

# Konsys. Das neue Tool für die Konstanzer Bibliothekssystematik

Ralph Hafner

*KonSys*<sup>1</sup> steht für Konstanzer Systematik und bezeichnet das neue Tool zum Browsen und Recherchieren in der systematischen Aufstellung der Bibliothek der Universität Konstanz. Zudem dient es dem Editieren der Konstanzer Systematik, also z. B. dem Hinzufügen neuer Systemstellen oder dem Korrigieren von Systemstellen. Damit ist es ein zentrales Tool für Fachreferent:innen, Teammitarbeiter:innen und Hilfskräfte, die die Neuerwerbungen unserer Bibliothek systematisieren, also nach der Konstanzer Bibliothekssystematik inhaltlich erschließen. Seit Anfang November 2021 ist

*KonSys* im Produktivbetrieb<sup>2</sup>. Es löst unsere bisherige „thematische Suche“ in unserem lokalen Bestand ab, die bis dahin mit dem Tool *SIS*<sup>3</sup> (Schlagwort-Informationssystem) durchgeführt wurde.

## **Never change a running system. Warum wechseln?**

*SIS* lief doch gut, warum also haben wir mit einigem Aufwand<sup>4</sup> ein neues System entwickelt? Zunächst: *SIS* war mit rund 20 Jahren für ein IT-System ziemlich alt. Für die dem *SIS* zugrundeliegende Datenbankversion

gab es seit Jahren keinen Support mehr vom Anbieter.<sup>5</sup> Das Ganze war also seit längerer Zeit etwas wackelig.<sup>6</sup> Aber der Hauptgrund für den Wechsel war das Fernziel<sup>7</sup>, zu einer (Teil-)Automatisierung der Sacherschließung zu kommen. Das wäre mit *SIS* nicht zu erreichen gewesen. Warum nicht? Die Systematikdaten in *SIS* sind nicht interoperabel, d.h. sie können nicht mit anderen Sacherschließungsdaten in Bezug gesetzt werden. Die Schlagwörter in *SIS* sind zwar zu einem großen Teil GND-konform, liegen in *SIS* aber nur als Zeichenketten vor; sie haben keine Referenz-ID zu einer Normdatei wie der GND.<sup>8</sup> Außerdem sind die Schlagwörter in *SIS* nicht eindeutig oder nicht *disambiguiert*, das bedeutet, dass *SIS* nicht zwischen zwei Begriffen mit gleicher Zeichenfolge unterscheiden kann. So sind für *SIS* beispielsweise das Werk *Antigone* von Sophocles (496-406 v.Chr.) und die *Antigone* des französischen Schriftstellers Jean Anouilh (1910-1987) dasselbe, obwohl es sich eindeutig um verschiedene Werke handelt.

Auch gibt es in *SIS* keine direkt auswertbare hierarchische Struktur, obwohl die Konstanzer Systematik als monohierarchische Bibliothekssystematik angelegt ist. Das bräuchte es aber, um unsere Systematik mit anderen Klassifikationen in Beziehung setzen zu können. Diese beiden Gründe, die nicht eindeutige Referenzierbarkeit der Schlagwörter und die nicht auswertbare Hierarchie machen eine maschinelle Vergleichbarkeit mit anderen Inhaltserschließungssystemen wie der RVK<sup>9</sup>, der DDC<sup>10</sup> oder auch der verbalen Sacherschließung mit dem Normvokabular der GND unmöglich.

### Ziele von *KonSys*

*KonSys* ist für unsere Sacherschließung der erste Schritt auf dem Weg raus aus der Konstanzer Insellösung. Seit über 50 Jahren erschließen wir in der Bibliothek / im KIM unsere neuen Medien nach unserer hauseigenen Klassifikation, ohne dabei von der Sacherschließungsarbeit anderer Bibliotheken profitieren zu können.<sup>11</sup> Das soll sich dank *KonSys* nun ändern. In *KonSys* sind knapp 80 % der Schlagwörter mit IDs der GND verknüpft.<sup>12</sup> Das wird es uns ermöglichen, sowohl Konstanzer Notationen abzuleiten für Titel, die mithilfe der GND verbal erschlossen wurden als auch für Titel, die mit der RVK erschlossen wurden, da die Begriffe der RVK ebenfalls auf die GND gemappt sind.

Die in *KonSys* bereinigte hierarchische Struktur unserer Konstanzer Systematik ermöglicht es uns zudem, Notationen der RVK und der DDC – unter Einbeziehung der jeweiligen Systematikbäume – mit unseren Notationen zu vergleichen und daraus Rückschlüsse auf entsprechende Notationen in unserer Systematik zu ziehen. In *KonSys* sind unsere Systematikdaten jetzt überwiegend interoperabel. Das zu erreichen hat ehrlich gesagt deutlich länger gedauert als ich mir vorgestellt hatte: Den *SIS-Visualizer* zur Visualisierung der Hierarchiestruktur unserer Systematik hatten wir 2013 entwickelt und die „Schlagwortwaschmaschine“ *SiGMaMat* in 2015.<sup>13</sup> Seitdem haben die Fachreferent:innen einige Zeit investieren müssen, um die Struktur und die Begriffe der Systematik auf Vordermann zu bringen und die in rund 50 Jahren<sup>14</sup> entstandenen Ungereimtheiten auszubügeln.

In 2022 kann nun damit begonnen werden, die Konstanzer Systematikdaten mit anderen Sacherschließungsdaten in Beziehung zu setzen. Dazu werden wir unsere Systematikdaten einerseits in unseren Bibliotheksverbund *K10plus*<sup>15</sup> einbringen und andererseits in die Webanwendung *Cocoda*, die verschiedene Inhaltserschließungssysteme aufeinander abbildet<sup>16</sup>.

### Funktionen von *KonSys*

*KonSys* sollte nicht nur die Automatisierung der Sacherschließung ermöglichen, sondern auch als modernes Recherche- und Systematisierungstool neue Möglichkeiten für die Arbeit mit unserer Systematik bieten. So soll es z.B. durch die integrierte Funktion von verbaler Suche und Browsen in der Systematik das – nach wie vor auch notwendige intellektuelle – Systematisieren<sup>17</sup> vereinfachen.<sup>18</sup>

### Neue Recherchemöglichkeiten

In *KonSys* gibt es die Möglichkeit, sich auf unterschiedliche Weisen der Systematik zu nähern, um sowohl beim Suchen nach Literatur als auch beim Systematisieren zu den passenden Systemstellen zu kommen.

1. Verbale Suche: Man gibt einen oder mehrere Begriffe in *KonSys* ein und kommt so zu den passenden Systemstellen und darüber zu der passenden Literatur.

Z.B.: Suche nach Literatur zur französischen Liebeslyrik:

**Schlagwörter Suchen**

---

**Ergebnisse**

Fachgebiete	Systemstelle	Schlagwortkette
Allgemeines und Kulturgeschichte (3)	frz 573:171	<b>Französisch</b> / Literaturwissenschaft / Literaturtheorie / Literaturgattung / Geschichte / <b>Lyrik</b> / Einzelne Arten und Formen / Liebeslyrik Verweisungen: <b>Liebe</b> <Motiv> / Kuss <Motiv> GND Verweisungen: Gedicht <Lyrik> / Poem <Lyrik> / ...
Altertumswissenschaft (81)		
Altökzitanisch (81)		
Amerikanistik (779)	frz 698.08:171	<b>Französisch</b> / Literaturwissenschaft / Literatur / Geschichte / Gesamtzeitraum / Anthologie / Literaturgattung / <b>Lyrik</b> / Thema / Motiv / <b>Liebe</b> <Motiv> GND Verweisungen: Gedicht <Lyrik> / Poem <Lyrik> / ...
Anglistik (1221)		
Anthropologie (3)		

2. Man browsst durch die Systematik und findet beim Aufklappen des Systematikbaums seines Faches die passenden Systemstellen. Z.B. zur Geschichte der Fotografie in Frankreich:

**Systemstellen Browsen u. Editieren**

```

mwa - mwz = Medienwissenschaft
├── mwa = Allgemeines
├── mwc = Medien / Geschichte
│   ├── mwc 5 - mwc 99 = Mediengeschichte <Fach> / Allgemeines
│   ├── mwc 100 - mwc 199 = Mediengeschichte <Fach> / Darstellung
│   ├── mwc 200 - mwc 899 = Mediengeschichte <Fach> / Einzelmedi
│   ├── mwc 200 - mwc 299 = Schrift
│   ├── mwc 300 - mwc 399 = Visuelle Medien
│   │   ├── mwc 300 - mwc 309 = Allgemeines
│   │   ├── mwc 360 - mwc 369 = Comic
│   │   ├── mwc 370 - mwc 379.99 = Fotografie
│   │   │   ├── mwc 370 - mwc 370.09 = Gesamtzeitraum
│   │   │   ├── mwc 370.40 - mwc 370.49 = Geschichte 1800-1900
│   │   │   │   ├── mwc 370.40 = Darstellung
│   │   │   │   ├── mwc 370.41 = Afrika
│   │   │   │   ├── mwc 370.42 = Asien
│   │   │   │   ├── mwc 370.43 = Australien
│   │   │   │   ├── mwc 370.44 = Europa
│   │   │   │   │   ├── mwc 370.44:b25 = Belgien
│   │   │   │   │   ├── mwc 370.44:d29 = Deutschland
│   │   │   │   │   ├── mwc 370.44:f71 = Frankreich
│   │   │   │   │   └── mwc 370.44:g76 = Großbritannien
                
```

Systemstelle an und entdeckt über die umliegenden Stellen weitere interessante Literatur zu seinem Thema.

Der größte Unterschied zum Vorgängersystem beim Suchen ist der, dass wir von einer Suche

mit dem Boole'schen Operator AND<sup>19</sup> und der Anzeige in alphanummerischer Ordnung auf die Suchmaschinenlogik umgestiegen sind, die mit dem Boole'schen Operator OR<sup>20</sup> und einem Ranking arbeitet statt mit einer alphanummerischen Liste. Das unterscheidet sich schon rein optisch deutlich vom Vorgängersystem in der Ergebnisdarstellung und erfordert für diejenigen, die an SIS gewöhnt waren, eine Umstellung. Die Gründe für diesen Umstieg sind:

1. Es handelt sich um den aktuellen Standard. Nahezu alle Suchsysteme arbeiten mit dem Boole'schen Operator OR und einem Ranking. Das entspricht also dem, was unsere Nutzer:innen aus

3. Man kombiniert die beiden Wege. Man beginnt mit einer verbalen Suche und wechselt dann bei einer passenden Systemstelle in den Suchergebnissen mit einem Klick in den Systematikbaum, schaut sich den Kontext der

anderen Systemen gewöhnt sind.

2. Es gibt weniger Suchanfragen, die null Treffer liefern. Wenn Nutzer:innen mit mehreren Begriffen suchen und es kommen nicht alle

Begriffe in Kombination in der Systematik vor, gibt es trotzdem Treffer. Kommen am Anfang der Trefferliste keine Suchergebnisse, die alle Suchbegriffe enthalten, weiß man, dass es diese Kombination von Suchbegriffen nicht gibt in der Systematik.

3. Durch das Ranking kommen die besten Treffer nach oben.<sup>21</sup> In *S/S* musste man ggfs. länger durch die Ergebnisliste blättern, wenn man nach einem Begriff gesucht hat, der in vielen Fächern vorkam und man selbst in einem Fach gesucht hat, das im Alphabet weiter hinten steht wie beispielsweise die Philosophie.<sup>22</sup>

### Unterstützung bei der Suche

*KonSys* bietet folgende neue Features, die die Nutzer:innen bei der Suche unterstützen.

1. Eine Suggest-Funktion, die beim Eingeben des Suchbegriffs mithilfe des Indexes Vorschläge macht, die weiterhelfen können, wenn man nicht genau weiß wie ein Begriff heißt oder wenn man nicht genau weiß wie eine Person geschrieben wird.<sup>23</sup> Diese Vorschläge können angenommen oder ignoriert werden.

2. Eine Fuzzy Search = unscharfe Suche, so dass es auch Treffer gibt, wenn man einen Tippfehler in seinen Suchbegriffen hat.

3. Integration der GND-Verweisungen. In *KonSys* sind zu den in unserer Systematik verwendeten GND-Schlagwörter die Synonyme und abweichenden Schreibweisen in die Suche eingebaut, was ebenfalls die Trefferwahrscheinlichkeit erhöht. Man muss nicht genau den Begriff kennen, der an der Systemstelle steht,

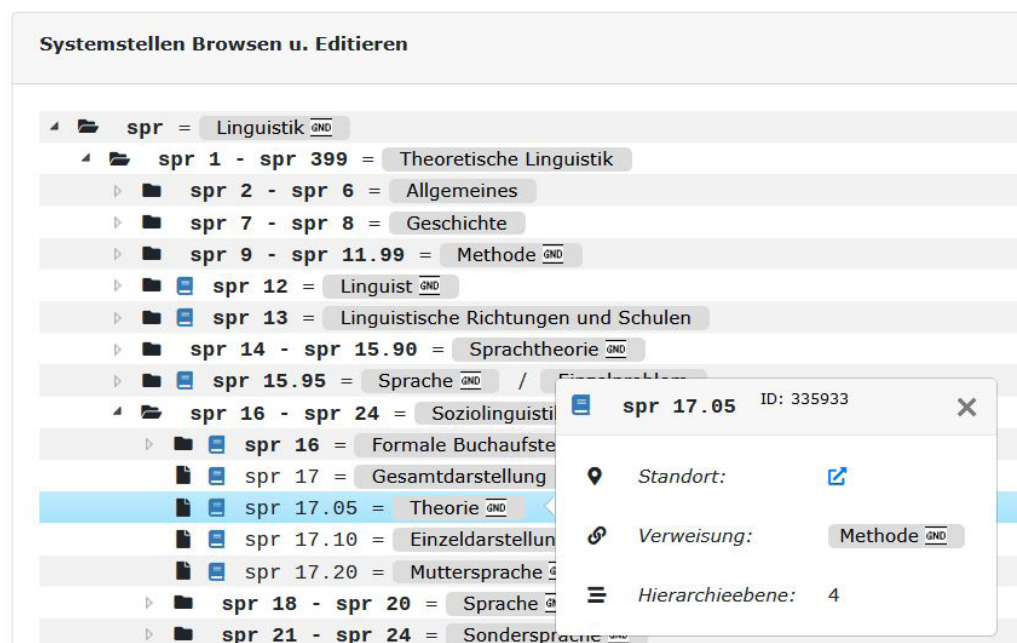
sondern wird auch mit einem synonymen Begriff oder einer abweichenden Schreibung fündig.

### Konsistenz der Systematikdaten

Unsere Bibliothekssystematik hat in *KonSys* eine viel höhere Konsistenz der Systematikdaten<sup>24</sup> als im Vorgängersystem und das auf mehreren Ebenen.

Auf der Ebene der hierarchischen Struktur ist jetzt jede Hierarchieebene durch (mindestens) einen Begriff repräsentiert.

Das folgende Beispiel, das das veranschaulichen soll, kommt aus der Linguistik: Die Theorie der Soziolinguistik (Systemstelle spr 17.05) befindet sich auf Hierarchieebene 4, s. Abbildung.



Im Systematikbaum zeigen wir an jedem Knoten immer nur den diesen Knoten spezifizierenden Begriff an – also den letzten Begriff der Schlagwortfolge. In der Ergebnisanzeige der verbalen Suche zeigen wir zu jedem Treffer die komplette Schlagwortfolge – vom Wurzelknoten des Faches bis zum den Treffer bezeichnenden Begriff an (s. folgende Abbildung).

Systemstelle	Schlagwortkette
spr 17.05	Linguistik / Theoretische Linguistik / Soziolinguistik / Theorie

In *S/S* wurde immer nur ein Teil der Begriffe aus der Schlagwortfolge angezeigt. Manchmal fehlten dadurch auch Begriffe für die Suche.

Titel	Systemstelle	Schlagwortkette	Fachgebiet
<a href="#">Titel aufrufen</a>	ame 17.05	Amerikanisches Englisch / Soziolinguistik / Theorie	Amerikanistik
<a href="#">Titel aufrufen</a>	spr 17.05	Soziolinguistik / Theorie	Linguistik



Insgesamt waren die Schlagwortfolgen in *SIS* durch ihre Unvollständigkeit kürzer, was die Ergebnislisten etwas übersichtlicher machte.

Auch auf der Ebene der Schlagwörter haben wir eine höhere Konsistenz erreicht. Dasselbe Schlagwort heißt jetzt immer gleich. Man kann in *KonSys* keine Tippfehler mehr bei der Eingabe von Begriffen machen, weil man den Begriff nicht mehr eintippt, sondern aus dem Index auswählt. Z.B. ist das Schlagwort „Darstellung“ jetzt immer „Darstellung“ und nicht mal „Dastellung“, mal „Darstelllung“ und mal „Darstelung“. Auf der anderen Seite können wir jetzt zwischen unterschiedlichen Entitäten, die die gleiche Zeichenfolge haben, unterscheiden, z.B. zwischen den oben bereits genannten Werken *Antigone* von Sophocles und *Antigone* von Jean Anouilh.

### Editieren

Alle Informationen, die zum Ausbau der Systematik benötigt werden, sollen in Zukunft in *KonSys* abgelegt sein, und nicht – wie es bisher war – ein Teil in einer Kladde, ein weiterer Teil in einer Word-Datei und ein dritter Teil im Systematik-Tool.

*KonSys* bietet auch beim Editieren der Systematik Unterstützung, um die Konsistenz der Systematik zu erhalten.

Durch Validierungen bei der Eingabe neuer Begriffe und neuer Systemstellen in die Systematik werden Brüche und Inkonsistenzen in der Systematik verhindert, die dazu führen, dass Systemstellen bei der Recherche nicht gefunden werden können. Validierungen haben wir für folgende Bereiche eingebaut:

1. Schlagwörter: Sie werden aus dem Index ausgewählt statt einfach eingetippt.

2. Hierarchische Struktur: Beim Anlegen und Verschieben von Systemstellen erfolgt eine Prüfung in Bezug auf die Richtigkeit der Einordnung in die Hierarchie des Strukturbaums.

3. Syntax der Notation: Beim Anlegen und Korrigieren von Systemstellen erfolgt eine Prüfung der Syntax der Notation. Notationen mit falscher Syntax lassen sich nicht speichern.

### Dank und Ausblick

Beschließen möchte ich den Artikel mit meinem Dank an alle, die dazu beigetragen haben, dass wir dieses neue Tool jetzt im produktiven Einsatz haben. Herzlichen Dank an das Projektteam für die vielen guten Ideen und die intensiven Diskussionen darüber, wie wir die vielfältigen Aufgaben lösen können, die die Neuentwicklung eines Systematiktools und die Migration der Systematik mit sich bringen. Mein besonderer Dank geht an Sorin Gheorghiu, der alles umgesetzt hat. Den Kolleg:innen im Fachreferat danke ich sehr für das aufwendige, zeit- und nervenraubende Aufräumen der Systematiken und das geduldige Mappen der Schlagwörter im *SiGMaMat*. Vielen Dank auch an die Teammitarbeiter\*innen und Hilfskräfte für das Aufspüren von Fehlern, die durch den Migrationsprozess in den Daten entstanden sind und für das Bereinigen dieser Fehler. Der Aufwand sowohl des Bereinigens unserer Systematik als auch des Entwickelns eines neuen Programms hat sich aus meiner Sicht gelohnt. Wir sind jetzt anschlussfähig mit unseren Sacherschließungsdaten und werden mittelfristig die Sacherschließungsarbeit anderer Bibliotheken automatisiert nachnutzen können. Mit ersten Schritten in diese Richtung werden wir in diesem Jahr beginnen.

### Fußnoten:

1 <https://konsys.uni-konstanz.de/>

2 Ein erstes Mockup zu *KonSys* wurde in 2014 erstellt, die erste funktionierende Version mit Elasticsearch Indices und mit Systematikdaten, die über Migrationsskripte aus *SIS* eingespielt wurden, stand 2018 bereit.

3 [http://sis.ub.uni-konstanz.de/sis/stammsatz\\_verbal.html](http://sis.ub.uni-konstanz.de/sis/stammsatz_verbal.html). *SIS* war seit Juli 2000 im Einsatz.

4 Das Projektteam bestand aus Bernd Schelling (IT-Leitung des Projekts), Ralph Hafner (inhaltliche Leitung), Sorin Gheorghiu (Hauptentwickler der Tools), Sebastian Danisch (beratender Entwickler) und Uwe Jochum (inhaltliche Beratung).

5 Die verwendete Oracle-Datenbank, die unter Sun OS lief, war damals ein teures Produkt. Damals standen im Gegensatz zu heute keine kostenlosen Datenbankprodukte wie MySQL oder Postgres zur Verfügung.

6 Zur Beruhigung: Die Systematikdaten wurden jede Nacht gesichert.

7 Jetzt hoffentlich nicht mehr so fern :-).

8 GND = Gemeinsame Normdatei der Deutschen Nationalbibliothek. Normiertes Vokabular zur verbalen inhaltlichen Erschließung im deutschsprachigen Raum.

9 RVK = Regensburger Verbundklassifikation, die in Deutschland am weitesten verbreitete Universalklassifikation

10 DDC = Dewey-Dezimalklassifikation, die international am weitesten verbreitete Universalklassifikation.

- 11 Wenn in einer Bibliothek ein Titel bereits z.B. mit einer RVK-Notation erschlossen wurde, kann eine andere Bibliothek, die ebenfalls die RVK zur Erschließung nutzt, die Notation einfach übernehmen und muss keine eigene Sacherschließungsarbeit mehr leisten. Sie spart dadurch also Arbeit und Zeit.
- 12 79,4 % der Schlagwörter in KonSys sind mit der deutschen Normdatei GND verknüpft. Weitere Schlagwörter sind mit der internationalen Schlagwortnormdatei VIAF, der Wikipedia oder mit Fachlexika verknüpft. Dazu haben wir Form- und Zeitschlagwörter. 13 % der Schlagwörter in KonSys haben keine Verknüpfung. Das sind z.B. Personen, Werke oder Sachbegriffe, für die es keine Normdaten gibt.
- 13 Wie mithilfe der Tools SiGMaMat für KonSys die Schlagwörter bereinigt wurden und mithilfe des SIS-Visualizers die hierarchische Struktur unserer Systematik und ihre Fehler sichtbar gemacht wurden, kann man in folgendem Artikel nachlesen: Hafner, Ralph: SiGMaMat, SIS-Visualizer und KonSys. Die Zukunft der Konstanzer Sacherschließung. In: KIM kompakt (2017) H. 101. S. 26–28.
- 14 Gut 50 Jahre, in denen es deutlich weniger technische Unterstützung gab, um Strukturfehler oder Dopplungen zu erkennen und zu vermeiden.
- 15 K10plus: gemeinsame Datenbank von BSZ und VZG mit etwa 200 Millionen Bestandsnachweisen.
- 16 Cocoda ist im Rahmen des VZG-Projekts coli-conc entwickelt worden.
- 17 Natürlich wird das automatisierte Systematisieren nur bei einem Teil der Titel erfolgreich sein, den Rest wird man nach wie vor intellektuell erschließen müssen. Dazu, wie groß dieser Rest ist, werden wir in den nächsten Jahren mehr sagen können. Je mehr Bibliotheken gute inhaltliche Erschließung mit gut strukturierten Daten bieten, desto besser gelingt automatisierte Sacherschließung. Es ist ein Geben und Nehmen.
- 18 Wie gut KonSys im Unterschied zu SIS das Systematisieren unterstützt, wird im Fachreferat unterschiedlich eingeschätzt.
- 19 AND = es werden nur die Systemstellen angezeigt, bei denen alle Suchbegriffe vorkommen.
- 20 OR = es werden alle Treffer angezeigt, bei denen einer der Suchbegriffe vorkommt. Die Treffer, bei denen alle Suchbegriffe vorkommen, werden nach oben gerankt. Das erzeugt natürlich Treffermengen, die deutlich höher sind als die in SIS. Es ist aber wie bei anderen Suchmaschinen auch, relevant sind üblicherweise die Treffer oben auf der ersten Seite, ggfs. noch die auf der zweiten Trefferseite, der Rest kann meistens ignoriert werden.
- 21 Jedenfalls meistens. An der Optimierung und der Transparenz des Rankings arbeiten wir aktuell noch.
- 22 Natürlich konnte man in SIS vorher das Fach auswählen, in dem man den Begriff sucht. Das haben die Expert:innen auch so gemacht.
- 23 Hätten Sie gewusst, wie man folgenden russischen Autor schreibt: Александр Исаевич Солженицын oder Aleksandr Solženicyn oder Alexander Solschenizyn oder Alexander Isaevich Solzhenitsyn, oder...? Egal, wie Sie ihn schreiben, in KonSys finden Sie ihn dank der Suggest-Funktion und dank der GND-Verweisungen. S.u.
- 24 Aktuell gibt es noch Fehler in den Daten, die von der Migration herrühren. Dass es Fehler geben wird, wussten wir, das wäre bei einem System dieser Größe mit ca. 120.000 Schlagwörtern und ca. 360.000 Systemstellen und der komplexen Übertragung und gleichzeitigen Bereinigung nur mit einem noch deutlich höheren Personalaufwand für die Bereinigung der Daten zu vermeiden gewesen. Aus Sicht des Projektteams ist die Fehlerquote sehr gering.