

# FORTSCHRITT DER PROMOVIEREN- DENERFASSUNG AN HOCHSCHULEN IN DEUTSCHLAND

ERGEBNISSE DER BEFRAGUNG 4 IM MAI 2019

## Inhalt

1	Promovierendenerfassung an Hochschulen in Deutschland.....	2
2	Methodisches Vorgehen und Beteiligung .....	2
3	Einsatz von Erfassungslösungen.....	2
	Aktuelle und zukünftige Erfassungslösung.....	3
	Verwendung verschiedener Komponenten einer Erfassungslösung.....	5
4	Datenerfassung und Datenlieferung.....	6
	Nacherfassung.....	6
	Nicht erfasste Teilgruppen.....	11
	Rückgemeldete Fehlerquote des jeweiligen Statistischen Landesamtes und Korrekturschleifen.	11
5	Auftretende Probleme bei der Bereitstellung der Daten für die Promovierendenstatistik .....	14
6	Fazit .....	18

## 1 PROMOVIERENDENERFASSUNG AN HOCHSCHULEN IN DEUTSCHLAND

Zur Verbesserung der Datenlage von Promovierenden sowie des Promotionsgeschehens werden nach Schaffung der entsprechenden Grundlagen durch die Novellierung des Hochschulstatistikgesetzes mit dem ersten Stichtag 1. Dezember 2017 Promovierendendaten durch die Hochschulen erhoben und übermittelt. Die UniWiND-Koordinierungsstelle Nachwuchsinformationen – UniKoN unterstützt alle Hochschulen mit Promotionsrecht in Deutschland, Systeme zur Erfassung ihrer Promovierenden einzurichten und Daten rund um das Promotionsgeschehen zu erheben.

Um den Stand der Hochschulen hinsichtlich der Einführung dieser Systeme abzubilden, führt UniKoN ein Monitoring durch wiederholte Befragungen der Hochschulen durch. Die hier präsentierten Ergebnisse sind Resultate der Befragung 4, die im Mai 2019 stattfand.

## 2 METHODISCHES VORGEHEN UND BETEILIGUNG

Die Daten wurden mittels eines Online-Fragebogens erhoben. Eingeladen wurden alle 155 Hochschulen mit Promotionsrecht, die zum Befragungszeitraum Mitglieder der Hochschulrektorenkonferenz (HRK) waren. Hierfür wurde das Präsidium bzw. Rektorat der Hochschule oder die Leitungen der jeweiligen Graduierteneinrichtungen persönlich per E-Mail angeschrieben. 50 Einrichtungen haben die Befragung beantwortet. Davon haben 28 auch an den Befragungen 1 bis 3 teilgenommen. Im Vergleich zur letzten Befragung lässt sich jedoch ein Teilnehmerrückgang erkennen. Damals hatten 76 Hochschulen an der Befragung teilgenommen.

Eingeladene Institutionen	Anzahl	Teilnehmende Hochschulen
alle Hochschulen mit Promotionsberechtigung	155	50 (32 %)
davon Universitäten	98	41 (42 %)
Mitglieder UniWiND e. V.	71	37 (52 %)

## 3 EINSATZ VON ERFASSUNGSLÖSUNGEN

Zunächst wurden die teilnehmenden Einrichtungen über ihre Erfassungslösung für die Promovierendendaten an ihrer Hochschule befragt. Im weiteren Verlauf der Befragung wurden zudem Informationen über einen möglicherweise geplanten Wechsel von der bisher genutzten Erfassungslösung hin zu einer neuen oder erweiterten Lösung erfragt. Wenn diese angestrebt wird, baten wir um die Nennung der zukünftig geplanten Erfassungslösung.

Die Systeme werden nachstehend anhand von Kategorien strukturiert, die im Rahmen von UniKoN-Workstattgesprächen entwickelt und diskutiert worden sind. Hierbei wird zwischen den folgenden Klassifizierungen unterschieden:

## **Campus Management Systeme oder andere Software zur Hochschulverwaltung (CMS):**

- **FlexNow**, fecher GmbH
- **CampusNet**, Datenlotsen Informationssystem GmbH
- **SAP** (nicht näher bezeichnet)
- **HISinOne**, HIS Hochschul-Informationssystem eG
- **HIS GX**, HIS Hochschul-Informationssystem eG
- **Weitere** (offene Frage)

## **Forschungsinformationssysteme (FIS):**

- **Converis**, Clarivate Analytics
- **Weitere** (offene Frage)

## **Spezialanwendung zur Promovierendenerfassung:**

- **Docata**, Divinus Soft GmbH
- **Weitere** (offene Frage)

## **Andere Erfassungslösung:**

- **nicht näher bezeichnet** (offene Frage)

## AKTUELLE UND ZUKÜNFTIGE ERFASSUNGSLÖSUNG

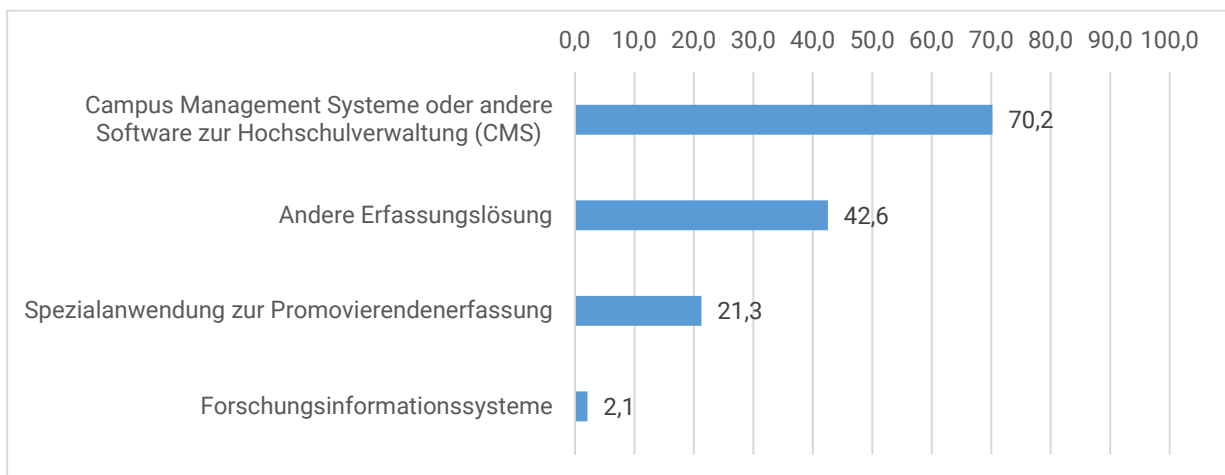
Die erste Frage lautete: **Der Stichtag für die zweite Datenerhebung für die Promovierendenstatistik war der 1. Dezember 2018. Welche Erfassungslösung haben Sie für diese zweite Erhebung der Promovierendenstatistik an Ihrer Hochschule verwendet?** Bei der Beantwortung der Frage waren Mehrfachantworten möglich. Mit dieser Frage konnten Informationen über die aktuell verwendete Erfassungslösung sowie die Kombinationen verschiedener Komponenten der verwendeten Erfassungslösung in Erfahrung gebracht werden.

Die Art der aktuellen Erfassungslösung an den einzelnen Einrichtungen unterscheidet sich stark. Die Variationsbreite betrifft sowohl das verwendete System als auch die Anzahl der genutzten Systemkomponenten von zum Teil verschiedenen Anbietern. Dieser Aspekt wird weiter unten näher beleuchtet.

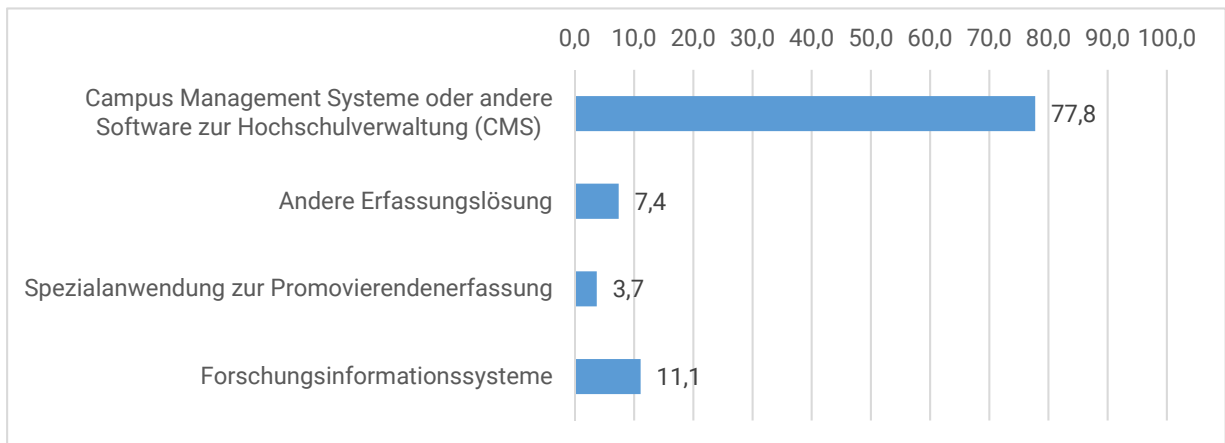
Die meisten Einrichtungen, 70 Prozent (33 Einrichtungen), verwenden derzeit ein Campus Management System oder andere Software zur Hochschulverwaltung (CMS), um die Datenerfassung zu realisieren. Zehn Einrichtungen (21,3 %, n=47) verwenden Spezialanwendungen zur Erfassung und zwei Hochschu-

len (2,1 %, n=47) nutzen ein Forschungsinformationssystem (FIS). 20 Einrichtungen (42,6 %, n=47) geben an, eine andere, nicht näher bezeichnete, Erfassungslösung in Anwendung zu haben. Die Ergebnisse können der Abbildung 1 entnommen werden.

**Abbildung 1** Aktuelle Komponenten der verwendeten Erfassungslösung (Angaben in %, Mehrfachantworten möglich, n=47)



**Abbildung 2** Zukünftige Komponenten der geplanten Erfassungslösung (Angaben in %, Mehrfachantworten möglich, n=27)



Auf die Frage **Welche andere Erfassungslösung ist zukünftig geplant?** gab es folgende Antworten:

55 Prozent (26 Einrichtungen, n=47) der befragten Einrichtungen gaben an, in Zukunft einen Wechsel der Erfassungslösung anzustreben. In Abbildung 2 sind die Erfassungslösungen abgebildet, die die Hochschulen in Zukunft beabsichtigen zu verwenden. Ebenso wie bei der aktuell verwendeten Erfassungslösung kommt auch hier die Kombination mehrerer Komponenten als Erfassungslösung vor. Dabei zeichnet sich ab, dass vor allem die Nutzung „andere Erfassungslösungen“ zugunsten des Einsatzes von CMS abnehmen wird. 21 (77,8 %, n=27) Einrichtungen planen zukünftig ein CMS zu verwenden. Drei

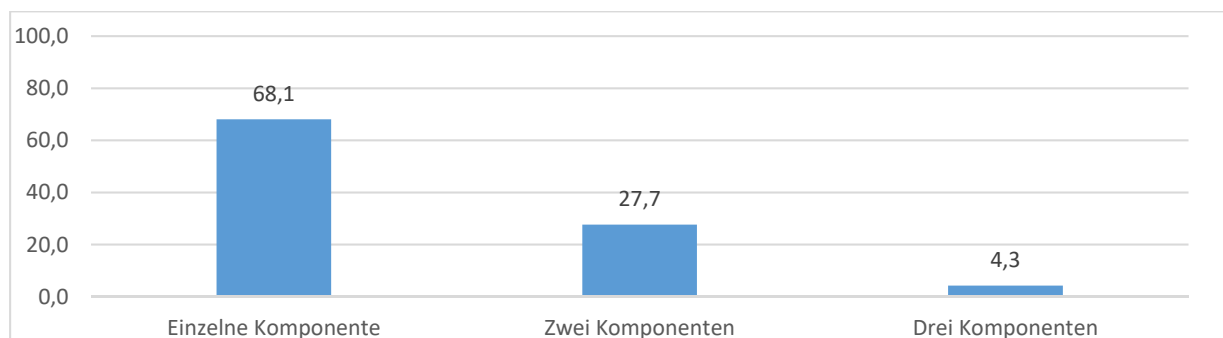
(11,1 %, n=27) nannten ein Forschungsinformationssystem und zwei (7,4 %, n=27) eine andere Erfassungslösung. Eine Einrichtung (3,7 %, n=27) nannte eine Spezialanwendung.

## VERWENDUNG VERSCHIEDENER KOMPONENTEN EINER ERFASSUNGSLÖSUNG

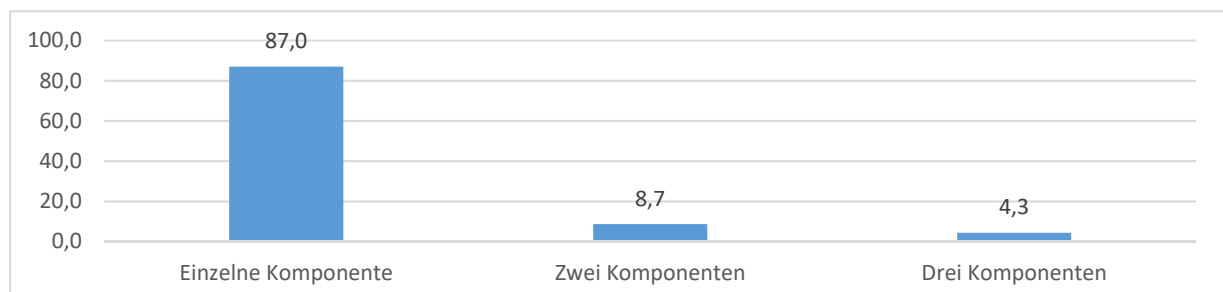
Wie bereits erwähnt, kann eine Erfassungslösung aus mehreren Komponenten bestehen, wie etwa aus einem CMS, z. B. HISinOne und einer Ergänzung durch Excel- oder Access-Tabellen.

Der Vergleich aktueller und künftig geplanter Erfassungslösungen findet sich in Abbildung 3 und 4. Sowohl bei der aktuellen als auch der zukünftigen Erfassungslösung lässt sich eine Tendenz zur Verwendung einer Erfassungslösung mit einer einzelnen Komponente erkennen. 68 Prozent der befragten Hochschulen (32 Hochschulen, n=47) verwenden aktuell nur eine Komponente. Beim Wechsel zu einer neuen Erfassungslösung planen 87 Prozent (20 Hochschulen, n=23) der befragten Hochschulen eine Komponente zu nutzen. Zwei Komponenten werden aktuell von 13 Hochschulen verwendet. Zukünftig planen dies nur noch zwei Hochschulen. Die Verwendung von drei Komponenten wurde aktuell von nur zwei Hochschulen angegeben, zukünftig bevorzugt dies nur noch eine der teilnehmenden Hochschulen.

**Abbildung 3** Aktuelle Erfassungslösung, Mehrfachverwendung von Komponenten innerhalb der Erfassungslösung (Angaben in %, n=47)



**Abbildung 4** Zukünftige Erfassungslösung, Mehrfachverwendung von Komponenten innerhalb der Erfassungslösung (Angaben in %, n=23)



Die Verteilung der derzeit verwendeten Erfassungslösungen ist aufgeschlüsselt nach konkreten Systemen bzw. Systemkomponenten. Die genutzten Lösungen der Einrichtungen sind in der jeweiligen Kombination unterschiedlicher Komponenten sehr vielfältig. Einrichtungen haben zum Teil eigene Lösungen

entwickelt, zum Teil werden die auf dem Markt verfügbaren Systeme verwendet. Bei unserer Befragung wurden verschiedene Komponenten, die momentan in Anwendung sind, genannt, teilweise ohne Nennung der exakten Lösung oder der Nennung von Kombinationen. Als CMS werden vor allem HIS GX, HISinOne, CampusNet und SAP genutzt. FIS werden nur von einer Hochschule verwendet und diese nutzt eine Eigenentwicklung. Docata ist bei den Spezialanwendungen die meistverwendete Komponente und bei den anderen Erfassungslösungen handelt es sich häufig um MS-Office-Anwendungen sowie um Eigenentwicklungen. Insgesamt gaben 13 Hochschulen an, dass sie eine Eigenentwicklung nutzen.

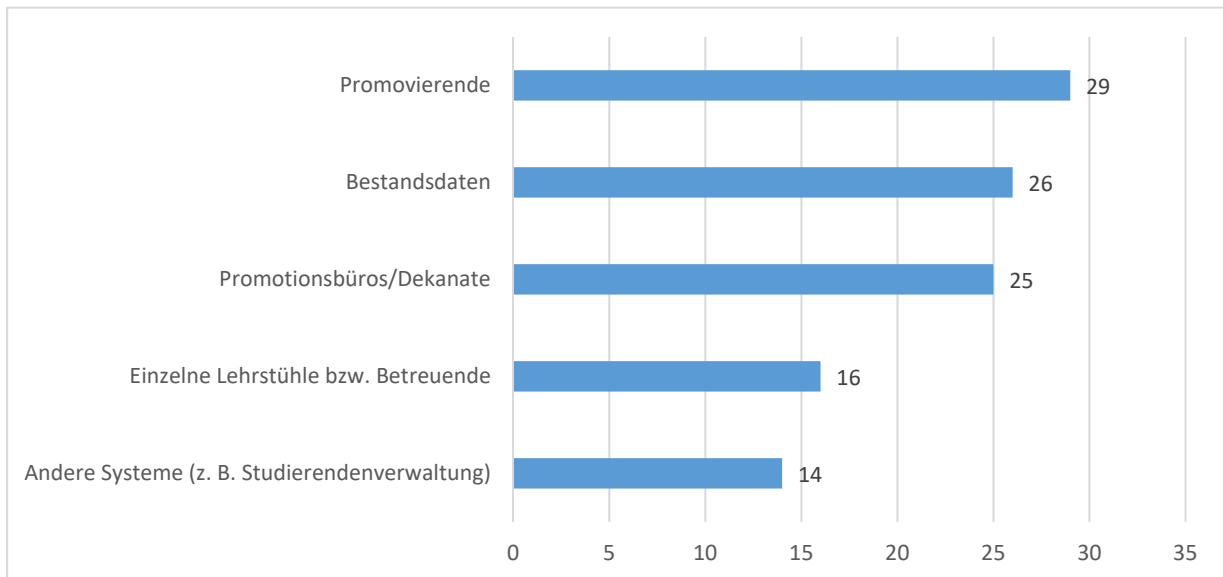
## 4 DATENERFASSUNG UND DATENLIEFERUNG

Ein weiterer Themenblock der Befragung beinhaltete Fragen zur Datenerfassung. Hintergrund dieser Fragen war die zweite Datenlieferung der Promovierendendaten mit Stichtag 1. Dezember 2018. Es wurden Fragen zu den Datenquellen, zur Aktualität der Daten, zu nicht erfassten Teilgruppen, zu Problemen bei der Datenbereitstellung und zur zurückgemeldeten Fehlerquote der statistischen Landesämter bei der Datenübertragung gestellt. Probleme der Datenbereitstellung werden im Kapitel 5 gesondert betrachtet.

### NACHERFASSUNG

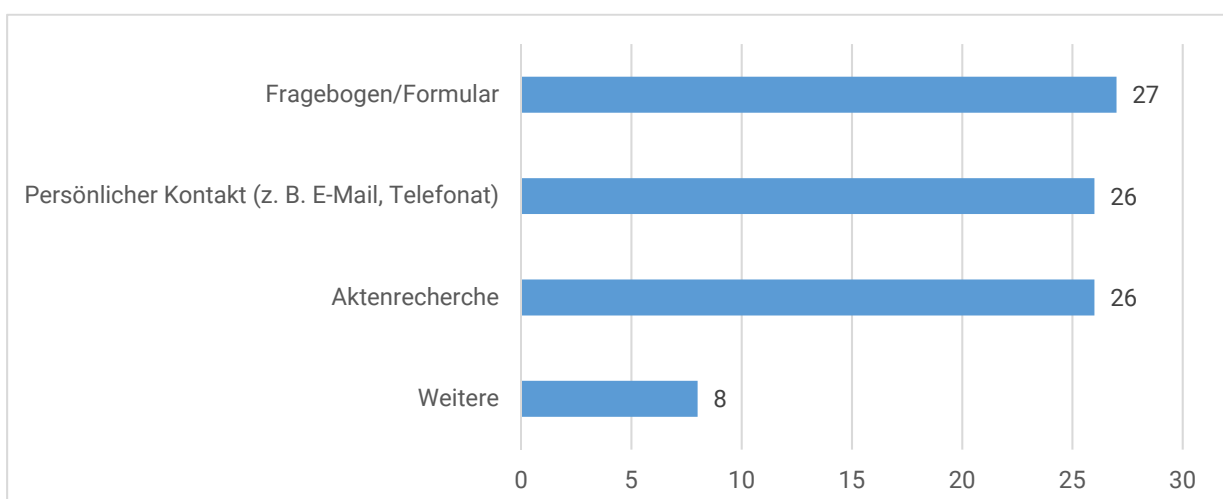
Die Ergebnisse der Antworten auf die Frage **Mit welchen Datenquellen haben Sie nacherfasst?** sind in Abbildung 5 dargestellt. Es zeigt sich, dass die befragten Einrichtungen die zum Teil notwendige Nacherfassung von Promovierendendaten mittels einzelner oder durch Kombination mehrerer Datenquellen durchgeführt haben. So fassten 26 Einrichtungen anhand von Bestandsdaten nach, 25 Einrichtungen befragten die Promotionsbüros bzw. Dekanate, 14 Einrichtungen verwendeten andere Systeme, wie z. B. Archive (eine nähere Erläuterung dazu findet sich in Abbildung 6). 29 Hochschulen befragten die Promovierenden persönlich, 16 Hochschulen gaben an, von Lehrstühlen bzw. Betreuenden Daten angefordert zu haben.

**Abbildung 5** Mit welchen Datenquellen haben Sie nacherfasst? (Anzahl Nennungen, Mehrfachantworten möglich, n= 49)



Nachfolgend sollten die Einrichtungen beschreiben, wie die Nacherfassung durchgeführt wurde und wie hoch der geschätzte Arbeitsaufwand dafür war (siehe Abbildung 6): 27 Hochschulen nutzten einen Fragebogen oder ein Formular zur Nacherfassung. Jeweils 26 Hochschulen kontaktierten die Promovierenden persönlich, z. B. per Mail oder per Telefon, oder führten eine Aktenrecherche durch. Acht Hochschulen verwiesen auf andere Quellen wie Datenbanken, Listen aus den Fakultäten, eine Hochschulmedienrecherche, eine Online-Recherche über die Bibliothek zu veröffentlichten Lebensläufen oder eine Recherche über soziale Netzwerke (ResearchGate, LinkedIn, Xing).

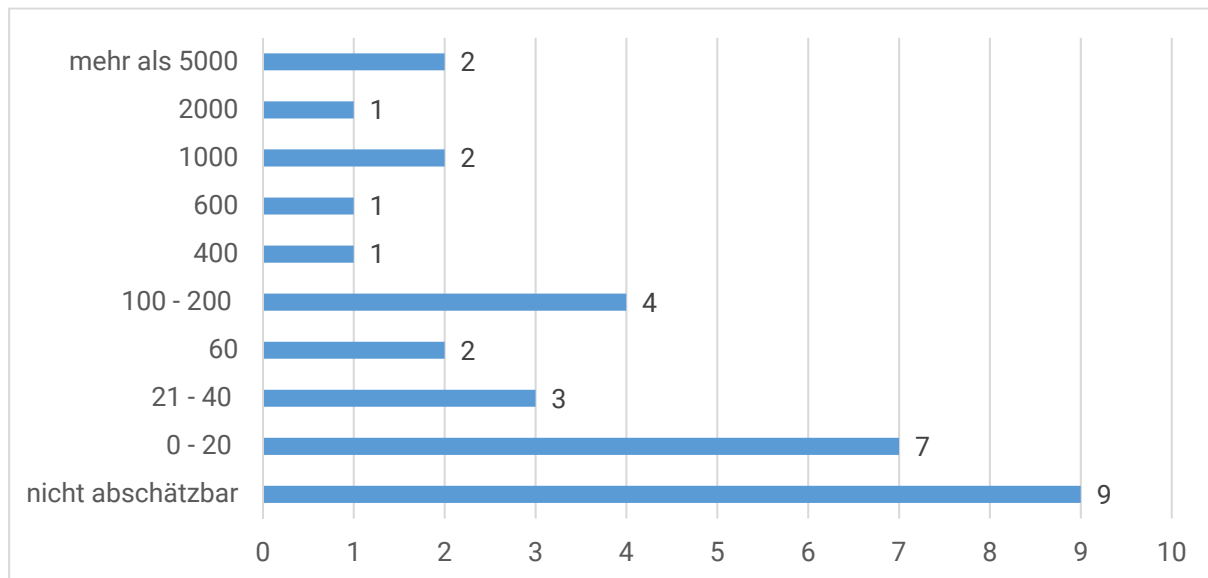
**Abbildung 6** Art der vorgenommenen Nacherfassung (Anzahl Nennungen, Mehrfachantworten möglich, n=49)



Die Frage nach dem Arbeitsaufwand (siehe Abbildung 7), den die Recherche für die Nacherhebung in Anspruch genommen hat, brachte sehr unterschiedliche Antworten hervor. Für viele Hochschulen war

der tatsächliche Aufwand nicht abschätzbar. Sieben Hochschulen gaben an bis zu 20 Stunden für die Nacherhebung aufgewendet zu haben. Sieben Einrichtungen gaben an mehr als 200 Stunden verwendet zu haben, zwei sogar mehr als 5000 Stunden.

**Abbildung 7** Arbeitsaufwand für die Nacherhebung in Stunden (Anzahl Nennungen, n=32)



16 Hochschulen beschrieben die Vorgehensweise der Nacherfassung konkret. Dabei variierte der Aufwand der Nacherfassung stark. Dieser war weniger abhängig von der Größe der Hochschule als von der Komplexität des Nacherhebungsprozesses. Viele Hochschulen gaben an, dass der Aufwand schwer abschätzbar ist, da unterschiedliche Einrichtungen beteiligt waren. Eine befragte Einrichtung gab zudem an, dass eine kontinuierliche Datenbankbetreuung erfolgt und damit eine „reine“ Nacherfassung nicht stattfand. Als Problem wurde häufig die Erreichbarkeit der Promovierenden genannt bzw. die Kooperation der Betreuenden. Weiterhin machten zwei Hochschulen Angaben zur Entwicklung eines neuen Erfassungssystems. Diese Arbeit erhöhte dementsprechend den angegebenen Aufwand in Stunden. Im Folgenden werden Zitate aus den Antworten dargestellt:

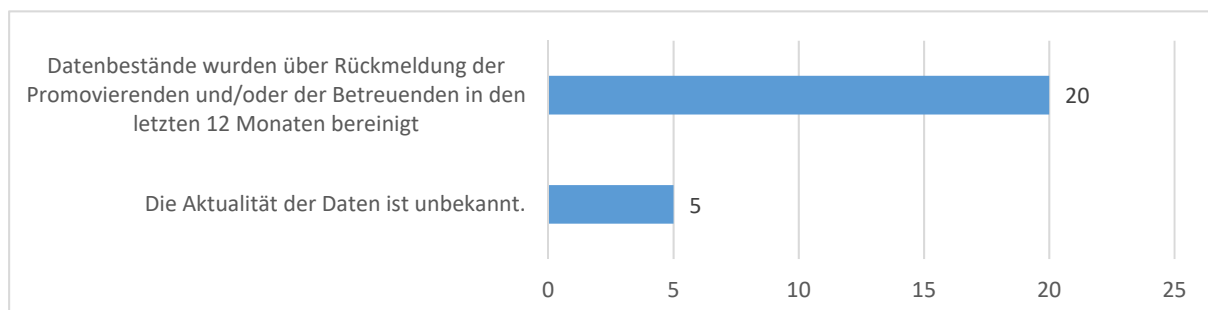
- Die Entwicklungsarbeit an der Software erforderte 3 PM E13, die übrigen Arbeiten für die Nacherfassung (Planung, Kommunikation, Datenpflege etc.) ca. 60 h SHK und ca. 150 h E13
- Wir haben für diese Lieferung eine höhere Fehlerquote in Kauf genommen. Priorität hatte die Vollständigkeit der Daten der abgeschlossenen Promotionen. Eine Befragung der Promovierenden hat in diesem Jahr nicht stattgefunden.
- Bei nicht eingeschriebenen Studierenden ist der Aufwand erheblich, da auch alle Stammdaten, wie z. B. Erst-HZB oder Ersthochschule nacherfasst werden müssen. Bei unseren eingeschriebenen Studierenden sind diese Daten bereits erfasst und plausibilisiert.
- Abfrage per Liste in den Fachbereichen, Entwicklung eines Fragebogens, Nacherfassung und Abgleich



- Der Zeitaufwand ist nicht abschätzbar – an der Nacherfassung waren Promotionsbüros von 18 Fakultäten beteiligt, die z. T. weitere Mitarbeiter zur Unterstützung hinzugezogen haben.
- Die Nacherhebung wurde als Anlass zur Umstellung der dezentralen Erfassung der Promotionsdaten in den Fakultäten auf eine zentrale Erfassung in einer Promotionsdatenbank genutzt. Durch diesen Schritt werden weitere Synergieeffekte erwartet.
- Nacherfassung gestaltet sich sehr schwierig, weil Promovierende trotz mehrmaliger Aufforderung nicht reagieren und/oder Betreuer mit dem Hinweis auf Datenschutz nicht kooperieren/Kontaktdaten ihrer Promovierenden nicht zur Verfügung stellen
- Nacherfassungen sind sehr schwierig, da die ehemaligen Promovierenden, deren Angaben fehlerhaft/lückenhaft sind, größtenteils nicht mehr erreichbar sind (E-Mail-Adressen nicht mehr aktiv, umgezogen, ...)
- Alle angenommene Promovierende, von denen E-Mail-Adressen in den Listen der Geschäftsstellen vorhanden waren, wurden individuell per E-Mail aufgefordert ihre Daten online einzugeben.
- Bei der Promovierendenerfassung für das Jahr 2018 mussten nur noch sehr geringe Daten nacherfasst werden. Dies war insbesondere bei Promovierenden der Fall, die ihre Promotion in 2018 abgeschlossen haben und bei den laufenden Promotionen noch nicht gemeldet wurden.
- nicht zu ermitteln, da verschiedene Personengruppen beteiligt sind
- kann nicht beziffert werden, da parallel zur allgemeinen Datenbankbetreuung
- allein 12 Stunden in der Verwaltung, der Aufwand in den Dekanaten kann nicht abgeschätzt werden
- Variiert stark nach Fakultät; wenige Stunden bis mehrere Stunden/pro Woche über mehrere Monate hinweg

Auf die Frage **Welche Aussage lässt sich zur Aktualität Ihrer Promovierendendaten treffen?** antworteten 48 Einrichtungen (siehe Abbildung 8). 20 Hochschulen bezeichnen die Datenbestände zum Zeitpunkt der Befragung als bereinigt, fünf Hochschulen bewerteten die Aktualität der Daten als unbekannt.

**Abbildung 8** Aktualität der Promovierendendaten (Anzahl Nennungen, n= 25)



23 Hochschulen beschrieben die Aktualität ihrer Daten offen. Dabei wiesen viele Hochschulen daraufhin, dass die Datenbestände neu zugelassener Promovierender aktuell sind. Die Bereinigung der Daten wird jedoch von einigen Einrichtungen als Problem angegeben. Auch innerhalb der Merkmale variiert die

Aktualität bei einigen Hochschulen. Beispielhaft wurde die Aktualität wie folgt beschrieben, hierbei handelt es sich um einen stichpunktartigen Auszug der Antwortzitate:

- Abbrüche wurden in Promotionsbüros nacherfasst; Langzeitpromovierende werden gemeinsam mit Promotionsbüros und Immatrikulationsamt überprüft
- Abschluss/Abbruch soweit im Dekanat bekannt sind aktuell. Neuerfassung geschieht laufend
- Aktualität wird von den Dekanaten, die die Promotionsverfahren verwalten, gewährleistet.
- ca. 90 % der Daten sind aktuell, Probleme mit nicht immatrikulierten Promovierenden
- Da die Promotion mit einer Immatrikulation verbunden ist, ein teilstrukturiertes Programm zu absolvieren ist und Jahresberichte anzufertigen sind, ist die Aktualität der Daten hoch.
- Datenbestände weitestgehend aktuell, v. a. im Hinblick auf die Gesamtzahl. Aktualität der Merkmale (z. B. Art der Dissertation) könnte variieren, da nicht jede Fakultät jährlich diese Merkmale bei laufenden Promotionen abfragt.
- Datenbestände wurden über Rückmeldung der Promovierenden und/oder der Betreuenden in den letzten 12 Monaten bereinigt, allerdings nicht vollständig. Es ist davon auszugehen, dass in den Daten noch eine Reihe nicht gemeldeter Abbrüche erfasst sind.
- Der Großteil der Daten wurde ergänzt. Die Datenbestände sind weiterhin nicht vollständig bereinigt.
- Die Aktualität der Daten für neue Promovierende ist hoch. Die Bereinigung der Daten ist allerdings noch ein Problem.
- Die Aktualität der Daten ist fakultätsspezifisch unterschiedlich von topaktuell bis unbekannt.
- Die Aktualität der Daten ist sehr heterogen. Wir gehen davon aus, dass die meisten Datensätze aktuell sind.
- Die Daten sind noch sehr unvollständig
- Die Daten wurden soweit möglich bereinigt und aktualisiert
- Durch die Systemumstellung wurden neue Prozesse der Datenaufnahme implementiert. Die Datenaktualität wird kontinuierlich verbessert. Bis zur Datenlieferung Ende 2019 wird die Qualität weiter deutlich steigen
- Laufende Prüfung von eingegebenen und nachgetragenen Daten, etwa 30 % aktuell
- Nach der Ersterfassung werden die Daten nicht wirklich bereinigt. Eine Rückmeldung erfolgt nur wenn eine Immatrikulation vorliegt. Hier wird aber auch nicht alle Merkmale überprüft bzw. aktualisiert
- Neu-Zugelassene: Daten aktuell; Alte: Aktualität unbekannt, da oftmals keine Rückmeldung
- Tagesaktuell sind die Promovierendendaten in den Fakultäten (Excel-Dateien), die gemeinsame Access-Datenbank zur Übermittlung der Promovierendendaten an das Stat. Landesamt wird fristgerecht zum Stichtag aktualisiert.
- z. T. bereinigte Daten, z. T. Altdaten

- Zahl der laufenden/abgeschlossenen Promotionen ist quasi zu 100 % valide, da bereinigt. Vollständigkeit der Daten aufgrund der Datenlage der zur Verfügung stehenden Unterlagen ist je nach Datenmerkmal noch sehr unterschiedlich (von 100 % hin zu ca. 50 %)

## NICHT ERFASSTE TEILGRUPPEN

Auf die Frage **Einige Universitäten haben Teilgruppen von Promovierenden, die in der Promovierendenstatistik nicht erfasst wurden. Wie sieht das bei Ihnen aus, haben Sie nicht erfasste Teilgruppen?** antworteten 57 Prozent (n=27) mit nein und 43 Prozent mit ja (n=20).

Von den Einrichtungen genannte nicht-erfasste Teilgruppen sind insbesondere Nicht-Immatrikulierte. Auch externe Promovierende oder Promovierende in der Anfangsphase werden benannt. Ein Spezialfall stellen Medizinstudierende, die ihre Promotion bereits begonnen haben, an einigen Hochschulen dar. Weiterhin wurden Promovierende mit unvollständigen Datensätzen sowie Altfälle benannt.

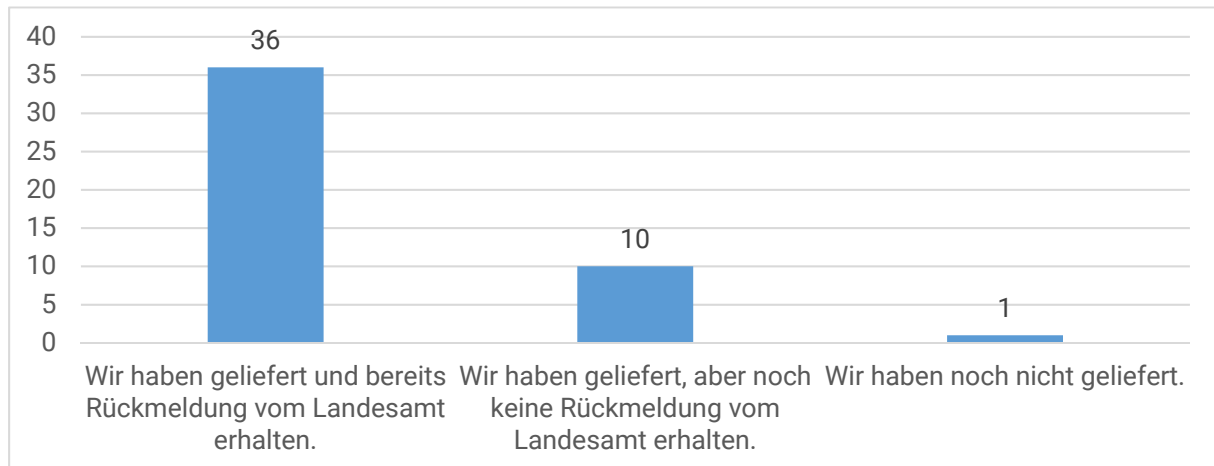
## RÜCKGEMELDETE FEHLERQUOTE DES JEWEILIGEN STATISTISCHEN LANDESAMTES UND KORREKTURSCHLEIFEN

Fragenkomplex: **Die Datenlieferung an das Statistische Landesamt zieht häufig Korrekturwünsche seitens des Landesamtes nach sich, so dass unter Umständen mehrere Lieferungen und Korrekturschleifen notwendig sind.**

Dieser Fragenkomplex beinhaltete Fragen zum Stichtag der Lieferung. Zudem wurde gefragt, in welcher Phase der Datenlieferung sich die Hochschule aktuell befindet. Weiterhin wurde erfragt, in der wievielten Korrekturschleife sich die Hochschule aktuell befindet und wieviel Prozent der Datensätze vom Statistischen Landesamt in der aktuellen bzw. letzten Korrekturschleife als fehlerhaft zurückgemeldet wurden. Auch nach einer Frist seitens des Landesamtes für die Korrektur der Daten wurde gefragt. Weiterhin konnten die Befragten weitere Anmerkungen zum Prozess der Datenmeldung machen.

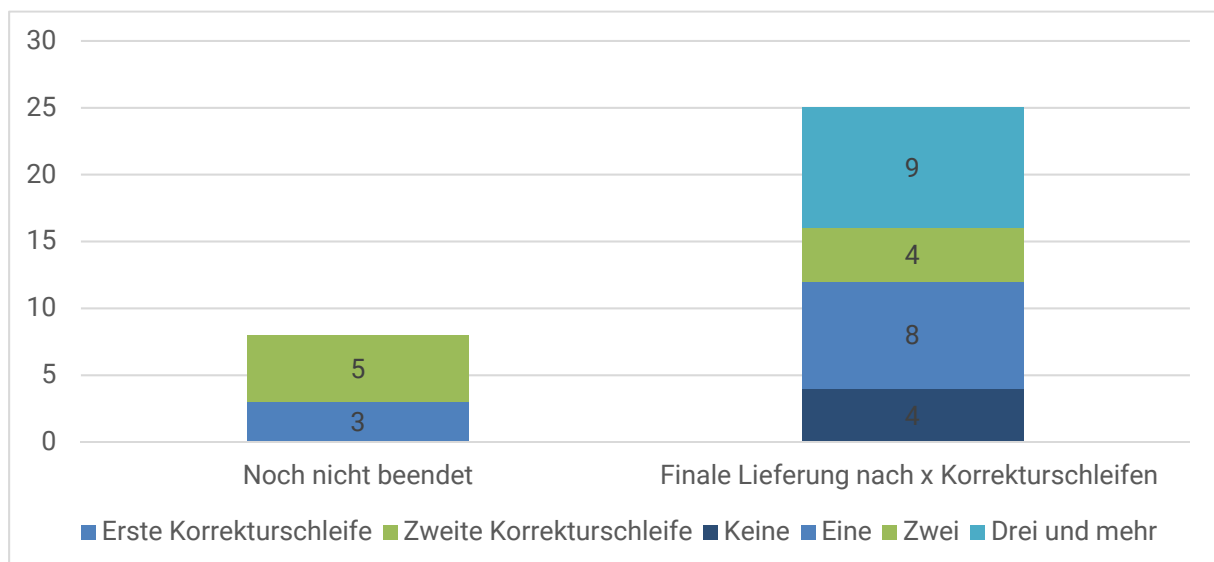
Die Hochschulen befanden sich zum Befragungszeitpunkt auf einem jeweils unterschiedlichen Stand des Verfahrens mit bspw. auf der einen Seite noch gar keiner oder auf der anderen Seite bereits mehreren Rückmeldungen durch das zuständige Statistische Landesamt. Der Stichtag zur Datenlieferung variierte dabei je nach Bundesland zwischen Dezember 2018 und Februar 2019. Für die meisten Hochschulen (n=28) lag der Stichtag im Januar. Vier Hochschulen gaben ein Datum im Dezember und fünf ein Datum im Februar als Stichtag an. 36 Hochschulen (77 %) hatten zum Zeitpunkt der Befragung bereits Daten geliefert und Rückmeldung vom Landesamt erhalten. Zehn Einrichtungen (21 %) hatten geliefert, aber noch keine Rückmeldung bekommen und eine Hochschule hatte noch nicht geliefert (siehe Abbildung 9).

**Abbildung 9** In welcher Phase der Datenlieferung befindet sich Ihre Hochschule aktuell? (Anzahl Nennungen, n=47)



Die Korrekturfrist seitens des Landesamtes variierte dabei je nach aktuellem Stand der Meldung und reichte von Mitte Januar bis Mitte Mai. Acht Hochschulen befanden sich zum Befragungszeitpunkt noch in der Korrektur. Davon drei in der ersten und fünf in der zweiten Korrekturschleife. 25 Hochschulen hatten bereits final geliefert. Vier Einrichtungen benötigten keine Korrekturschleife bei der Datenlieferung. Acht Hochschulen durchliefen eine, vier Hochschulen zwei und neun Hochschulen benötigten drei und mehr Korrekturschleifen bis zur finalen Lieferung (siehe Abbildung 10).

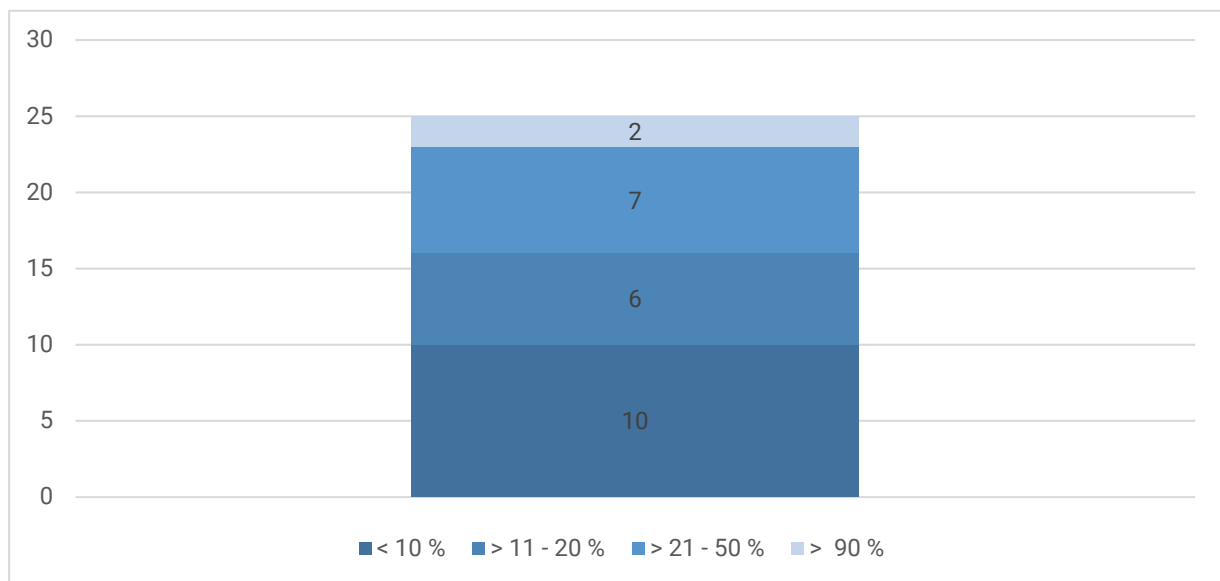
**Abbildung 10** Aktuelle Korrekturschleife bzw. Finale Lieferung nach x Korrekturschleifen (Anzahl Nennungen, n=33)



Es gibt Hochschulen, die zum Befragungszeitpunkt sehr geringe Fehlerquoten vom Statistischen Landesamt zurückgemeldet bekommen hatten, aber auch solche mit noch sehr hohen Fehlerquoten. Im Allgemeinen ist bekannt, dass die Fehlerquote mit der Anzahl der Rückmeldungen durch die Statistischen Landesämter abnahm. Abbildung 11 verdeutlicht die Fehlerquoten als momentane Bestandsaufnahme unabhängig von der Anzahl der durchlaufenen Korrekturschleifen. Eine Kategorisierung nach

Anteil der erhaltenen Fehlerrückmeldungen in Prozent zeigte folgendes Bild: Ein großer Teil, nämlich 10 von 25 Einrichtungen (40 %), gab an, eine Fehlerquote von bis zu zehn Prozent zurückgemeldet bekommen zu haben. Sechs Hochschulen (24 %) gaben an, Rückmeldungen über eine Fehlerquote von elf bis 20 Prozent erhalten zu haben. Sieben Hochschulen (28 %) hatten eine Fehlerquote von 21 bis 50 Prozent und zwei Einrichtungen haben eine sehr hohe Fehlerquote von über 90 Prozent zurückgemeldet bekommen.

**Abbildung 11** Wie viel Prozent der Datensätze Ihrer Datenlieferung wurden vom Statistischen Landesamt als fehlerhaft zurückgemeldet? (Anzahl Nennungen, n=25)



Zudem gab es eine Reihe von Anmerkungen der Befragten bezüglich der Fehler und Fehlerrückmeldungen, zur Zusammenarbeit mit den Statistischen Landesämtern sowie dem Statistischen Bundesamt. Dabei wurde die Zusammenarbeit mit dem Landesamt u. a. als kooperativ beschrieben. Bestimmte Merkmale führen dabei häufiger zu Fehlermeldungen, wie z. B. die HZB oder der Abschluss bzw. das Studienfach. Eine Hochschule gab zudem an, dass einige Merkmale nicht vorlagen. Und zwei Einrichtungen verwiesen auf technische Probleme bei der Datenlieferung (Antwortzitate):

- Fehlermeldungen sind überwiegend entstanden durch a) die Verbindung von Abschluss und Studienfach an anderen Universitäten: Das Statistische Landesamt prüft, ob die Kombination Abschluss/Studienfach an den jeweiligen Universitäten möglich ist. Hier gab es aufgrund der Fächer-schlüssel einige Rückfragen. b) Bei B.Ed./M.Ed. oder alten Lehramtsstudiengängen musste als Abschlussfach ein Schulfach stehen und nicht die Bildungswissenschaften c) Probleme bei der Interpretation der HZB: im Schlüsselverzeichnis wird die HZB definiert als den 1. Zugang zum deutschen Hochschulsystem. Das Statistische Landesamt will hier aber in jedem Fall den Schulabschluss. Hier kam es insb. bei Promovierenden, die die vorherigen Bildungswege nicht in Deutschland absolviert

haben, zu fehlerhaften Angaben, da als HZB der Abschluss eines Studiums im Ausland angegeben wurde statt des Schulabschlusses.

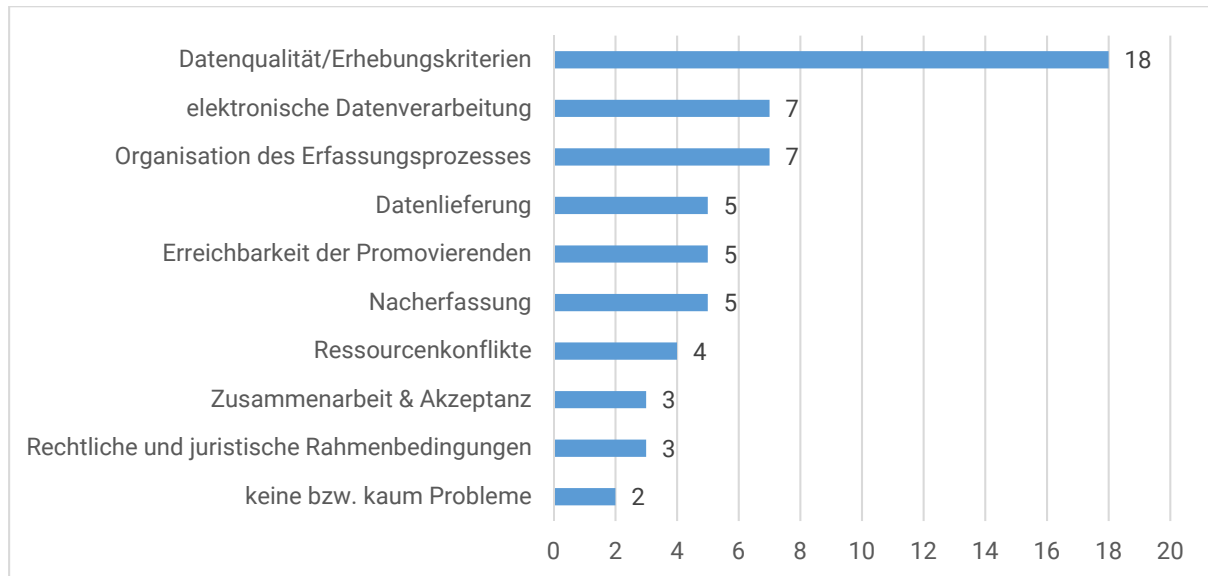
- Probleme der Kennzeichnung von nicht vorhandenen Informationen
- Das Hochladen der Daten ergab immer wieder Fehlermeldungen und die anschließend mit IT-Unterstützung hochgeladene Datei war dann im Nachhinein doch fehlerhaft.
- Der Austausch mit dem Landesamt war kooperativ
- Die Aufbereitung der Fehler seitens des StaLa war sehr hilfreich.
- Die zur Korrektur rückgemeldeten Daten sind per se nicht falsch. Die Daten wurden – wie von den Doktoranden geliefert – an das StaLa übermittelt. Leider entspricht nicht jede Angabe seitens der Doktoranden der Logik des StaLa, so dass Daten im Nachgang angepasst werden müssen. Letztlich sind es zu viele abgefragte Informationen, welche nicht eindeutig auseinandergelassen werden können.
- Für die Korrektur gab es keine harte Frist, die Bearbeiterin ist direkt mit uns in Kontakt getreten und hat mit uns gesprochen, in welchen Zeitrahmen wir die Korrektur schaffen können.
- Im Vergleich zur ersten Datenmeldung im Januar 2018 ist die Prozentzahl der fehlerhaften Datensätze deutlich gesunken.
- Keine konkrete Frist für die Nachlieferung vorhanden. Die hohe Fehlerquote liegt an den fehlenden technischen Voraussetzungen und sind größtenteils systembedingt.
- neue Form der Datenlieferung unter CORE war nicht erfolgreich. Alte Form unter IDEV war aber ok.
- ständiger Austausch mit dem Landesamt für Statistik findet statt; der relativ hohe Anteil der fehlerhaften Datensätze liegt darin begründet, dass einzelne Daten teilw. nicht vorliegen – d. h. wenn in einem Datensatz eines von 34 Merkmalen fehlt, wird dies als fehlerhafter Datensatz gewertet.
- Uns wurde eine Meldung in mehreren Teillieferungen gewährt, zu denen es z. T. dann auch Rückmeldung vom StatLA gab. Mittlerweile wurden alle Daten übermittelt aber es steht noch eine finale Rückmeldung aus.

## 5 AUFTRETENDE PROBLEME BEI DER BEREITSTELLUNG DER DATEN FÜR DIE PROMOVIERENDENSTATISTIK

In Form einer offenen Frage hatten die Befragten die Möglichkeit, spezielle Probleme zu benennen, die bei der Erhebung und Bereitstellung der Daten für die Promovierendenstatistik aufgetreten sind. Den Einrichtungen wurde folgende Frage gestellt: **Welche Probleme traten bei der Bereitstellung der Daten für die Promovierendenstatistik für den Stichtag 1. Dezember 2018 auf? Diese Probleme können von der Nacherhebung der Daten bis hin zu Problemen bei der Datenlieferung reichen.**

Anhand dieser Frage sollte der Beratungs- und Unterstützungsbedarf der Hochschulen ermittelt werden. Insgesamt haben 34 Hochschulen diese offene Frage beantwortet. Es konnten zwölf Problemfelder identifiziert werden (siehe Abbildung 12).

**Abbildung 12** Problemfelder bei der Bereitstellung der Daten für die Promovierendenstatistik (Anzahl Nennungen, n=34)



Der Punkt Datenqualität und Erhebungskriterien stellt das am häufigsten thematisierte Problemfeld dar. Aber auch Probleme mit den eingesetzten Erfassungssystemen und der elektronischen Datenverarbeitung sowie der Organisation des Erfassungsprozesses werden benannt. Fünf Hochschulen benannten auch den Prozess der Datenlieferung an das Landesamt als problematisch. Weiterhin ist der Bereich Nacherfassung der Daten und die Erreichbarkeit der Promovierenden ein Bereich, der bei einigen Hochschulen zu Problemen führte. Aber auch ein hoher Arbeitsaufwand bei der Bereitstellung der Daten (Ressourcenkonflikte) sowie interne Akzeptanzprobleme oder Probleme in der Zusammenarbeit mit den Fakultäten, Landesämtern etc. wurden benannt. Drei Hochschulen nannten rechtliche Probleme, z. B. bei der Nacherfassung von Daten. Zwei Einrichtungen gaben an, dass die Datenbereitstellung ohne Probleme erfolgte. Im Folgenden werden Zitate der Antworten zu den jeweiligen Problemfeldern aufgezeigt.

### **Datenqualität und Erhebungskriterien**

- Bestimmte Daten, auch die zur Pseudonymgenerierung erforderlichen, liegen nicht vollständig vor
- Bestimmte Daten, wie Ersteinschreibung als Studierender sind, besonders für ausländische Promovierende, unbekannte Konzepte und daher anfällig für Fehleingaben. Daraus ergeben sich wiederum oft Plausibilitätsfehler, die einzeln überprüft werden müssen.
- die Daten der immatrikulierten Promovierenden weisen teilweise Lücken auf, diese müssen identifiziert und nacherfasst werden
- Nach wie vor Probleme durch unvollständige Datensätze
- Altfälle, die eingeschrieben sind, den Dekanaten aber nicht bekannt sind und auch noch keinen Antrag eingereicht haben.
- nicht alle abgeschlossenen Promotionen können aufgrund fehlender Daten gemeldet werden. Hier gibt es dann Differenzen in den Statistiken.

- Bereinigung und Koordination, v. a. im Hinblick auf widersprüchliche Daten (z. B. Promovierende geben auf dem Formular nur einen Teil ihres Namens an, Tippfehler o. ä.)
- Dubletten (v. a. nach Namensänderung)
- Angaben der HZB fehlen bei etwa 40 %. Kumulative Dissertation vs. Monographie wurde bisher nicht erfasst.
- Die Datenqualität ist noch nicht zufriedenstellend, für 17 % der Datensätze wurden Fehler/Lücken zurückgemeldet. Bei drei Fakultäten steht eine systematische Prüfung auf Karteileichen aus.
- Einzelne Felder der Statistik konnten für die Datenlieferung nur schwer erhoben werden, z. B. HZB, Geburtsort etc.
- fehlerhafte Rückmeldungen internationaler Promovierender, die nicht verwendet werden können.

## **Elektronische Datenverarbeitung**

- keine vorhandene Datenbank für Promovierendenerfassung, Erfassung mit Excel-Tabellen und Access-Datenbank des statistischen Landesamtes sehr zeitaufwendig und fehleranfällig
- Datenlücken können in der Access-Datenbank des statistischen Landesamtes nicht entsprechend gekennzeichnet werden. Das bedeutet, dass nach der Datenlieferung eine Fehlerliste an uns zurückgesendet wird, in die händisch bei jeder fehlenden Information nv eingetragen werden muss, jedes Jahr aufs Neue
- Technische Probleme, z. B. kann die Frage nach dem promotionsberechtigten Abschluss nicht mit ja beantwortet werden, wenn keine Informationen zum Abschluss vorliegen.
- Das eingesetzte System hat noch Nachbesserungsbedarf bei der Bestimmung der zu liefernden Kohorte sowie bei der Festlegung einzelner Attribute.
- Das System zur flächendeckenden Erfassung der Promovierendendaten befindet sich noch in der Entwicklung.
- Ein Problem ist die fehlende Schnittstelle zu den Personaldaten. Des Weiteren können einige Informationen nicht in unseren Systemen gespeichert werden, z. B. Art der Dissertation oder die Teilnahme an strukturierten Promotionsprogrammen. Diese Daten müssen wir separat in einer Access-Datenbank erfasst werden.
- Lieferung über e-core funktionierte nicht auf Anhieb.
- Probleme mit und bei der Nutzung der Software

## **Organisation des Erfassungsprozesses**

- Zudem ist die Bearbeitung auf dezentraler Seite durch Sachbearbeiter\*innen organisiert, die zuvor nie mit Studierendenstatistiken o. ä. in Kontakt gekommen sind. Beim Befüllen der Datensätze müssen sie jedoch Entscheidungen über Eintragungen treffen, z. B. ob ein vorheriger Studiengang konsekutiv war oder nicht. Dies führt leider zu Falscheintragungen und dadurch zu Datenqualitätsproblemen, so dass hier noch viel Schulungsaufwand nötig ist, um die Daten zu verbessern.
- universitätsinterne Prozesse zur Erfassung von Kooperationen kommen nur langsam in Gang



- Die Nacherhebung der Daten und die personelle Zuständigkeit waren problematisch.
- Ein Problem ist die fehlende Schnittstelle zu den Personaldaten.

## **Datenlieferung**

- Die Statistik-Lieferungen sind unnötig kompliziert und die Formatierung im csv-Datenformat sehr unpraktisch
- Datenlieferung, vor allem die Fehlerprüfung bei der Übermittlung ist nicht nachvollziehbar
- Es sind noch nicht alle Daten erfasst, so dass einige Abfragen frei bleiben mussten und andere händisch aus bestehenden Daten abgeleitet werden mussten.
- Wir hatten insbesondere Probleme bei der Datenlieferung. Die Art der Registrierung musste manuell angepasst werden, da dies aus SOS falsch exportiert wurde. Auch bei den Daten zur HZB und zur Promotion berechtigte Abschlussprüfung entstanden Probleme. Nach dem Datenexport aus SOS musste manuell einiges ausgebessert werden. Zusätzlich mussten wir nach der Lieferung die Daten nochmals ausbessern. Insgesamt verursachte die korrekte Meldung der Statistik rund 24 Arbeitsstunden

## **Erreichbarkeit der Promovierenden**

- Außerdem war es schwierig, die Promovierenden zu erreichen, um diese Daten zu erheben, die nicht an der Universität vorhanden waren.
- nicht vollständig ausgefüllte Datenblätter der Promovierenden
- Promovierende, die weder für uns noch für den Doktorvater/Doktormutter erreichbar sind um die Daten zu vervollständigen
- Daten waren z. T. nicht auffindbar, Promovierende haben auf Nachfragen nicht reagiert
- Nicht alle Promovierenden haben sich an der Erhebung beteiligt.

## **Nacherfassung**

- die Daten der immatrikulierten Promovierenden weisen teilweise Lücken auf, diese müssen identifiziert und nacherfasst werden
- großer Aufwand für alle Beteiligten bei der Nacherfassung der nicht immatrikulierten Promovierenden
- Die Nacherhebung der Daten und die personelle Zuständigkeit waren problematisch. Bereinigung der Karteileichen steht noch aus.
- Es existieren unverändert Datenlücken bei Promovierenden, die vor der Novelle des Hochschulstatistikgesetzes ihre Promotion begonnen haben und die sich an den erfolgten Nacherhebungen nicht beteiligt haben.

## **Ressourcenkonflikte**

- Der Aufwand zur Nachkorrektur steht in keinem guten Verhältnis zum erwarteten Ertrag
- Arbeitsaufwand enorm hoch
- Zusammenführung der Daten extrem personal- und zeitintensiv

## **Zusammenarbeit & Akzeptanz**

- universitätsinterne Prozesse zur Erfassung von Kooperationen kommen nur langsam in Gang
- Kommunikation des Landesamtes ist verbesserungswürdig. Nach Datenmeldung im Januar bis heute keine Informationen
- Die, die für die zentrale Erfassung zuständig sind, sind auf die Mitarbeit/Datenlieferung der Fakultäten angewiesen. Hier ist die Zusammenarbeit aber schwierig, Zuständige nicht entsprechend motiviert...
- Die Kommunikation mit dem Landesamt für Statistik war wenig intensiv (vermutlich gibt es auch dort Kapazitätsprobleme).

## **Rechtliche und juristische Rahmenbedingungen**

- Die Klärung der rechtlichen Frage, inwieweit für die Promotionsverlaufsstatistik, die ja eine Sekundärstatistik ist, Daten nacherfasst werden müssen, die bislang nicht erfasst wurden, die aber nicht für die Promotion erforderlich sind.
- Immer noch Datenlücken vorhanden, die wir aus datenschutzrechtlichen Gründen bei den Promovierenden nicht nacherfassen durften
- Es können nicht alle Daten für die Promovierendenstatistik angegeben werden, da nicht alle Daten vorliegen (da diese für die eigene Promovierendenerfassung bisher nicht relevant waren). Für eine Nacherfassung aber die rechtliche Grundlage noch fehlt, da das Hochschuldatenschutzgesetz bisher noch nicht angepasst wurde.

## **6 FAZIT**

Die vierte Befragung der Hochschulen zum Stand der Promovierendenerfassung zeigt ein differenziertes Bild. Dabei gibt es bei einigen Themen eindeutige Tendenzen, bei anderen unterscheiden sich die Einrichtungen stark voneinander. Zudem war der Rücklauf der Hochschulen geringer im Vergleich zur dritten Befragung. 50 statt zuletzt 76 Hochschulen beteiligten sich an der aktuellen Befragung.

Derzeit tendiert die Mehrheit der befragten Einrichtungen zur Verwendung eines Campus Management Systems zur Promovierendenerfassung. In Zukunft plant auch eine Mehrheit der Hochschulen, die einen Wechsel der Erfassungslösung anstreben, ein CMS zu verwenden. 13 Hochschulen gaben zudem an, Eigenlösungen zur Erfassung zu verwenden.

Die Hochschulen befanden sich zum Befragungszeitpunkt auf einem jeweils unterschiedlichen Stand der Datenerfassung. Dabei bezeichnete nur ein Teil der Einrichtungen (n=20) ihre Daten als aktuell und

bereinigt. Die meisten Hochschulen nutzten zur Nacherfassung der Daten Abfragen über die Promovierenden oder Promotionsbüros bzw. Dekanate. Für viele Hochschulen war der tatsächliche Aufwand der Nacherfassung nicht abschätzbar. Dies könnte auf eine Verstetigung des Prozesses hinweisen, bei dem Datenbanken eher kontinuierlich gepflegt werden und eine „reine“ Nacherfassung damit an Gewicht verliert.

Auch die Fragen zur Meldung an die Statischen Landesämter zeigten ein wenig einheitliches Bild. Die Stichtage der Lieferung, aber auch die Korrekturfristen variierten dabei deutlich. Einige Einrichtungen hatten noch gar keine und andere bereits mehrere Rückmeldungen durch das zuständige Statistische Landesamt erhalten. Die Fehlerquoten lagen dabei zum Zeitpunkt der Erhebung im Schnitt bei zehn bis 15 Prozent. Es gab jedoch auch Einrichtungen, die noch aktuelle Fehlerquoten von über 90 Prozent angaben. Der Punkt Datenqualität und Erhebungskriterien stellt das am häufigsten thematisierte Problemfeld bei der Bereitstellung der Daten dar und viele Hochschulen benannten in der Befragung konkrete Merkmale, die zu Problemen führen. Insgesamt lässt sich jedoch aus Sicht der Hochschulen eine Verbesserung der Datenerhebung und -lieferung im Vergleich zum Vorjahr feststellen.