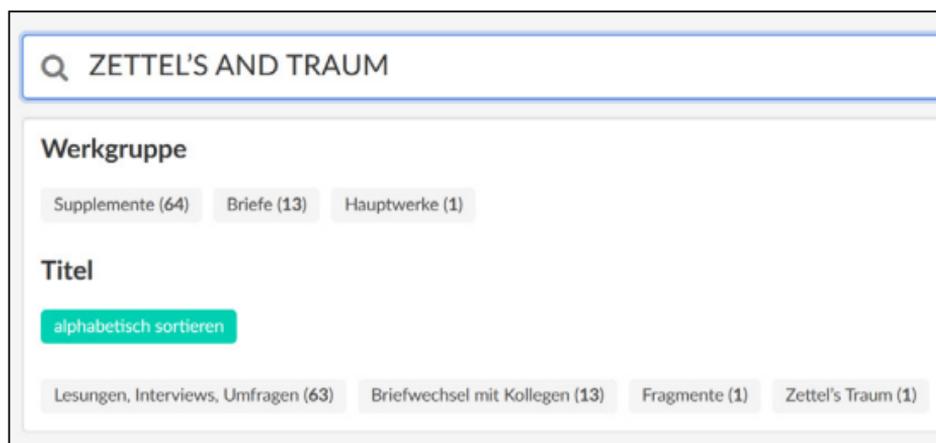
6. Der Apostroph wird wie ein Buchstabe behandelt!

Gleich zu Beginn eine besondere Kuriosität: Die Suche nach **"Zettel's Traum"** findet kein Vorkommen in *Zettel's Traum*! Das Werk wird schon beim Filtern *nicht* angezeigt:



The screenshot shows a search bar with the text "ZETTEL'S TRAUM". Below the search bar, there is a section titled "Werkgruppe" with two filters: "Supplemente (64)" and "Briefe (13)". Underneath, there is a section titled "Titel" with a green button labeled "alphabetisch sortieren". At the bottom, there are three filters: "Lesungen, Interviews, Umfragen (63)", "Briefwechsel mit Kollegen (13)", and "Fragmente (1)".

Bei der Suche nach **Zettel's AND Traum** taucht *Zettel's Traum* dann zwar in der Filterliste auf:



The screenshot shows a search bar with the text "ZETTEL'S AND TRAUM". Below the search bar, there is a section titled "Werkgruppe" with three filters: "Supplemente (64)", "Briefe (13)", and "Hauptwerke (1)". Underneath, there is a section titled "Titel" with a green button labeled "alphabetisch sortieren". At the bottom, there are four filters: "Lesungen, Interviews, Umfragen (63)", "Briefwechsel mit Kollegen (13)", "Fragmente (1)", and "Zettel's Traum (1)".

Es wird jedoch die Stelle »ZETTEL's Träumung« auf ZT7 1261 gefunden, nicht der Titel!

Die Abfrage **"Zettels Traum"** findet auch nur die Schreibweisen ohne Apostroph; um die mit Apostroph zu finden, reicht nicht der einfache Apostroph auf der Tastatur ('), da wird nichts gefunden; man muß den auf Tastaturen nicht vorhandenen *typographischen* Apostroph verwenden ('), nur dann funktioniert die Abfrage **"Zettel's Traum"**. Auf

einem Windows-PC muß man dafür die Zeichenkombination **Alt+0146** auf der Zehnertastatur eingeben!

»Es ist irr=ratio=nal.«

Es kommt in der BA sowohl die Schreibweise »Detâ'j« als auch »Detâ"j« vor. Die erste kann ich mit **Detâ'j** (mit typographischem Apostroph!) finden, es ist mir jedoch bisher noch nicht gelungen, ein Abfrage zu formulieren, die die zweite SdA 72 findet. Die Abfrage **Detâ"j** selbst liefert 150 Ergebnisse, die jedoch rein gar nichts mit dem Abgefragten zu tun haben.

Das Wort »lesn« kommt sowohl in dieser Schreibweise vor als auch mit Apostroph: »les'n«. Die sollten nun bei der Suche nach **lesn** alle gefunden werden, was jedoch nicht geschieht. (Da die Software »lesn« natürlich nicht als Verb erkennt, findet sie auch keine Flexionsformen, s.o.)

Um »les'n« zu finden, muß ich auch exakt **les'n** abfragen. Ich muß also schon vorher die exakte Schreibweise dessen kennen, was ich finden will.

»Es ist irr=ratio=nal.«

Apostrophe sollten bei der Eingabe der Abfrage ignoriert werden und auch bei der Erstellung der normalisierten Wörter! Ein Apostroph steht (meist) für das Wegfallen von etwas, so sollte er auch behandelt werden.

7. Die Wildcards ? und * funktionieren seltsam

Ein Zitat von Christiane Vulpius/von Goethe kommt in der BA dreimal vor:

1. »Seit'ch den Geheem' Rat geheirat' hab', hab'ch keene Schtunde Ruhe mehr gehabt!«
2. »Seit'ch den Härrn Geheem' Rat geheirat' hab, hab'ch keene ruhje Schtunde mehr gehabt.«

3. »Seit ich den Herren Geheemen Rat geheiratet hab, hab' ich keene ruhige Stunde mehr gehabt«.

Man sollte meinen, mit **"Geheem* Rat geheirat*"** würde man all drei finden, was bei meiner E-BA-Engine auch der Fall ist:

Geheem' Rat geheirat' hab', hab'ch keene Schtunde Ruhe mehr	fDSe	S/1	0274	:	006
Geheem' Rat geheirat' hab, hab'ch keene ruhje Schtunde mehr	dVsp	2/2	0139	:	019
Geheemen Rat geheiratet hab, hab' ich keene ruhige Stunde me	iGOZ	S/2	0112	:	021

Doch mit der eBA klappt es nicht, die findet die Variante mit »geheiratet« nicht; die Abfrage **Geheem* AND Rat AND geheirat*** findet dagegen die beiden Varianten mit »geheirat'« nicht. Offensichtlich sind in Phrasen weder Wildcards noch Fuzzy-Suche erlaubt, welch eine Einschränkung! Ich habe noch weitere Abfragen ausprobiert, es ist mir jedoch mit der eBA nicht gelungen, eine Abfrage zu formulieren, die alle drei Stellen findet.

Abfragen wie **"zweit* programm*"** oder **"wi* si* si* so*"** funktionieren bei der eBA also nicht: die E-BA-Engine liefert bei der letzteren Suche 4 Funde, darunter die gesuchte Stelle, »die schöne Häuptlinx=Tochter <Wih=Sih=Sih=Soh>« in *Kaff*.

"Nichts Niemand Nirgends Nie" liefert 7 brauchbare Funde und einen seltsamen (im Anhang zu *Schwänze*); nun wird die Phrase von Schmidt allerdings mehrfach verschrieben, die Suche nach **"ni* nie* nir* nie*"** wäre also sinnvoller. Die E-BA-Engine liefert 11 Stellen:

Nichts Niemand Nirgends Nie! : Nichts Niemand Nirgens Nie! :	Kaf	1/3	0011	:	007
<Nix Niemann Nirgns Nieh> greinte es maschinen noch Uns hinn	Kaf	1/3	0012	:	015
Nichts Niemand Nirgends Nie! / Umdient von 4 Lemuren in fahl	Kaf	1/3	0076	:	028
Nichz=Niemant=Nirgnz=Nih : Nirr=Wahna : <Tao>.) / Ich beugte	Kaf	1/3	0143	:	022
Nichts Niemand Nirgends Nie! : Da hatten die Ärzte immer die	Kaf	1/3	0171	:	004
<Nichz Niemannt Nirgnz Nie> : <Nichz Niemannt Nirgnz Nie> :	Kaf	1/3	0213	:	030
<Nichts Niemand Nirgends Nie>; das, je nach <Beleuchtung>, s	Kaf	1/3	0547	:	023
»Nichts Niemand Nirgends Nie«¶Tipp: recte Tip¶Akademie: Am 4	bES	B/3	0155	:	005
nichts, niemand, nirgends, nie; nichts, niemand, nirgends, n	iFJM	S/2	0016	:	004
>Nichts Niemand Nirgends Nie<. Beziehungsweise, wie ich einm	bES	B/3	0240	:	002
: »Nichts Niemand Nirgends Nie! -« (winslte Sie) : »- disheiß	ZT3	4/1	0461	:	034

Die eBA liefert mit dieser Abfrage »Keine Ergebnisse«, weil sie Wildcards in Phrasen nicht unterstützt, und es ist mir auch nicht gelungen, eine Abfrage zu formulieren, mit der ich alle 11 auf einmal fände. (Der Operator **exakt:** ist auch hier wieder keine Lösung.)

Auch bei Abfragen mit Abstandsabgabe wie "**unbequemer Autor**"~5 funktionieren Wildcards, z.B. in der Form "**unbequem* Auto***"~5, nicht.

Probleme mit den Wildcards wurden schon in anderen Beispielen aufgezeigt, es ist an der Handhabung von * und ? irgend etwas grundlegend kaputt, eventuell kommen sich Wildcards und andere Abfragemöglichkeiten ins Gehege, z.B. die sog. »Varianzen«. Das ist durch Ausprobieren kaum einzugrenzen, da die Ergebnisse (was bei Computern eigentlich unmöglich sein sollte) geradezu eigensinnig, launisch und grillenhaft wirken.

In vielen Fällen stellte sich bei mir der Eindruck ein, daß * nur funktioniert, wenn man mindestens einen Buchstaben, manchmal auch zwei, zu wenig angibt:

Schiller* AND Kulisse* findet »Schillers Gärten von Aranjuez haben bloße *Kulissen*flachheit«, aber nicht »Bei Schiller und Wieland spielt die Natur höchstens *Kulissen*rolle«, das wird erst gefunden, wenn ich das * bei »Schiller« wegnehme: **Schiller AND Kulisse***, dann fehlt aber natürlich die erste Fundstelle.

»Es ist irr=ratio=nal.«

Ähnlich liefert **Alltagsleben*** kein Ergebnis, erst **Alltagsleb*** ergibt 5 Seiten mit allen »Alltagleben(s)«.

Oppolzer* findet 2 Stellen, **Oppolze*** ebenso, erst die Verkürzung der Abfrage auf **Oppolz*** findet alle 11 Seiten.

Muschg* dagegen funktioniert problemlos.

Die eBA scheint bei Wörtern, die z.B. auf »en« oder »er« enden, ihre Flexionsmaschinerie anzuschmeißen und sich damit selbst durcheinanderzubringen.

Die Phrase »gingen durch die Sümpfe/Symphe« aus Paula Dehmels Gedicht *Die wilden Beinchen* wird mit der E-BA-Engine und der Abfrage **s*mp?e AND gingen** viermal gefunden:

die gingen durch die	Symphe((jajá : 'tsmí) »Ein Buch in Spanisch also; aber gesch
die gingen durch die	Symphe((Sie gabs(sofort zu : !). / (Nun den Saum ab=lösn. Da
die gingen durch die	Sümpfe(: ? - / : »Du bist ein Wasser=Fisch« / : »UnDDù der
die gingen durch die	Symphe ...«; (hm=hm, murmlte Er. Er massierte Seine StirnHau

In der eBA liefert diese Abfrage nichts, **s*mp?e* AND gingen** ein völlig wirres Ergebnis, wobei jedoch einmal die Schreibweise »Symphe« dabei ist, »Sümpfe« aber fehlt; **sümp?e* AND gingen** und **sümpfe* AND gingen** liefern überhaupt keine Ergebnisse. Erst mit **sümpfe AND gingen** wird die Schreibweise »Sümpfe« einmal gefunden, allerdings dazu noch 9 Nonsensergebnisse.

»Es ist irr=ratio=nal.«

8. »Kleinkram«

Die Nachbemerkung zu ZT, ein hochwichtiger Text, fehlt!

Für effektives Arbeiten ist die Oberfläche der eBA nicht geeignet. Die Trefferliste sollte erst aus der Textanzeige bestehen, in der man nach einem Klick auf eine Stelle oder eine verkleinerte Seitendarstellung die Seite sieht, damit wäre zügiges Arbeiten möglich (s.u. das Beispiel von walter-benjamin.online).

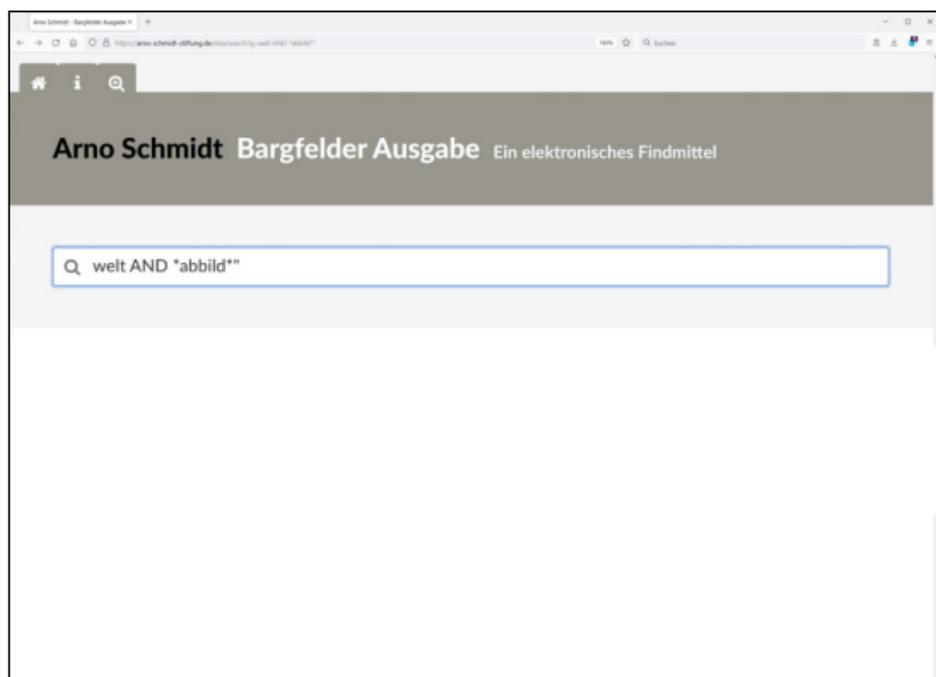
Nach Anklicken eines Treffers in der Textanzeige wird die entsprechende Seite angezeigt, die Trefferliste verschwindet; wenn man sie weiter durchgehen möchte, muß man wieder von vorn anfangen und bis zur (hoffentlich) gemerkten Seite vorschrollen.

Man kann aus der Text-Trefferliste nichts direkt kopieren.

Die Fundstellen werden nicht chronologisch angezeigt, was (wie in der E-BA-Engine) den leichten Nachweis einer ersten Nennung eines Autoren oder frühesten Zitats erlauben würde.

Die Fundstellen sind nur in der Textansicht numeriert, das sollte auch bei der Seitenansicht so sein, und besser in der Form 3/91 (Nr. 3 von 91), sonst hat man keinen Überblick, wie lange das Sichten der (meist unnütz langen) Trefferliste wohl noch dauern mag.

Es gibt kaum Fehlermeldungen. Wenn man z.B. versucht, in Phrasen Wildcards oder die Fuzzy-Funktion zu verwenden, was nicht geht, kommt keine Fehlermeldung. Bei dem kleinsten Fehler in der Abfragesyntax wird noch nicht einmal »Keine Ergebnisse« angezeigt.



Es fehlen die Möglichkeiten der E-BA-Engine für Konkordanzen und Worthäufigkeitslisten.

In der E-BA-Engine wird standardmäßig der ganze Text durchsucht, ich kann die Abfrage aber auch einschränken auf »nur AS« oder »nur die Fremdtex te« (vom Herausgeber (d.h. Anmerkungen), von Alice Schmidt sowie Brief- und Gesprächspartnern), auch das fehlt in der eBA.

Bei den bibliographischen Daten wird kurioserweise die ISBN angegeben, aber Informationen zur Entstehungszeit der Werke fehlen, obwohl sie in meinen Daten enthalten waren, da sie natürlich von großem Nutzen sind!

Arno Schmidt: Aus dem Leben eines Fauns. In: Arno Schmidt: Bargfelder Ausgabe, Werkgruppe I, Bd. 1. Zürich 1987: Haffmans. ISBN: 3-251-80001-9, S. 348
Auch muß man sich daraus erst mühsam die in Arbeiten über Schmidt längst üblich gewordene Kurzangabe »BA I/1, S. 348« extrahieren.

Im Menü »Zitieren« sollte es auch einen Link auf die komplette Abfrage geben, also ohne die Angabe von Werkgruppe/Band/Seite:

.../search?q=August%20Lafontaine#p2_3_202

ZT ist groß, beim Filtern sollte auch die Auswahl nach Büchern möglich sein: ZT1, ZT2 usw.

Das Ampersand & ist nicht suchbar. Da es gewöhnlich anstelle von »und« verwendet wird, sollte es beim Aufbau der Datenbank auch als das Wort »und« gespeichert werden. Man könnte also "**Potz Dittografie und Homöoteleton**" suchen und würde »Potz Dittografie & Homöoteleton« finden. Jetzt muß man "**Potz Dittografie Homöoteleton**" eingeben.

In den *Ländlichen Erzählungen* werden Wörter, die in der BA in kleinen Kapitälchen gesetzt sind (statt Versalien), im Text klein geschrieben, anstelle von »sagte sie

DIESE JUNGEN DINGER« liefert die eBA z.B. »sagte sie diese jungen dinger«:

- | | | |
|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|
| 1 | [...] »Erzählt ihr ja nichts von dem, was Ihr geträumt habt. Die konstruiert Euch einen ›Sinn‹ hinein, daß Ihr Kopf steht!«. Aber Ida lächelte nur mütterlich. »Das'ss ja alles so kinderleicht,« sagte sie diese jungen dinger , »und zumal Männer so erstaunlich einfach gefügt – dazu bedarf es nun wahrlich keines Dokortitels; ein wirklich intelligenter Mensch weiß[...] | Kundisches Geschirr,
Abt. 1, Bd. 3, S. 381 |
|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|

(Warum die Funde wohl zusätzlich zur orangefarbenen Hervorhebung auch noch kursiviert werden?)

Schmidt schreibt meist »ETA Hoffmann«, gelegentlich aber auch die einzelnen Initialen und mal mit, mal ohne Leerzeichen. Die Abfrage "**E.T.A. Hoffmann**" liefert ein anderes Ergebnis als "**E._T._A. Hoffmann**" (mit Leerzeichen)!

»Es ist irr=ratio=nal.«

Die Abfrage "**Himmels Lampen löschen**" zeigt die richtige S. 161 an, aber mit falscher Markierung, vmtl. weil die Zeilen auf dieser Seite »vuno« zu lesen sind.

Die Abfrage **welt AND *abbild*** liefert einige kuriose Markierungen mit nur einem passenden Wort auf der Seite.

Das Beispiel zur »Einschränkung nach Datierung eines Textes«, **held AND datum:[1933 TO 1945]**, liefert auch alle Vorkommen von »Held*« in den Fragmenten, von denen keines aus dieser Zeit stammt.

Die Programmierer der eBA scheinen es allgemein für ein wichtiges positives Merkmal von Suchsoftware gehalten zu haben, daß möglichst *viele* Stellen gefunden werden!

»Es ist irr=ratio=nal.«

Ableger

Die Trefferliste ist übrigens auf walter-benjamin.online, wo ebenfalls eine Lösung der Pagina GmbH verwendet wird, besser gelöst:

<https://www.walter-benjamin.online/suche/edition?q=Wort>

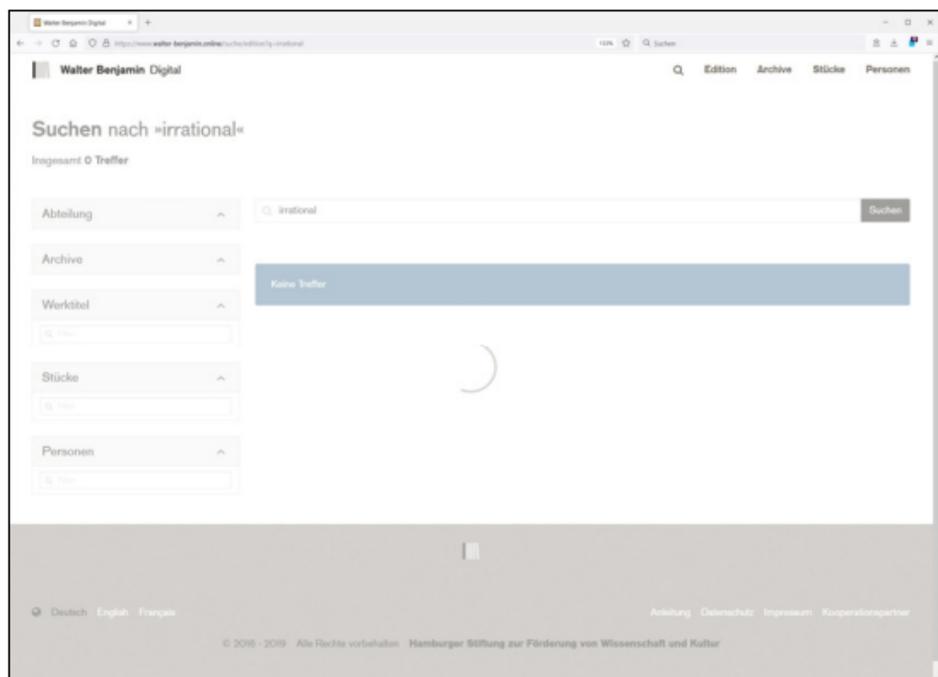
Text plus kleine Darstellung der Seite, die man dann anklicken und vergrößern kann.

The screenshot shows the search results for the query 'Wort' on the website walter-benjamin.online. The page is titled 'Walter Benjamin Digital' and features a search bar with the query 'Wort' and a 'Suchen' button. The results are displayed in a list format, with each entry showing a thumbnail of a manuscript page, the title, and a short text snippet. The first result is 'Ein Weihnachtseigel' with a snippet: 'Es war nur ein Wind, so daß die **Wörter** die sich auf meinen Lippen bildeten, wie Falten waren, die ein träges Siegel plötzlich vor einer frischen Brise wirt: „Alle Jahre wieder / Kommt das Christkind / Auf die Erde nieder! Wo wir Menschen sind“ – mit diesen **Wörtern** hatte der Engel, der in ihnen sich zu bilden begonnen hatte, sich auch verflüchtigt.' The second result is 'Markthalle Magdeburger Platz' with a snippet: 'Nein, man sprach „Mark=Thalle“ und wie diese beiden **Wörter** in der Gewohnheit des Sprechens verschliffen waren, daß keines seinen ursprünglichen Sinn beibehielt, so waren in der Gewohnheit dieses Ganges verschliffen alle Bilder, welche er gewährte, so daß ihrer keines sich dem ursprünglichen Begriff von Einkauf oder Verkauf darbot. Und dennoch schien er uns X verheißungsvoller als in ihr, wie das verlorne **Wort**, das wir noch eben auf hatten, nun, wenn wir es besäßen uns die Zunge zu demosthenischer Beredsamkeit scheint lösen zu wollen. Wie das verlorne **Wort** ... zu demosthenischer uns Befähigung lösen würde, so scheint das Vergessene schwer vom ganzen gelebten Leben, das es uns verspricht.' The third result is '([Der] Turnhalle [zu])' with a snippet: 'Da fand einmal ein Schüler namens Schulze, der ein langer und hübscher Bengel war, ein guter Turner, und ganz und gar vom Schlag der schlechten Schüler, die jederzeit das Ohr der Klasse haben, ein **Wort** das meinen ganzen Eckel vor der sklavischen Hetzerei vollendet formte.'

Bei der eBA wäre es dann allerdings sinnvoll, die Funde formatiert anzuzeigen, d.h. mindestens mit Kursivierungen, Unterstreichungen, Versalien und Kapitalchen. Auch hätte sich dann all der gequälte Unfug erledigt, der nun in der eBA dazu aufgewendet wird, die kopierbare Textmenge einzuschränken.

Allerdings verhält sich walter-benjamin.online beim Nichtfinden des Gesuchten, als sei das System abgestürzt:

<https://www.walter-benjamin.online/suche/edition?q=irrational>

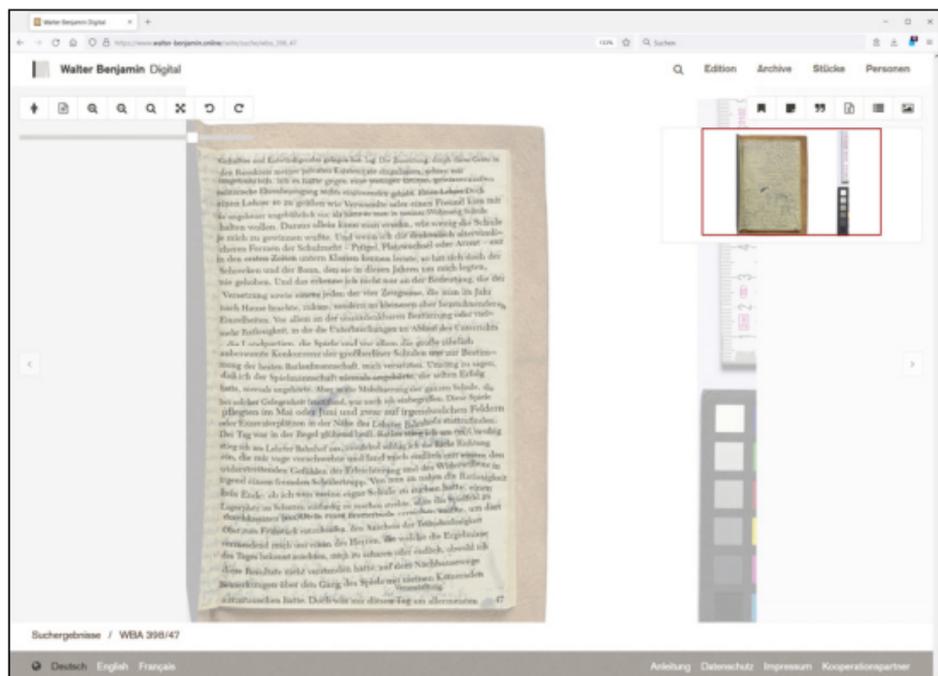


Der Hinweis »Kein Treffer« und die sofortige Möglichkeit der Neueingabe eines Suchbegriffs wäre hier die Lösung gewesen, so ist man erst einmal »gefangen« und kommt nur über »Seite zurück« des Browsers oder Klick auf einen Menüpunkt wieder raus.

»Es ist irr=ratio=nal.«

Ansonsten hat walter-benjamin.online zwar einen geringeren Funktionsumfang, aber im Wesentlichen dieselben Probleme, vor allem die Unmöglichkeit des Findens von Wörtern und Phrasen über den Seitenwechsel hinaus: die Phrase »am allermeisten verhaßt« gibt es zwar in der *Berliner Chronik*, sie ist jedoch als Phrase oder mit **AND** unauffindbar, weil sie halb auf S. 47, halb auf S. 48 steht. Bei den wenigen Dokumenten läßt sich die Stelle jetzt noch leicht mit einer Suche nach »allermeisten« finden, doch nach dem Anklicken des Funds kann man von der angezeigten Seite nicht weiterblättern, auch wenn's hier absolut notwendig wäre:

https://www.walter-benjamin.online/seite/suche/wba_398_47



»Es ist irr-ratio-nal.«

Fazit

Die eBA ist mit Sicherheit die schlechteste Implementierung eines für Literatur gedachten Suchsystems in der uns bekannten Welt. Man muß in vielen Fällen aufs I-Tüpfelchen genau die Schreibweise und akkurat die umgebende Interpunktion von dem kennen, was man doch eigentlich erst finden möchte. Man muß meist sämtliche Fehler und Unzulänglichkeiten des Systems sowie die exakten Schreibweisen bei Schmidt kennen, um eine Suche mit dem erwünschten Suchergebnis formulieren zu können.

Dies ist vor allem bemerkenswert, wenn man bedenkt, daß die Pagina GmbH eine fertige E-BA-Engine (für die

Serverseite) bekommen hat, und es nur notwendig war, die Clientseite (das Webinterface) zu erstellen. Auch erhielt die Firma ein fertig installiertes Testsystem, mit dem sie ja wenigstens die Ergebnisse ihrer eigenen Anpasserei einer Open-Source-Business-Software hätte überprüfen können. Doch die E-BA-Engine, das darin vorhandene Know-how und das monatelange Feintuning wurden völlig ignoriert und statt dessen eine von vornherein zum Scheitern verurteilte, vorgeblich »eigene Lösung« implementiert.

Unter Programmierern gibt es eine Bezeichnung für solche Software, die es sogar schon in die englischen Dictionaries geschafft hat:

kludge /klu:dz/

‘An ill-assorted collection of poorly-matching parts, forming a distressing whole’ (Granholm); esp. in Computing, a machine, system, or program that has been improvised or ‘bodged’ together; a hastily improvised and poorly thought-out solution.

Verbesserungen sind bei der eBA an mehreren Stellen gut machbar, etwa durch Rausschmeißen von Flexionen- und Stoppwörterunsinn (womit auch **exakt:** überflüssig würde), Implementierung solider vertrauter Wildcards – auch in Phrasen, Anpassung der Fuzzy-Suche usw., aber die Hauptprobleme sind systemimmanent, »gut« werden eBA und walter-benjamin.online deshalb nie werden, sondern wegen des unüberwindbaren Seitenwechsels immer »ungenügend« und für wissenschaftliches Arbeiten ungeeignet bleiben. Den Einsatz einer wie beschrieben überarbeiteten Pagina-Software könnte man guten Gewissens nur für Projekte aus sehr kleinen Textstücken

empfehlen, etwa Gedichtsammlungen, einfache Nachschlagewerke oder Tagebücher ohne längere Beilagen.

Daß dieser Pfuscher drei Jahre unentdeckt bleiben konnte, sagt einiges über die sog. Schmidt-»Forschung« aus: Es gibt offensichtlich nur wenige Nutzer der eBA und diese wenigen haben keine Ahnung von Computern (weil die meisten von ihnen auf Haptik schwörende Ewiggestrige sind) und/oder kennen sich im Werk Schmidts nicht aus.

Mich persönlich würden noch die Antworten auf zwei Fragen interessieren:

1. Wie konnte es dazu kommen, daß die Pagina GmbH ohne Erfahrung im Programmieren von literarischer Suchsoftware und in völliger Unkenntnis der Schwierigkeiten des Schmidt'schen Werkes sich dazu entschloß (und damit durchkam!), ein ausgetestetes und auf Schmidt zugeschnittenes Suchmodul wegzuwerfen, ja vorher nicht einmal genauer anzuschauen, und statt dessen eine für den Business-Einsatz mit Webseiten gedachte Software für eine bessere Grundlage für etwas Neues zu halten?
2. Wann begann den Beteiligten zu schwanen, daß ihr Ansatz, das Werk in völlig isolierte Einzelseiten zu splitten, für den wissenschaftlichen Einsatz mit literarischen Werken von weniger als ein, zwei Seiten völlig ungeeignet ist, und wer entschied dann, das zu verschweigen und den Murks trotzdem zu verwenden? (Denn daß die Folgen für Phrasen- und AND-Suche katastrophal waren, dürfte einem versierten Programmierer – und ich hoffe doch, die Pagina GmbH hat mindestens 1 solchen – wohl nicht lange unentdeckt geblieben sein.)

Nachtrag 10.12.2021 mittags via Twitter:



Arno Tator @Arnotationen · 10. Dez.

Noch'n kapitaler Bock in der eBA!

In 'Kaff' kommt 52-mal *potz* vor. Die eBA sagt "42 Ergebnisse", meint allerdings Seiten:

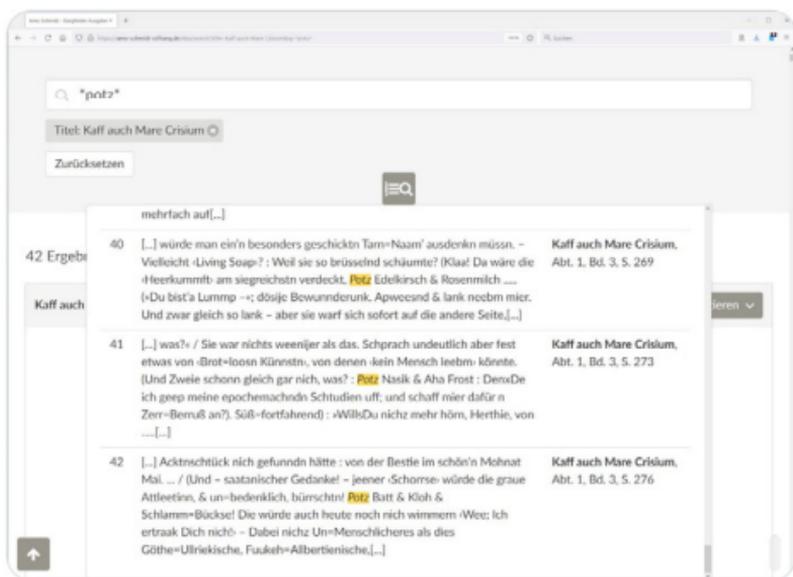
arno-schmidt-stiftung.de/eba/search?tit...

Und bei der Durchsicht in der SEITENAnzeige stimmt's auch.

In der SUCHTREFFERliste sind es jedoch nur noch 42 Fundstellen: Bei mehreren auf einer Seite wird immer nur eine Stelle angezeigt! Was soll man damit dann anfangen? Potz Stellenschwund & 18%.

»Es ist irr=ratio=nal.«

arnotationen.de/pdfs/Juergensm...



Nachtrag 24.12.2021:

Die erste öffentliche Stellungnahme der Arno Schmidt Stiftung durch Susanne Fischer in der ASML am 20.12.21 (via blog.digithek.ch):

Günter Jürgensmeier hat seine Kritikpunkte an der eBA der Arno Schmidt Stiftung gegenüber mit demselben Furor und weitgehend demselben Wortlaut vor rund 2 Monaten geäußert; ich habe ihm dafür gedankt und ihm mitgeteilt, dass ich die Kritikpunkte mit den Pagina-Mitarbeitern besprechen und dann schauen werde, was sich wie mit welchem Aufwand verbessern lässt. Genau das passiert gerade.

Auch nach zwei Monaten fehlt ein Warnhinweis auf der Webseite, dass das von Pagina GmbH Publikationstechnologien entwickelte »elektronische Findmittel« eBA aufgrund zahlreicher Unzulänglichkeiten für den wissenschaftlichen Einsatz nicht geeignet ist.



Arno Tator @Arnotationen · 24. Dez. ...

[.@zettelstraum](#) Wenn bei den alten Römern jemand drei- oder viermal sein hehres Haupt schüttelte, dann verwendeten sie dafür niemals das Wort 'furor', sondern sagten z.B. "terque quaterque concutiens illustre caput".

blog.digithek.ch
Kritik an der elektronischen Bargfelder Ausgabe (e...
Günter Jürgensmeier liefert unter dem Titel «Kaputt by Design» eine Kritik an der elektronischen ...

Nachtrag 26.12.2021:

Giesbert Damaschke am 25.12.2021 auf asml.de:

Die eBA und reguläre Ausdrücke

Samstag, 25. Dezember 2021

Günter Jürgensmeier hat sich die eBA genauer angeschaut (PDF) und fällt ein vernichtendes Urteil:

Die eBA ist mit Sicherheit die schlechteste Implementierung eines für Literatur gedachten Suchsystems in der uns bekannten Welt.

Nun ist die eBA sicherlich nicht perfekt und die von Jürgensmeier angeführten Kritikpunkte sind durchaus valide – allein: Seine Schlussfolgerung, das System sei irreparabel und »kaputt by design« scheint mir dann doch etwas übertrieben. Ich habe zu wenig Hintergrundwissen, um das wirklich beurteilen zu können, habe aber so meine Zweifel, ob das Problem des »unüberwindbaren Seitenwechsels« wirklich so unüberwindbar ist. Aber es mag jeder selbst entscheiden, ob er oder sie Jürgensmeier in seinem harschen Urteil folgt oder nicht. Darum soll es hier auch gar nicht gehen.

Die Kritik an der eBA ist mir nur ein willkommener Anlass, kurz auf den Einsatz von »regulären Ausdrücken«, wie sie die eBA bietet, einzugehen.

Nachtrag 29.12.2021:

In der eBA fehlen auch noch die Auszüge aus Alice und Arno Schmidts Tagebüchern im *Briefwechsel mit Kollegen*, BA B/5, S. 365–376, während die aus den *Briefwechseln mit Eberhard Schlotter* und *Wilhelm Michels* enthalten sind, aus Letzterem sogar Bernd Rauschenbachs *Gespräch mit Wilhelm Michels*.