

Heider-Tauf- und Sterberegister: Benutzeranleitung

zur SW-Einrichtung, OCR-Umwandlung, Daten-Formatierung und -Prüfung der Heider-Register

- 295 -

Tag	Name	Eltern, Beruf, Ort	Tom.
5.12.1670	Stockmayer Nikolaus	Thomas - Martha Niederpuemberg	I 159
20. 3.1671	Steininger Joachim	Johann - Christina zu Püreth	I 162
7. 9.1671	Stockinger Tobias	Jakob - Eva an der Donau	I 163
6. 5.1672	Stockmayer Tobias	Tobias - A. Maria zu Harbach	I 167
17. 8.1672	Stöger Regina	Johann - Eva an der Lohrmühl	I 168
16. 9.1672	Steininger Elisabeth	Johann - Christina zu Püreth	I 169
28. 9.1672	Stelzer Martha	Adam - Maria Pruckmühl	I 169
29.11.1673	Steininger Andreas	Johann - Christina zu Püreth	I 173
20. 7.1674	Stockmayer Andreas	Andreas - A. Maria zu Harbach	I 175
19. 3.1675	Stöger Josef	Johann - Eva an der Lohrmühl	I 177
9. 7.1675	Stöckl Heinrich	Georg - Eva	I 178
11. 3.1676	Stockmayer Magdalena	Stefan - Katharina	I 182
11. 3.1677	Stockmayer Josef	Tobias - Maria	I 185
3. 8.1677	Stöger Clara	Johann - Eva an der Lohrmühl	I 187
5. 2.1678	Stöckl Mathias	Georg - Eva	II 1
31. 8.1679	Stockmayer Adam	Tobias - Maria	II 10
1. 9.1680	Stöckl Rosina	Georg - Eva	II 17
22.10.1680	Stöger Barbara	Johann - Eva	II 17
3. 9.1682	Stockmayer Nikolaus	Tobias - Anna Maria	II 29
28. 7.1683	Stöger Stefan	Johann - Eva	II 36
8. 9.1683	Stöckl Maria	Georg - Eva am Stöckl	II 36
26. 1.1684	Strasser Euphrosina	Wolfgang - Sara Wizersdorf	II 38
17. 5.1684	Stockmayer Anton	Tobias - Anna Maria zu Harbach	II 40
26. 5.1684	Starnberger Justina	Joachim - Sophia Wizersdorf	II 40
9. 2.1686	Strasser Susanna	Wolfgang - Sara Riezersdorf	II 49
21. 4.1686	Starnberger Katharina	Joachim - Sophia Wizersdorf	II 50

Meine Dokumentationen zu den Themen Kirchenmatriken, Digitalisierungsprojekt der Heider-Register und Josef Heider auf den Netzseiten der Familia Austria

- *Lebenslauf von Prof. Josef Karl Heider*
mit einleitenden Gedanken zur Familienforschung und zwei Fotos, zur Verfügung gestellt von seinen Erben
Autor: Josef Heider
- *Einführung zu den Matrikeln*
Autor: Josef Heider, Wien 1983
- *Verzeichnis der Register zu den Matrikeln der Pfarren des Mühlviertels und des Salzkammergutes*
mit Angaben über den Inhalt der einzelnen Bände.
Autor: Josef Heider, Wien 1983
- *Heider-Register und Digitalisierungsprojekt: Einführung*
Was sind die Heider-Register und wozu dient ihre Digitalisierung durch den Verein Familia Austria zusammen mit dem Oberösterreichischen Landesarchiv?
Autor: Sepp Asanger
- *Heider-Tauf- und Sterberegister: Benutzeranleitung*
Benutzeranleitung zum Einrichten der OCR-Software und des PC-Editors KEDIT, zur OCR-Umwandlung, Formatierung und Prüfung der Heider-Register
Autor: Sepp Asanger
- *Heider-Projekt: technische Dokumentation*
beschreibt technische Grundlagen, Dateiverwaltung, Namensschemata, Nachbearbeitung von Scan-Dateien
Autor: Sepp Asanger



Sepp Asanger

sepp.asanger@aon.at

INHALTSVERZEICHNIS

1	Einleitung.....	6
2	Die Tauf- und die Sterberegister von Josef Heider	7
2.1	Aufbau der Taufregister	7
2.2	Aufbau der Sterberegister	8
2.3	Erklärungen zu den vier Spalten Tag, Name, Eltern... und Tom.....	9
2.3.1	Tag-Spalte	9
2.3.2	Namens-Spalte	9
2.3.3	Spalte Eltern, Beruf, Ort (EBO).....	9
2.3.4	Tom.-Spalte	9
2.4	Verarbeiten der Registerdaten nach dem Scannen in drei Verarbeitungsschritten.....	10
3	OCR-Schrifterkennung mit ABBYY FineReader.....	11
3.1	Verwendete Software.....	11
3.2	Einrichten des FineReaders	11
3.3	Öffnen der Scans (Registerseiten) zu Schrifterkennung	11
3.4	Bereich der Texterkennung.....	12
3.5	OCR-Erkennung	14
3.6	Arbeiten mit den umgewandelten Registerseiten	14
3.6.1	Zoom-Fenster	14
3.6.2	blau markierte Zeichen, die nicht korrigiert werden müssen	15
3.6.3	Manuelle Prüfung der umgewandelten Texte	16
3.6.4	Prüfung von Namen und Begriffe mit dem Wörterbuch von Abbyy FineReader.....	16
3.6.5	Leerzeichen.....	16
3.6.6	Trennung von Begriffen (Namen, Beruf, Ort) auf zwei aufeinanderfolgende Zeilen.....	17
3.6.7	Bearbeitung rückgängig machen	17
3.6.8	Speichern der bearbeiteten Text-Datei (Registerseite)	17
3.6.9	Unterbrechung der OCR-Bearbeitung	17
4	Formatierung und Datenprüfung der Registerseiten mit dem Texteditor KEDIT.....	18
4.1	KEDIT	18
4.2	KEDIT-Profile	18
4.2.1	Startposition der 4 Spalten (Blöcke).....	19
4.2.2	Reihenfolge von tomus und pagina	19
4.2.3	Liste der tomus-Nummern	19
4.3	Aufruf von Funktionen für die Tauf- und Sterberegister über Funktionstasten.....	20
4.4	Starten der Formatierung und Prüfung einer Registerseite	20
4.5	Lesen der vorhergehenden Seite.....	21
4.6	Seitennummer der Heider-Registerseite	21
4.7	Überschriftzeile	21
4.8	Leerzeilen	22

4.9	Eintragung	22
4.9.1	Name	22
4.9.2	Vertauschte Vor- und Familiennamen	22
4.9.3	Sonderfälle beim Namen.....	23
4.10	Eltern, Beruf, Ort (EBO).....	23
4.11	Zahlenangaben in der EBO-Spalte.....	24
4.12	tomus-Block mit tomus und pagina	24
4.12.1	tomus-Block beim Taufregister	24
4.12.2	tomus-Block und Altersangabe beim Sterberegister	25
4.12.3	Manuelle Trennung der Altersangabe beim Sterberegister.....	25
4.12.4	Automatische Trennung der Altersangabe beim Sterberegister.....	30
4.12.5	Aufruf der Version manuell oder automatisch.....	32
4.13	falsche Zeilen	32
4.14	Manuelle Korrekturen in den Heider-Registern durch Benutzer	32
4.15	Blockzeile in der Text-Datei	33
4.16	Zwillinge und mehrfache Vornamen	34
4.17	Fehlermeldungen, Kontrolle.....	35
4.18	Nachtrag	36
4.19	Rückgängig machen von Änderungen.....	36
4.20	Anhang.....	37
5	OCR-Schrifterkennung mit ABBYY FineReader.....	38
5.1	Die FineReader-Optionen zum Einrichten und Anpassen	38
5.1.1	Dokument	38
5.1.2	Scannen/Öffnen.....	38
5.1.3	Lesen: Gründlich, Benutzermuster	39
5.1.4	Speichern: TXT	40
5.1.5	Anzeigen: Schriftart für Text	40
5.1.6	Erweitert.....	41
5.1.7	Zwischenspeichern der OCR-Ergebnisse	41
5.2	Benutzerwörterbuch von FineReader.....	41
5.2.1	Benutzerwörterbuch als wichtiges Hilfsmittel zur Datenprüfung.....	41
5.2.2	Benutzerwörterbuch erstellen.....	42
5.2.3	Neue Wörter im Benutzerwörterbuch speichern	43
5.2.4	Benutzerwörterbuch exportieren/importieren.....	43
6	KEDIT Texteditor: Installation und Editierfunktionen.....	45
6.1	Installation der KEDIT-Software und der Verarbeitungs-Programme (Macros)	45
6.2	Die wichtigsten Editor Zeilenbefehle	45
6.2.1	Einfache Befehle	45
6.2.2	Befehle mit Zahlenangaben.....	46
6.2.3	Verdoppelung der Befehle	46

6.3	Arbeiten mit Textblöcken.....	46
6.3.1	Textblock in den Kommentar verschieben	47
6.3.2	Verschieben der Altersangaben in den Sterberegistern	47
7	Abbildungsverzeichnis	49

1 Einleitung

In diesem Dokument geht es nicht darum, das Heider-Projekt insgesamt zu beschreiben. Die allgemeine Beschreibung der Heider-Register, des Digitalisierungsprojektes und den gesamten Ablauf kann man in der Dokumentation ***Heider-Register und Digitalisierungsprojekt Einführung*** auf den FA Netzseiten nachlesen. Diese Doku ist eine Anleitung für:

- das Umwandeln der Bilddateien der gescannten Tauf- und Sterberegister von Josef Heider in Textdateien,
- das Berichtigen von Umwandlungsfehlern durch die Texterkennung, die Erstellung einer spaltengenaue, einheitliche Form der Eintragungen durch meine Programme und das Ergänzen der Daten für die spätere Aufgliederung der Daten für die Übernahme in die FA-Datenbank,
- die Kontrolle der fertig umgewandelten und formatierten Registerseiten und Einfügen allfälliger Kommentare.

Die Beschreibung geht davon aus, dass die Registerseiten von den Heider-Büchern bereits eingescannt, gerade ausgerichtet und mit einem Fotoprogramm nachbearbeitet wurden, die Einzeldateien richtig durchnummeriert und in einem Scan-Ordner gespeichert sind. Informationen dazu gibt es im Dokument ***Heider-Projekt: technische Dokumentation***.

In eigenen Kapiteln gibt es Hinweise zum Einrichten der erforderlichen Software-Komponenten ABBYY Fine Reader und KEDIT.

2 Die Tauf- und die Sterberegister von Josef Heider

Den allgemeinen Aufbau der Heider-Register und Hinweise zum Alphabet und zum Schriftbild ist im Dokument *Heider-Register und Digitalisierungsprojekt Einführung* beschrieben. In den folgenden Kapiteln gibt es kurze Wiederholungen und zu beachtende Sonderheiten mit Beispielen.

2.1 Aufbau der Taufregister

- 36 -

Tag	Name	Eltern, Beruf, Ort	Tom.
17. 9.1720	Pesenböck Katharina	Simon - Eva Ebretschlag	II 370
7.11.1720	Pamer Elisabeth	Johann - Anna Maria Amessschlag	II 371
7.11.1720	Pachner Leopold	Michael - Katharina Bürger	II 371
27. 2.1721	Prieschl Elisabeth	Michael - Rosina Ebretschlag	II 373
19. 3.1721	Pachner Maria Josefa	Joh.Michael - A.Magdal. Ratsbürger	II 374
29. 3.1721	Pesenböck Anna	Wolfgang - Sophia Amessschlag	II 375
8. 5.1721	Pichler Franz	Josef - Elisabeth Bürger	II 376
27. 5.1721	Pichler Johann	Andreas - Maria Oberstiftung	II 377
7. 9.1721	Pernauer Adam Heinr.&	Joh.Mich. Josef-Kath. Bürger	II 379
9.10.1721	Pogner Martin	Georg - Apollonia Silberhartschlag	II 379
13.10.1721	Brungraber Leopold	Thomas - Salome Bürger	II 379
16.11.1721	Proier Maria	Mathias - Maria Langbruck	II 380
19. 1.1722	Prandstetter A.Maria	Thomas - Susanna Mauern, Dietrichschlag	II 382

Abb. 1: Seite eines Taufregisters

Zum Verständnis sei zunächst das Format der Heider-Register erklärt. Im Beispiel sehen wir den grundsätzlichen Aufbau einer Registerseite für **Taufen** mit einer Seitennummer oben zwischen zwei – Zeichen. Die Seitennummer kann in der Mitte oder am rechten Rand stehen. Nach ein bis zwei Leerzeilen kommt zwischen 2 Querlinien die Spaltenüberschrift und danach beginnen die Registereintragungen in 4 Spalten (Blöcken):

- **Tag** Datum der Taufe (Geburt) bzw. Sterbedatum
- **Name** Familien- und Vorname(n) des Kindes bzw. der Kinder oder des Verstorbenen (Proband)
- **Eltern, Beruf, Ort (EBO)**, gemischte Angaben zu den Probanden
- **Tom.** Buchnummer (tomus) und Seitennummer (pagina) des Kirchenbuches

Heider hat seine Register einseitig auf sehr dünnem Papier mit Schreibmaschine geschrieben und war dabei auf eine Seitenbreite von etwas weniger als dem A4-Format beschränkt. Bei längeren Namen wurde die Einteilung in die 4 Spalten daher nicht immer eingehalten. Bei den Zwillingen am 7.9.1721 ist zu sehen, wie die Vornamen der Kinder in die EBO-Spalte hineinragen. Beim manuellen Lesen ist das zwar meistens klar erkennbar, doch die automatische Erkennung der Spalten durch ein Standardprogramm am PC versagt, weil dort eine fixe Spalteneinteilung Voraussetzung ist.

Weitere Einsparungen im Platzbedarf hat Heider durch Weglassung von trennenden Leerzeichen erreicht. Wir sehen das bei den erwähnten Zwillingen (Adam Heinr.& Joh.Mich.), als auch beim Bindestrich zwischen

den Vornamen der Eltern. Bei der Eintragung vom 19.3.1721 fehlt das Leerzeichen sowohl bei den Vornamen der Eltern, als auch als auch zwischen der Mutter und dem tomus (A.Magdal.II 374).

Die *pagina* im tomus-Block kann fehlen. Die Reihenfolge *tomus-pagina* ist bei manchen Registerbüchern umgedreht.

2.2 Aufbau der Sterberegister

- 2 -

Tag	Name	Eltern, Beruf, Ort	Tom.
18.10.1674	Aichinger Stefan	Aigen	II 247
20. 8.1674	Autengruber Michael	Rudolfing	II 247
13. 5.1675	Autengruber Anna		II 248
11. 9.1675	Autengruber Martha	Totengräber, Wächter	II 248
4. 6.1678	Amprosin Maria	Praitenstein	II 254
6. 6.1682	Almesberger Kind	Florian, Schindlau	II 261
23. 3.1684	Autengruber Andreas	Rudolfing	II 265
2.12.1688	Almesberger Adam	Aigen	II 274
14. 9.1715	Almesberger Kaspar	Schindlau	6o III 496
22. 6.1716	Almesberger M.Magdalena	Schindlau	5 III 498
9. 6.1718	Almesberger Maria	Wurmbrand	6 III 505
6.12.1719	Asanger Johann	Knecht zu Aigen	3o III 509
5. 5.1723	Aigner Lukas	Landgerichtsdienner zu Aigen	6o IV 589
24. 1.1725	Almesmüller Philipp	Wurmbrand	7o IV 595
26. 8.1726	Almesberger Magdalena	Hans Georg, Schindlau	IV 603
9. 7.1787	Aigner Hans Georg	ledig, Aigen	15M 28 IV 609
6. 2.1729	Asanger Maria	Michael, ledig zu Unterneudorf	11W IV 613
7. 4.1729	Almesberger Gertrude	Hans Georg	7 IV 615
22.12.1729	Asanger Margarete	Holzhäuseln	6o IV 619
10. 3.1730	Aigner Eva Maria	Landgerichtsdienner zu Aigen	71 IV 621
17. 9.1730	Aigner M. Katharina	Kaspar, Landgerichtsdiener zu Aigen	7T IV 624
9. 7.1731	Andexlinger Maria	Rudolfing	51 IV 627
29. 9.1731	Aigner M. Magdalena	Kaspar, Landgerichtsdiener zu Aigen	low IV 628
27.12.1731	Aigner Anna Maria	Landgerichtsdienner zu Aigen	24 IV 629

Abb. 2: Seite 2 des Sterberegisters der Pfarre Aigen

Die Sterberegister haben denselben grundsätzlichen Aufbau, aber eine neue Information kommt hinzu: das **Alter**. Die besondere Schwierigkeit besteht dabei darin, dass die Altersangabe in keiner eigenen Spalte steht. Sie steht zwar vor dem tomus, aber innerhalb des Elternblocks an dessen Ende. Außerdem ist sie optional und fehlt sehr häufig. Wenn der Platz zu eng wurde, steht das Alter manchmal in der Folgezeile. In obigem Beispiel sehen wir das bei den Eintragungen vom 26.8.1726 und vom 29.9.1731. Auch bei den Sterberegister kann die *pagina* fehlen oder die Reihenfolge *tomus-pagina* umgedreht sein. Wir sind hier also mit einigen Unsicherheiten konfrontiert. Zwingend ist damit nur die tomus-Angabe, die uns vor allem bei der Erstformatierung, als ‚Ankerpunkt‘ dient, von der die Prüfungen ausgehen.

2.3 Erklärungen zu den vier Spalten Tag, Name, Eltern... und Tom.

2.3.1 Tag-Spalte

Form und Inhalt der Spalten können stark variieren. So steht etwa in Spalte **Tag** in der Regel ein gültiges Datum, manchmal aber fehlen Tag und/oder Monat. Auch Angaben wie Ostern oder Fasching statt des konkreten Datums kommen vor.

Das Format des Datums ist im Normalfall **tt.mm.jjjj** und beginnt Stelle (column) 1, wenn der Tag 2-stellig ist. Ein einstelliger Tag beginnt auf der Stelle 2. Es kommt aber immer wieder vor, dass auch der einstellige Tag auf der Stell 1 beginnt. Das muss durch Einfügen einer Leerstelle korrigiert werden (siehe [Abb. 4: Auswahlrahmen erstellen](#) auf Seite 13).

2.3.2 Namens-Spalte

Die Namens-Spalte ist relativ einheitlich, nur manchmal sind Familien- und Vornamen vertauscht. Die Namen müssen dann manuell umgedreht werden. Vornamen von Mehrlingen sind durch das **&**-Zeichen gekennzeichnet bzw. verbunden. Vor allem bei den seltenen Drillingen (Vierlinge sind mir bisher nur einmaluntergekommen) reichen die Namen weit in die nächste Spalte, die EBO-Spalte, hinein.

Eine Abweichung kommt auf den Registerseiten im Anhang am Ende der Registerbücher vor. Dort hat J. Heider unvollständige Eintragungen zusammengefasst und es kommen daher keine Familiennamen, sondern nur Vornamen vor und manchmal fehlt sogar der.

2.3.3 Spalte Eltern, Beruf, Ort (EBO)

Diese Spalte kann sehr gemischte Informationen enthalten, die manchmal für den Leser, aber noch viel mehr von einem Programm, nicht einfach zu interpretieren sind. Die Vornamen der Eltern sind mit einem Bindestrich verbunden, vor dem und nach dem eine Leerstelle sein muss (im Unterschied zu Doppelnamen u.ä.). Bei Platznot wurden aber oft die Leerstellen weggelassen: Josef-Kath., Ratsbürger, Fleischhacker. Die Begriffe können auch abgekürzt oder abgeteilt und in der nächsten Zeile fortgesetzt sein:

Witwer, Gemeinbraumei-
ster

Schwierig wird die Interpretation vor allem bei ledigen Elternteilen, die durch die Zusätze **led.**, **ledig**, **ill.** oder **fil. ill.** gekennzeichnet werden (=Familienstand). Nach **fil. ill.** folgen z.B. oft Namen der Großeltern(-teile) und fallweise deren Ort und/oder Beruf des Großvaters. Eine andere Form der Angabe von **ledigen** Eltern ist die Kennzeichnung mit **V. :** bzw. **M. :**. Hier wird, nachdem der eine Elternteil beschrieben wurde, der zweite angeführt, manchmal mit dem Vermerk **verheiratet**, also ein außereheliches Kind.

Es gibt dann noch andere Vermerke wie **Soldat**, meist mit Angabe des Regimentes, **Findelkind** und mehr.

2.3.4 Tom.-Spalte

In der Spalte **Tom.** steht die Nummer (*tomus*) und die Seitennummer (*pagina*) des Kirchenbuches der Eintragung. In dieser Beschreibung verwende ich den Begriff *pagina* oder *pag* für die Seitennummer des Kirchenbuches, um sie von der Seitennummer der Heider-Registerbücher zu unterscheiden.

Josef Heider hat in seinem Verzeichnis der von ihm registrierten Kirchenbücher die Originalbezeichnung der Kirchenbücher (*tomus*) und den Zeitbereich tabellarisch zusammengestellt. In der Regel hat er diese Originalbezeichnung in den Registern übernommen, aber nicht immer. Z.B. sind die Kirchenbücher von Altmünster mit den römischen Ziffern XIX bis XXIV bezeichnet. Der vorletzte Band, den Heider registriert hat, hat die Bezeichnung XXIII, also eine fünfstelligen römischen Nummer. Da Heider in seinen Sterberegistern die *pagina* mit max. 4 Stellen begrenzt hat, hat er hier statt der römischen Nummern lateinische verwendet. In *matricula* wiederum wurde ein eigenes Nummerierungsschema eingeführt, das weder mit dem römischen *tomus* der Kirchenbücher, noch mit den lateinischen Nummern von Heider übereinstimmt. In unseren digitalisierten Registern haben wir die von Josef Heider vergebenen *tomi* (Pl.) unverändert übernommen.

Tomus und *pagina* sollten innerhalb eines Registerbuchstaben aufsteigend sein. Eine Abweichung von dieser Regel entsteht bei den Kirchenbüchern, die nach Ortschaften geführt wurden. Da beginnen innerhalb desselben Anfangsbuchstabens des Namens die Matriken je Ortsteil datumsmäßig wieder von vorne. Dadurch springen

die Seitennummern im Register, das ja nach dem Anfangsbuchstaben des Namens und nach Datum sortiert ist, ständig auf und ab und können daher durch das Prüfprogramm nicht kontrolliert werden.

2.4 Verarbeiten der Registerdaten nach dem Scannen in drei Verarbeitungsschritten

Das Verarbeiten der Registerdaten erfolgt in zwei voneinander unabhängigen Schritten:

- Optische Schrifterkennung (OCR) mit ABBY-Fine Reader,
- Korrektur von Zeichen- und Schreibfehlern, Formatierung und Datenprüfung mit KEDIT.
- Prüfung und Klärung offener Fragen

In jedem der drei aufeinanderfolgenden Verarbeitungsschritte können die Registerdaten verändert bzw. korrigiert werden. Aufgrund der spezifischen Prüfungen und Rückmeldungen ergeben sich aber Schwerpunkte für das Eingreifen und Korrigieren.

1. Die OCR-Umwandlung markiert unsichere Zeichen mit blauer Farbe und unbekannte Wörter werden rot unterstrichen, daher sind diese die ersten Kandidaten zur Überprüfung bzw. zur Korrektur. Dazu gehört auch das Einfügen fehlender Leerzeichen zur Worttrennung und das Berichtigen von offensichtlichen Schreibfehlern.
2. Die KEDIT-Maros formatiert die Daten sauber in 4 Spalten und ermittelt die dafür erforderliche Größe, was die optische Prüfung erleichtert. Sie prüfen Datum, tomus und pagina und drehen die Reihungen pagina-tomus um. Gefundene Fehler werden gemeldet.
3. Im letzten Schritt werden ungeklärte Namen und Bezeichnungen sowie Fehler in der Reihung der Eintragungen an Hand der Originalbücher geklärt.

Manchmal hat Josef Heider besonders interessante Texte aus den Matrikeneinbüchern übernommen. Das sind dann oft ausführlichere Texte, die quer über die Namens- und EBO-Spalte hinweg geschrieben sein können und die als Kommentar in die Datenbank übernommen werden sollten. Dazu eignet sich am besten die Funktion Textblöcke im KEDIT. Siehe Beispiel *Textblock in den Kommentar verschieben* auf Seite 47.

3 OCR-Schrifterkennung mit ABBYY FineReader

3.1 Verwendete Software

Als OCR-Software (Optical Character Recognition, optische Buchstabenerkennung) verwenden wir den ABBYY FineReader 9.0 Professional Edition von Adobe. Der Verein Familia Austria hat für das Projekt 3 Lizenzen angeschafft. Kauf am 20.6.2011 von CH.S-Shop, Kunden-Nr.: 179729878, Serien-Nr. FPEF-9010-0001-3425-7224-7882. Davon verwende ich eine auf meinem PC. Eine zweite Installation ist auf einem Rechner von Frau Nadja Kicking. Die 3. Installation ist dzt. ungeklärt. Für einen neuen Benutzer müsste eine Lizenz deaktiviert (de-installiert) werden, um sie frei zu bekommen.

Ende 2014 habe ich die Version 12 versuchsweise probiert. Die bestehenden Probleme mit Version 9.0 (Wörterbuch) wurden aber nicht verbessert. Außerdem ergab sich ein neues Problem, weil einstellige Leerzeichen manchmal durch eine Leerstelle (wie bisher), manchmal aber auch durch ein Tabulatorzeichen ersetzt werden. Damit verliert eine Textzeile völlig seine Form und das Ergebnis war unbrauchbar. Das Problem konnte damals mit der Supportstelle von Abbyy weder erklärt noch gelöst werden.

FineReader versucht, aus der Bilddatei (die digitalisierten Buchseiten sind computertechnisch nur Bilder) die Texte zu erkennen und damit für den Computer lesbar zu machen. Aus den Bilddateien entstehen dadurch Textdateien, die die Grundlage für die weitere Verarbeitung sind. Die Erkennungsquote der Texterkennung hängt wesentlich von der Schrift- und von der Scanqualität der Registerseite ab. Die Scan-Qualität kann vor allem bei sehr hellen Schriften durch Erhöhung des Kontrastes verbessert werden, was zu einer deutlichen Verbesserung der Texterkennung führen kann. Ich benütze dazu die Fotobearbeitung in meinem Foto-Programm. Siehe auch das Kapitel *Erfassen und Nachbearbeiten der gescannten Registerseiten* im Dokument *Heider-Projekt technische Dokumentation*.

3.2 Einrichten des FineReaders

Für ein erfolgreiches Arbeiten muss ABBYY FineReader an die speziellen Bedürfnisse dieses Projektes angepasst werden. Dies geschieht vor allem mit der Funktion *Extras/Optionen...* Das Einrichten des FineReader wird in dem Kapitel, *5.1 OCR-Schrifterkennung mit ABBYY FineReader* auf Seite 38 beschrieben.

3.3 Öffnen der Scans (Registerseiten) zu Schrifterkennung

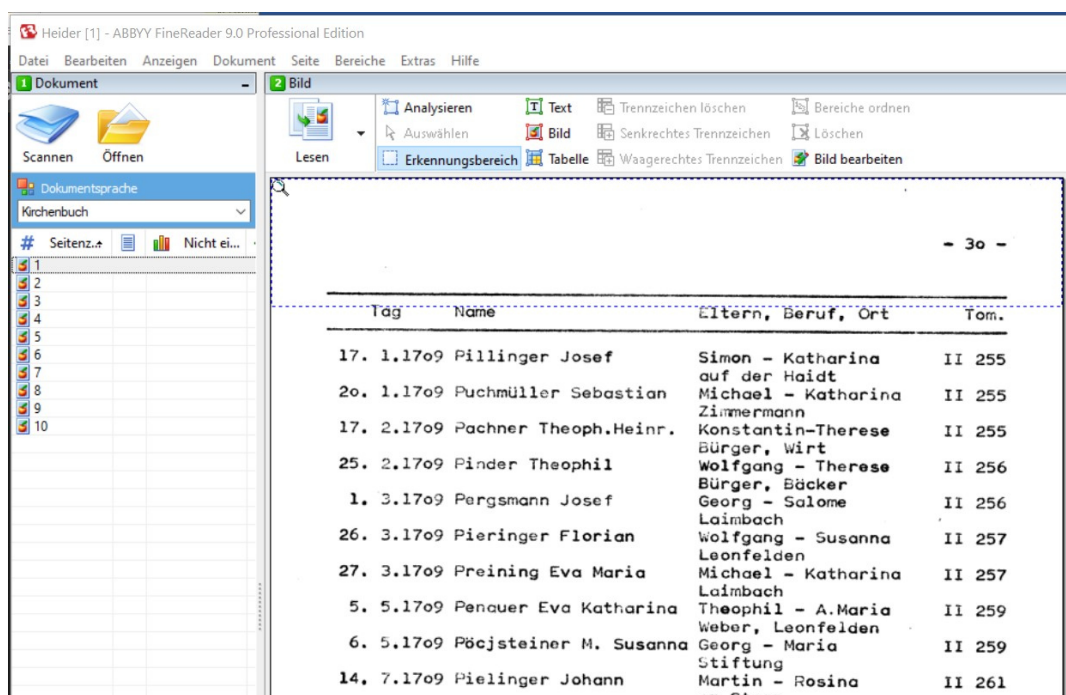


Abb. 3: Öffnen von 10 Registerseiten

Im Fenster **1 Dokument** des FineReader sind 2 Symbole dargestellt: Scannen und Öffnen. Nachdem die Heider-Register bereits als Scans vorliegen, brauchen wir nur die Funktion für das Öffnen.

Unter dem Begriff **Dokument** sind im FineReader alle Seiten (Dateien) zu verstehen, die gleichzeitig zu einem Erkennungsvorgang geöffnet werden. Ob diese Seiten zu einem einheitlichen Thema gehören oder voneinander unabhängig sind, ist für diesen Vorgang irrelevant.

Mit der Funktion **Öffnen** wird das aktuelle Verzeichnis (Ordner) mit den darin enthaltenen Scan-Dateien angezeigt. In unserem Fall sollte der Name des Verzeichnisses Leonfelden_GH_A-G_Scan sein. Über das Icon mit dem Pfeil nach oben neben dem Suchfeld kann man ein anderes Verzeichnis suchen. Ob man jeweils Einzeldateien (Registerseiten) oder gleich mehrere Seiten (=Dateien) zur Verarbeitung öffnet, ist eine freie Entscheidung des Bearbeiters. Für mich ist die optimale Menge 10 Seiten. In obiger Ansicht habe ich vom Taufregister A-G von Bad Leonfelden die Seiten 30 bis 39 mit der Funktion **Öffnen** in den FineReader geladen. Mit der Lösch-Taste können einzelne oder alle Seiten wieder entfernt werden. Im Fenster **1 Dokument** des FineReader sieht man die Anzahl der geöffneten Seiten und den Verarbeitungsfortschritt (siehe [Abb. 4: Auswahlrahmen erstellen](#) auf Seite 13).

3.4 Bereich der Texterkennung

ABBYY FineReader kann auf eine Textstruktur eingehen. So können beispielsweise spaltenorientierte Texte auch spaltenweise bearbeitet werden. Wie wir sehen, ist der Registerseite in 4 Spalten gegliedert. Weil Heider die Spaltengrenzen jedoch nicht immer eingehalten hat, ist eine automatische spaltenweise Texterkennung nicht sinnvoll. Daher wird die ganze Registerseite bzw. der ganze Scan als **ein** Textbereich definiert und verarbeitet. Das genaue Aufteilen der erkannten Texte in die einzelnen Spalten erfolgt erst nachträglich mit einem eigenen Texteditor (siehe [Starten der Formatierung und Prüfung einer Registerseite](#) auf Seite 20).

Der Erkennungsbereich kann im Fenster **2 Bild** auf 2 Arten definiert werden:

- Klicken auf das grüne Symbol für **1 Text** und aufziehen des Texterkennungsrahmens mit der Maus.
- Laden eines vordefinierten Bereiches unter **Bereiche/Bereichsvorlage laden...** Vorher muss ein passender Rahmen mit der Funktion **Bereichsvorlage speichern...** gespeichert worden sein.

Nach Aufziehen bzw. nach Laden des Bereichsrahmens muss dieser in der Regel noch genau justiert werden. Dabei ist darauf zu achten, dass

- alle Texte innerhalb des Bereichsrahmens liegen,
- der Rahmen möglichst genau am linken Rand, also bei der ersten Stelle des Datums, beginnt. Ist dort der Abstand zwischen Rahmen und erstem Zeichen zu groß, fügt Abbyy FineReader unerwünschte Leerzeichen ein, wobei der FineReader bei einem nach rechts geneigten Bild empfindlicher reagiert als bei einem nach links geneigten. Bei diesem kann am unteren Ende einer Seite die Eintragung um gut eine Stelle weiter rechts beginnen, ohne dass unerwünschte Leerzeichen eingefügt werden. Durch Verzerrungen durch die Schreibmaschine und/oder durch den Scan-Vorgang können die Scans schief sein. Das Ausrichten von zu schiefen Scans kann mit einem Fotoprogramm erfolgen, darauf wird hier nicht eingegangen.

Die genaue Einstellung des Rahmens ist nur am **linken** Rand wichtig. Rechts, oben und unten können freie Bereiche bleiben. Graue Linien vom Seitenrand der Registerseite oder Abbilder von Fingern, mit denen die Seite beim Scannen gestrafft wird, sollten jedoch möglichst außerhalb des Rahmens liegen, weil sie Störung verursachen können. Normalerweise werden sie aber ignoriert.

Der Rahmen kann als Ganzes durch Gedrückt-halten der Taste **Steuerung** und gleichzeitiges Bewegen der Maus verschoben werden. Die Ränder können aber auch einzeln verschoben werden wie bei einem normalen Windows-Fenster. Vor allem bei einem vordefinierten Auswahlrahmen sollte man nicht vergessen zu kontrollieren, ob unten alle Zeilen innerhalb des grünen Auswahlfeldes liegen. Manche Buchseiten können länger sein als üblich und dann sind die Zeilen möglicherweise nicht mehr innerhalb des Auswahlfeldes und würden bei der Texterkennung verloren gehen.

Der Erkennungsrahmen lässt sich desto genauer justieren, je größer das Bild-Fenster ist. Am unteren Rand des Erkennungsfensters kann man dessen Darstellungsgröße einstellen. Mit der **F6**-Taste nütze ich die ganze Breite des **2 Bild** Fensters. Man muss dann zwar mehr herumschrollen, um die Ränder zu kontrollieren, aber mit dem Mausekranz geht das ja sehr flott und der Mehraufwand lohnt sich.

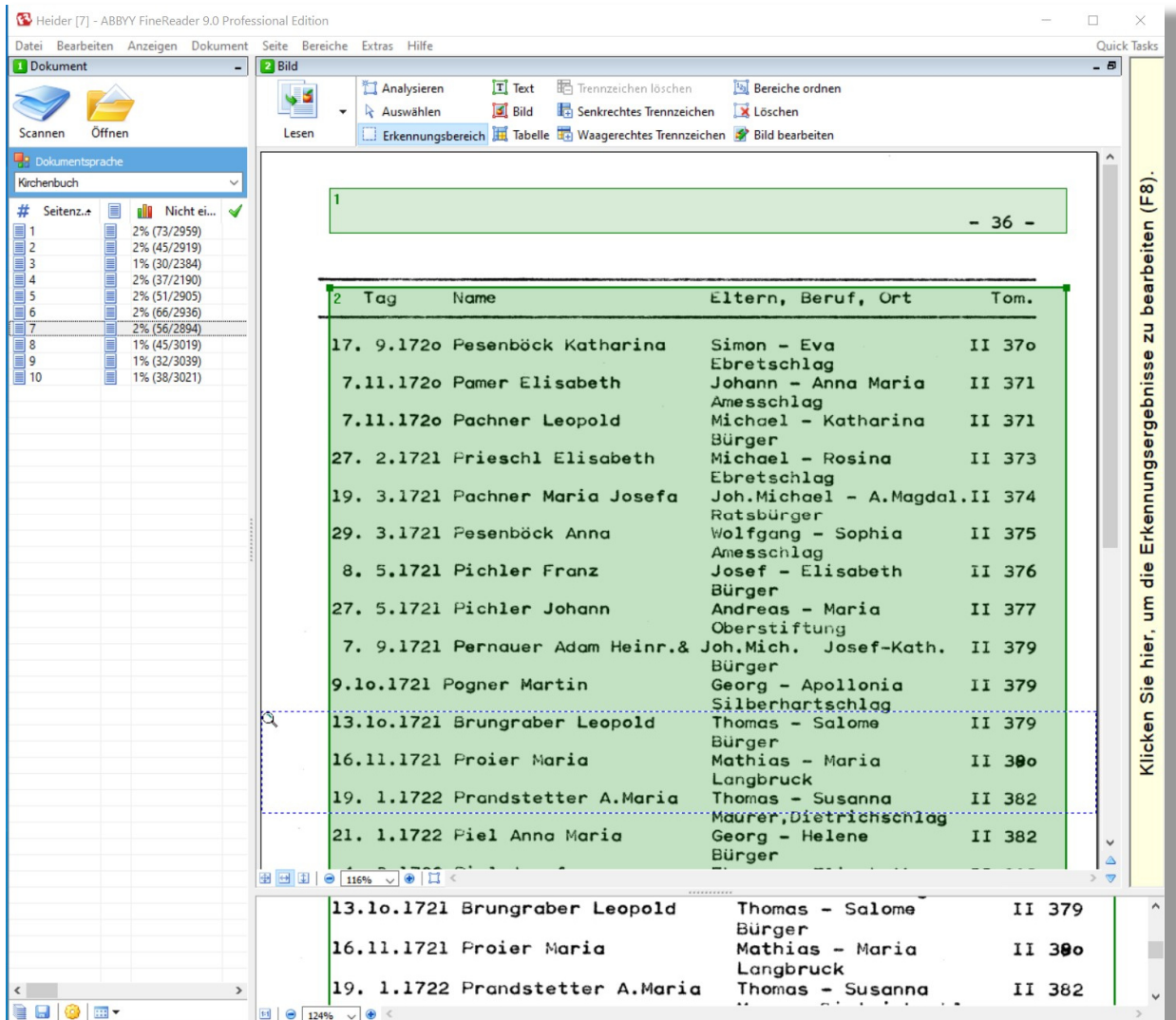


Abb. 4: Auswahlrahmen erstellen

Manchmal beginnt das Datum um eine Stelle zu weit links. Hier ist es besser, den Rahmen so zu ziehen, dass diese eine Stelle aus dem Rahmen herausfällt. Damit wird vermieden, dass bei allen anderen richtigen Zeilen eine falsche Leerstelle eingefügt wird. Bei der Kontrolle ist so ein Datum um 1 Stelle verschoben und fällt dadurch optisch auf. Durch Einfügen des abgeschnittenen Zeichens wird das Problem rasch behoben. Siehe [Abb. 5: Tag außerhalb des Erkennungsrahmens](#) auf der nächsten Seite.

In obigem Beispiel hingegen beginnt das Datum **9.10.1721** durch einen Tippfehler um eine Stelle zu früh. Hier ist nach der OCR-Erkennung ein Leerzeichen einzufügen, um die ganze Zeile um eine Stelle nach rechts zu schieben.

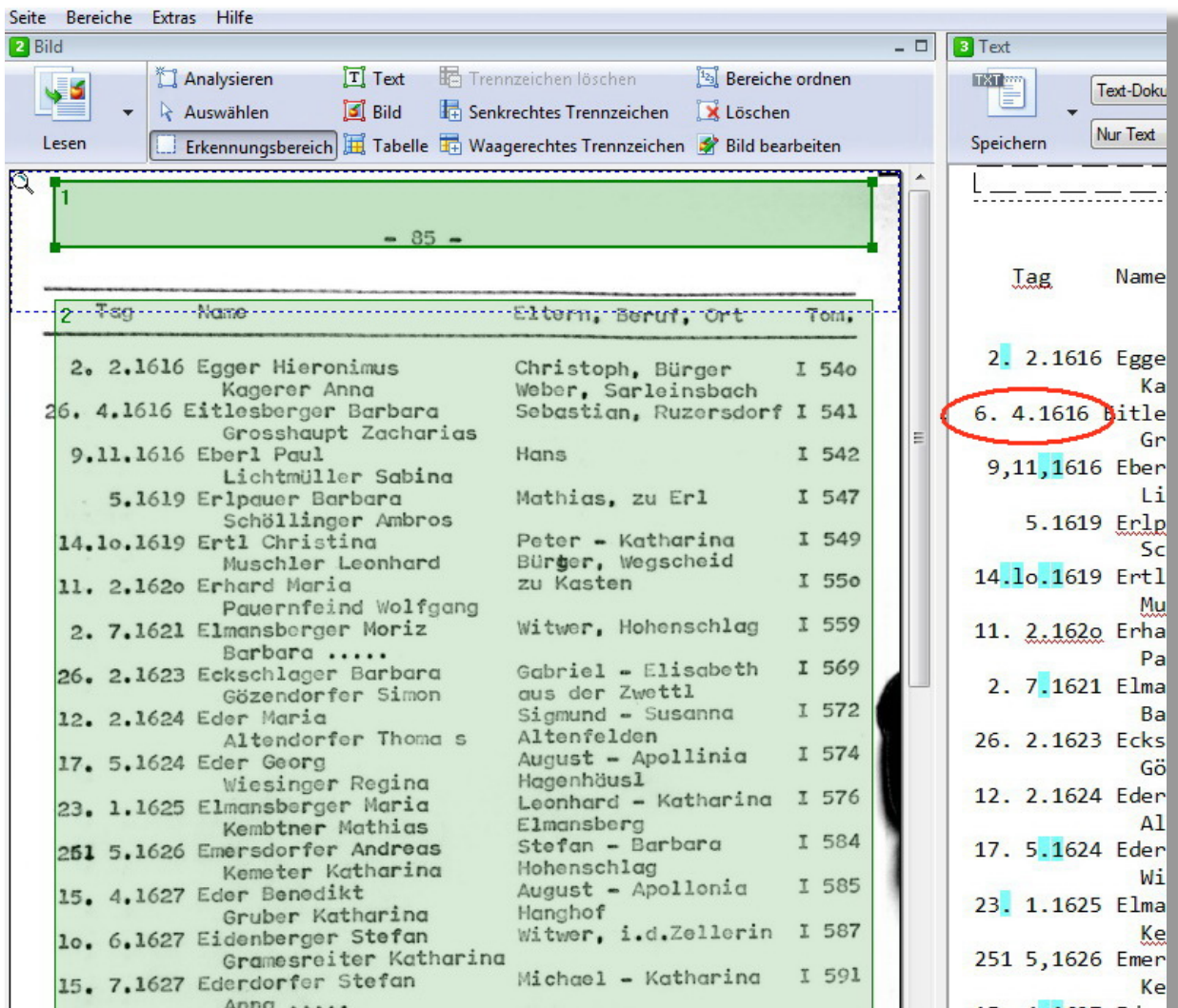


Abb. 5: Tag außerhalb des Erkennungsrahmens

3.5 OCR-Erkennung

Durch Klick auf die Funktion **Dokument Lesen** (alle geöffneten Seiten) bzw. **Seite Lesen** (nur ausgewählte Seiten) wird die OCR Schrifterkennung gestartet. Im Fenster **1 Dokument** kann man den Fortschritt des OCR-Prozesses mitverfolgen (siehe *Abb. 6: lesen und bearbeiten der Registerseiten* auf Seite 15). Je bearbeiteter Seite wird die Gesamtzahl der Zeichen angezeigt. Die unsicheren Zeichen werden in absoluter Zahl und als Prozentsatz angezeigt. Ein Prozentsatz von 1 bis 3 % unsichere Zeichen ist ein guter Wert. Nach Ende des OCR-Prozesses werden durch Drücken der **F8**-Taste Seiten mit den umgewandelten Texten im Erkennungsfenster **3 Text** angezeigt. Die unsicheren Zeichen sind blau unterlegt, unbekannte Wörter sind rot unterstrichen. Es sind dann zwar oft viele Zeichen blau unterlegt, aber nur wenige sind tatsächlich falsch und müssen korrigiert werden. Die meisten blau markierten Zeichen sind entweder richtig oder werden durch das nachfolgende Prüfprogramm automatisch korrigiert. Siehe Kapitel *blau markierte Zeichen, die nicht korrigiert werden müssen* auf Seite 15.

3.6 Arbeiten mit den umgewandelten Registerseiten

3.6.1 Zoom-Fenster

Für das Bearbeiten des Textes nach der OCR-Erkennung sollte unbedingt das Zoomfenster eingeschaltet sein. Es befindet sich standardmäßig im unteren Teil des Bildschirms und wird über **Anzeigen/Zoomfenster/Zoomfenster anzeigen** aktiviert. Das wird aber bereits beim Einrichten des FineReaders so eingestellt (siehe *OCR-Schrifterkennung mit ABBYY FineReader* auf Seite 40).

Sobald im Bild- oder Textfenster ein Zeichen angeklickt wird, wird parallel der Cursor im Zoomfenster am Originaltext platziert. Damit kann der Originaltext mit dem Ergebnis der Umwandlung sofort verglichen werden. Die Vergrößerung der Darstellung lässt sich unabhängig von der Vergrößerung im Fenster **2 Bild** bzw. Fenster **3 Text** einstellen. Handschriftliche Korrekturen, die ja berücksichtigt werden sollen, sind oft nur im **2 Bild** -Fenster zu lesen. In solchen Fällen schalte ich mit der **F6**-Taste auf das Bildfenster um und kehre danach mit der **F8**-Taste wieder zum Textfenster zurück.

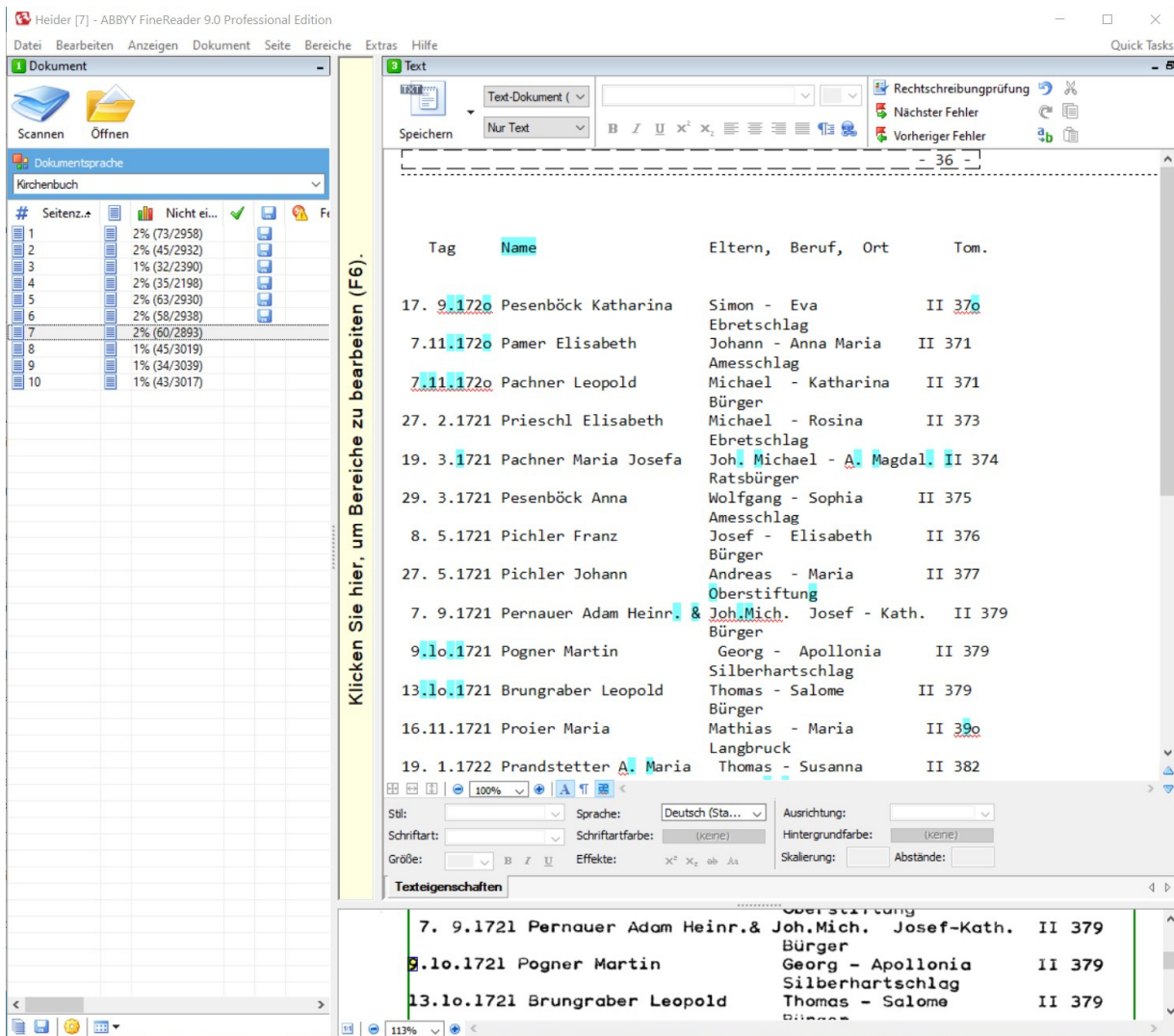


Abb. 6: lesen und bearbeiten der Registerseiten

3.6.2 blau markierte Zeichen, die nicht korrigiert werden müssen

Folgende unsichere (**blau markierten**) Zeichen benötigen keine Korrektur, weil sie entweder richtig sind oder im nächsten Bearbeitungsschritt durch das Programm automatisch korrigiert werden:

- alle richtig umgewandelten Zeichen, die als unsicher markiert sind
- folgenden Zeichen werden vom Formatierungsprogramm (KEDIT-Macro) automatisch korrigiert:

umwandel	von Zeichen	in
allgemeiner Text	— _ ~	-
	;	:
Dateinummer	O o @	0
	I L I	1
Datum	O o c @	0
	I L I	1
tomus	1 i I	I
	e	E
	x	X

	Y y v ¥ 7	V
	G	0
	3 ß	B
pagina	O o @	0
	I I i	1

3.6.3 Manuelle Prüfung der umgewandelten Texte

Bei der Prüfung der umgewandelten Texte gehe ich spaltenweise vor, d. h. ich konzentriere mich zuerst nur auf die Datumsspalte, dann auf die Namen, danach auf die EBO-Spalte und zuletzt auf die Angaben für tomus und pagina. Diese Vorgehensweise hat sich für mich als die geeignetste herausgestellt, andere mögen ein anderes Vorgehen vorziehen. Beim tomus sollte man prüfen, ob er auch in jeder Zeile angegeben ist. Vor allem bei langen Texten in der EBO-Spalte hat Heider ihn öfters weggelassen, wodurch das Prüfprogramm die pagina, so sie vorhanden ist, als tomus auffasst und ins Schleudern kommt. Bei der manuellen Einfügung gibt es insofern eine kleine Erleichterung, als die Buchstaben I, V, E, X als Kleinbuchstaben eingegeben werden können, weil sie vom Prüfprogramm auf Großbuchstaben umgewandelt werden. Das spart das lästige Betätigen der Umschalttaste auf der Tastatur.

Abhängig von der Schriftqualität der Registerseite gibt es bestimmte Zeichen, die besonders anfällig für Fehlinterpretationen sind bzw. bei denen das OCR-Programm eine falsche Sicherheit vorspielt. Typisch dafür ist die 3, die ich daher oft überprüfe, unabhängig davon, ob FineReader sie blau markiert hat oder nicht. Auch die 5 ist zwar nicht so oft aber doch auch ein Problemzeichen. Beim Text sind m und rn oft verkehrt interpretiert. e, o, c sind weitere Kandidaten, die oft Korrekturen erfordern. Offenbar abhängig von der verwendeten Schreibmaschine werden manchmal bestimmte Typen ganz schwach bis gar nicht angedruckt. In den Sterbematriken von St. Johann am Wimberg z.B. sind die Typen von 8 und W sehr stark betroffen.

3.6.4 Prüfung von Namen und Begriffen mit dem Wörterbuch von Abbyy FineReader

Viele Namen und Begriffe werden als falsch erkannt, weil sie im Wörterbuch von Abbyy FineReader nicht vorkommen. Das Arbeiten mit diesem Wörterbuch und seine Pflege ist daher ganz wichtig und wird in einem eigenen Kapitel beschrieben. Siehe *Benutzerwörterbuch von FineReader* auf Seite 41.

3.6.5 Leerzeichen

Häufig fehlen Leerzeichen zur Trennung von zusammengesetzten Vornamen, z.B. nach einem abgekürzten Vornamen wie A.Maria oder nur um Platz zu sparen (Joh.Michael). Fehlende Leerzeichen bei Vornamen werden bei der späteren Aufteilung der Daten in den Programmen von Dr. Peter Haas richtig erkannt. Es entstehen aber durch zusammengezogene Vornamen ungültige Begriffe, die Abbyy FineReader rot unterstreicht. Man kann zwar solche zusammengezogene bzw. abgekürzte Namen als eigene Begriffe ins Benutzerwörterbuch speichern, aber ich füge fehlende Leerzeichen immer ein, weil es den Text übersichtlicher macht und die visuelle Kontrolle erleichtert. Dasselbe gilt auch für unterschiedliche Begriffe, die durch einen Beistrich getrennt sind, z.B. Beruf und Ort: Maurer,Dietrichschlag. Auch hier dient ein eingefügtes Leerzeichen der Erkennung der Einzelbegriffe und der Übersichtlichkeit. Ebenfalls muss bei den Vornamen der Eltern vor und nach dem Bindestrich ein Leerzeichen stehen, um sie von einem Doppelnamen zu unterscheiden:

Joh.Michael-A.Magdal. → Joh. Michael - A. Magdal.

Genauso muss auch zwischen unterschiedlichen Begriffen ohne trennenden Beistrich, z.B. zwischen Name und tomus, ein Leerzeichen eingefügt werden: A.Magdal.II 374 → A. Magdal. II 374 (siehe Beispiel *Abb. 4: Auswahlrahmen erstellen* auf Seite 13, Eintragung vom 19.3.1721)

Keine Korrektur ist erforderlich, wenn ein EBO-Text um eine Stelle vor der Startposition der Elternspalte beginnt, sofern vor dieser Stelle ein Leerzeichen steht, also der Text nicht Teil der Namensspalte ist. Da dies öfters vorkommt, wird dieser Fehler vom nachfolgenden Formatierungsprogramm toleriert bzw. korrigiert,

Josef Heider hat manchmal zusätzliche Leerzeichen eingetippt. Sie entstehen aber auch durch die OCR-Umwandlung selbst, weil bei langen Abfolgen von Leerstellen die Texterkennung außer Tritt geraten kann. Das führt bei den letzten beiden Textspalten oft zu Verschiebungen und sie beginnen dann nicht mehr exakt an der vorgesehenen Position. Mehrfache Leerzeichen zwischen Wörtern und führende Leerzeichen werden durch die KEDIT-Macros automatisch entfernt und müssen daher nicht manuell gelöscht werden.

3.6.6 Trennung von Begriffen (Namen, Beruf, Ort) auf zwei aufeinanderfolgende Zeilen

Ein Begriff kann auf zwei aufeinanderfolgende Zeilen mit einem Abteilungsstrich aufgeteilt sein. Das Programm von Dr. Haas entfernt den Abteilungsstrich und fügt dadurch beide Teile zusammen, aber nur, wenn der Teil nach dem Abteilungsstrich in der nächsten Zeile mit einem Kleinbuchstaben beginnt. Ist dort ein Großbuchstabe, so handelt es sich nicht um eine Worttrennung, sondern um einen Doppelnamen.

3.6.7 Bearbeitung rückgängig machen

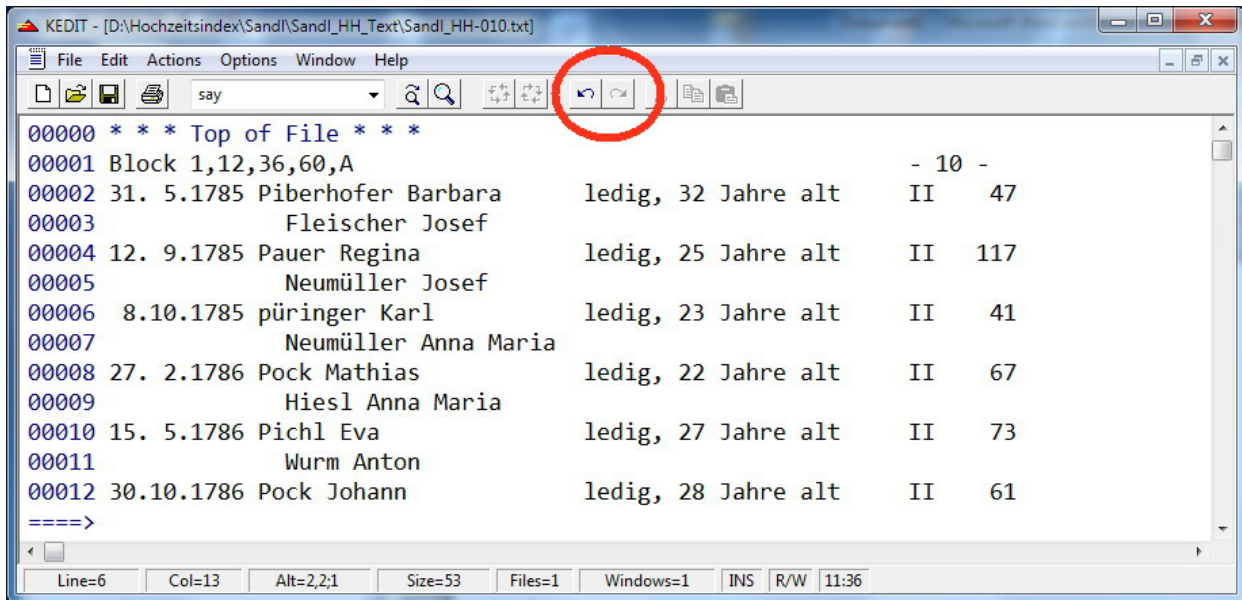


Abb. 7: OCR Textänderungen rückgängig machen

Änderungen des Textes können durch Klick auf die Icons rechts oben rückgängig gemacht bzw. wieder aktiviert werden.

3.6.8 Speichern der bearbeiteten Text-Datei (Registerseite)

Nach der manuellen Korrektur der Textseite muss diese für den nächsten Verarbeitungsschritt gespeichert werden. Das geschieht mit der Funktion **Speichern**. Wie beim Öffnen muss man auch hier zumindest beim ersten Mal das richtige Verzeichnis aufrufen, bevor man die Datei speichert. In unserem Beispiel ist das das Verzeichnis **Leonfelden_GH_A-G_Text**. Der Name der Datei ist ident mit dem Namen der Scan-Datei. Deshalb muss im untersten Auswahlfeld Dateioptionen **Dateien als Quell-Bilddateien benennen** ausgewählt sein. Der Dateityp muss **.txt** sein. Wichtig ist noch, dass das Kästchen **Dokument nach Speichern öffnen** angehakt ist. Alle diese Einstellungen sollte durch das Einrichten des FineReader bereits fix so definiert sein (in den Optionen). Der Cursor steht bereits auf dem **Speichern**-Button, sodass ein Mausklick genügt. Die Seite wird mit demselben Namen, aber mit dem Dateityp txt gespeichert und mit KEDIT geöffnet. Die Formatierung und Datenprüfung mit den KEDIT-Macros wird dann durch einfaches Drücken der Enter-Taste gestartet.

3.6.9 Unterbrechung der OCR-Bearbeitung

In *Abb. 6: lesen und bearbeiten der Registerseiten* auf Seite 15 sieht man das Ergebnis bzw. den Fortschritt der OCR-Erkennung der 10 geöffneten Scan-Dateien und welche der Dateien schon gespeichert wurden (Diskettensymbol). Wegen Müdigkeit oder aus anderer Gründe möchte man vielleicht die Arbeit unterbrechen. Das ist jederzeit möglich, weil Abbyy FineReader in einer internen Datei den aktuellen Zustand zwischenspeichert. Beim nächsten Aufruf des Programmes wird der aktuelle Stand angezeigt und man kann die Arbeit ohne Informationsverlust an der abgebrochenen Stelle fortsetzen.

4 Formatierung und Datenprüfung der Registerseiten mit dem Texteditor KEDIT

4.1 KEDIT

Ich habe schon erwähnt, dass die 4 teilweise unregelmäßig aufgeteilten Spalten der Registerseiten in eine einheitliche Aufteilung gebracht werden müssen und dass die Daten zu prüfen sind. Am Beginn des Projektes mit den Heiraten habe ich das manuell gemacht, aber z.B. die Prüfung des Datums ist so eintönig, dass ich schon nach wenigen Seiten Fehler nicht mehr gesehen habe. Es entstand daher sehr schnell die Idee, für diese Tätigkeiten ein geeignetes Werkzeug einzusetzen. Das fand ich in dem programmierbaren PC-Editor **KEDIT**. Damit kann man sowohl die Daten im KEDIT-Fenster an PC editieren, als auch mit der sehr umfangreichen Programmiersprache Makros zur Prüfung und Manipulation der auf einem Datenträger gespeicherten Daten erstellen. Die wichtigsten Funktionen, die man zum Editieren der Daten im KEDIT-Fenster braucht, sind in einem eigenen Kapitel beschrieben (*Die wichtigsten Editor Zeilenbefehle* auf Seite 45).

Mit zunehmender Kenntnis der Möglichkeiten der Makrosprache wurden die Prüf- und Formatierungsprogramme immer umfangreicher. Das war auch deshalb notwendig, weil Josef Heider seine Register in oft immer neuen Varianten erfasst hat. Dieser Variantenreichtum hat sich erst im Fortschritt des Projektes gezeigt und erforderte immer wieder Anpassungen der Programme. Diese Anpassungen führten zu einer immer besseren Fehlererkennung der Registerdaten. Es entstand aber auch ein ‚Konglomerat‘ an Programmen mit schlechter Strukturierung, das ihre Wartung sehr erschwerte. Um die von mir entwickelten Makros zu benützen, muss der PC-Editor KEDIT auf dem PC installiert sein und die Makros müssen auf dem eigenen PC gespeichert werden. Siehe *Installation der KEDIT-Software und der Verarbeitungs-Programme (Macros)* auf Seite 45). Auf die Makro-Sprache gehe hier nicht ein.

4.2 KEDIT-Profile

```

00036 * =====
00037 *   allgemein gültige Makros starten
00038 * =====
00039 'define f2.kex'      /* toggle Bottom Toolbar */
00040 'define f3.kex'      /* Suchfunktion */
00041 'define f4.kex'      /* Kommentareintragungen suchen */
00042 'define f5.kex'      /* XEDIT Profile ändern */
00043 'define f8.kex'      /* Formalprüfung */
00044 'define f10.kex'     /* zurückblättern */
00045 'define f11.kex'     /* vorwärtsblättern */
00046 'define Enter.kex'  /* Formalprüfung */
00047 if MatrikelTyp ='HH' then do /* nur für die Kategorie Hochzeiten */
00048   'define f6.kex'     /* Elternblock rechts verschieben */
00049   'define f7.kex'     /* pagina nach rechts verschieben nur bei Matrikeltyp=Heirat */
00050   'define f9.kex'     /* Gegeneintragung suchen (nur bei Matrikeltyp=Heirat) */
00051 end
00052 if MatrikelTyp ='SH' then 'set toolbar on both'
00053 * Die folgenden Variablen definieren Start- (und End-) positionen der 4 Blöcke (Spalten),
00054 * nach der OCR-Umwandlung der Registerseiten.
00055 * Bei Beginn der Verarbeitung eines neuen Registerbuches sind einige Werte zu überprüfen,
00056 * weil sie je Buch variieren können.
00057 'editv set DatumBlock 1' /* Beginn des Datums */
00058 'editv set NameBlock 12' /* Beginn des Namen-Blocks (Proband) */
00059 'editv set ElternBlock 35' /* diese Position schwankt von Registerbuch zu Registerbuch und muss angepasst werden */
00060 'editv set tomBlock 58' /* muss auch an die Spalteneinteilung der Registerseite angepasst werden */
00061 'editv set tomBlockEnde 70' /* an einer freien Stelle hinter pagina (oder tomus) und vor einem allfälligem Kommentar */
00062 * Die folgenden Angaben sind die Mindest-Positionen für die formatierten Seiten (Datei).
00063 'editv set ElternBlockMin 40' /* neue Position des ElternBlock, darf nicht < ElternBlock sein! */
00064 'editv set tomBlockMin 70' /* neue Position des tomBlock */
00065 'editv set ParkPos 100' /* auf diese Position werden der Eltern- und tomBlock temporär zwischengeparkt */
00066 *
00067 * Die Reihenfolge von Tomus und Pagina ist in manchen Büchern verkehrt und wird hier definiert
00068 if MatrikelTyp ='HH' then 'editv set TomSequence C' /* Pag unterhalb von Tom oder im Elternteil nur bei Heiraten*/
00069 /* kommt bei Sterbe nur bei einzelnen Eintragungen vor und muss händisch korrigiert werden */
00070 'editv set TomSequence B' /* Reihenfolge: Pag Tom bzw. Pag/Tom, manchmal auch Pag.Tom */
00071 'editv set TomSequence A' /* Reihenfolge: Tom Pag (Normalfall) */
00072 * In der folgenden Tabelle sind alle tomus-Nummern definiert, die im Registerbuch, das bearbeitet wird, vorkommen.
00073 * Die Sortierfolge der Bücher wird nicht durch die tomus-Nummer definiert, sondern durch die Reihenfolge in der Variablen.
00074 * Die tomus-Nummern (römischen Nummern, Buchstaben etc.) müssen durch eine Leerstelle getrennt sein.
00075 AllTom='I II'
00076 AllTom='A B C D E'
00077 'editv put AllTom'
00078 end
00079 * * * End of File * * *

```

Abb. 8: Ausschnitt aus dem KEDIT-Profile

Für die Formatierung und für die Prüfung der Daten sind einige Informationen notwendig, die in Form von Parametern im KEDIT-Profil mit dem Namen WINPROF.KEX gespeichert sind. Einige davon sind spezifisch für die zu bearbeitenden Registerseite und müssen vor Beginn des Prozesses an die Form der Registerseiten angepasst werden. Die Datei WINPROF.KEX wird beim Aufruf von KEDIT durchlaufen und speichert dabei die Informationen in globalen Variablen, die dann von den Verarbeitungs-Makros benützt werden.

4.2.1 Startposition der 4 Spalten (Blöcke)

In obiger Abbildung von WINPROF.KEX wird in den Zeilen 00057 bis 00058 definiert, auf welchen Positionen die 4 Spalten (Blöcke) beginnen. Beim ElternBlock und tomBlock kann es, abhängig von der Einstellung der Schreibmaschine, mit der Josef Heider die Registerseiten geschrieben hat, leichte Verschiebungen geben. Diese Werte sind daher gegebenenfalls anzupassen.

4.2.2 Reihenfolge von tomus und pagina

In den Heider-Büchern gibt es drei Varianten, wie tomus und pagina angeordnet sein können. In der Variablen TomSequence wird die aktuelle Reihenfolge definiert. Im Profile sieht dies wie folgt aus:

```
'editv set TomSequence A'      /* Reihenfolge: Tom Pag (Normalfall) */
'editv set TomSequence B'      /* Reihenfolge: Pag Tom */
'editv set TomSequence C'      /* Pagina unterhalb von Tom */
```

In der Version C gibt es noch Unterformen, die vom Programm analysiert werden müssen, um die pagina herauszufiltern. Diese verschiedenen Formen sind allerdings bisher nur bei den Heiraten vorgekommen. Bei den Sterberegistern ist bis zu Ende 2025 nur in der Pfarre Bad Aussee die Form B vorgekommen.

4.2.3 Liste der tomus-Nummern

Ebenfalls anzupassen ist die Liste der tomus-Nummern in der Variablen AllTom. Das sind die Nummern der Matrikenbücher einer Kategorie einer Pfarre (Geburt, Heirat, Sterbe) in chronologisch aufsteigender Reihenfolge. Sie sind oft auf einer der Anfangsseiten eines Registerbuches angegeben, auf alle Fälle aber im Verzeichnis der Kirchenbücher *Heider-Register: Verzeichnis der erfassten Matrikenbücher nach Pfarren*, das Josef Heider angelegt hat und auf den Netzseiten der Familia Austria aufgerufen werden kann ¹.

```
AllTom='I II III IV'
```

Die tomus-Nummern sind ein- oder mehrstellige Zeichen bzw. Begriffe, die durch eine Leerstelle getrennt sind. Meistens sind es römische Nummern, aber nicht immer. In obigem Beispiel sind die tomus-Nummern alphabetisch aufsteigend. Durch besondere tomus-Bezeichnungen wie z.B. E für evangelische ist diese Sortierfolge aber nicht gewährleistet. In jedem Fall sind sie aber so einzugeben, wie sie Josef Heider in seiner Aufstellung je Kategorie definiert hat, denn so hat Heider sie auf den Registerseiten unter Tom. angeführt. (Die Chronologie ist dann unterbrochen, wenn sich Kirchenbücher zeitlich überlappen, z.B. bei Führung eines eigenen Buches für uneheliche Kinder.). Die Variable AllTom kann im WINPROF.KEX mehrfach definiert werden. Das kann nützlich sein, wenn man zwischen zwei oder mehr Pfarren öfters wechselt. Wirksam ist jeweils die letzte Eintragung der Variablen.

```
AllTom='I II III IV'          /* Walding Taufen */
AllTom='A I B C II III IIIb' /* Rohrbach Taufen */
```

Die Profile-Datei WINPROF.KEX befindet sich im selben Verzeichnis wie die KEDIT-Makros und kann von dort direkt aufgerufen und modifiziert werden. Der Aufruf ist aber auch bei einer mit KEDIT geöffneten Register-Datei mit der **F5**-Taste möglich. Nach Veränderung der Profile-Datei nicht vergessen, die Registerseite zu schließen und dann neu zu öffnen, damit die Profile-Datei neu durchlaufen wird und die Änderungen wirksam werden!

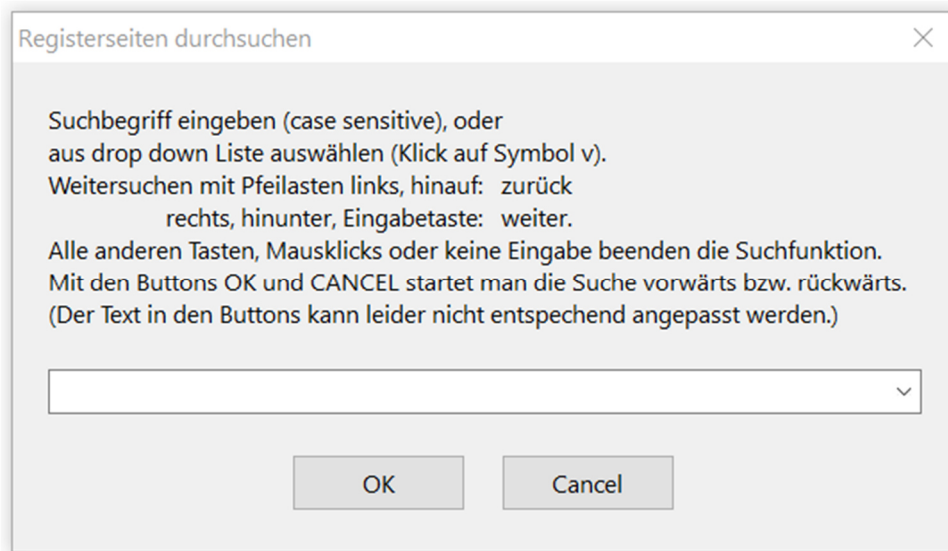
¹ <https://www.familia-austria.at/index.php/component/search/?searchword=heider&searchphrase=all&Itemid=110> (Link erstellt Ende 2024)

4.3 Aufruf von Funktionen für die Tauf- und Sterberegister über Funktionstasten

Es gibt 7 Funktionen, die über Funktionstasten aufgerufen werden.

F2 Ein-/Ausschalten bottom toolbar (der top toolbar ist immer sichtbar), siehe *Arbeiten mit Textblöcken* auf Seite 46. Bei Sterberegistern wird der bottom toolbar durch das Profile eingeschaltet.

F3 Mit dieser Funktion kann man ein ganzes Registerbuch oder Teile davon **Seite für Seite** nach einem Begriff durchsuchen (kann auch aus mehreren Wörtern bestehen), um ihn zu überprüfen oder auch, um ihn zu ändern. Die Suche ist case-sensitiv, der Suchbegriff muss also mit Groß- und Kleinbuchstaben korrekt im Suchfenster eingegeben werden. Mit dem V-Zeichen in der Suchzeile kann man aus der drop down Liste auch früher verwendete Suchbegriffe selektieren. Mit den Pfeiltasten links und rechts bzw. oben und unten auf der Tastatur wird nach dem nächsten oder vorhergehenden Begriff gesucht. Mit der Eingabetaste sucht man in der aktuellen Richtung weiter. Die Suche wird mit Mausklick auf das KEDIT-Fenster oder mit irgendeiner Taste beendet.



Mit OK startet man die Suche von der Curser-Position vorwärts, mit Cance sucht man rückwärts. Die Buttons OK und Cancel sollten eigentlich mit **vorwärts** und **rückwärts** oder + und – bezeichnet sein, aber leider kann man die Beschreibungen nicht ändern.

F4 Nach Kommentaren suchen.

Damit kann man rasch prüfen, ob es Registereintragungen mit einem Kommentar gibt, die einer weiteren Nachprüfung, z.B. über matricula, bedürfen. Mit den Pfeiltasten links und rechts kann man die Suchrichtung bestimmen.

F5 XEDIT-Profil WINPROF.KEX aufrufen.

Damit können Variable im Profile auf die Werte gesetzt werden, wie sie für die Textprüfung benötigt werden (siehe *KEDIT-Profil* auf Seite 18).

F6, F7 nur bei Heiratsregistern

F8, Enter Formatierung und Prüfung einer Registerseite

F10 rückwärtsblättern im aktuellen Registerbuch

F11 vorwärtsblättern im aktuellen Registerbuch

4.4 Starten der Formatierung und Prüfung einer Registerseite

Nach der OCR-Umwandlung, manuellen Korrektur von Fehlern und anschließendem Speichern der Registerseite wird diese automatisch mit KEDIT zur weiteren Bearbeitung aufgerufen. Voraussetzung dafür ist, dass Dateien mit dem Dateityp **txt** mit der Anwendung **KEDIT** geöffnet werden. Das ist im Betriebssystem mit der Funktion **Standard-Apps nach Dateityp auswählen** festzulegen. Eine gespeicherte Registerseite kann aber auch später jederzeit geöffnet und bearbeitet werden. Beim ersten Aufruf hat die Seite noch das Format, wie sie

Josef Heider mit der Schreibmaschine erstellt hat inkl. der Korrekturen, die nach der OCR-Umwandlung vorgenommen wurden (siehe [Abb. 6: lesen und bearbeiten der Registerseiten](#) auf Seite 15). Durch Drücken der Eingabe- oder F8-Taste startet der Formatierungsprozess.

Im Unterschied zu den Taufregistern müssen bei den Sterberegistern die Altersangaben aus dem EBO-Block herausgelöst werden. Dafür gibt es 2 verschiedene Verfahren. Siehe die genaue Beschreibung im Kapitel [tomus-Block und Altersangabe beim Sterberegister](#) auf Seite 25.

4.5 Lesen der vorhergehenden Seite

Für manche Prüfungen ist die letzte Registereintragung der vorhergehenden Seite mit dem gleichen Registerbuchstaben maßgeblich. Am Beginn des Prüfprozesses wird daher diese Eintragung gesucht. Wird keine vorhergehende Seite gefunden, weil z.B. die aktuelle Seite die erste des Registerbuches ist oder hat die vorhergehende Registerseite einen anderen Index-Buchstaben, entfällt die Prüfung auf gleichen Anfangsbuchstaben und eine Nachricht weist darauf hin.

4.6 Seitennummer der Heider-Registerseite

Die Seitennummer steht zwischen 2 Minus-Zeichen (die erste allein genügt). Die Seitennummer ist zwingend notwendig und muss mit der Seitennummer im Dateinamen übereinstimmen. Bei fehlender Seitennummer oder bei Nichtübereinstimmung bricht das Programm mit einer Nachricht ab, ohne dass Daten verändert werden.

Abweichungen zwischen Seitennummer des Registerbuches und der Dateinummer in den Scan- und Text-Dateien entstehen durch fehlerhafter Seitennummerierung im Heider-Register, z.B. durch Doppelvergabe derselben Seitennummer oder Auslassung von Seitennummern. Da die Dateinummer einmalig ist, kann sie nicht mehrfach vorkommen und muss in solchen Fällen abweichend von der Heider-Nummer vergeben werden. Damit man auch im Nachhinein weiß, welche Seitennummer Heider ursprünglich vergeben hat, ist diese in der Textdatei nach der berichtigten Seitennummer zwischen runden Klammern anzugeben.



Abb. 9: Die Seite 171B wurde eingeschoben, weil Heider die Nr. 171 zweimal vergeben hat

Mehr zum Thema Seitennummer findest du im Dokument [Heider-Projekt technische Dokumentation](#) im Kapitel [Fehler in der Nummerierung der Heider-Registerseiten und ihre Behandlung im Dateinamen](#).

4.7 Überschriftzeile

Die Überschrift über den Spalten beginnt mit dem Wort **Tag**. Diese Zeile wird automatisch erkannt und gelöscht. Die Zeile kann aber auch zur Gänze fehlen. Ein manueller Eingriff ist nur erforderlich, wenn die Überschriftzeile nicht mit dem Schlüsselwort **Tag** beginnt, weil es z.B. durch das OCR-Programm falsch übersetzt wurde. In diesem Fall muss das Schlüsselwort manuell korrigiert oder die ganze Zeile manuell gelöscht werden. Die Texte nach **Tag** sind irrelevant.

4.8 Leerzeilen

Am Beginn einer Registerseite sind bis zum Beginn der ersten Eintragung ein oder mehrere Leerzeilen. Diese werden durch das Programm automatisch gelöscht. Auch Leerzeilen zwischen den Eintragungen werden vom Programm gelöscht und benötigen daher keinen manuellen Eingriff.

4.9 Eintragung

Eine Eintragung besteht in der Regel aus 2 Zeilen, von denen die zweite oft leer ist. Bei umfangreicheren Eintragungen können es auch 3 oder mehr Zeilen sein. In der ersten Zeile steht das Datum des Vorganges (Geburt etc.) und sie endet mit der tomus-Spalte. In den Folgezeilen kann der Text noch in den tomus-Block hineinreichen.

4.9.1 Name

Der Name beginnt auf der Position 12 oder später. Der Anfangsbuchstabe muss ein Großbuchstabe sein und mit dem Indexbuchstaben der Seite übereinstimmen. Vor dem Namen kann ein *v.*, *von*, *de*, oder eine (stehen. Der relevante Anfangsbuchstabe steht dann entsprechend weiter rechts. Bei der Prüfung auf Übereinstimmung ist das spezielle Alphabet im Heider-Index zu beachten, siehe im Dokument *Heider-Register und Digitalisierungsprojekt Einführung* im Kapitel *Alphabetischer Index*.

Ein Register-fremder Anfangsbuchstabe wird durch eine Fehlermeldung und durch gelbe Einfärbung der Zeile angezeigt. Nichtübereinstimmung kann entweder durch einen Fehler in der OCR-Umwandlung entstehen, die dann zu korrigieren ist, oder Josef Heider hat einen Namen erwischt, der zu einer anderen Registerseite gehört. Das passiert z.B. manchmal, wenn er einen Namen mit *St* unter *S* oder *Sch* etc. einordnet oder die Eintragung schlicht verreicht hat. In diesem Fall tragen wir einen Kommentar ein und lassen die Eintragung unverändert stehen.

Eine besondere Situation entsteht, wenn der Anfangsbuchstabe in der **ersten** Eintragung der Registerseite nicht mit dem Index übereinstimmt. In diesem Fall werden alle anderen, richtigen Namen mit der Fehlermeldung versehen, sie würden mit einem falschen Anfangsbuchstaben beginnen und die Eintragungen werden gelb markiert. Damit gehen echte Fehler in den vielen gelb markierten Zeilen unter. Daher übertippe ich den Anfangsbuchstaben in der ersten Eintragung mit dem aktuellen Indexbuchstaben, prüfe die Registerseite wie üblich und mache die ‚falsche Korrektur‘ des Anfangsbuchstaben in der ersten Eintragung erst vor dem Speichern der Seite wieder rückgängig.

4.9.2 Vertauschte Vor- und Familiennamen

Es kommt vor, dass Josef Heider in der Namensspalte die Reihenfolge Familienname – Vorname vertauscht hat. Hier ist die Reihenfolge richtig zu stellen. Dasselbe gilt auch, wenn er die Namens- und die EBO-Spalte vertauscht hat. Mit Selektieren und Verschieben mit gedrückter Maus-Taste geht das sehr schnell.

4.9.3 Sonderfälle beim Namen

4.9.3.1 Alternative Schreibweise des Familiennamens

- 32 -

Tag	Name	Eltern, Beruf, Ort	Tom.
21.10.1597	Bell Eva	Georg, Eckersberg	AI 3
9. 2.1598	Pausch Johann	Wolfgang, Inwohner	AI 4
15. 3.1598	Peiter Georg	Mathias, Altenfelden	AI 4
4. 7.1598	Perntaller Simon	Adam, Weber, zu Edt	AI 6
8. 8.1698	Pausch Maaria	Stefan, Azesberg	AI 7
11.10.1598	Pernecker Barbara	Paul, von Lanfeldt	AI 8
3. 5.1599	Pocksrucker Florian	Georg, Kazenbach	AI 14
2. 9.1599	Plaumauer Maria	Mathias, Altenfelden	AI 16
14. 9.1599	B(W)iblinger Mathias	Urban, Altenfelden	AI 17
17. 6.1600	Pauer Thomas	Hauns, Neundling	AI 25

Abb. 10: alternative Schreibweise des Familiennamens

Heider hat manchmal verschiedene Schreibweisen des Familiennamens entweder direkt hintereinander angeführt oder, wie in der vorletzten Zeile in obiger Abbildung, mit einem anderen Anfangsbuchstaben in Klammern angegeben. Der Familienname kann in dieser Form mit einer Suchfunktion nur schwer gefunden werden. Daher ist die 2. Namensvariante in voller Länge, aber in Klammern, nach der ersten Variante anzuführen:

14. 9.1599 Biblinger (Wiblinger) Mathias Urban, Altenfelden AI 17

4.9.3.2 Vulgo-Namen, 2 Familiennamen

Manchmal sind beim Kind der Familienname und ein Vulgo-Name angegeben:

4.11.1629	vulgo Schramm gen. Puderer Thomas	Leopold	A
20. 5.1630	vulgo Schwarzbart gen. Wallmann Johann	Hans - Ursula	A

Josef Heider hat in diesen Fällen das Kind, so wie bei unehelichen Kindern, unter beiden Namen in seinem Register verzeichnet. Damit findet man es praktischerweise unter sowohl als auch.

Es gibt auch Fälle, in denen beim Kind 2 Familiennamen ohne vulgo angeführt werden:

5. 5.1633 Schwarzbart oder Wallmann Tobias Hans - Ursula A

Vermutlich handelt es sich auch hier um einen vulgo-Namen ohne es explizit anzugeben. Diese Fälle werden bei der Aufteilung der Daten behandelt bzw. sind durch den Forscher oder die Forscherin zu klären.

4.9.3.3 Sohn, Tochter, Kind

Vor allem bei Totgeburten oder kurz nach der Geburt verstorbener Kindern finden sich die Bezeichnungen **Sohn**, **Tochter** oder **Kind**. Mit Sohn und Tochter ist zumindest das Geschlecht bestimmt. Den Begriff **Kind** bitte durch 5 Punkte (.....) ersetzen.

4.10 Eltern, Beruf, Ort (EBO)

In der Elternzeile kann es bereits zu kleinen Verschiebungen des Textes kommen, sodass der Text schon eine Stelle vor der definierten Position des Elterntextes stehen kann, sofern davor ein Leerzeichen ist (sonst wäre es ja Teil des Namens). Da eine einstellige Linksverschiebung des Elterntextes öfters vorkommt, wird das vom

Programm toleriert. Beginnt der Elterntext weiter rechts, werden die entstandenen Leerstellen eliminiert, genauso wie mehrfache Leerstellen zwischen den Textteilen. Nach erfolgter Formatierung beginnt der Elterntext genau an der vorgesehenen Position. Abgekürzte Namen kann man in eindeutigen Fällen ergänzen, z.B. Schwertbg. in Schwertberg, um sie bei der DB-Suche leichter auffindbar zu machen.

4.11 Zahlenangaben in der EBO-Spalte

Manchmal kommen in der EBO-Spalte Zahlen vor, z.B. ein Sterbedatum. Weil das Programm nicht wissen kann, ob es sich um eine Ziffer oder um einen Buchstaben handelt, werden die Zeichen l und o vom Formatierungsprogramm nicht verändert. Sollte es sich tatsächlich um Ziffern und nicht um Buchstaben handeln, müssen dies händisch korrigiert werden.

4.12 tomus-Block mit tomus und pagina

tomus, die Nummer des Buches, ist zwingend. Wenn der tomus fehlt, werden die nächsten Zeichen (*pagina*), so vorhanden, als tomus interpretiert. Das ergibt dann normalerweise eine ungültige tomus-Nummer und führt zu einer entsprechenden Fehlermeldung.

Heider hat besonders bei den Heiraten oft keine *pagina* in seinen Registern angegeben. Die Ursache kann sein, dass das Kirchenbuch keine Seitennummern enthält, oder weil Heider sich die Arbeit ersparen wollte. Wenn bei der manuellen Kontrolle der Register zufällig eine Eintragung im Originalbuch mit einer Seitennummer gefunden wird, die Heider nicht in seine Registereintragung übernommen hat, tragen wir diese in der Registerseite nach. Das ist sozusagen unser Service für die Benutzer.

Die normale Form ist *tomus pagina* oder nur *tomus*, wenn die Seitennummer nicht angegeben ist. Manchmal ist die Reihenfolge umgedreht: *pagina/tomus*. In der Regel zieht sich die Form *tomus pagina* oder *pagina tomus* durch das ganze Registerbuch. Das wird vom Programm korrekt behandelt, sofern diese umgedrehte Form vorher durch die Variable *tomSequence* in WINPROF.KEX definiert wurde. In Fällen, in denen Matrikenbücher derselben Kategorie unterschiedlich paginiert wurden, kann die Reihenfolge mitten in einer Registerseite wechseln. Dann sind *tomus* und *pagina* durch einen Schrägstrich oder durch einen Punkt getrennt. Hier ist ein manueller Eingriff zur Korrektur notwendig.

4.12.1 tomus-Block beim Taufregister

Zunächst muss das Programm wissen, wo der Tom.-Block beginnt. Bei den Taufen wird bei noch nicht formatierten Eintragungen die im WINPROF.KEX definierte Zahl in der Variablen *tomBlock* verwendet. Diese Zahl definiert also die Position der tomus-Spalte. Diese Positionsangabe ist aber nicht absolut zu sehen, da durch die OCR-Umwandlung Verschiebungen in den Eintragungen passieren können. Das Programm geht daher zunächst von der definierten Position 3 Stellen weiter nach links und sucht von dort nach rechts das erste Zeichen, das kein Leerzeichen ist. Der dort beginnende Text wird als tomus aufgefasst. Durch diese Vorgehensweise wird auch ein tomus erkannt, der weiter rechts steht, als in der Variablen *tomBlock* definiert ist. Voraussetzung ist, dass der Cursor außerhalb des Bereiches vom tomus-Block steht, also vor Position *tomBlock*-4 oder nach *tomBlock*+2. Ist bei einer Seite der tomus zu weit links, weil z.B. FineReader durch sehr viele Leerstellen in der EBO-Spalte Zeichen verschluckt hat, kann man alternativ den Cursor an den Beginn des am weitesten links-stehenden tomus auf der Registerseite stellen. Das Suchen nach der tomus-Nummer beginnt dann von dieser Position aus. Man kann das leicht ausprobieren, da man Änderungen jederzeit rückgängig machen kann (siehe [Rückgängig machen von Änderungen](#) auf Seite 36). Weiter rechtsstehender tomus wird gefunden, sofern kein überlanger Text aus der EBO-Spalte hineinragt. In diesem Fall muss man händisch eingreifen.

Ist die Zeile bereits formatiert, stehen in der ersten Zeile (mit Schlüsselwort Block) die Positionen der 4 Blöcke (siehe [Blockzeile in der Text-Datei](#) weiter unten). Die Suche nach dem tomus beginnt dann auf der dort angegebenen Position für den tomus-Block. Die Position des Cursors wird ignoriert.

Da in der *pagina*, die nach dem tomus folgt, Kleinbuchstaben vorkommen können, wird **c nicht in 0** umgewandelt. Ist tomus absteigend oder ungültig oder die *pagina* innerhalb desselben tomus absteigend, wird eine Fehlermeldung ausgegeben und die Zeile gelb markiert.

Ein absteigender **tomus** ist in der Regel ein Tippfehler, vereinzelt auftretende absteigende **pagina** sind meist durch Verreihung der Eintragungen entstanden oder ebenfalls durch Tippfehler. Gehäuft hin- und herspringende *pagina* treten als Folge der unseligen Reform der Kirchenbücher von 1784 auf, wonach die Eintragungen

nach Ortschaften zu erfolgen hatte. Diese Bestimmung hatte nur Nachteile und wurde bald wieder zurückgenommen. Ihre Anwendung führt dazu, dass im Heider-Register die pagina ständig hin- und herspringt. Jeder Rücksprung in der pagina führt zu einer Fehlermeldung und das können dann auf einer Seite sehr viele sein. Durch Eingabe des Buchstaben **P** bzw. **p** (pagina) in der Befehlszeile (gelb eingefärbte Zeile am unteren Fensterrand nach dem ==>) werden diese Fehlernachrichten unterdrückt. Andere Fehler wie z.B. absteigendes Datum werden aber weiterhin angezeigt. Die Unterdrückung der Fehlermeldungen wird in der Blockzeile der Registerseite vermerkt und gilt damit auch bei einer späteren Bearbeitung dieser Seite.

4.12.2 tomus-Block und Altersangabe beim Sterberegister

Die Besonderheit in den Sterberegistern gegenüber den Heiraten und den Taufen besteht in der Angabe des Alters, das am Ende des EBO-Blocks stehen kann (die Altersangabe ist optional, kann also fehlen). Sie besteht aus ein bis drei Ziffern und dem Buchstaben **T** für Tag, **W** für Wochen und **M** für Monate. Ist kein Buchstabe angegeben, sind mit der ein- bis dreistelligen Zahl **Jahre** gemeint. Diese Altersangabe muss nun vom Prüf- und Formatierungs-Macro von der EBO-Information herausgelöst werden. Für diese Auftrennung stehen zwei Versionen des KEDIT-Makros zur Verfügung:

- ☐ Format_Sterbe_manuell
- ☐ Format_Sterbe_auto

In der manuellen Version muss die Trennung der Altersangabe von der EBO-Spalte manuell gemacht werden, in der auto-Version erfolgt die Trennung durch das Programm. Das funktioniert aber dzt. nur bei Registerbüchern mit der Reihenfolge **tomus pagina** bzw. nur **tomus** (TomSequence='A'). Für Registerbücher mit umgedrehter Reihenfolge **pagina tomus** (TomSequence='B') muss die manuelle Version verwendet werden.

4.12.3 Manuelle Trennung der Altersangabe beim Sterberegister

Zur Erklärung des Verschiebevorganges habe ich mit Aigen Registerseite 2 ein etwas komplexeres Beispiel gewählt. Nach dem OCR-Prozess und den Korrekturen sehen wird In *Abb. 11: Aigen Seite 2 im KEDIT-Editor nach der OCR-Umwandlung* auf Seite 26 von Zeile 00008 bis 00050. In den ersten Eintragungen gibt es keine Altersangaben, in den Zeilen 00032 ragt die Ortsbezeichnung Schindlau in den Bereich der Altersangaben hinein und in der Zeile 00044 sind Orts- und Altersangabe nicht getrennt. Außerdem stehen in den Zeilen 00033 und 00049 die Altersangaben in der Folgezeile und nicht in der ersten Zeile der Eintragung. Diese Fälle müssen wir zuerst einmal händisch korrigieren. Beim Alter sollten wir noch die Buchstaben kontrollieren. Ein T kann vom OCR-Programm als I oder als 1 und ein M als W bzw. umgekehrt umgewandelt worden sein und das ist nur durch manuelle Kontrolle erkennbar. Man könnte nun Alter, tomus und pagina von jeder Eintragung durch Einfügen von Leerstellen so weit nach rechts verschieben, dass sie weit genug von den EBO-Texten entfernt sind.

Schneller als die zeilenweise Verschiebung ist es, den Bereich von Alter bis pagina als ganzen Block auszuwählen und diesen nach rechts zu verschieben. Dazu ist die Funktion *Arbeiten mit Textblöcken* wie auf Seite 46 beschrieben, sehr hilfreich. Zuerst stellen wir den Cursor in der ersten Eintragung vor die am weitesten links stehende Altersangabe der Seite. Das ist in unserem Fall die Position 54. Dann ziehen wir bei gedrückter Alt-Taste den Cursor nach rechts unten, sodass ein schwarzer Auswahlblock entsteht. Die Verschiebung des ausgewählten Blockes nach rechts muss nun so weit erfolgen, dass der Abstand zwischen dem Ende der EBO-Information und dem tomus mindestens 6 Stellen beträgt. In der bottom toolbar drücken wir dazu entsprechend oft auf das rote Feld mit dem Pfeil nach rechts. Die bottom toolbar wird beim Aufrufen eines Sterberegisters mit KEDIT bereits automatisch eingeblendet, kann aber auch mit der F2-Taste ein- und ausgeblendet werden.

EBO-Informationen, die bis in den Bereich der Altersangabe reichen, werden allerdings bei der Definition des Verschiebblockes mit ausgewählt und somit mitverschoben. In unserem Beispiel ist das in den Zeilen 10, 32, 44 und 48 der Fall. In *Abb. 12: Bereich mit Alter und tom-Block nach rechts verschoben* auf Seite 27 ist das gut zu sehen. Diese müssen nun wieder zurückgeschoben werden, was durch Auswahl und Ziehen mit der Maus allerdings sehr rasch geht. Auch müssen Altersangaben, die in der Folgezeile stehen, auf dieselbe Art an die richtige Stelle der ersten Eintragszeile geschoben werden.

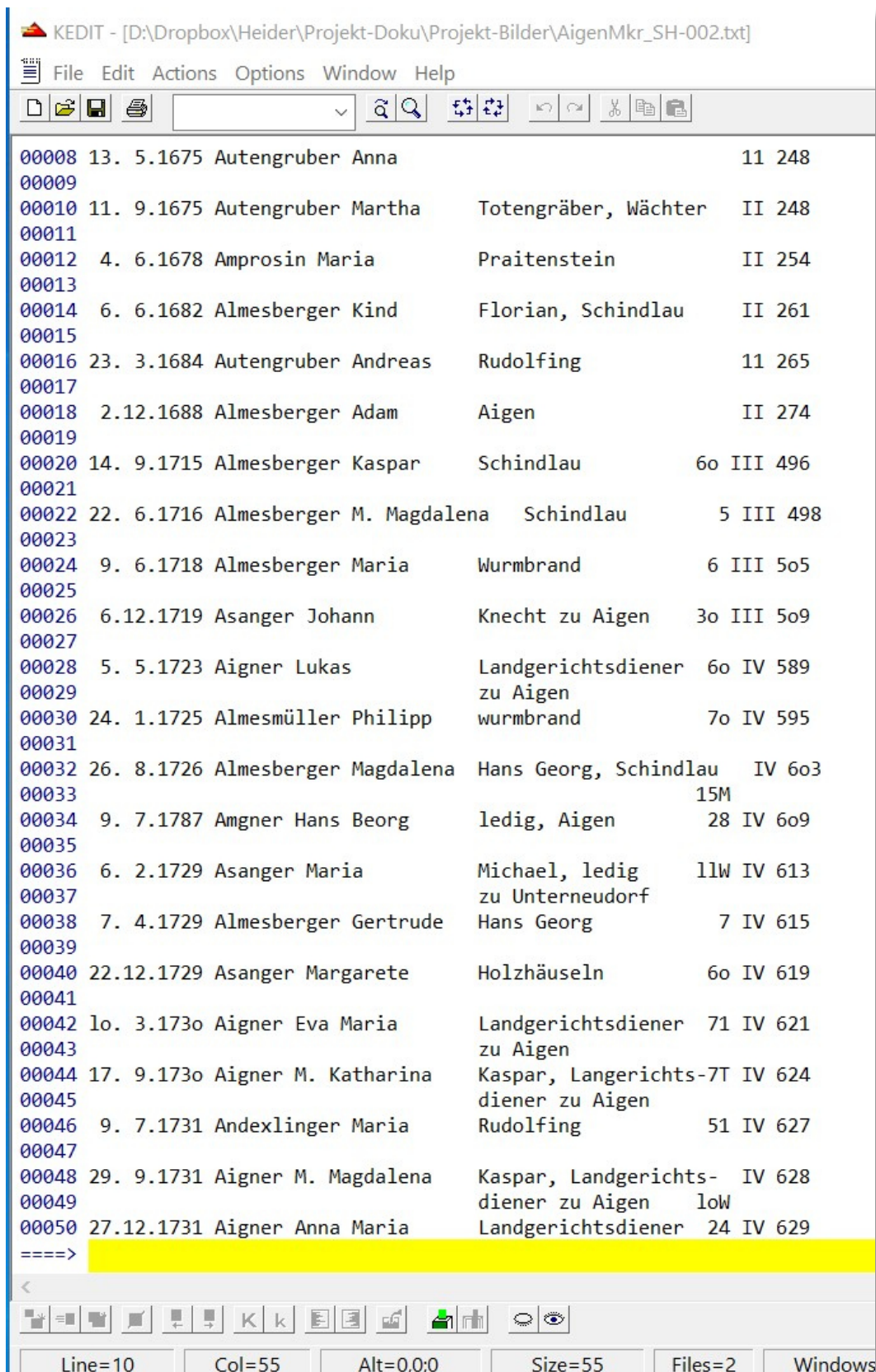


Abb. 11: Aigen Seite 2 im KEDIT-Editor nach der OCR-Umwandlung

Es sind die Zeilen 00008 bis 00050 dargestellt. Am unteren Ende ist die Bottom Tool Bar eingeblendet.

00001	Tag	Name	Eltern, Beruf, Ort	Tom.
00002				
00003				
00004	18.10.1674	Aichinger Stefan	Aigen	II 247
00005				
00006	20. 8.1674	Autengruber Michael	Rudolfing	II 247
00007				
00008	13. 5.1675	Autengruber Anna		11 248
00009				
00010	11. 9.1675	Autengruber Martha	Totengraber, Wächte	r II 248
00011				
00012	4. 6.1678	Amprosin Maria	Praitenstein	II 254
00013				
00014	6. 6.1682	Almesberger Kind	Florian, Schindlau	II 261
00015				
00016	23. 3.1684	Autengruber Andreas	Rudolfing	11 265
00017				
00018	2.12.1688	Almesberger Adam	Aigen	II 274
00019				
00020	14. 9.1715	Almesberger Kaspar	Schindlau	6o III 496
00021				
00022	22. 6.1716	Almesberger M. Magdalena	Schindlau	5 III 498
00023				
00024	9. 6.1718	Almesberger Maria	Wurmbrand	6 III 5o5
00025				
00026	6.12.1719	Asanger Johann	Knecht zu Aigen	3o III 5o9
00027				
00028	5. 5.1723	Aigner Lukas	Landgerichtsdieners zu Aigen	6o IV 589
00029				
00030	24. 1.1725	Almesmüller Philipp	Wurmbrand	7o IV 595
00031				
00032	26. 8.1726	Almesberger Magdalena	Hans Georg, Schindl	au IV 6o3
00033				15M
00034	9. 7.1787	Aigner Hans Georg	ledig, Aigen	28 IV 6o9
00035				
00036	6. 2.1729	Asanger Maria	Michael, ledig zu Unterneudorf	11W IV 613
00037				
00038	7. 4.1729	Almesberger Gertrude	Hans Georg	7 IV 615
00039				
00040	22.12.1729	Asanger Margarete	Holzhäuseln	6o IV 619
00041				
00042	10. 3.1730	Aigner Eva Maria	Landgerichtsdieners zu Aigen	71 IV 621
00043				
00044	17. 9.1730	Aigner M. Katharina	Kaspar, Langericht diener zu Aigen	s-7T IV 624
00045				
00046	9. 7.1731	Andexlinger Maria	Rudolfing	51 IV 627
00047				
00048	29. 9.1731	Aigner M. Magdalena	Kaspar, Landgerich diener zu Aigen	ts- IV 628
00049				low
00050	27.12.1731	Aigner Anna Maria	Landgerichtsdieners zu Aigen	24 IV 629
00051				
00052	7. 2.1732	Asanger Thomas	Johann, Zaglau	7W IV 63o
00053				
00054	28, 3.1732	Almesberger Rosina	Hans Georg, zu Schindlau	4 IV 63o
00055				
00056	* * * End of File * * *			

====>

Abb. 12: Bereich mit Alter und tom-Block nach rechts verschoben

Da das Programm Alter und tomus trennen soll, muss sie Grenze zwischen beiden Bereichen eindeutig und definiert sein. In Zeile 00043 ist die Altersangabe 7T bereits im Bereich von tomus. Wir löschen deshalb davor so viele Leerzeichen, bis das Alter genügend weit nach links gerutscht ist. (Die Positionen der Altersangaben und der tomi dürfen sich auf derselben Seite nicht überlappen.)

Mit dieser Vorbereitung können wir die erste Formatierung starten. Die tomus-Position definieren wir für das Programm, indem wir den Cursor auf den (die) am weitesten linksstehenden tomus stellen. Z.B. in Zeile 99920, 00024 oder 00026. Ich habe sie in der folgenden Darstellung durch die rote Linie markiert. Sie trennt Alter und tomus. Die Altersangabe ist innerhalb der blauen und der roten Linie isoliert.

00010	11. 9.1675	Autengruber Martha	Totengräber, Wächter	II	248
00011					
00012	4. 6.1678	Amprosin Maria	Praitenstein	II	254
00013					
00014	6. 6.1682	Almesberger Kind	Florian, Schindlau	II	261
00015					
00016	23. 3.1684	Autengruber Andreas	Rudolfing	11	265
00017					
00018	2.12.1688	Almesberger Adam	Aigen	II	274
00019					
00020	14. 9.1715	Almesberger Kaspar	Schindlau	6o III	496
00021					
00022	22. 6.1716	Almesberger M. Magdalena	Schindlau	5 III	498
00023					
00024	9. 6.1718	Almesberger Maria	Wurmbrand	6 III	5o5
00025					
00026	6.12.1719	Asanger Johann	Knecht zu Aigen	3o III	5o9
00027					
00028	5. 5.1723	Aigner Lukas	Landgerichtsdieners zu Aigen	6o IV	589
00029					
00030	24. 1.1725	Almesmüller Philipp	wurmbrand	7o IV	595
00031					
00032	26. 8.1726	Almesberger Magdalena	Hans Georg, Schindlau	15M IV	6o3
00033	9. 7.1787	Amgner Hans Georg	ledig, Aigen	28 IV	6o9
00034					
00035	6. 2.1729	Asanger Maria	Michael, ledig zu Unterneudorf	11W IV	613
00036					
00037	7. 4.1729	Almesberger Gertrude	Hans Georg	7 IV	615
00038					
00039	22.12.1729	Asanger Margarete	Holzhäuseln	6o IV	619
00040					
00041	1o. 3.173o	Aigner Eva Maria	Landgerichtsdieners zu Aigen	71 IV	621
00042					
00043	17. 9.173o	Aigner M. Katharina	Kaspar, Langerichts- diener zu Aigen	7T IV	624
00044					
00045	9. 7.1731	Andexlinger Maria	Rudolfing	51 IV	627
00046					
00047	29. 9.1731	Aigner M. Magdalena	Kaspar, Landgerichts- diener zu Aigen	1oW IV	628
00048					

Abb. 13: EBO-Information und Altersangabe entflechtet

Nach Drücken die Eingabetaste wird am Beginn der Datei eine Blockzeile erzeugt und die Seite formatiert, wie wir es von den Geburtsregistern kennen inkl. der tomus-Spalte. Zwischen der EBO-Spalte und der tomus-Spalte sehen wir die Altersangabe: Alles, was unmittelbar nach oder auch weiter rechts von der Cursor-Position steht (rote Linie), wird als tomus und pagina, was links davon steht, wird als Altersangabe interpretiert. Fehler,

die das Programm beim Formatieren und Prüfen erkennt wie falsches oder absteigendes Datum etc. werden wie gewohnt farblich markiert und können anschließend korrigiert werden (siehe *Abb. 16: formatierte Seite des Sterberegisters mit Altersangaben*).

Line	Date	Name	Location	Page
00000	* * * Top of File * * *			
00001	Block 1,12,40,70,A, - 2 -			
00002	18.10.1674	Aichinger Stefan	Aigen	II 247
00003	20. 8.1674	Autengruber Michael	Rudolfing	II 247
00004	13. 5.1675	Autengruber Anna		II 248
00005	11. 9.1675	Autengruber Martha	Totengräber, Wächter	II 248
00006	4. 6.1678	Amprosin Maria	Praitenstein	II 254
00007	6. 6.1682	Almesberger Kind	Florian, Schindlau	II 261
00008	23. 3.1684	Autengruber Andreas	Rudolfing	II 265
00009	2.12.1688	Almesberger Adam	Aigen	II 274
00010	14. 9.1715	Almesberger Kaspar	Schindlau	60 III 496
00011	22. 6.1716	Almesberger M. Magdalena	Schindlau	5 III 498
00012	9. 6.1718	Almesberger Maria	Wurmbrand	6 III 505
00013	6.12.1719	Asanger Johann	Knecht zu Aigen	30 III 509
00014	5. 5.1723	Aigner Lukas	Landgerichtsdieners	60 IV 589
00015			zu Aigen	
00016	24. 1.1725	Almesmüller Philipp	Wurmbrand	70 IV 595
00017	26. 8.1726	Almesberger Magdalena	Hans Georg, Schindlau	15M IV 603
00018	9. 7.1787	Amgner Hans Beorg	ledig, Aigen	28 IV 609
00019	6. 2.1729	Asanger Maria	Michael, ledig	11W IV 613
00020			zu Unterneudorf	
00021	7. 4.1729	Almesberger Gertrude	Hans Georg	7 IV 615
00022	22.12.1729	Asanger Margarete	Holzhauseln	60 IV 619
00023	10. 3.1730	Aigner Eva Maria	Landgerichtsdieners	71 IV 621
00024			zu Aigen	
00025	17. 9.1730	Aigner M. Katharina	Kaspar, Langerichts-	7T IV 624
00026			diener zu Aigen	
00027	9. 7.1731	Andexlinger Maria	Rudolfing	51 IV 627
00028	29. 9.1731	Aigner M. Magdalena	Kaspar, Landgerichts-	10W IV 628
00029			diener zu Aigen	
00030	27.12.1731	Aigner Anna Maria	Landgerichtsdieners	24 IV 629
00031			zu Aigen	
00032	7. 2.1732	Asanger Thomas	Johann, Zaglau	7W IV 630
00033	28. 3.1732	Almesberger Rosina	Hans Georg, zu	4 IV 630
00034			Schindlau	
00035	* * * End of File * * *			

Abb. 14: formatierte Seite des Sterberegisters mit Altersangaben

Der Block mit Alter, tomus und pagina wird bei den Sterberegistern etwas anders behandelt als bei den Heiraten und Geburten. Das Programm fügt zwischen EBO-Block und der Altersangabe das Trennzeichen | ein. Das ist eine Hilfe beim Auseinanderhalten vom EBO-Text und der Altersangabe, vor allem, wenn diese fehlt. Was rechts vom Trennzeichen und vor dem tomus steht, ist eine Altersangabe, was links vom Trennzeichen steht, gehört zum EBO-Text. Nun kann der EBO-Text beliebig manipuliert werden. Dabei kann sich das Trennzeichen nach links oder rechts verschieben, darf aber nicht gelöscht oder überschrieben werden. Fehlt es bei neuerlichen Formatierung, erfolgt eine Fehlermeldung. Man muss dann das fehlende Trennzeichen händisch einfügen und die Formatierung neu durchführen. Die formatierte Registerseite (-datei) wird mit den Trennzeichen an die nächste Verarbeitungsstufe weitergegeben. Die Angabe für tomBlock in der Blockzeile (ersten

Zeile der formatierten Registerseite) ist auch hier die Position des tomus-Blockes (und nicht die der Altersangabe). Siehe *Blockzeile in der Text-Datei* auf Seite 33.

4.12.4 Automatische Trennung der Altersangabe beim Sterberegister

Ob die Trennung der Altersangabe manuell zeilenweise oder blockweise erfolgt, in beiden Fällen ist sie ein zeitraubender Vorgang, der viel Aufmerksamkeit erfordert und ermüdend ist. Ich habe daher nach einer automatischen Lösung per Programm gesucht, die wegen der Variabilität der Daten in dem Bereich und im Zusammenhang mit OCR-Umwandlungsfehlern schwierig zu finden war. Dabei beginnt das Programm vom Ende des EBO-Blockes die letzten drei Datenfelder zu analysieren. Wenn das letzte oder vorletzte Datenfeld nicht ein gültiger tomus ist, wird die Zeile markiert und nicht weiterbearbeitet. Für die erfolgreiche automatische Trennung gilt daher folgende Bedingungen bzw. Voraussetzung:

- Im tomus-Block können tomus und pagina oder nur tomus stehen.
- Die Reihenfolge der tomus-Felder kann **tomus** oder **tomus pagina** sein, die umgekehrte Reihenfolge **pagina tomus** führt zu einem Fehler.
- Das letzte oder vorletzte Datenfeld im tomus-Block muss ein gültiger tomus sein. Kleinbuchstaben werden als Großbuchstaben interpretiert (dadurch entfällt das lästige Umschalten auf der Tastatur, wenn ein tomus korrigiert werden muss).
- Am Ende des EBO-Blockes, also vor dem tomus-Block, kann ein Alter sein. Das Alter ist entweder eine ein- bis 3-stellige Zahl oder eine ein- bis zweistellige Zahl mit den Buchstaben M, T oder W an der rechten Stelle.
- Wenn das Alter nicht in der Datumszeile sondern in der Folgezeile einer Eintragung steht, muss es manuell an die richtige Position in die erste Zeile verschoben werden.
- Alter, tomus und pagina müssen alleinstehende, also durch eine oder mehrere Leerstellen getrennte Datenfelder sein.

Die berichtigte Seite von unserem Beispiel sehen wir in *Abb. 15: die für die automatische Trennung vorbereitete Seite 2 des Sterberegisters Aigen* auf Seite 31. Die manuellen Korrekturen sind rot eingrahmt. Nach Drücken der Eingabetaste wird die Seite formatiert und geprüft. Die Position des Cursors ist nicht von Relevanz, denn wo das Ende des EBO-Blockes ist, ist in WINPROF.KEX definiert und ist daher dem Programm bekannt. Das Ergebnis der Erstformatierung sieht dann genauso aus, wie bei der manuellen Trennungsversion des Programmes (siehe *Abb. 14: formatierte Seite des Sterberegisters mit Altersangaben* auf Seite 29), mit einer Ausnahme: Wenn die automatische Trennung keinen gültigen tomus in der ersten Zeile einer Eintragung findet, bleibt die Zeile so lange unformatiert, bis ein gültiger tomus gefunden wird. Erst dann kann das Programm eindeutig die Daten zu den Feldern Datum, tomus und pagina zuordnen,

In *Abb. 16: formatierte Seite des Sterberegisters mit ungültigem tomus* auf Seite 32 habe ich in der Zeile 00004 den tomus statt auf ii (Kleinbuchstaben sind erlaubt und werden auf große umgewandelt) ix eingegeben. Diese tomus-Angabe ist bei Aigen nicht gültig. Wir sehen daher neben den beiden bekannten Datumsfehlern in den Zeilen 00003 und 00019, dass die Zeile 00004 unbearbeitet geblieben ist.

In beiden Versionen der Trennung EBO von Alter und tomus-Block müssen natürlich Tipp- und OCR-Fehler wie üblich korrigiert werden. Aber das Auseinanderziehen in der manuellen Version erfordert genaues Schauen und entfällt in der automatischen Version. Es ist nur wichtig, dass in der noch unformatierten Datumszeile das letzte Datenfeld im tomus-Block (pagina oder tomus) nicht später als auf Position 70 beginnt (tomBlockEnde im WINPROF.KEX) und das ist in der Regel immer der Fall.

```

00000 * * * Top of File * * *
00001   Tag      Name      Eltern,  Beruf, Ort      Tom.
00002
00003
00004 18.10.1674 Aichinger Stefan      Aigen      II 247
00005
00006 20. 8.1674 Autengruber Michael  Rudolfing  II 247
00007
00008 13. 5.1675 Autengruber Anna      ii 248
00009
00010 11. 9.1675 Autengruber Martha  Totengraber, Wächter  II 248
00011
00012 4. 6.1678 Amprosini Maria      Praitenstein  II 254
00013
00014 6. 6.1682 Almesberger Kind      Florian, Schindlau  II 261
00015
00016 23. 3.1684 Autengruber Andreas  Rudolfing  ii 265
00017
00018 2.12.1688 Almesberger Adam      Aigen      II 274
00019
00020 14. 9.1715 Almesberger Kaspar  Schindlau  6o III 496
00021
00022 22. 6.1716 Almesberger M. Magdalena  Schindlau  5 III 498
00023
00024 9. 6.1718 Almesberger Maria      Wurmbrand  6 III 505
00025
00026 6.12.1719 Asanger Johann      Knecht zu Aigen  3o III 509
00027
00028 5. 5.1723 Aigner Lukas      Landgerichtsdieners  6o IV 589
00029      zu Aigen
00030 24. 1.1725 Almesmüller Philipp  Wurmbrand  7o IV 595
00031
00032 26. 8.1726 Almesberger Magdalena  Hans Georg, Schindlau 15M IV 603
00033 9. 7.1787 Aigner Hans Georg      ledig, Aigen  28 IV 609
00034
00035 6. 2.1729 Asanger Maria      Michael, ledig  11W IV 613
00036      zu Unterneudorf
00037 7. 4.1729 Almesberger Gertrude  Hans Georg  7 IV 615
00038
00039 22.12.1729 Asanger Margarete  Holzhäuseln  6o IV 619
00040
00041 10. 3.1730 Aigner Eva Maria      Landgerichtsdieners  71 IV 621
00042      zu Aigen
00043 17. 9.1730 Aigner M. Katharina  Kaspar, Landgerichts- 7T IV 624
00044      diener zu Aigen
00045 9. 7.1731 Andexlinger Maria      Rudolfing  51 IV 627
00046
00047 29. 9.1731 Aigner M. Magdalena  Kaspar, Landgerichts- 10W IV 628
00048      diener zu Aigen
00049 27.12.1731 Aigner Anna Maria      Landgerichtsdieners  24 IV 629
00050      zu Aigen
00051 7. 2.1732 Asanger Thomas      Johann, Zaglau  7W IV 630

```

Abb. 15: die für die automatische Trennung vorbereitete Seite 2 des Sterberegisters Aigen

Zeile 3: Datum kleiner als vorhergehendes Datum: 18.10.1674
 Zeile 4: kein gültiger tomus gefunden:
 Zeile 19: Datum kleiner als vorhergehendes Datum: 09.07.1787

00000	* * * Top of File * * *				
00001	Block 1,12,40,70,A, - 21 -				
00002	18.10.1674	Aichinger Stefan	Aigen	II	247
00003	20. 8.1674	Autengruber Michael	Rudolfing	II	247
00004	13. 5.1675	Autengruber Anna	ix 248		
00005	11. 9.1675	Autengruber Martha	Totengraber, Wächter	II	248
00006	4. 6.1678	Amprosini Maria	Praitenstein	II	254
00007	6. 6.1682	Almesberger Kind	Florian, Schindlau	II	261
00008	23. 3.1684	Autengruber Andreas	Rudolfing	II	265
00009	2.12.1688	Almesberger Adam	Aigen	II	274
00010	14. 9.1715	Almesberger Kaspar	Schindlau	60 III	496
00011	22. 6.1716	Almesberger M. Magdalena	Schindlau	5 III	498
00012	9. 6.1718	Almesberger Maria	Wurmbrand	6 III	505
00013	6.12.1719	Asanger Johann	Knecht zu Aigen	30 III	509
00014	5. 5.1723	Aigner Lukas	Landgerichtsdieners zu Aigen	60 IV	589
00015					
00016	24. 1.1725	Almesmüller Philipp	Wurmbrand	70 IV	595
00017	26. 8.1726	Almesberger Magdalena	Hans Georg, Schindlau	15M IV	603
00018	9. 7.1787	Aigner Hans Georg	ledig, Aigen	28 IV	609
00019	6. 2.1729	Asanger Maria	Michael, ledig	11W IV	613
00020			zu Unterneudorf		
00021	7. 4.1729	Almesberger Gertrude	Hans Georg	7 IV	615
00022	22.12.1729	Asanger Margarete	Holzhauseln	60 IV	619
00023	10. 3.1730	Aigner Eva Maria	Landgerichtsdieners zu Aigen	71 IV	621
00024					
00025	17. 9.1730	Aigner M. Katharina	Kaspar, Landgerichtsdieners zu Aigen	7T IV	624
00026					
00027	9. 7.1731	Andexlinger Maria	Rudolfing	51 IV	627
00028	29. 9.1731	Aigner M. Magdalena	Kaspar, Landgerichtsdieners zu Aigen	10W IV	628
00029					
00030	27.12.1731	Aigner Anna Maria	Landgerichtsdieners zu Aigen	24 IV	629
00031					
00032	7. 2.1732	Asanger Thomas	Johann, Zaglau	7W IV	630
00033	28. 3.1732	Almesberger Rosina	Hans Georg, zu Schindlau	4 IV	630
00034					
00035	* * * End of File * * *				

Abb. 16: formatierte Seite des Sterberegisters mit ungültigem tomus

4.12.5 Aufruf der Version manuell oder automatisch

Der Aufruf der jeweiligen Programmversion – manuell oder automatisch – erfolgt im Makro F8.kex. Bei Registern mit der Reihenfolge **pagina tomus** (tomSequence='B') wird die manuelle Version aufgerufen, sonst die automatische. Bei der Kontrolle eines bereits formatierten Registerbuches ist die Reihenfolge immer **tomus pagina** und daher wird dort immer die automatische Version aufgerufen. Durch einen manuellen Eingriff im Makro F8.kex kann man das bei Bedarf ändern.

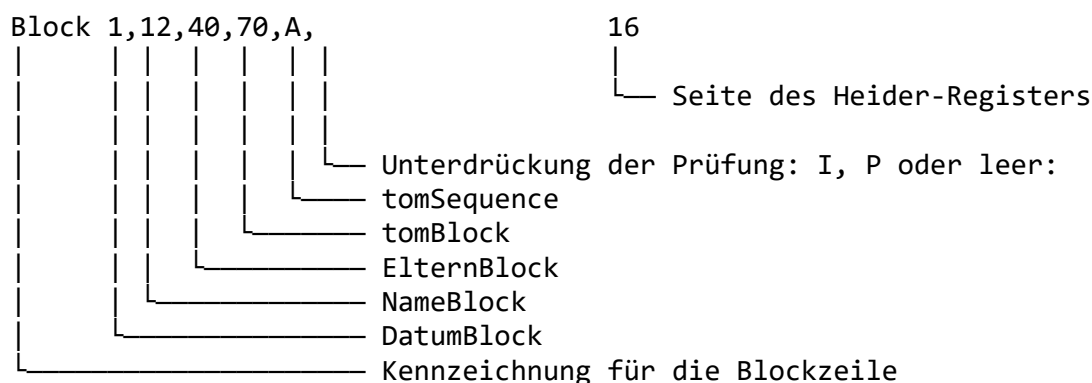
4.13 falsche Zeilen

Manchmal sind Zeilen im Heider-Buch mit dem Wort falsch gekennzeichnet, weil schon beim Schreiben der Eintragung ein Fehler erkannt wurde. Solche Zeilen sind zur Gänze zu löschen. Im Kommentarfeld der richtigen Zeile ist ein Hinweis einzutragen, z.B. vorhergehende Zeile falsch, gelöscht.

4.14 Manuelle Korrekturen in den Heider-Registern durch Benutzer

Häufig findet man auf den maschinengeschriebenen Seiten der Heider-Bücher händische Eintragungen. Das sind Korrekturen von Benutzern, die Fehler gefunden haben. Wir übernehmen diese Korrekturen nur dann, wenn wir ihre Richtigkeit durch Vergleich mit der Originaleintragung in matricula geprüft haben.

4.15 Blockzeile in der Text-Datei



Nach dem Speichern einer mit dem Abbyy FineReader umgewandelten Registerseite hat diese noch das Format, wie sie Josef Heider mit der Schreibmaschine erstellt hat inkl. der Korrekturen, die nach der OCR-Umwandlung vorgenommen wurden. Bei der ersten erfolgreichen Formatierung und Datenprüfung mit den KEDIT-Macros wird in der ersten Zeile ein Informationsblock erzeugt mit dem Schlüsselwort **Block** am Zeilenbeginn. Die Informationen in dieser Zeile werden beim wiederholten Formatieren mit KEDIT benützt und auch bei der weiteren Bearbeitung durch das Programm von Dr. Peter Haas verwendet. Die Zahlen definieren die Startposition der verschiedenen Blöcke, der Buchstabe am Ende definiert die Anordnung von *tomus pagina*. Bei fertigen Seiten ist dieser immer A, weil durch die Bearbeitung mit KEDIT die Normalform *tomus pagina* hergestellt wird. Dahinter kann noch der Buchstabe für die Fehlerunterdrückung folgen.

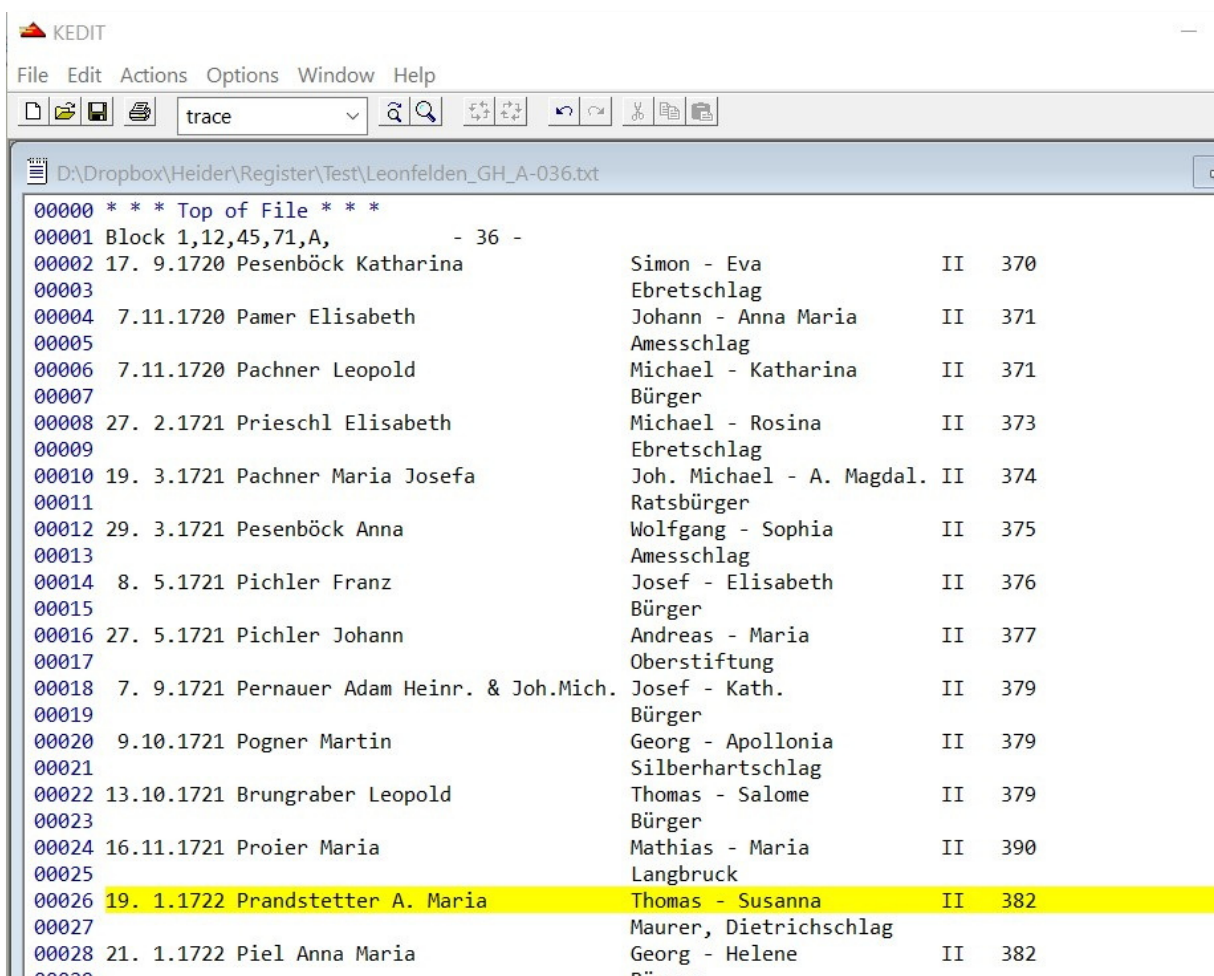


Abb. 17: formatierte Registerseite mit eingefügter Blockzeile

Dem aufmerksamen Leser fällt vielleicht auf, dass die Startpositionen von EBO und von **Tom.** nicht den Minimalpositionen von 40 und 70 entsprechen, wie sie im WINPROF.KEX definiert sind. Ich gestehe, dass ich bei diesem Beispiel gemogelt habe, indem ich beim 2. Zwilling Joh.Mich. am 7.8.1721 **kein** Leerzeichen eingefügt habe. Außerdem ist die Zeile **00026** gelb hervorgehoben. Die zugehörige Fehlermeldung habe ich

schon entfernt (man hätte sonst die Block-Angaben nicht gesehen). Sie hat darauf hingewiesen, dass die pagina absteigend ist. Der Grund ist der Tippfehler in der vorhergehenden Zeile, wo die pagina 390 mit einer 8 über-tippt wurde, FineReader die Zahl aber trotzdem mit 390 übersetzt hat. Dieser Fall war schnell erklärt, bei den Zwillingen muss ich etwas weiter ausholen.

4.16 Zwillinge und mehrfache Vornamen

Es gilt zwar die Regel, dass Begriffe in der 3. Spalte zum EBO-Block gehören – mit einer Ausnahme: Vorna-men, die auf Grund ihrer Anzahl und/oder Länge in den 3. Block ragen, gehören zum Namen-Block des Kin-des. Das passiert oft bei Zwillingen, die Josef Heider mit einem & verbindet (manchmal fälschlicher Weise mit +, mit dem aber normalerweise ein Sterbedatum markiert ist). Alle Vornamen, die mit einem & verbunden sind, platziert KEDIT in der Namensspalte, wodurch sie manchmal länger wird. Es verschiebt dann die EBO- Spalte so weit nach rechts, dass die Namen Platz haben. Bei diesem Prozess darf aber der letzte Vorname nach einem & kein doppelter sein, weil dann nicht klar ist, ob ein Teil des letzten Vornamens nicht schon zum Vater gehört. Aus diesem Grund habe ich bei Joh.Mich. kein Leerzeichen eingefügt und damit KEDIT vorgeschwindelt, es handle sich nur um **einen** Namen. Man darf dann aber nach der Formatierung der Seite nicht vergessen, die Leerstelle nachträglich wieder einzufügen, sobald die Vornamen in der Namensspalte sind. Dadurch wird dann die Namensspalte noch einmal um eine Stelle breiter, aber auch das erfolgt automatisch bei neuerlichem For-matieren.

Die Verlängerung der Namensspalte muss einheitlich für alle Eintragungen einer Seite sein. Das Programm benötigt daher 2 Durchläufe, um die längste Namensspalte (und auch die längste EBO-Spalte) zu ermitteln. Dabei kann es allerdings zu einer sehr großen, unschönen und unpraktischen Breite der Spalten kommen. Durch eine manuelle Verschiebung des Vornamens bzw. der Vornamen in die nächste Zeile (siehe Zeile 18, es dürfen auch mehrere Zeilen sein) lässt sich das wieder rückgängig machen. In der nächsten Abbildung sieht man, dass das Formatierungsprogramm durch die Verschiebung die EBO- und die tomus-Spalte auf die Mini-mumpositionen 40 und 70 zurückgegangen sind. Die Leerstelle zwischen Joh.Mich. ist dann noch einzufü-gen. Eine ähnliche Situation tritt auf, wenn es sich nicht um Zwillinge (oder Mehrlinge) handelt, sondern um viele und lange Vornamen eines Einzelkindes, wie wir sie manchmal bei Adeligen (oder Möchtegern-Adeli-gen) finden. Hier ist dann immer ein manueller Eingriff notwendig. In jedem Fall immer optisch kontrollieren!

Block	Date	Name	Page
00001	Block 1,12,40,70,A,	- 36 -	
00002	17. 9.1720	Pesenböck Katharina	II 370
00003		Simon - Eva Ebretschlag	
00004	7.11.1720	Pamer Elisabeth	II 371
00005		Johann - Anna Maria Amessschlag	
00006	7.11.1720	Pachner Leopold	II 371
00007		Michael - Katharina Bürger	
00008	27. 2.1721	Prieschl Elisabeth	II 373
00009		Michael - Rosina Ebretschlag	
00010	19. 3.1721	Pachner Maria Josefa	II 374
00011		Joh. Michael - A. Magdal. Ratsbürger	
00012	29. 3.1721	Pesenböck Anna	II 375
00013		Wolfgang - Sophia Amessschlag	
00014	8. 5.1721	Pichler Franz	II 376
00015		Josef - Elisabeth Bürger	
00016	27. 5.1721	Pichler Johann	II 377
00017		Andreas - Maria Oberstiftung	
00018	7. 9.1721	Pernauer Adam Heinr. & Joh.Mich.	II 379
00019		Josef - Kath. Bürger	
00020	9.10.1721	Pogner Martin	II 379
00021		Georg - Apollonia Silberhartschlag	

Abb. 18: Vornamen in die Folgezeile verschoben

Eine andere, sehr einfache Lösung mit überlangen oder vielen Vornamen besteht darin, nach dem letzten Vornamen des Probanden das Trennzeichen | händisch zu setzen. Zwischen Text und Trennzeichen muss in diesem Fall keine Leerstelle sein. Das Programm ordnet dann die Daten links vom Trennzeichen der Namensspalte und jene rechts davon – so vorhanden – der EBO-Spalte zu und entfernt das händisch eingefügte Trennzeichen. In der folgenden Abbildung habe ich der Eintragung vom 17. 2.1785 zwischen A. Maria und Michael das | eingefügt.

2. 9.1784	Pötscher Rosalia	Martin - Elisabeth	V 127
16.10.1784	Preining Anna Maria	Bernhardschlag	V 65
6. 1.1785	Praunschmidt Emerentia	Josef, Soldat,	V 2
		M.: Magdalena Feimayer	
		"nach Sage"	
17. 2.1785	Praunschmidt Joh. & A. Maria Michael - Magdalena	Weissenbach	V 3
22. 2.1785	Pötscher Mathias	Mathias - Anna	V 128
		Bernhardschlag	

Abb. 19: Trennung von Namen durch Einfügen des Trennzeichens |

4.17 Fehlermeldungen, Kontrolle

Wir hatten gerade einen Fall mit einem Fehler in einer Zeile. Da es sich hier um einen simplen Tippfehler handelte, wird der stillschweigend ausgebessert. Auch bei eindeutig verschriebenen Vornamen halte ich das so. Es gibt aber Situationen, bei denen erst durch Nachschau im Matrikenbuch eindeutig geklärt werden kann, ob tatsächlich ein Fehler vorliegt. So kann z.B. ein absteigendes Datum ein Tippfehler, oder der Eintrag von Josef Heider beim Einsortieren verreicht worden oder ein Nachtrag im Matrikenbuch sein. Das Nachschauen in matricula kostet aber viel Zeit und stört den Arbeitsfluss. Dzt. schreibe ich nur einen Kommentar in der Kommentarspalte, denn das Ergebnis meiner Arbeit wird von Kontrolllesern bzw. Kontrollleserinnen (dzt. Marianne Reisinger und Georg Pribas) kontrolliert. Damit entdecken sie Fehler, die mir durch die Lappen gegangen sind und prüfen Zweifelsfälle. Um das Nachschauen in matricula zu beschleunigen, sortiert Marianne einen Block von Eintragungen so um, dass sie nach Matrikenbüchern geordnet sind und sie nicht bei jedem Buchstabenwechsel von vorne mit dem Suchen beginnen muss.

In Zeile **00014** sehen wir einen Kommentar nach der **Tom.-**Spalte. Das Prüfprogramm hat ein absteigendes Datum erkannt und gemeldet. Zunächst würde man vermuten, dass das Jahr **1603** oder größer heißen sollte und der Fehler würde nach Korrektur verschwinden. Ob die Vermutung **1603** zutrifft, lässt sich allerdings nur durch Nachschau in der Originalseite mit Sicherheit sagen. Tatsächlich ist nach Aufruf der Seite über matricula zu sehen, dass im Kirchenbuch eine nachträgliche Taufeintragung für das Jahr **1602** vorgenommen wurde. Bei der Kontrolle diese Seite wurde daher ein Kommentar hinzugefügt, um der nachfolgenden Datenaufteilung die Korrektheit der Eintragung mitzuteilen.

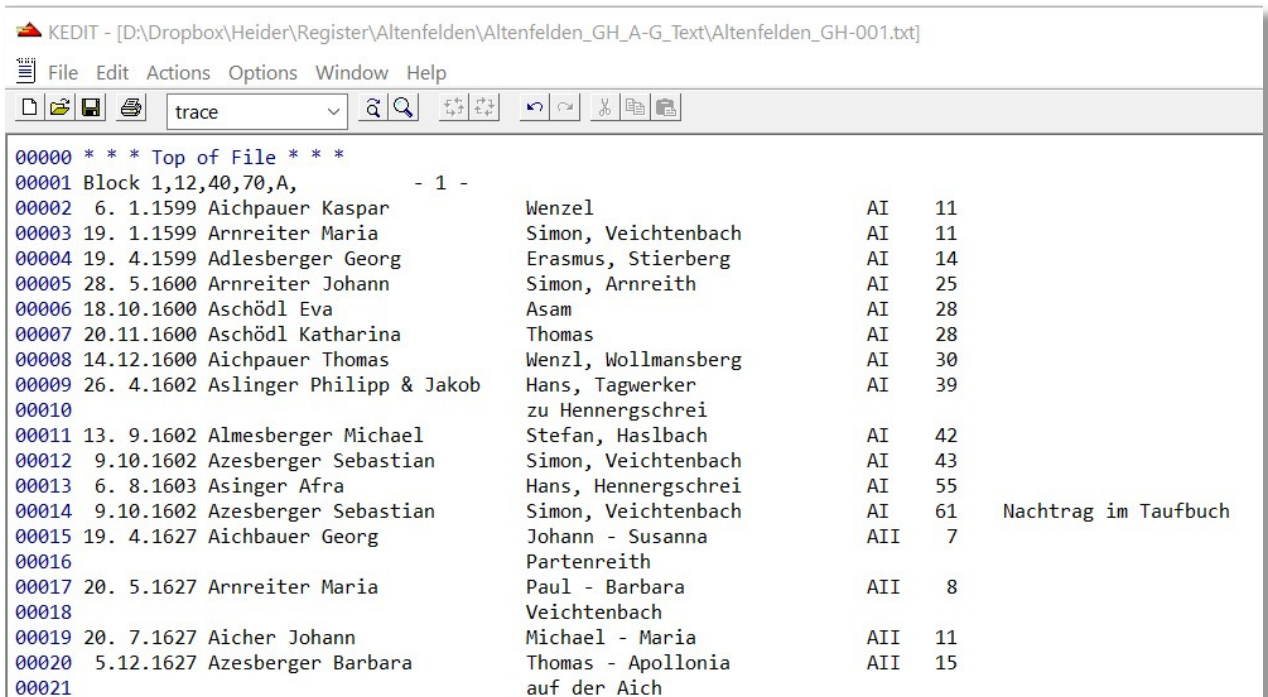


Abb. 20: Meldung im Kommentarfeld

4.18 Nachtrag

Auf der letzten Seite eines Buchstabens steht oft ein Nachtrag, gekennzeichnet durch den Text **Nachtrag**: Die danach folgenden Registereintragungen werden auf gleiche Weise, wie oben beschrieben, geprüft. Datums- und tomus-Prüfungen beginnen also wieder von vorne.

4.19 Rückgängig machen von Änderungen

Änderungen, die manuell oder über ein Programm (Macros) gemacht wurden, können durch Klick auf das linke Pfeilsymbol wieder rückgängig gemacht werden. Z.B. bei fehlendem tomus, den man übersehen hat, kommt die Formatierung etwas durcheinander. Es ist dann oft schneller, die Formatierung rückgängig zu machen, den fehlenden tomus nachzutragen und die Formatierung zu wiederholen.

Mit jedem Klick auf das Symbol werden Änderungen in zeitlicher Reihenfolge rückgängig gemacht. Das ist sehr praktisch, weil man damit Änderungen von mehreren Durchgängen Stufe für Stufe zurücknehmen kann. Ist man eine Stufe zu weit gegangen, kann man das mit Klick auf das rechte Pfeilsymbol wieder korrigieren.

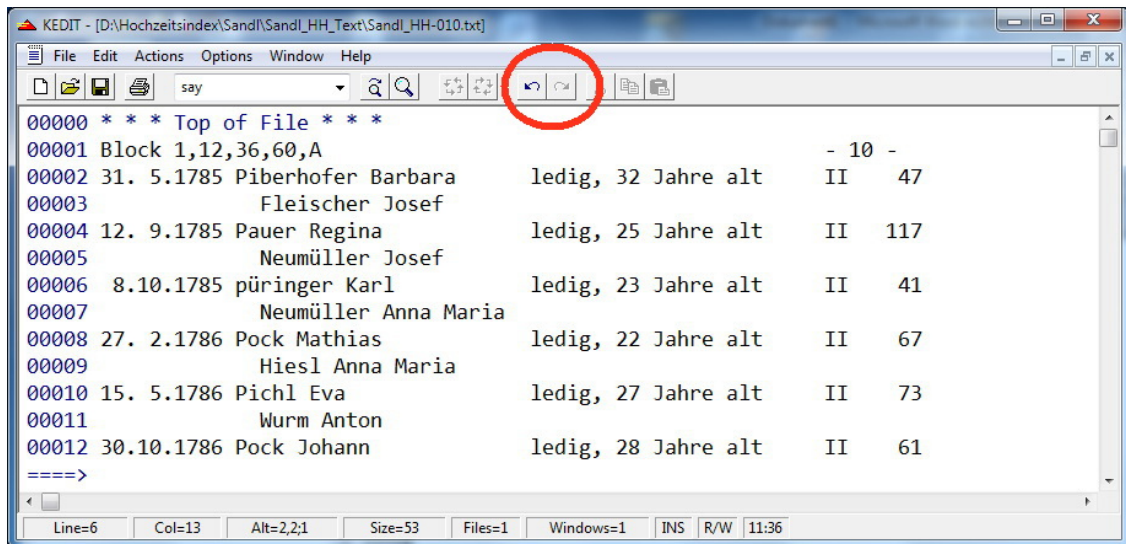


Abb. 21: Rückgängig machen von Änderungen in KEDIT

4.20 Anhang

Am Ende eines Register-Buches gibt es meist eine nicht nummerierte Buchseite mit dem Text **Anhang** und danach ein oder mehrere Seiten mit Eintragungen, die in den Matrikelbüchern fehlerhaft, unleserlich oder unvollständig sind und die Heider nicht richtig zuordnen konnte. Nach der Prüfung sind fast alle Zeilen als fehlerhaft markiert, weil hier kaum bzw. gar keine Familiennamen, sondern nur Vornamen aufgelistet sind und die Anfangsbuchstaben nicht mit dem Indexbuchstaben übereinstimmen. Diese Prüfung und die Fehlermeldungen sind im Grunde unsinnig und daher zu ignorieren. Mit der Eingabe **i** oder **I** (initial) in der Befehlszeile wird die Prüfung des Anfangsbuchstabens unterdrückt. Andere Fehler wie z.B. absteigendes Datum werden aber weiterhin angezeigt. Die Unterdrückung der Fehlermeldungen wird in der Blockzeile der Registerseite vermerkt und gilt damit für alle weiteren Aufrufe zur Formatierung bzw. Prüfung dieser Seite.

5 OCR-Schrifterkennung mit ABBYY FineReader

5.1 Die FineReader-Optionen zum Einrichten und Anpassen

ABBYY FineReader muss an die speziellen Bedürfnisse dieses Projektes angepasst werden. Dies geschieht vor allem mit der Funktion **Extras/Optionen...** Nach Installation des FineReader sind die nachfolgenden Optionen über die Reiter in der obersten Zeile zu wählen.

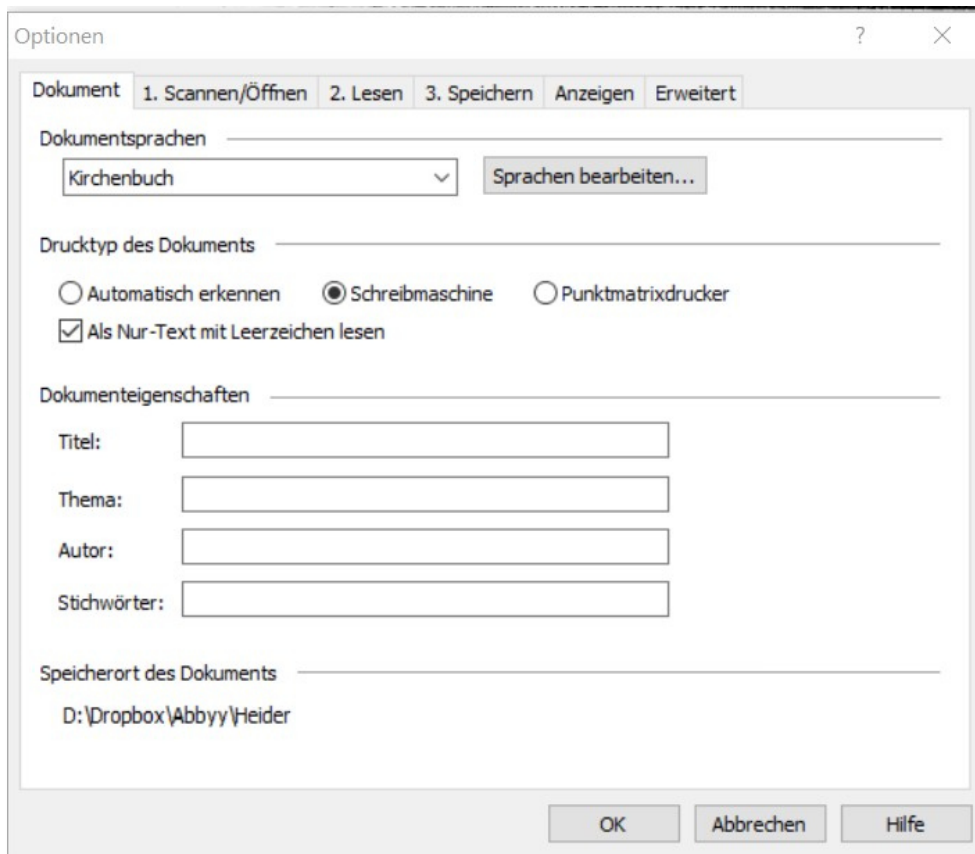


Abb. 22: FineReader-Option Dokument

5.1.1 Dokument

Zum Begriffe **Dokument** siehe Begriffserklärung *Dokument* auf Seite 12.

Zu **Dokumentsprachen** gibt es ein eigenes Kapitel (siehe *Benutzerwörterbuch* auf Seite 41). Beim **Drucktyp** bewirkt das Häkchen bei der Auswahl **Als Nur-Text mit Leerzeichen** eine möglichst positionsgenaue Textdarstellung ohne Tabulatoren. Voraussetzung dafür ist die Verwendung einer monospace Schrift für die Darstellung des übersetzten Textes, siehe *Anzeigen: Schriftart für Text* auf Seite 40.

Unter **Speicherort des Dokumentes** wird das Verzeichnis angegeben, in dem FineReader unter anderem die aktuellen Seiten des Dokumentes mit ihrem Bearbeitungsstand speichert. Der Ordner kann nicht explizit ausgewählt werden, aber unter der Funktion **Datei/FineReader-Dokument speichern...** kann man einen eigenen Ordner angeben. Ab diesem Zeitpunkt scheint dieser Ordner als Speicherort auf und FineReader speichert das Dokument mit allen Seiten in diesen bzw. liest von diesem Ordner.

5.1.2 Scannen/Öffnen

Im Reiter **1. Scannen/Öffnen** kann aus drei Möglichkeiten ausgewählt werden, wann der Prozess der Texterkennung beginnen soll. Hier ist die Variante **Erfasste Seitenbilder nicht automatisch lesen...** zu wählen. Andernfalls beginnt der Prozess bereits unmittelbar nach dem Laden der Scans, also noch bevor der Erkennungsbereich manuell definiert werden kann.

Die Option **Bildneigung korrigieren** sollte nicht ausgewählt werden. FineReader richtet die Bilder waagrecht aus. Da durch Verzerrungen der Seite die Zeilen manchmal schief sind, sind die Zeilen zwar waagrecht, aber der linke Rand wird dabei schief. Entscheidend ist jedoch, dass die Zeilen senkrecht untereinanderstehen. Schiefe Seiten sollten vor dem Öffnen mit einem Fotoprogramm geradegerichtet werden.

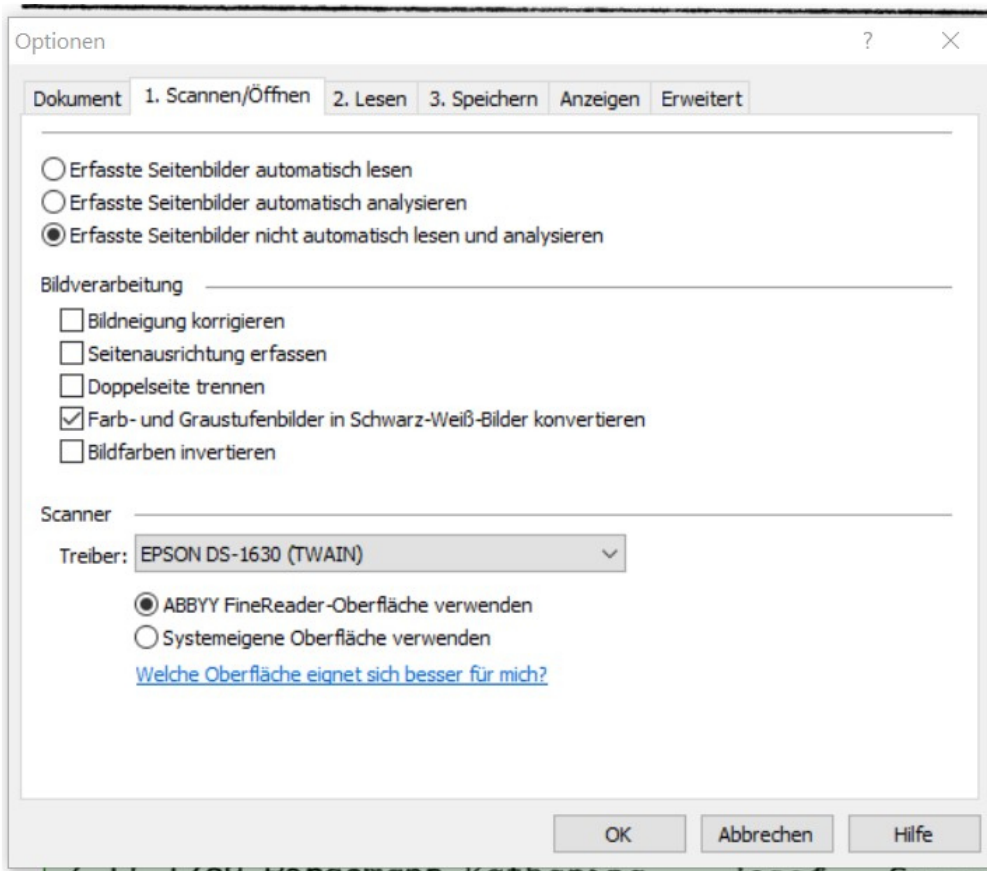


Abb. 23: FineReader-Option Scannen/Öffnen

5.1.3 Lesen: Gründlich, Benutzermuster

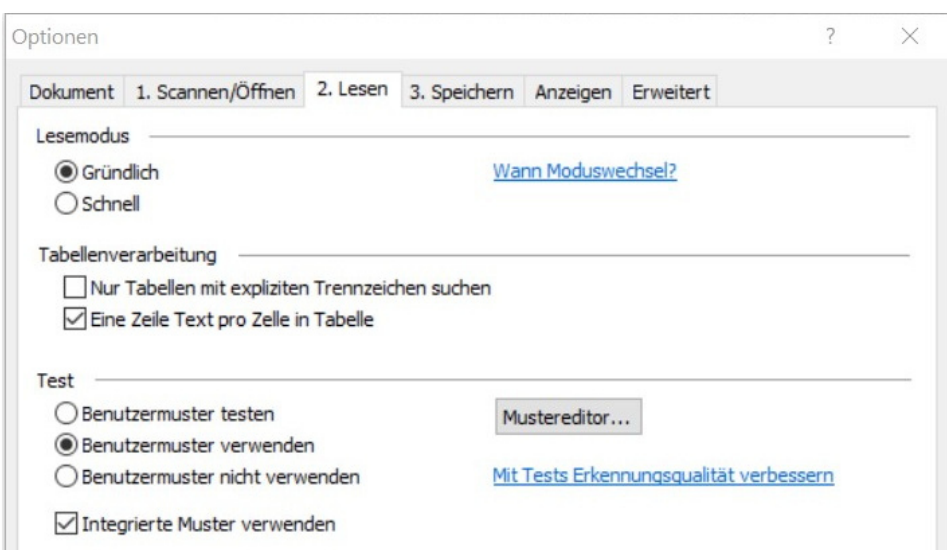


Abb. 24: FineReader-Option Lesen

5.1.4 Speichern: TXT

Der übersetzte Text kann in verschiedenen Formaten abgespeichert werden (RTF, XLS, HTML, PDF, ...) Wir verwenden beim Reiter **3. Speichern** das Format **TXT**.

Kopf- und Fußzeile beibehalten ist auszuwählen, weil FineReader die Seitennummer zu einer Kopfzeile macht, die mit dem Text mitgespeichert werden muss.

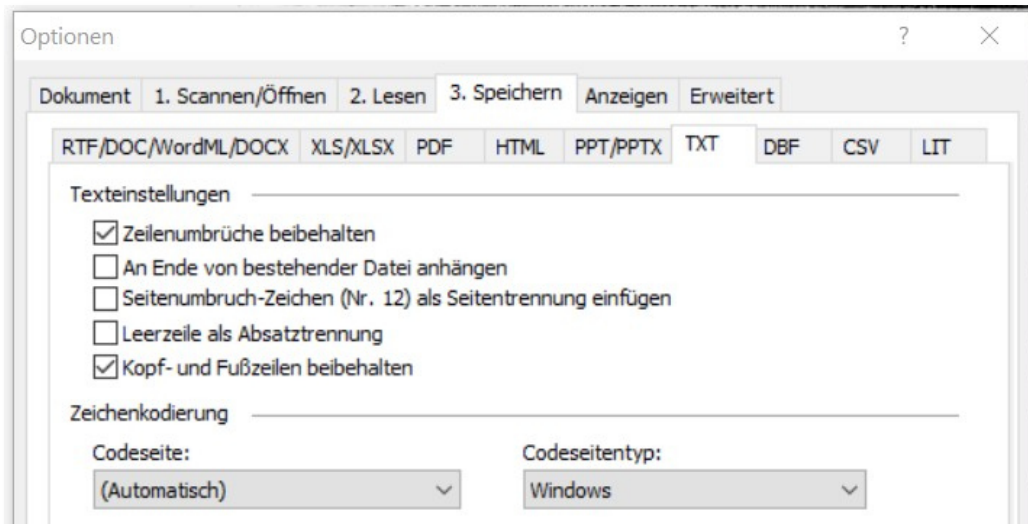


Abb. 25: FineReader-Option Speichern

5.1.5 Anzeigen: Schriftart für Text

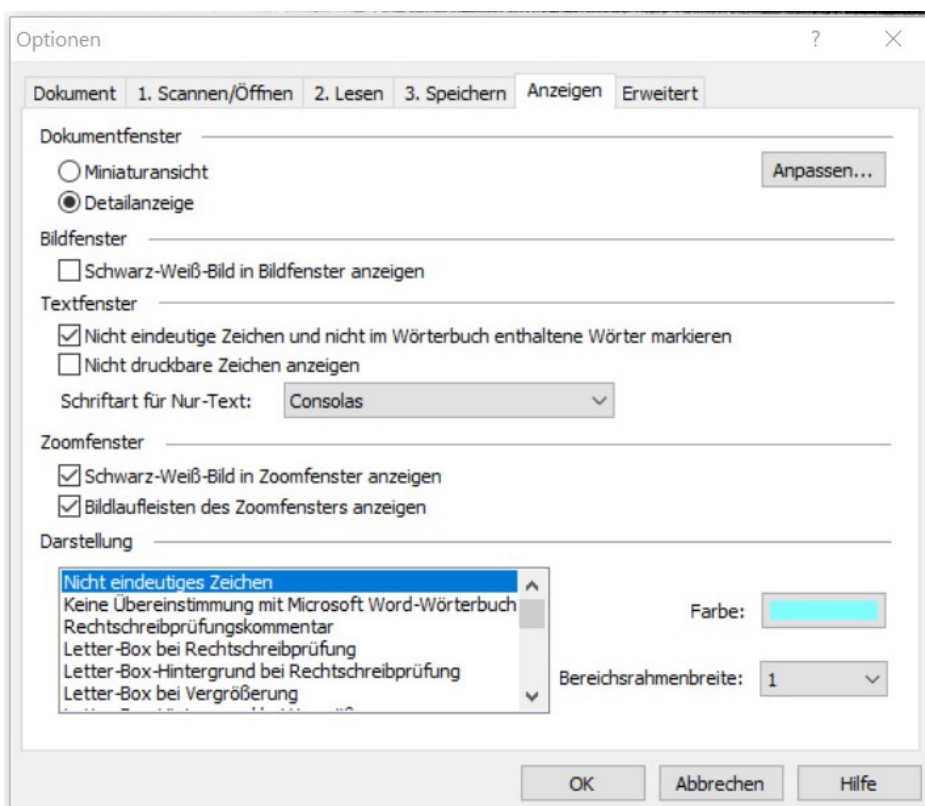


Abb. 26: FineReader-Option Anzeigen

Die von Heider verwendeten Schreibmaschinen sind vom Typ monospace, (jedes Zeichen hat gleiche Buchstabenbreite). Im Reiter **Anzeigen** muss daher für die Textanzeige ebenfalls eine monospace-Schriftart gewählt werden, weil sonst die fixe Spalteneinteilung optisch verloren geht. Ich habe im Reiter **Anzeigen** die Schriftart **Consolas** ausgewählt.

5.1.6 Erweitert

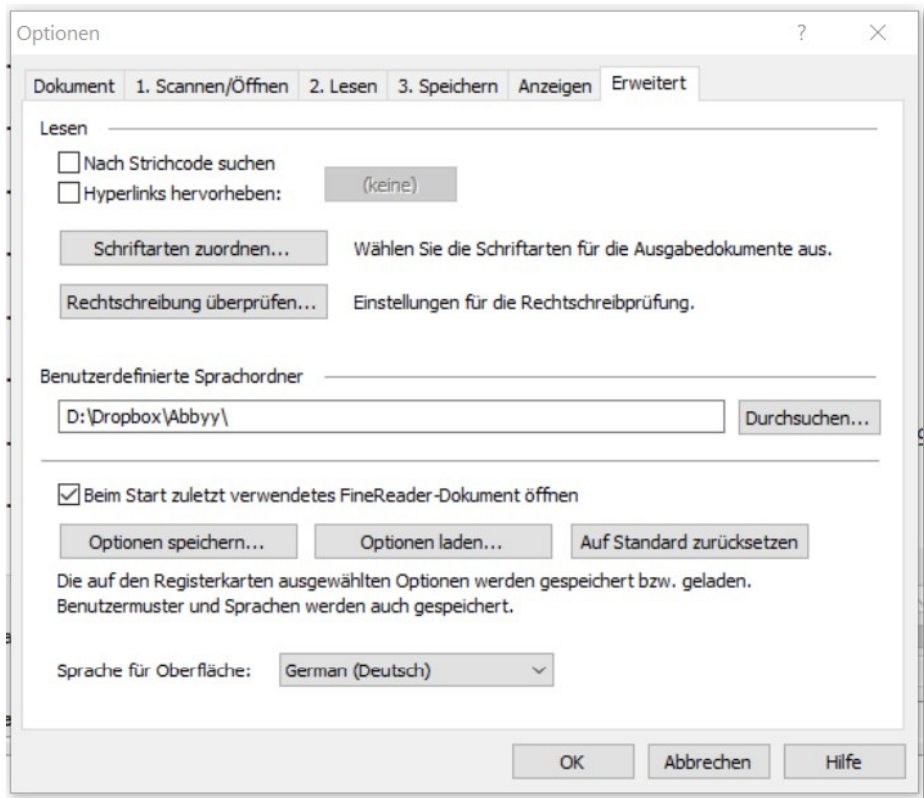


Abb. 27: FineReader-Option Erweitert

Hier gibt es Einstellungen für einige Optionen für Speicherort. Im **Benutzerdefinierte Sprachordner** speichert FineReader das Benutzerwörterbuch (Dateierweiterung .amd). Mit der Export-Funktion des Benutzerwörterbuches exportiere ich es unter **Kirchenbuch.txt** in diesen Ordner und kann es damit extern bearbeiten oder an andere Benutzer weitergeben (siehe [5.2.4 Benutzerwörterbuch exportieren/importieren](#) auf Seite 43). Auch die vordefinierten Bereichsvorlagen speichere ich dort ab (Dateierweiterung .blk). Für das Speichern des Dokuments mit allen seinen Seiten verwende ich ebenfalls diesen Ordner (siehe [5.1.1 Dokument](#) auf Seite 38).

5.1.7 Zwischenspeichern der OCR-Ergebnisse

Im Reiter **Erweitert** gibt es das Kästchen **Beim Start zuletzt geöffnetes FineReader Dokument öffnen**. Ein Häkchen bewirkt, dass nach Unterbrechung der Arbeit beim nächsten Start des FineReader die in Arbeit befindlichen Seiten wieder geladen werden. Man kann also das Bearbeiten der Seiten jederzeit unterbrechen und beim nächsten Start diese weiterbearbeiten. Beim erstmaligem Schließen des FineReaders mit geöffneten Scan-Seiten wird man daher gefragt, ob man die geöffneten Seiten speichern will und unter welchem Namen. Ab diesem Zeitpunkt verwendet FineReader diesen Ordner zum Speichern des Dokuments mit allen seinen Seiten im aktuellen Bearbeitungsstand. Zum Ändern dieses Ordner siehe [5.1.1 Dokument](#) auf Seite 38.

5.2 Benutzerwörterbuch von FineReader

5.2.1 Benutzerwörterbuch als wichtiges Hilfsmittel zur Datenprüfung

FineReader kennzeichnet Wörter, die nicht bekannt sind, mit einer roten Unterstreichung. Dafür gibt es eine Reihe von verschiedenen Sprachen mit integrierten Wörterbüchern. Da es sich bei den Kirchenmatrikeln fast ausschließlich um Eigennamen und alten Berufsbezeichnungen handelt, ist die Verwendung des internen Wörterbuches mit den Flexionen wie Mehrzahl, Fall, Zeit etc. sinnlos bzw. sogar hinderlich. Sinnvoll ist nur ein eigenes Benutzerwörterbuch, in dem die als korrekt eingestuft Namen gespeichert werden.

Die rote Unterstreichung ist aber auch eine gute Möglichkeit, um durch das OCR-Programm Wörter bewusst hervorzuheben. Ich habe deshalb das Wort **ledig** nicht in das Wörterbuch aufgenommen. Die rote Unterstrei-

chung macht mich darauf aufmerksam, dass nach dem Wort *ledig* immer ein Beistrich zur Worttrennung gesetzt sein muss. Auch das *und* ist dafür ein Kandidat. Es verbindet oft 2 Berufe, die ebenfalls durch einen Beistrich zu trennen sind.

5.2.2 Benutzerwörterbuch erstellen

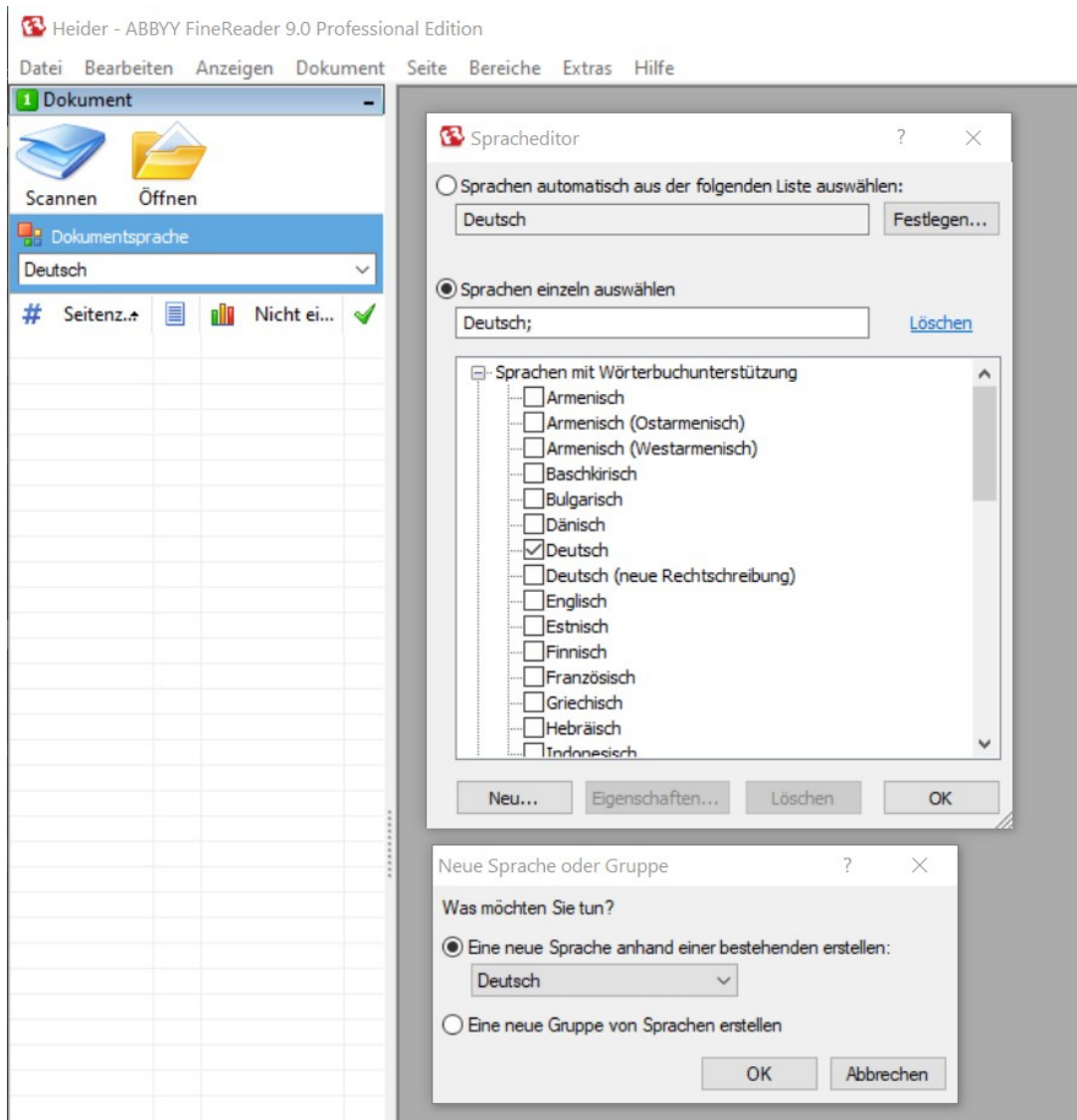


Abb. 28: FineReader-Option Spracheditor

Mit dem Einrichten eines eigenen Wörterbuches hatte ich immer wieder Schwierigkeiten und ich hoffe, dass ich nachfolgend den Ablauf richtig beschreibe (ich habe sehr lange herumexperimentiert, bis ich einen Weg gefunden habe).

Im Dokumentfenster von ABBYY FineReader ist am Anfang als **Dokumentsprache Deutsch** zu sehen. Mit **Extras/Spracheditor...** kann man eine oder mehrere andere Sprachen auswählen, oder mit **Neu...** neue Sprachen anlegen. Im nächsten Fenster ist **Eine neue Sprache anhand einer bestehenden erstellen:** auszuwählen. Ich habe Deutsch gewählt. Damit wird aber nur die alphabetische Sortierfolge festgelegt.

Im nächsten Fenster **Spracheigenschaften** gibt man unter **Sprachenname** einen geeigneten Namen an. Ich habe **Kirchenbuch** gewählt und als Wörterbuch **Benutzerwörterbuch**. Damit hat man ein neues, leeres Wörterbuch ohne Integration erstellt, das man jetzt im Dokumentfenster von ABBYY FineReader unter **Dokumentsprache** auswählen kann. Mit **Extras/Wörterbuch anzeigen...** sollte man das eben erstellte Benutzerwörterbuch auswählen und sicherstellen, dass das Kästchen für **Microsoft Word-Wörterbuch verwenden** **abgewählt** ist.

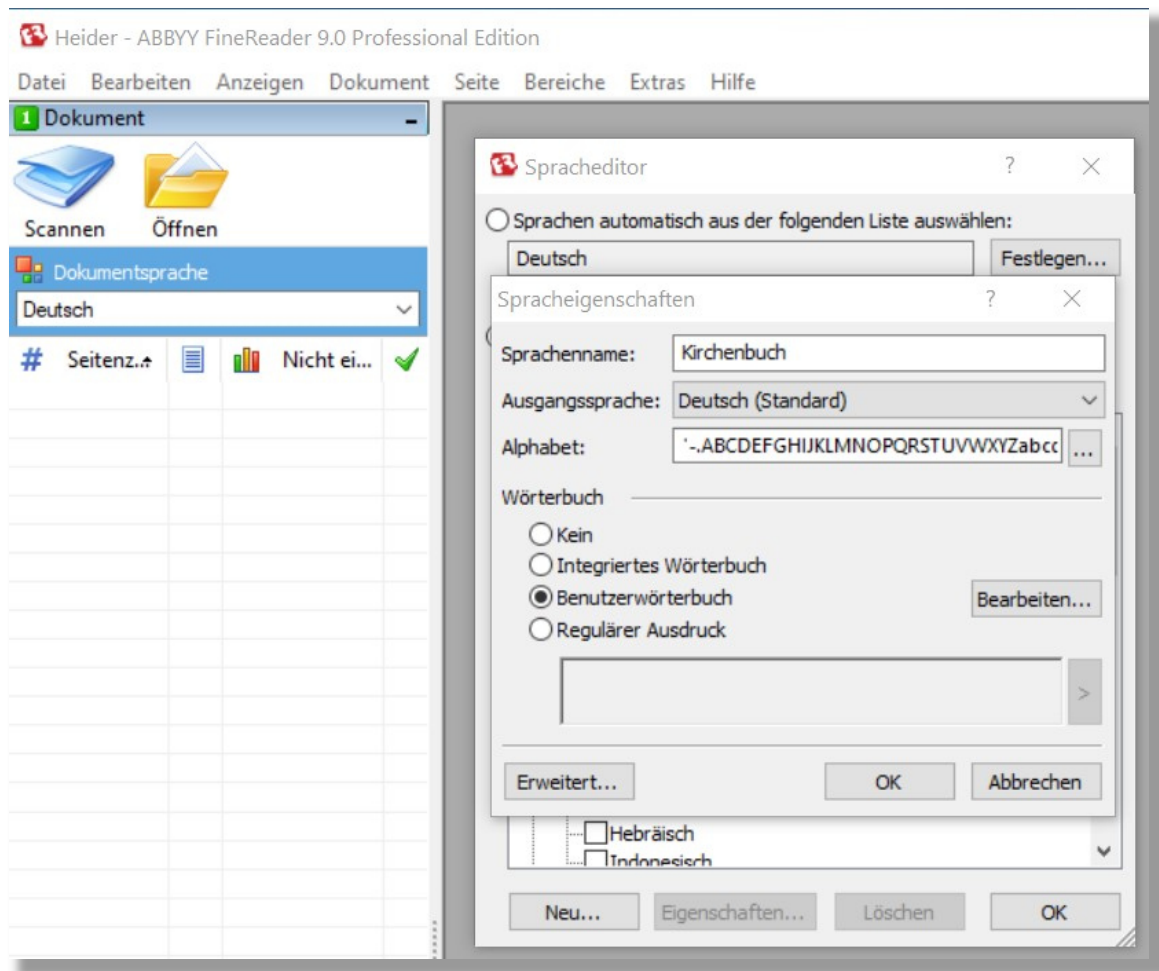


Abb. 29: FineReader: Neues Benutzewörterbuch anlegen

5.2.3 Neue Wörter im Benutzerwörterbuch speichern

Beim Auftreten neuer Begriffe ist zu überlegen, ob sie in das Benutzerwörterbuch übernommen werden sollen. Bei Standardnamen und -berufen ist das eindeutig. So wird man z.B. sehr bald die wichtigsten Vornamen und Berufe erfasst haben. Bei Familien- und Ortsnamen gibt es aber oft sehr viele Varianten und es ist manchmal nicht leicht zu entscheiden, ob eine Variante nur ein Schreibfehler in der Matrikel oder bei Heider ist. Auch sollte man prüfen, ob eine Variante vielleicht nur ein einmaliger Ausreißer ist oder doch öfter vorkommt. Je mehr Varianten im Wörterbuch gespeichert sind, desto größer die Gefahr, dass fehlerhaft geschriebene Namen nicht als Fehler angezeigt werden. Werden umgekehrt durch die OCR-Erkennung zu viele Begriffe durch die rote Unterstreichung markiert, besteht wieder die Gefahr, dass die Markierungen durch ihre Vielzahl nicht mehr ins Auge springen und übersehen werden. Außerdem hat das Wörterbuch Einfluss auf die OCR-Erkennung. Manchmal übersetzt der FineReader ein Wort in eines, das er im Wörterbuch findet, obwohl in der Registerseite etwas Anderes steht.

Ein Mangel des Benutzerwörterbuches ist die Speicherung der Begriffe ohne Großschreibung. Wird z.B. der Name Georg ins Wörterbuch aufgenommen, steht dort nur georg. Dadurch sind für den FineReader georg und Georg gültige Schreibweisen. Meine Macros prüfen daher zumindest den ersten Buchstaben des Namens und in der Namens-Spalte und erkennen diesen Fehler.

5.2.4 Benutzerwörterbuch exportieren/importieren

Namen sind oft sehr ortsspezifisch und kommen dann nur in bestimmten Pfarren vor. Es ist daher sinnvoll, die gespeicherten Begriffe von Zeit zu Zeit zu überprüfen und ausgefallene Begriffe und Schreibweisen wieder aus dem Wörterbuch zu löschen und damit dessen Umfang zu reduzieren. Leider werden die Begriffe im Fenster des Wörterbuches mit einem relativ kleinen Font angezeigt, was mir das Bearbeiten des Wörterbuches schwermacht. Außerdem ist das Aufsuchen eines Begriffes nur durch Verschieben des seitlichen Schiebers möglich und nicht z.B. durch Eingabe der ersten Buchstaben. Man kann sich daher nur sehr mühsam an den

gesuchten Begriff herantasten. Wenn man umfangreichere Änderungen am Wörterbuch vornehmen möchte, ist es daher einfacher, dieses zu exportieren und extern zu bearbeiten. Beim Exportieren wird eine Datei mit der dreistelligen Erweiterung **txt** gespeichert, die man mit KEDIT bearbeiten kann. Das veränderte Wörterbuch kann man mit der Importfunktion wieder zurückspeichern. Zuvor muss man aber das Benutzerwörterbuch im FineReader löschen, sonst tauchen die extern gelöschten Begriffe wieder auf.

In der Abbildung sind die 3 Schritte mit den zugehörigen Fenstern nebeneinander in einem einzigen Bild dargestellt.

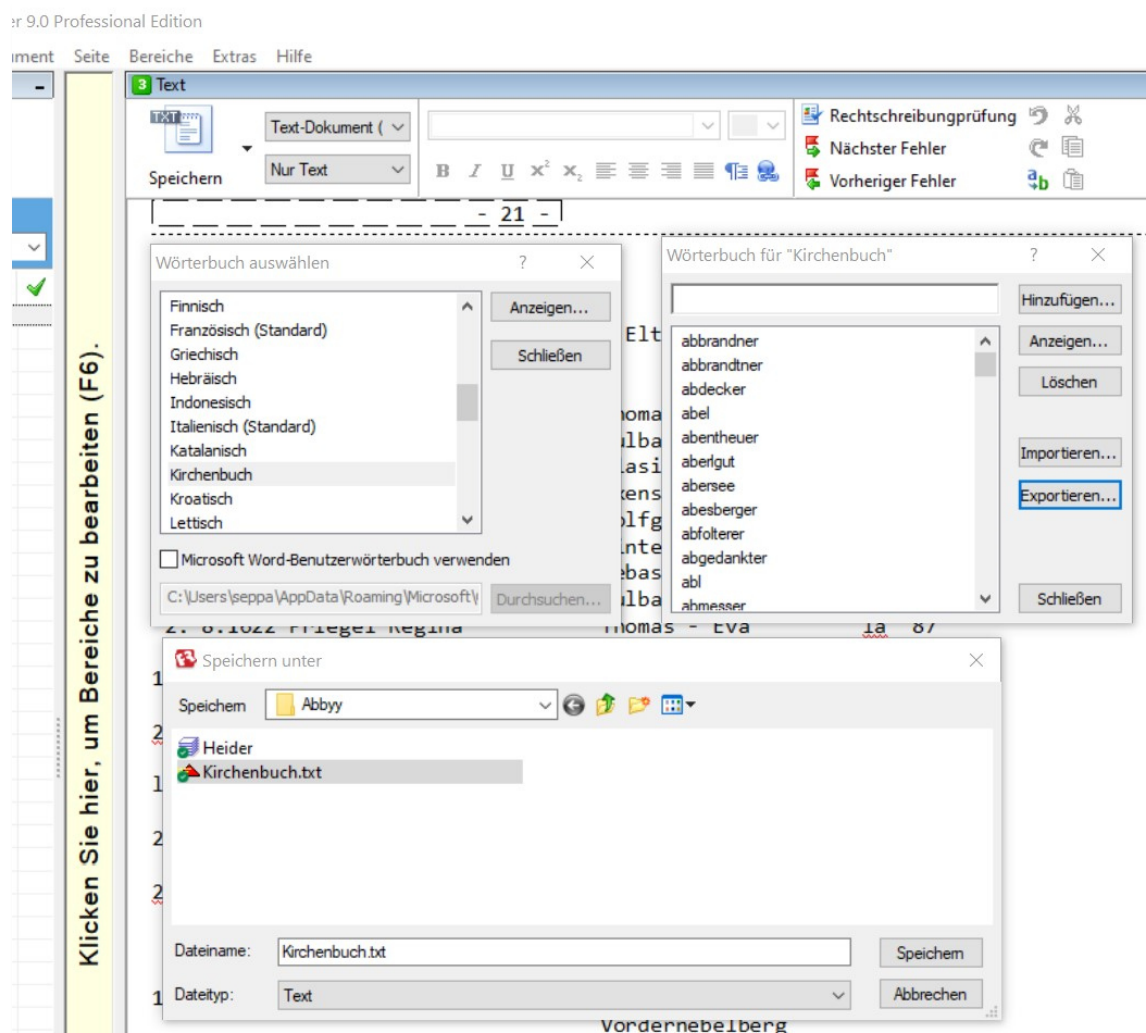


Abb. 30: Exportieren und importieren des Benutzewörterbuches

6 KEDIT Texteditor: Installation und Editierfunktionen

6.1 Installation der KEDIT-Software und der Verarbeitungs-Programme (Macros)

KEDIT ist eine Nachbildung des IBM-XEDIT-Editors im VM-Betriebssystem von IBM Großrechnern. Es ist ein kostenpflichtiges Produkt von der Firma Mansfield Software Group, Inc. Auf der Homepage des Herstellers gibt es umfassende Hinweise zum Produkt: <http://www.kedit.com>. Von dieser Seite kann das Herunterladen von KEDIT gestartet werden. Auch die Dokumentationen (*User's Guide und Reference Manual*) können von dort heruntergeladen werden. Ich habe von der Support Group oft Unterstützung bekommen und kann KEDIT offiziell kostenlos für das Projekt verwenden. Hier noch eine Antwort von der KEDIT Support Group im Zusammenhang meiner Installation auf einen neuen PC:

Sepp: Yes, your existing license can be used on the new PC. Here is a link you can use on the new PC to download the KEDIT install program: <https://www.kedit.com/kedit.downup.installer.html> - you won't need to enter a license key. Kevin Kearney Mansfield Software Group, Inc. -----Original Mess...

KEDIT sucht nach den von Benutzern entwickelten Macros in verschiedenen Verzeichnissen. Das Standard-Verzeichnis dafür ist **C:\Benutzer\user-name\Dokumente**. Zur Weitergabe der Macros habe ich diese in meiner Dropbox gespeichert und sie können von dort mit dem Link https://www.dropbox.com/scl/fo/w6ipxtwign4b9lj8sr819/AMESTIYFAM_Zhnr9TbY6mr8?rlkey=uxp6f3xm95ncwhy1o7d8l3yx8&dl=0 Heruntergeladen und in das oben angegebene Verzeichnis gespeichert werden (bitte den Link in den Browser kopieren).

6.2 Die wichtigsten Editor Zeilenbefehle

Das Formatieren und Prüfen einer Registerseite mit den von mir entwickelten KEDIT-Macros wurde bereits beschrieben und damit kommt man in der Regel auch aus. Es gibt aber noch eine Reihe von Editorbefehlen zum Löschen, Verschieben, Kopieren von Texten und ganzen Zeilen und Textblöcken, auf die ich hier für den Bedarfsfall eingehe.

Durch die Einstellungen im Profile für die Arbeit an unserem Projekt wird das KEDIT-Fensters mit der sog. prefix area dargestellt (Inhalte und Aussehen des Edit-Fensters lassen sich konfigurieren). Die prefix area ist die 5-stellige Nummer am Beginn der Editor-Zeilen, wie sie in den verschiedenen KEDIT-Beispielen dieser Doku zu sehen ist. In einem oder mehreren dieser prefix area kann man Befehle eingeben, die nach Drücken der **PF12-Taste** ausgeführt werden.

6.2.1 Einfache Befehle

- **a** (add) oder **i** (insert): eine Leerzeile wird eingefügt
- **d** (delete): eine Zeile wird gelöscht
- **c** (copy) Zeile wird vor die Zeile mit einem **p** (previous) oder nach der Zeile mit **f** (following) kopiert
- **f** (following) kodierte oder verschobene Zeile wird danach eingefügt
- **m** (move) Zeile wird vor Zeile mit **p** oder nach der Zeile mit **f** verschoben
- **p** (previous) kodierte oder verschobene Zeile wird davor eingefügt
- **l** (lowercase) alle Buchstaben werden auf Kleinbuchstaben gesetzt
- **u** (uppercase) alle Buchstaben werden auf Großbuchstaben gesetzt
- **„** Zeile wird dupliziert
- **<** Zeile wird um eine Stelle nach links verschoben
- **>** Zeile wird um eine Stelle nach rechts verschoben

Das Verschieben mit den **< >** Befehlen betrifft immer die ganze Zeile. Um einen Teil des Textes in einer oder in mehreren Zeilen zu verschieben verwendet man die Block-Funktion. Siehe *Arbeiten mit Textblöcken* auf Seite 46.

6.2.2 Befehle mit Zahlenangaben

Vor oder nach den Befehl **a**, **i**, **d**, **c**, **m**, **l**, **u**, **,**, kann eine Zahl angegeben werden. Damit wird die Anzahl der Zeilen angegeben, für die der Befehl wirksam sein soll. bzw. um wieviel Stellen ein Text verschoben werden

- **a3** füge 3 Zeilen ein
- **d4** lösche die nächsten 4 Zeilen

Beim Verschieben Text definiert die Zahl die Anzahl der zu verschiebenden Stellen:

- **>5** verschiebe den Text der Zeile um 5 Stellen nach rechts

6.2.3 Verdoppelung der Befehle

Als Alternative zu den Zahlenangaben kann man auch Anfang und Ende der zu bearbeitenden Zeilen angeben. Dazu gibt man in der prefix area der ersten und der letzten Zeile den verdoppelten Befehles ein: **cc**, **dd**, **ll**, **mm**, **uu**, **<<**, **>>**

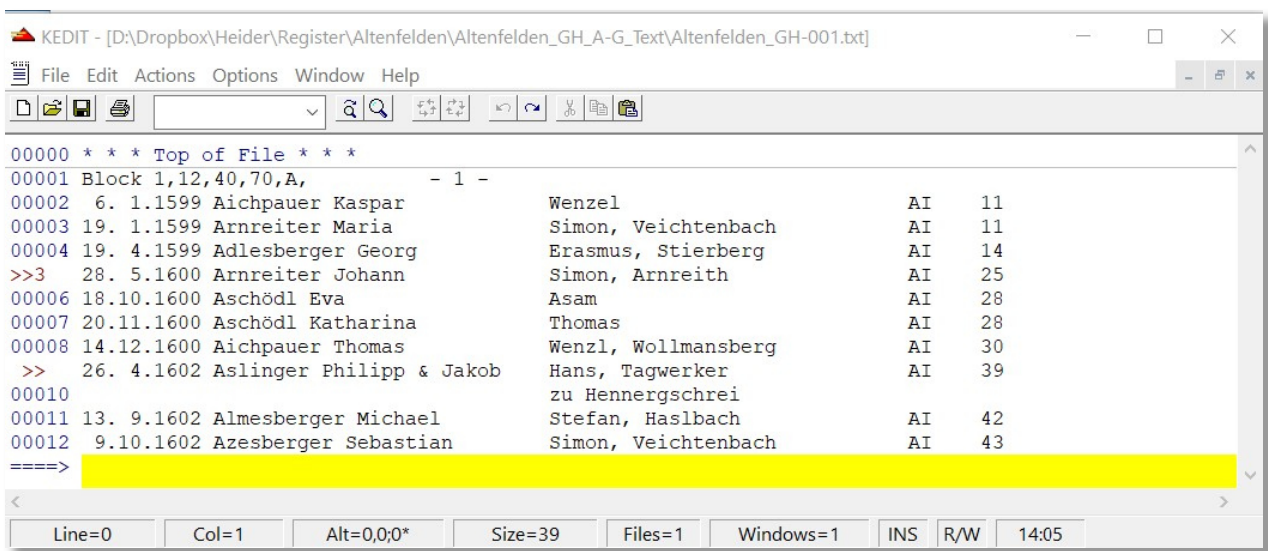


Abb. 31: Auswahl von mehreren Zeilen

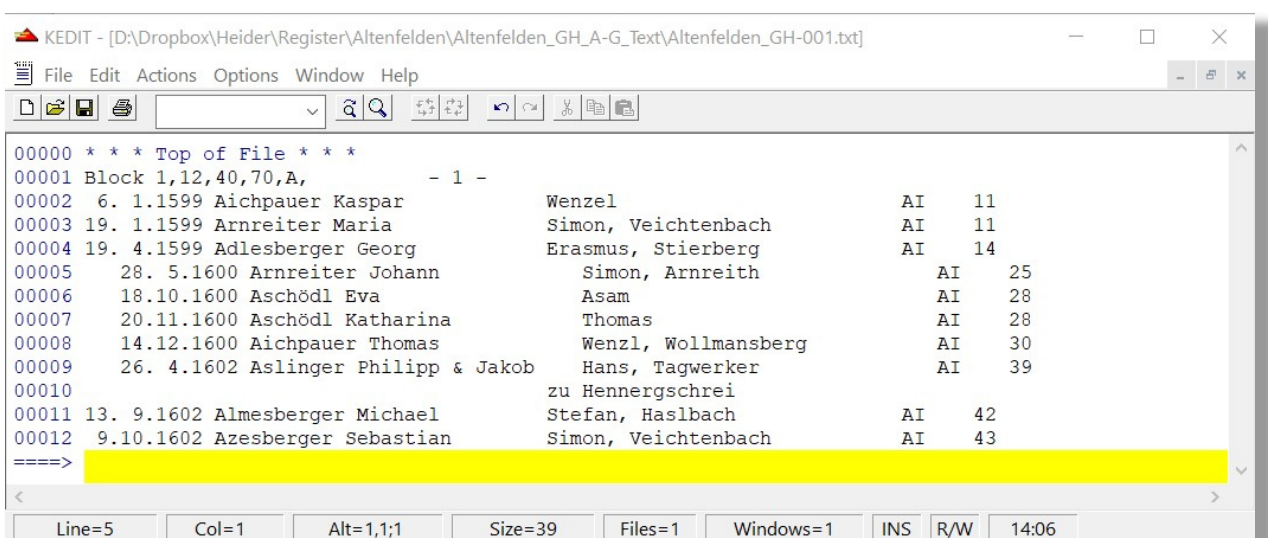


Abb. 32: verschiebe Zeilen 5 bis 9 um 3 Stellen nach rechts

6.3 Arbeiten mit Textblöcken

KEDIT kann auch mit Textblöcken arbeiten. Einen Block wählt man aus, indem man den Cursor auf den ersten Buchstaben des Blockes stellt und dann die Maus mit gedrückter Alt-Taste nach rechts über den auszuwählenden Bereich zieht. Nun gibt es eine Reihe von Manipulationen, die sich auf einen ausgewählten Textblock

beziehen. Für sie gibt es kleine Icons in der bottom toolbar. Die bottom toolbar wird bei Aufruf von Sterbematriken automatisch aktiviert, kann aber auch durch Drücken der **F2**-Taste (toggle bottom toolbar) am unteren Rand des Fensters eingeblendet werden. Sobald ein Textblock ausgewählt wird, werden die Icons aktiviert. Die folgenden zwei Beispiele sollen das Arbeiten mit Textblöcken veranschaulichen.

6.3.1 Textblock in den Kommentar verschieben

In der Registerseite 20 des Taufbuches von St. Johann am Wimberg hat J. Heider einen Originaltext aus der Taufmatrikel übernommen. Durch Auswahl des Textblockes und Ziehen desselben mit der Maus lässt sich der Text als Ganzes an die Position des Kommentares verschieben. Nach Drücken der Eingabetaste werden dann auch wie bei allen anderen Texten mehrfache Leerzeichen entfernt.

00041	26. 5.173o Plank Johann	Johann, Glasmacher	III 151
00042			
00043		aus Frankfurth am Main gebürtig, laut	
00044		seiner Matrimonial-Attestation in recht-	
00045		mässiger Ehe, Vagus. Mater Anna Maria	
00046		von Stift Fulda im frankenland gebürtig.	

Abb. 33: Textblock von Namens- und EBO-Spalte

26. 5.1730 Plank Johann	Johann, Glasmacher	III 151	aus Frankfurth am Main gebürtig, laut seiner Matrimonial-Attestation in recht- mässiger Ehe, Vagus. Mater Anna Maria von Stift Fulda im frankenland gebürtig.
-------------------------	--------------------	---------	--

Abb. 34: in den Kommentarbereich verschieben

6.3.2 Verschieben der Altersangaben in den Sterberegistern

Das Verschieben von Textblöcken ist vor allem bei den Sterberegistern eine sehr nützliche Funktion. Das in der folgenden Abbildung gezeigte Beispiel ist die 2. Seite des Sterberegisters von Aigen im Mühlkreis. Unter *tomus-Block und Altersangabe* beim Sterberegister auf Seite 25 wird erklärt, warum dieser Textblock verschoben wird. In diesem Beispiel wurde der Cursor auf die Position 54 gesetzt, der Textblock mit der Maus ausgewählt und mit der Funktion **Shift Block Right** um 3 Stellen nach rechts verschoben.

```

00000 * * * Top of File * * *
00001   Tag      Name                      Eltern, Beruf, Ort      Tom.
00002
00003
00004 18.10.1674 Aichinger Stefan           Aigen                   II 247
00005
00006 20. 8.1674 Autengruber Michael        Rudolffing              II 247
00007
00008 13. 5.1675 Autengruber Anna            11 248
00009
00010 11. 9.1675 Autengruber Martha          Totengräber, Wächte r  II 248
00011
00012 4. 6.1678 Amprosin Maria               Praitenstein            II 254
00013
00014 6. 6.1682 Almesberger Kind             Florian, Schindlau       II 261
00015
00016 23. 3.1684 Autengruber Andreas          Rudolffing              11 265
00017
00018 2.12.1688 Almesberger Adam             Aigen                   II 274
00019
00020 14. 9.1715 Almesberger Kaspar          Schindlau               6o III 496
00021
00022 22. 6.1716 Almesberger M. Magdalena    Schindlau               5 III 498
00023
00024 9. 6.1718 Almesberger Maria            Wurmbrand               6 III 5o5
00025
00026 6.12.1719 Asanger Johann               Knecht zu Aigen         3o III 5o9
00027
00028 5. 5.1723 Aigner Lukas                  Landgerichtsdieners zu Aigen 6o IV 589
00029
00030 24. 1.1725 Almesmüller Philipp          wurmbrand               7o IV 595
00031
00032 26. 8.1726 Almesberger Magdalena       Hans Georg, Schindl au IV 6o3
00033
00034 9. 7.1787 Amgner Hans Beorg            ledig, Aigen            28 IV 6o9
00035
00036 6. 2.1729 Asanger Maria                Michael, ledig zu Unterneudorf 11w IV 613
00037
00038 7. 4.1729 Almesberger Gertrude         Hans Georg              7 IV 615
00039
00040 22.12.1729 Asanger Margarete           Holzhäuseln            6o IV 619
00041
00042 10. 3.173o Aigner Eva Maria             Landgerichtsdieners zu Aigen 71 IV 621
00043
00044 17. 9.173o Aigner M. Katharina          Kaspar, Langerichtsdieners zu Aigen -7T IV 624
00045
00046 9. 7.1731 Andexlinger Maria            Rudolffing              51 IV 627
00047
00048 29. 9.1731 Aigner M. Magdalena          Kaspar, Landgerichtsdieners zu Aigen s- IV 628
00049
00050 27.12.1731 Aigner Anna Maria           Landgerichtsdieners zu Aigen low 24 IV 629
00051
00052 7. 2.1732 Asanger Thomas               Johann, Zaglau          7w IV 63o
00053
00054 28, 3.1732 Almesberger Rosina          Hans Georg, zu Schindlau 4 IV 63o
00055
00056 * * * End of File * * *
=====

```

Abb. 35: Textblock verschieben



7 Abbildungsverzeichnis

Abb. 1: Seite eines Taufregisters	7
Abb. 2: Seite 2 des Sterberegisters der Pfarre Aigen	8
Abb. 3: Öffnen von 10 Registerseiten.....	11
Abb. 4: Auswahlrahmen erstellen	13
Abb. 5: Tag außerhalb des Erkennungsrahmens.....	14
Abb. 6: lesen und bearbeiten der Registerseiten.....	15
Abb. 7: OCR Textänderungen rückgängig machen	17
Abb. 8: Ausschnitt aus dem KEDIT-Profil	18
Abb. 9: Die Seite 171B wurde eingeschoben, weil Heider die Nr. 171 zweimal vergeben hat	21
Abb. 10: alternative Schreibweise des Familiennamens.....	23
Abb. 11: Aigen Seite 2 im KEDIT-Editor nach der OCR-Umwandlung	26
Abb. 12: Bereich mit Alter und tom-Block nach rechts verschoben.....	27
Abb. 13: EBO-Information und Altersangabe entflechtet	28
Abb. 14: formatierte Seite des Sterberegisters mit Altersangaben.....	29
Abb. 15: die für die automatische Trennung vorbereitete Seite 2 des Sterberegisters Aigen.....	31
Abb. 16: formatierte Seite des Sterberegisters mit ungültigem tomus	32
Abb. 17: formatierte Registerseite mit eingefügter Blockzeile	33
Abb. 18: Vornamen in die Folgezeile verschoben	34
Abb. 19: Trennung von Namen durch Einfügen des Trennzeichens 	35
Abb. 20: Meldung im Kommentarfeld	36
Abb. 21: Rückgängig machen von Änderungen in KEDIT	37
Abb. 22: FineReader-Option Dokument.....	38
Abb. 23: FineReader-Option Scannen/Öffnen	39
Abb. 24: FineReader-Option Lesen	39
Abb. 25: FineReader-Option Speichern	40
Abb. 26: FineReader-Option Anzeigen	40
Abb. 27: FineReader-Option Erweitert	41
Abb. 28: FineReader-Option Spracheditor	42
Abb. 29: FineReader: Neues Benutzewörterbuch anlegen.....	43
Abb. 30: Exportieren und importieren des Benutzewörterbuches.....	44
Abb. 31: Auswahl von mehreren Zeilen	46
Abb. 32: verschiebe Zeilen 5 bis 9 um 3 Stellen nach rechts.....	46
Abb. 33: Textblock von Namens- und EBO-Spalte	47
Abb. 34: in den Kommentarbereich verschieben.....	47
Abb. 35: Textblock verschieben	48