

# FDZ *Datenreport*

---

Dokumentation zu Arbeitsmarktdaten

Nr. 3/2009 (DE)

## Stichprobe der Integrierten Erwerbsbiografien

IEBS

*Dirk Oberschachtsiek, Patrycja Scioch,*

*Christian Seysen, Jörg Heining*

Handbuch  
für die IEBS in der Fassung 2008



Bundesagentur für Arbeit

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung .....</b>	<b>5</b>
<b>2</b>	<b>Kurzbeschreibung des Datensatzes.....</b>	<b>7</b>
2.1	Übersicht .....	8
2.2	Mengengerüst und Füllgrad im Zeitablauf.....	9
2.3	Unterschiede zu IABS und BA-Beschäftigtenpanel.....	10
<b>3</b>	<b>Datenaufbereitung.....</b>	<b>11</b>
3.1	Datenquellen .....	11
3.1.1	Beschäftigten-Historik des IAB (BeH) .....	11
3.1.2	Leistungsempfänger-Historik des IAB (LeH) .....	11
3.1.3	Maßnahme-Teilnehmer-Gesamtdatenbank (MTH) .....	12
3.1.4	Arbeitsuchendenstatus aus dem Bewerberangebot (BewA) .....	13
3.2	Verknüpfung der Daten .....	13
3.3	Bereinigungen und Filter .....	15
3.3.1	BeH .....	15
3.3.2	LeH.....	15
3.3.3	MTH .....	17
3.3.4	BewA.....	18
3.3.5	Gesamtdatei.....	18
3.4	Rechtlich unzulässige zeitgleiche Zustände.....	19
3.5	Datenqualität und Probleme.....	20
3.5.1	BeH .....	21
3.5.2	LeH.....	21
3.5.3	MTH .....	22
3.5.4	ASU / BewA .....	22
3.6	Episodensplitting .....	24
3.7	Datenanonymisierung und sensible Merkmale .....	25
3.8	Testdaten und Hinweise für Datenauswertungen .....	26
<b>4</b>	<b>Variablenbeschreibung.....</b>	<b>26</b>
4.1	Allgemeiner Teil.....	27
4.1.1	Merkmalsübersicht.....	27
4.1.2	Fehlende Werte .....	29
4.2	Identifikationsnummern .....	29
4.2.1	Personennummer .....	29
4.2.2	Satznummer.....	30
4.2.3	Betriebsnummer.....	30
4.3	Spellanfang und -ende .....	31
4.3.1	Beginndatum Originalspell .....	31
4.3.2	Enddatum Originalspell .....	31
4.3.3	Beginndatum der gesplitteten Episode .....	31
4.3.4	Beginn- und Enddatum .....	31
4.4	Generierte technische Merkmale .....	32
4.4.1	Quelle der Information .....	32
4.4.2	Kombination der Quellen.....	32

4.4.3	Spellzähler pro Konto .....	33
4.4.4	Anzahl der Spells pro Konto .....	33
4.4.5	Spellzähler pro Episode .....	34
4.4.6	Anzahl der zeitgleichen Spells .....	34
4.4.7	Spellzähler pro Episode und Quelle .....	34
4.4.8	Anzahl der zeitgleichen Spells pro Quelle .....	34
4.5	<i>Personenstatus vor, während und nach dem aktuellen Spell</i> .....	35
4.5.1	Erwerbsstatus vor Arbeitsuche .....	35
4.5.2	Erwerbsstatus: Personengruppe, Leistungsart, Maßnahmeart, Arbeitsuche-Status .....	35
4.5.3	Grund des Spellendes/Beendigungsgrund .....	37
4.5.4	Status nach Abgang .....	38
4.6	<i>Persönliche Merkmale</i> .....	38
4.6.1	Geburtsjahr .....	38
4.6.2	Geschlecht .....	39
4.6.3	Staatsangehörigkeit .....	39
4.6.4	Schwerbehindertenstatus .....	40
4.6.5	Schulabschluss .....	40
4.6.6	Ausbildung .....	40
4.6.7	Qualifikation .....	41
4.6.8	Familienstand .....	42
4.6.9	Anzahl der Kinder im Haushalt .....	42
4.6.10	Gesundheitliche Einschränkungen .....	43
4.7	<i>Angaben zu Beschäftigungsverhältnis und Arbeitsuche</i> .....	43
4.7.1	Stellung im Beruf und Arbeitszeit .....	43
4.7.2	Beruf .....	44
4.7.3	Tagesentgelt / täglicher Leistungssatz .....	44
4.7.4	Gleitzone .....	45
4.7.5	Wirtschaftszweige (73, 93, 03) .....	45
4.7.6	Bereitschaft zu bundesweiter Stellensuche .....	46
4.7.7	BA-Kundengruppe .....	47
4.7.8	Art der Kündigung .....	47
4.7.9	Arbeitszeit des Stellengesuchs .....	47
4.7.10	Restanspruchsdauer Arbeitslosengeld .....	48
4.7.11	Geplantes Beginndatum .....	48
4.7.12	Geplantes Enddatum .....	48
4.8	<i>Angaben zum Beschäftigungsbetrieb</i> .....	49
4.8.1	Anzahl der hauptbeschäftigten Mitarbeiter .....	49
4.8.2	Gründungsjahr des Betriebes .....	49
4.9	<i>Ortsangaben</i> .....	49
4.9.1	Arbeitsort Bundesland (Kreis) .....	49
4.9.2	Wohnort Bundesland (Kreis) .....	50
4.9.3	Arbeitsort Regionaldirektion (Agenturbezirk) .....	51
4.9.4	Wohnort Regionaldirektion (Agenturbezirk) .....	51
4.9.5	Wohnort Regionaltyp der Arbeitsagentur .....	52
4.9.6	Wohnort im Ausland .....	53
<b>5</b>	<b>Literatur</b> .....	<b>54</b>
<b>6</b>	<b>Anhang</b> .....	<b>55</b>
6.1	<i>Abkürzungsverzeichnis</i> .....	55
6.2	<i>Merkmalsliste alphabetisch</i> .....	56

---

**Datenverfügbarkeit**

Der in diesem Beitrag beschriebene Datensatz ist für die Fachöffentlichkeit zugänglich. Nähere Informationen dazu erhalten Sie auf der Internetseite <http://fdz.iab.de> unter der Rubrik „Personendaten“.

**Danksagung**

Für ihre umfangreichen Vorarbeiten und Hilfestellungen bedanken wir uns bei den Kolleginnen und Kollegen des Servicebereichs ITM des Instituts für Arbeitsmarkt- und Berufsforschung.

Für die Inhalte der FDZ-Daten- und Methodenreporte sind die jeweiligen Autoren verantwortlich.

---

## 1 Einleitung

Bei der IEBS handelt es sich um eine Zufallsstichprobe aus den **Integrierten Erwerbsbiografien (IEB)** des IAB. Die IEB sind nicht als geschlossener Datensatz zu verstehen, sondern als Verfahren der Zusammenspielung von Daten aus vier unterschiedlichen Quellen zur Prüfung der Datenqualität, sowie zur Ziehung von Stichproben wie der IEBS. Bei den vier Datenquellen handelt es sich um

- die **Beschäftigten-Historik (BeH)** mit Spells zu sozialversicherungspflichtiger Beschäftigung aus dem Meldeverfahren,
- die **Leistungsempfänger-Historik (LeH)** mit Spells zum Empfang von Arbeitslosengeld, Arbeitslosenhilfe und Unterhaltsgeld,
- die **Maßnahme-Teilnehmer-Historiken (MTH)** mit Spells zu Maßnahmeteilnahmen und
- **Arbeitsuchenden und Bewerberangebotsdaten (BewA)** mit Spells zur Arbeitsuche.

Die IEB wurden im Rahmen der Projekte Biografische Daten und colAB entwickelt und dabei von einem Beirat aus den IAB-Forschungsbereichen begleitet. Die IEBS 2008 basiert auf der IEB-Version 7.02.

Wichtige Hinweise zur Datenqualität:

- IEB-Daten stehen immer unter dem Vorbehalt ihrer späteren Korrektur durch die datenverursachenden Institutionen. Sowohl im Meldeverfahren, als auch in den Geschäftsdaten der BA finden Korrekturen mitunter Jahre nach dem Ende eines Spells statt. Der Anteil der betroffenen Spells ist jedoch gering.
- IEB-Daten stehen vergleichsweise am Anfang ihrer Entwicklung. Erfahrungsgemäß kommt es dabei zu technisch bedingten Einschränkungen der Datenqualität. Die Freigabe der Daten erfolgt unter diesem Vorbehalt.

In der Regel sind die in den Daten auftretenden Inkonsistenzen nicht eindeutig zu bereinigen. Oftmals hängt die Art der Bereinigung vom Forschungsziel ab. In IEB und IEBS werden bekannte Inkonsistenzen daher nur in eindeutigen Fällen bereinigt, in anderen lediglich dokumentiert. Der Aufwand an Datenaufbereitung ist daher höher als beispielsweise in der IABS, die umfassendere Aufbereitungen durchlaufen hat. Bitte melden Sie Auffälligkeiten und Unstimmigkeiten sowohl in den Daten als auch in dieser Dokumentation dem FDZ.

Informationen zur IEBS und zum Datenzugang finden Sie auf der FDZ-Internetseite (<http://fdz.iab.de>).

Änderungen und Besonderheiten gegenüber der IEBS in der Fassung 2005:

Die Datenlage hat sich seit dem Ladestand der IEBS 2005 erheblich geändert. Grundsätzlich sind alle Meldungen aktuell nachgeladen worden. Zudem sind neue Ausprägungen beim Erwerbsstatus (insbesondere durch neue Maßnahmeteilnahmen) und im Status nach Abgang, sowie im Abgangsgrund entstanden. In einigen Fällen (siehe unten) konnten die bisherigen Ausprägungen nicht vollständig beibehalten werden. Die wichtigsten Änderungen gegenüber der IEBS in der Fassung 2005 sind nachfolgend aufgeführt:

Aktualisierung des Ladestandes: Beschäftigtenmeldungen bis 12/2007; Meldungen zur Maßnahmeteilnahme und Arbeitssuche bis 09/2008 und Informationen zum Leistungsempfang bis 10/2008.

Hinzunahme neuer Variablen: Familienstand, Qualifikation des Bewerbers, Anzahl der Kinder im Haushalt, gesundheitliche Einschränkungen, Wirtschaftszweige, Mobilitätsbereitschaft, BA-Kundengruppenzuordnung, Art der Beendigung der letzten Tätigkeit, Anzahl der Beschäftigten im Beschäftigungsbetrieb, Gründungsjahr des Beschäftigungsbetriebes, Wohnort im Ausland, Arbeitszeit des Stellengesuchs, Restanspruchsdauer Arbeitslosengeld, geplantes Beginn- und Enddatum einer Maßnahmeteilnahme, Wohnort im Ausland

Wegfall bisheriger Variablen: Maßnahmennummer<sup>1</sup>, Bereinigungskennzeichen<sup>2</sup>, Status des Enddatums<sup>3</sup>, Fortsetzung der Arbeitslosigkeit nach Arbeitsunfähigkeit<sup>4</sup>, Rentenversicherungsträger<sup>5</sup>, Beitragsgruppe<sup>6</sup>, Beginn der Arbeitslosigkeit<sup>7</sup>, Dauer der Arbeitslosigkeit<sup>8</sup>, Geringfügigkeitskennzeichen<sup>9</sup>

Die Variable „grund“ ist im Merkmalspektrum umgeschlüsselt.

Die fehlenden Werte sind einheitlich auf den Wert -7 umgeschlüsselt (unbekannte oder fehlende Angabe, Fehler im Ursprungswert sowie fehlende Zuordnung).

Kreisgebietsreformen in Sachsen-Anhalt und Thüringen führen zu neuen Kreisnummern ab 2007. Dies ist bei Auswertungen auf Kreisebene (MTH und BewA) zu beachten.

---

<sup>1</sup> Die Maßnahmennummer ist als Identifikator für eine zeitlich begrenzte Fördermaßnahme ungeeignet. Die Praxis zeigt, dass eine Abgrenzung gegenüber Dauer-Projekten nicht möglich ist.

<sup>2</sup> In Anlehnung an die IABS 2005 wird dieses Merkmal nicht mehr geführt.

<sup>3</sup> Bei den Enddaten der Quellen BewA, MTH ist z.T. davon auszugehen (vorwiegend die Maßnahmegruppen FF und ESF betreffend), dass diese Planungscharakter haben. Für in der Zukunft liegende LeH-Enddaten (relativ zum Ladestand) kann von einer Rechtszensurierung ausgegangen werden.

<sup>4</sup> Diese Variable ist nicht mehr in den Ladedaten für die IEBS enthalten.

<sup>5</sup> Der Informationsgehalt dieser Variablen wurde als zu gering/ungeeignet für diesen Datensatz eingestuft (vgl. u.a. IABS und BAP, siehe <http://fdz.iab.de>)

<sup>6</sup> Der Informationsgehalt dieser Variablen wurde als zu gering/ungeeignet für diesen Datensatz eingestuft (vgl. u.a. IABS und BAP, siehe <http://fdz.iab.de>).

<sup>7</sup> Kann selbst erzeugt werden.

<sup>8</sup> Kann selbst erzeugt werden.

<sup>9</sup> Kann selbst erzeugt werden (umfasst die Erwerbscodes 109 und 209 der Beschäftigtenmeldungen).

Besonderheiten

Der vorliegende Datensatz beinhaltet keine SGB II-Meldungen.

Mit 2004/2005 wurden in den Arbeitsagenturen zeitlich und regional gestaffelt neue Fachverfahren eingeführt (Verbis / CoSachNT). Dies führt ggf. zu Veränderungen in der Datenmenge bei den durch Verbis erfassten Personenmerkmalen während dieses Zeitraumes und bei den Maßnahmeteilnahmen, welche über CoSachNT erfasst werden. Durch die Umstellung werden einzelne Variablen nicht mehr oder anders gefüllt. Soweit möglich wird darauf noch einmal an entsprechender Stelle hingewiesen. Mit der Einführung des Verfahrens CoSachNT wurde auch die Architektur der historisierten Datenhaltung der Maßnahmeteilnahmen geändert. Damit erhöhte sich die Anzahl der Teilnahmen, was entsprechende Auswirkungen auf das Mengengerüst hatte.

Mit 2006 wurde das Fachverfahren coLei zur Erfassung der Leistungsempfangsdaten auf das Verfahren COLIBRI umgestellt (siehe 4.1.2 sowie 4.5.2).

## 2 Kurzbeschreibung des Datensatzes

Die IEB entsprechen grundsätzlich einer Vollerhebung.<sup>10</sup> Die Ziehung der Stichprobe IEBS erfolgt über die zufällige Auswahl von acht Geburtstagen. Alle Personen mit diesen Geburtstagen werden gezogen, dies entspricht etwa 2,2% ( $\approx 8/365$ ) der Personen der IEB. Diese Geburtstags-Stichprobe ist ohne Gewichtung repräsentativ bezogen auf die zum betreffenden Zeitpunkt in den IEB enthaltenen Personengruppen (bzw. Datenquellen). Mit anderen Worten: Das Verhältnis der Bestände einzelner Personengruppen in der Stichprobe entspricht zu jedem Zeitpunkt recht genau deren Verhältnis in den IEB und somit auch der Realität.<sup>11</sup> Hier liegt ein Unterschied der IEBS zu den IAB-Beschäftigtenstichproben IABS. Die dortige Zufallsauswahl erfolgt durch eine zufällige Auswahl von Sozialversicherungsnummern der Beschäftigten.<sup>12</sup> Personen, die im Zeitraum 1975 bis 2001 nicht sozialversicherungspflichtig beschäftigt sind, sind nicht in der IABS enthalten – allerdings in der IEBS.<sup>13</sup>

---

<sup>10</sup> Im Zuge der Datengenese können einzelne Spells verloren gehen. So werden beispielsweise Spells gelöscht, die keiner Person zugeordnet werden können (siehe Abschnitt 3.3). Die Zahl dieser Ausfälle ist jedoch gering.

<sup>11</sup> Eine Einschränkung der Repräsentativität ergibt sich dadurch, dass Personen, deren Geburtsdatum bei der Vergabe ihrer Sozialversicherungsnummer nicht bekannt war, ein bestimmtes Geburtsdatum zugewiesen bekamen, das nicht in den acht Geburtsdaten der Stichprobe enthalten ist. Bei diesen Personen handelt es sich vorwiegend um ältere Ausländer, weshalb diese Personengruppe in der IEBS leicht unterrepräsentiert sein dürfte.

<sup>12</sup> Die Ziehung in der IABS erfolgt geschichtet nach Ost/West, sowie nach Ausländern/Deutschen.

<sup>13</sup> Es gibt keine Auszählungen zu dieser Selektivität in den Daten der IABS. Die Selektivität dürfte aber nur in den ersten Jahren von Bedeutung sein, da Leistungsempfang grundsätzlich vorhergehende sozialversicherungspflichtige Beschäftigung voraussetzt.

## 2.1 Übersicht

<b>Inhaltliche Charakteristika</b>	
Themen/Merkmalgruppen	<p><i>Soziodemographische Merkmale:</i> Geschlecht, Geburtsjahr, Ausbildung, Staatsangehörigkeit, Regionaltyp, Schwerbehindertenstatus, Schulabschluss, Qualifikation des Bewerbers, Familienstand, Anzahl der Kinder im Haushalt, gesundheitliche Einschränkungen, Wohn- und Arbeitsort</p> <p><i>Beschäftigung:</i> Beruf, Tagesentgelt, Stellung im Beruf und Arbeitszeit, Wirtschaftszweig, Gleitzone</p> <p><i>Leistungsbezug:</i> Arbeitslosengeld, Arbeitslosenhilfe, Unterhaltsgeld</p> <p><i>Maßnahmen der aktiven Arbeitsmarktförderung:</i> Maßnahmentyp: ABM, SAM, FBW, TM, DSL, FF, ESF_BA, PSA, JUMP, JUMP_PLUS, Beauftragung Dritter) Förderzeiten, geplantes Beginn- und Enddatum</p> <p><i>Arbeitsuche:</i> Status der Arbeitsuche, Erwerbsstatus vor Arbeitsuche, Art der Beendigung der letzten Tätigkeit, Bereitschaft zu bundesweiter Stellensuche, BA-Kundengruppe, Arbeitszeit des Stellengesuchs, Restanspruchsdauer Arbeitslosengeld</p> <p><i>Betrieb:</i> Wirtschaftszweig, Anzahl Hauptbeschäftigte, Gründungsjahr</p>
Untersuchungseinheit	Sozialversicherungspflichtig Beschäftigte (ab 1999 auch geringfügig Beschäftigte), Leistungsempfänger, Arbeitssuchende, Teilnehmer an arbeitsmarktpolitischen Maßnahmen
Fallzahlen	1.487.835 Personen 23.667.828 Originalspells, 30.693.764 überschneidungsfreie Spells
Zeitraum	Der Zeitraum hängt von der Datenquelle ab. <sup>14</sup> Beschäftigung 1/1990-12/2007 (BeH) Leistungsempfang 1/1990-10/2008 (LeH) Maßnahmeteilnahme 1/2000-09/2008 (MTH) Arbeitsuche 1/2000-09/2008 (BewA)
Zeitbezug	kontinuierliche Erwerbsgeschichte, tagesgenau
Regionale Gliederung	Arbeitsort: West-/Ostdeutschland, Bundesländer, Regionaldirektionen, Wohnort: West-/Ostdeutschland, Bundesländer, Regionaldirektionen (nicht für alle Datenquellen und Zeiträume) weitere Merkmale siehe unter „sensible Merkmale“
Gebietsstand	BeH: 31.12.2007 (gebietsstandsaktualisiert) LeH, MTH, BewA: Originalgebietsstand; keine Bereinigung
<b>Methodische Charakteristika</b>	
Erhebungsdesign	ca. 2,2 %ige Zufallsauswahl der Personen aus den IEB
An der Erhebung beteiligte Institutionen	Sozialversicherungsträger, Bundesagentur für Arbeit
Frequenz der Datensammlung	laufend, Aktualisierung der Datenbasis (IEB): alle 9 Monate
Dateiformat (-größe)	STATA (2,9 GB), SPSS und SAS auf Anfrage
Datenzugang	Gastaufenthalt am FDZ der BA im IAB

<sup>14</sup> Eine genauere Beschreibung der abgedeckten Zeiträume findet sich in Abschnitt 2.2.



## 2.2 Mengengerüst und Füllgrad im Zeitablauf

Fallzahlen der IEBS	vor Splitting	nach Splitting
BeH	16.817.514	19.152.561
LeH	2.869.572	4.420.896
MTH	569.130	1.029.709
davon ABM, SAM, usw.	145.830	331.056
FBW, TM, DSL	298.302	457.309
FF	42.749	77.450
ESF-BA	12.584	26.138
PSA	3.575	7.882
JUMP	18.894	43.302
JUMP-PLUS	2.917	5.785
Beauftr. Dritter	44.279	80.787
BewA	3.503.691	6.201.840
Spells insgesamt	23.759.907	30.805.006
Personen	1.487.835	

Anmerkung: Zur Verteilung der Spells siehe Variable „Erwerbsstatus: Personengruppe, Leistungsart, Maßnahmentyp, Arbeitsuche-Status“ in Abschnitt 4.5.2.

Bei der Erstellung der IEBS wird ein Episodensplitting durchgeführt. Dabei werden bei Überschneidungen von Zeiträumen innerhalb eines Kontos die Spells so geschnitten und dupliziert, dass nur vollständig parallele und somit überschneidungsfreie Spells entstehen. Die Anzahl der Spells erhöht sich dadurch. Eine ausführliche Beschreibung des Episodensplittings erfolgt in Abschnitt 3.6.

Die Häufigkeiten der Spells der vier Datenquellen finden sich in den Auszählungen zu den Merkmalen zum Spellanfang und -ende in Abschnitt 4.3.4. Dort sind Brüche des Füllgrades im Zeitablauf markiert, die im Folgenden erläutert werden:

**BeH:** Aufgrund der Einführung des Meldeverfahrens in den neuen Bundesländern kann bei BeH-Spells erst ab 1993 von einer hinreichenden Vollständigkeit der Meldungen ausgegangen werden. Der Anstieg der Zahl der BeH-Spells ab 1999 folgt aus der Meldepflicht für geringfügig Beschäftigte ab 01.04.1999.

**LeH:** Beginnendaten vor 1990 erklären sich dadurch, dass alle Spells enthalten sind, die 1990 oder später enden. Die wenigen Enddaten vor 1990 entstehen durch spätere Bereinigungen. Für die neuen Bundesländer sind die Spells erst ab 1992 vollständig. Ist das Ende der 14.10.2008 oder später, handelt es sich i.d.R. um das Ende des Bewilligungszeitraums, nicht unbedingt um das des Leistungsbezugs.

**MTH:** Maßnahme-Spells sind ab Beginndatum 2000 vollständig. Erfasst werden alle Spells, die bei der Statistik der BA ab 2000 eingehen und deren Enddatum 1995 oder später ist. Enddaten vor 2000 entstehen, wenn Agenturmitarbeiter Maßnahmeteilnahmen verzögert eingeben oder korrigieren und durch Eingabefehler. Bei Enddaten nach 2008 handelt es sich um Planungsdaten.

**BewA:** BewA-Spells sind ab Enddatum 2000 enthalten. Ältere Spells sind nur für Arbeitsuchende enthalten, die ab 2000 und vorher unter derselben Kundennummer geführt wurden. Sehr frühe Spells (vor 1990) und Enddaten nach 2008 dürften jedoch vorwiegend Tippfehler sein.

### 2.3 Unterschiede zu IABS und BA-Beschäftigtenpanel

Es stehen verschiedene Stichproben der BA-Personendaten zur Verfügung, die sich für unterschiedliche Analysezwecke eignen.<sup>15</sup> In der folgenden Tabelle wird die IEBS mit den zum Teil auf denselben Datenquellen beruhenden Datensätzen IABS und BA-Beschäftigtenpanel verglichen.

Charakteristika	IEBS 2008	IABS-R75-04	BA-Beschäftigtenpanel
Zeitraum nach Spelltyp:			
Beschäftigung	1990-2007	1975-2004	1998-2006
Leistungsempfang	1990-2008	1975-2004	
Maßnahmeteilnahme	2000-2008		
Arbeitsuche	2000-2008		
Erhebungsdesign	2%-Stichprobe	2%-Stichprobe	2%-Stichprobe
Zeitbezug	überschneidungsfreie Spells	überschneidungsfreie Spells	Stichtagpanel, Quartalsweise, 36 Wellen
spätere Korrekturen	werden eingearbeitet	werden eingearbeitet	werden nicht eingearbeitet (Meldungen werden z.T. fortgeschrieben)
Bereinigungen von Lücken und Inkonsistenzen	Bereinigungen nur in eindeutigen Fällen	weitergehende Bereinigungen zur Erleichterung der Auswertung wie z.B. die Erzeugung künstlicher Beschäftigungsspells	wenige Bereinigungen von Inkonsistenzen, fortgeschriebene Meldungen bleiben erhalten
Besonderheiten	Spells zu Maßnahmeteilnahme und Arbeitsuche	Merkmale zur Beschäftigtenstruktur der Betriebe, großer Nutzerkreis, gute Kenntnis der Datenqualität	Merkmale zur Beschäftigtenstruktur der Betriebe, Aggregationen entsprechen den Werten der BA-Statistik
Datenzugang	Gastaufenthalt	Gastaufenthalt, SUF	Gastaufenthalt, Datenfernverarbeitung, SUF

<sup>15</sup> Einen aktuellen Überblick über das Datenangebot des FDZ der BA im IAB erhalten Sie unter <http://fdz.iab.de>.

## 3 Datenaufbereitung

### 3.1 Datenquellen

#### 3.1.1 Beschäftigten-Historik des IAB (BeH)

Die BeH enthält die Entgeltmeldungen der Betriebe an die Sozialversicherungsträger nach der Datenerfassungs- und -übermittlungsverordnung (DEÜV).<sup>16</sup> Die Entgeltmeldungen bestehen aus Jahres-, Ab-, Unterbrechungs- und Änderungsmeldungen für jede sozialversicherungspflichtige Beschäftigung und ab 01.04.1999 auch für geringfügige Beschäftigungen.

Für die IEBS wird aus der BeH der Zeitraum 1990 bis 2007 verwendet. Ab 1999 geht in die BeH das so genannte Jahres-Zeitraum-Material (JZM) der Statistik ein.

Das verwendete Jahres-Zeitraum-Material (JZM) der Statistik unterscheidet sich in den Jahresscheiben hinsichtlich der berücksichtigten Zeiträume von Nachmeldungen:

- 1999-2003: 36-Monatsversion
- 2004-2005: 18-Monatsversion
- 2006: 12-Monatsversion (offiziell nicht freigegeben)
- 2007: 6-Monatsdatei (offiziell nicht freigegeben)

Für 2007 ist von einem Füllgrad von knapp 90% auszugehen. Diese Daten haben bereits verschiedene Bereinigungsverfahren durchlaufen. Dazu siehe Abschnitt 3.3.

#### 3.1.2 Leistungsempfänger-Historik des IAB (LeH)

Für die IEBS wird aus der LeH der Zeitraum 1990 bis zum aktuellen Rand (14.10.2008) verwendet. Die LeH enthält die Abmeldungen, die die Krankenkassen nach Zeiten des Leistungsbezugs von der BA erhalten, sowie Meldungen zu laufenden Bezügen am aktuellen Rand. In früheren Versionen stammten die Daten ausschließlich aus dem Altverfahren coLei und wurden am aktuellen Rand aus der Zahlungsnachweisdatei (ZND) ergänzt. In der aktuellen LeH stammen die Daten ab 2004 aus dem Data Warehouse (DWH), wo Daten aus dem Nachfolgeverfahren COLIBRI bis zum aktuellen Rand tagesgenau erfasst sind.

---

<sup>16</sup> Die DEÜV wird umfassend und verständlich von verschiedenen Krankenkassen dokumentiert, siehe z.B. Ratgeber zur Sozialversicherung 2009, Broschüre des Deutsche BKK Arbeitgeberservice, [http://www1.deutschebkk.de/Service/Arbeitgeberservice/Wissen\\_von\\_A-Z/5B2D7F649D116FBDC1256DC90\\_0510B3E/RatgeberSV\\_2009\\_Internet.pdf](http://www1.deutschebkk.de/Service/Arbeitgeberservice/Wissen_von_A-Z/5B2D7F649D116FBDC1256DC90_0510B3E/RatgeberSV_2009_Internet.pdf), S.32 (Link vom 04.05.2009).

Der Übergang vom Verfahren coLei zu COLIBRI fand schrittweise von 2004 bis 2006 statt. Seit dem 01.07.2004 wurden die Daten in der Arbeitsagentur Gotha mit COLIBRI erfasst. Ab Dezember 2005 wurden die restlichen Agenturen umgestellt. Es erfolgte keine komplette Migration der Daten sondern die Leistungssätze wurden Satz für Satz vom Sachbearbeiter aus dem Altsystem abgemeldet und ins neue System übernommen. Bis 31.12.2006 war die Umstellung erfolgt. Die seit dem 01.01.2005 eingeführten SGB II-Leistungen werden in COLIBRI nicht erfasst. Im Jahr 2004 konnten jedoch noch ALHI- und UHG-Leistungen, die zum 01.01.2005 abgeschafft wurden, in COLIBRI erfasst werden. Die zeitlich uneinheitliche Umstellung der Verfahren wirkt sich daher bei Auswertungen der LeH auf regionaler Ebene aus.

### 3.1.3 Maßnahme-Teilnehmer-Gesamtdatenbank (MTH)

Die MTH des IAB enthält Daten zu Maßnahmen der aktiven Arbeitsmarktförderung in Kontenform. Aus einer komplexen Datenbank, die historisiert Informationen von allen statistisch erfassten Bewegungen enthält, werden mehrere Dateien erzeugt:

ABM	Maßnahmen der aktiven Arbeitsmarktpolitik (ABM/SAM, aber nicht einmalige Leistungen wie z. B. MOBI-Fahrtkosten)
FF	Freie Förderung
FbW	Maßnahmen der beruflichen Weiterbildung, Trainingsmaßnahmen (TM) und Deutschlehrgänge (DSL)
ESF-BA	Maßnahmen des ESF-BA-Programmes (ohne Leistungsart Sozialpädagogische Betreuung)
PSA	Personal-Service-Agenturen
JG	Jugendsofortprogramm
AfL/JP	Arbeit für Langzeitarbeitslose/Jump Plus
P37	Beauftragung Dritter mit der Vermittlung nach § 37 SGB III
AGH	Arbeitsgelegenheiten
ESG	Einstiegsgeld nach § 29 SGB II
SWL	Sonstige weitere Leistungen nach § 16 Abs. 2 Satz 1 SGB II

Diese Dateien, in denen jede Maßnahmeteilnahme mit zusätzlichen Merkmalen aus dem Bewerberangebot BewA verknüpft ist (z.B. dem Schwerbehindertenstatus), bilden den MTH-Anteil der IEBS.

Die aus der MTH verwendeten Spells sind für den Zeitraum von 1/2000 bis 09/2008 vollständig, teilweise sind auch Maßnahmeteilnahmen vor 2000 enthalten. Maßnahmeteilnahmen mit Eintrittsdatum vor 1995 wurden generell ausgeschlossen. Für Maßnahmeeintritte am aktuellen Rand ist von einem geringeren Füllgrad auszugehen. Aufgrund der veränderten Datenquelle (IAB-Forschungsdatenbank ISAAK statt MTG) ist davon auszugehen, dass der Füllgrad für Maßnahmeteilnahmen in 2000 mit Beginndatum vor 2000 wesentlich größer ist als in der IEBS 1.0.

Da in den Quellverfahren zu einer Maßnahmeteilnahme mehrere, z. T. sehr viele Datensätze geliefert werden und in den Daten eindeutige Kriterien dafür fehlen, ob es sich um einzelne oder mehrfache Teilnahmen handelt, wurde für die Forschungsdatenbank des IAB ein fachlich abge-

stimmtes Regelsystem für die Identifizierung einer Teilnahme entwickelt. Weitere Informationen zu diesem Regelsystem sind auf Anfrage erhältlich.

#### 3.1.4 Arbeitssuchendenstatus aus dem Bewerberangebot (BewA)

Die Daten zum Arbeitssuchendenstatus werden ursprünglich aus der DWH-Schicht des Bewerberangebots (BewA) extrahiert. In dieser Datenbank werden Arbeitssuchendenstatus und Bewerberangebote unter verschiedenen Identifikatoren geführt; der Arbeitssuchendenstatus unter der Kundennummer, das Bewerberangebot unter der BewA-Eingabenummer. In der Forschungsdatenbank des IAB werden alle Bewerberangebote den entsprechenden Kundennummern zugeordnet. Auch in dieser Datenbank liegen die Daten historisiert in Kontenform vor. Für die IEB werden die Daten der Arbeitssuche (ASU) aus dem ASU-Auswertefile der ISAAK-Anwenderschicht geholt, in dem Arbeitssuchenden-Spells mit Kundendaten und Daten aus dem Bewerberangebot verknüpft sind.

Die Daten liegen für den Zeitraum von 1/2000 bis 09/2008 vollständig vor. Sofern Arbeitssuchende vor dem Jahr 2000 unter derselben Kundennummer geführt wurden, sind auch diese Daten enthalten.

### 3.2 Verknüpfung der Daten

Die vier Datenquellen werden über die Sozialversicherungsnummern und BA-Kundennummern verknüpft. Dabei ergeben sich verschiedene Probleme. Eine Vielzahl der Beschäftigten kommt nie mit der BA in Kontakt und hat entsprechend keine Kundennummer. Manche Beschäftigte haben zeitweise mehr als eine Sozialversicherungsnummer. Dies wird durch den Verband deutscher Rentenversicherungsträger (VDR) korrigiert, sobald es entdeckt wird. Der VDR erstellt eine Korrekturdatei, die in die BA- und IAB-Daten eingearbeitet wird, indem eine Versicherungsnummer auf alle Spells der Person übertragen wird. Für Leistungsempfänger liegen im Allgemeinen sowohl die Sozialversicherungsnummer, als auch eine – aber häufig auch mehrere – Kundennummern vor. Erst seit 1999 wird die Kundennummer eindeutig vergeben, wobei die Vergabe seit 2003 bundesweit durch das Verfahren der zentralen Personendatenverwaltung (zPDV) koordiniert wird. MTH und BewA enthalten häufig nur Kundennummern.

Für die IEBS gelten folgende Verknüpfungsregeln:

Aus den Kundeninformationen im DWH der BA wird die Relation KNR-VSNR erstellt, die für jede Kundennummer (KNR) genau eine Versicherungsnummer (VSNR) enthält. Bei der Erstellung dieser Relation werden folgende Regeln beachtet:

- Zu jeder Kundennummer wird die „aktuellste“ Versicherungsnummer geholt,
-

- zu jeder Kundennummer wird das Geburtsdatum und Geschlecht aus den Kundeninformationen des DWH geschrieben und zu jeder Versicherungsnummer das Geburtsdatum und Geschlecht aus den Versicherteninformationen des DWH,
- die Versicherungsnummer wird dann verworfen, wenn sich Vorname **und** Zuname **und** Geburtsdatum in den Versichertendaten und Kundendaten unterscheiden,
- anschließend wird die VDR-Datei mit Stilllegungen von Versicherungsnummern eingearbeitet,
- falls durch die Einarbeitung die Versicherungsnummer geändert wurde, müssen Geburtsdatum und Geschlecht neu berechnet werden.

Für Sätze aus den Quellen LeH, LHG, MTH und ASU werden Versicherungsnummer, Geburtsdatum und Geschlecht aus dieser Relation (KNR-VSNR) zugespielt, wenn die Kundennummer gefüllt ist.

LeH-Daten sind ursprünglich nach BA-Kundennummern organisiert. Bei einem sehr hohen Prozentsatz der Daten (99,7 %) ist jedoch auch die Versicherungsnummer enthalten. Soweit vorhanden, wird bei LeH-Spells die Relation aus Kunden- und Versicherungsnummer für Spells mit einem Enddatum ab 01.01.1999 verwendet. Für Spells vor 1999 wird sie nur dann verwendet, wenn die dort auftretende Kombination von Kundennummer und Versicherungsnummer auch in einem Datensatz ab 1999 vorkommt und für diese Kundennummer in der Relation eine Versicherungsnummer gefunden wird. Falls über die Relation in den Spells der LeH ab 1999 keine Eindeutigkeit hergestellt werden kann, d.h. wenn die Kundennummer in der Relation nicht enthalten ist und wenn es in der LeH zur Kundennummer mehrere Sozialversicherungsnummern gibt, wird die Versicherungsnummer in den Konten gelöscht.

Bei MTH- und ASU-Daten wird eine Person in ISAAK ursprünglich über die BA-Kundennummer (ca. 80 %) bzw. die BewA-Eingabenummer (ca. 20 %) identifiziert. Ein Kunde kann jedoch z. B. für die Zeit vor der Einführung der zentralen Personendatenverwaltung (zPDV) im Jahr 2003 nach einem Umzug mehrere verschiedene Kundennummern haben. Bei einem hohen Prozentsatz der Daten (MTH: über 90 %; ASU: über 80 %) kann über die Kundennummer oder die BA-Kunden-Daten eine Versicherungsnummer zugeordnet werden, bei sehr wenigen Datensätzen (MTH: unter 1 %, ASU: ca. 7 %) sind weder VSNR noch KNR vorhanden.

Die Qualität der Zuordnung von Versicherungsnummern zu Kundennummern wurde bei Einführung der Relation getestet:

Aus ca. 900.000 Kunden wurden diejenigen selektiert, die dieselbe Versicherungsnummer, jedoch unterschiedliche Vor- oder Nachnamen haben. 557 derartige Datensätze wurden gefunden.

Von 29.465.474 Zuordnungen von Versicherten zu Kunden, wurde bei ca. 8.200 Fällen festgestellt, dass es sich vermutlich um unterschiedliche Personen handelt. Diese werden bei Erstellung

der Relation durch die Bedingung "Vorname oder Zuname oder Geburtsdatum in den Versicherten- und den Kundendaten müssen identisch sein" ausgeschlossen.

Insgesamt kann festgestellt werden, dass die Relation Kunde–Versicherter im DWH eher zuviel zusammenfasst, aber von recht guter Qualität ist.

### 3.3 Bereinigungen und Filter

#### 3.3.1 BeH

Die BeH-Spells haben folgende Bereinigungs-schritte durchlaufen:

- Geringfügige Datumskorrekturen wurden durchgeführt.
- Im JZM ab 1999 werden (von der BA-Statistik) nur die Personengruppen 101, 102, 103, 105, 106, 112, 118, 119, 120, 140, 141, 142, 143, 201, 203 und 205 geliefert, sowie die beiden Gruppen 109 und 209, die geringfügig Beschäftigte kennzeichnen. Nicht geliefert werden u.a. ab 1999 kurzfristig Beschäftigte, d.h. Personengruppen 110, 202 und 210. Die BeH vor 1999 enthält alle Personengruppen, auch solche, die ab 1999 mit dem JZM nicht mehr geliefert werden.
- Bei der Unterscheidung zwischen sozialversicherungspflichtig und geringfügig Beschäftigten werden die Personengruppen 109, 110, 202, 209 und 210 als "geringfügig" eingestuft.
- Anmeldungen werden ausgefiltert.
- Storno- und Änderungsmeldungen werden eingearbeitet, Stornosätze werden gelöscht.
- Informationen zur Betriebsnummer (z.B. Arbeitsort, Wirtschaftszweig) werden ergänzt.
- Duplikate werden nach folgenden Kriterien entfernt: Versicherungsnummer, Betriebsnummer, Beginn- und Enddatum, Abgabegrund, Entgelt und Erstelldatum.
- Aus der BeH werden für die IEB nur Datensätze für den Zeitraum 1990 bis 2007 übernommen.
- Versicherungsnummern werden zum Stand vom 31.12.2006 aktualisiert, stillgelegte werden gelöscht.
- Geschlecht und Geburtsdatum werden aus der Versicherungsnummer übernommen.
- Der Gebietsstand für Arbeitsort und Wohnort wird zum Stand vom 31.12.2007 aktualisiert.

#### 3.3.2 LeH

Für LeH-Spells werden folgende Bereinigungs-schritte und Filterbedingungen ausgeführt, wobei zwischen Basisbereinigungen und sonstigen Bereinigungen unterschieden werden kann.

Unter die Basisbereinigungen fallen folgende Arbeitsschritte:

- Ausfiltern von Kranken-/Pflegeversicherungsbeiträgen (§207 A SGB III)
-



- Datumskorrekturen
- Zuspielen der Versicherungsnummer über die Kundennummer aus der Relation Kundennummer–Versicherungsnummer (mit VDR-Bereinigung der Stilllegungen). Datensätze der LeH vor 1999 ohne Versicherungsnummer werden ausgeschlossen, da hier nicht mit Sicherheit davon ausgegangen werden kann, dass es sich um eine korrekte Kundennummer handelt.

Im Folgenden werden diese Basisbereinigungen, sowie sonstige Bereinigungen im Einzelnen aufgeführt:

- Aus der LeH werden für die IEB nur Spells für den Zeitraum ab 1.1.1990 (Enddatum) bis zum aktuellen Rand 11.09.2008 übernommen.
- Zu beachten ist, dass bei noch laufendem Bezug das Enddatum dem Ende des Bewilligungszeitraums entspricht und nicht dem Ende des Bezugszeitraums. Spells mit leerem oder ungültigem Enddatum werden von der weiteren Verarbeitung ausgeschlossen, denn bei Datensätzen ohne Enddatum kann laut Auskunft der Fachseite nicht davon ausgegangen werden, dass überhaupt eine Zahlung geleistet wurde.
- Spells ohne gültiges Beginndatum werden von der weiteren Verarbeitung ausgeschlossen.
- Spells mit einem Enddatum, das vor dem Beginndatum liegt, werden ausgeschlossen.
- Liegt das Enddatum für die Arbeitslosenhilfe einen Tag vor dem Beginndatum, so wird das Enddatum um ein Jahr verschoben.
- Es werden nur Spells mit den gruppierten Leistungsarten 1 (ALG), 2 (ALHI), 3 (UHG) übernommen. Die Werte 5 (§166b AFG / §207a SGB III) und 9 (Fehlkodierung) werden ausgefiltert.
- Es werden nur formal korrekte Versicherungsnummern übernommen (laut Prüfungs-Kennzeichen "nicht gelöscht" und "korrekt").
- Versicherungsnummern werden zum Stand vom 30.06.2006 aktualisiert, stillgelegte werden nicht übernommen.
- Eine Gebietsstandsaktualisierung wird nicht durchgeführt, da nur die Arbeitsagentur des Wohnorts, aber nicht die dazu benötigte Gemeinde zur Verfügung steht.

- **Bereinigung sich überschneidender LeH-Spells**

Sich überschneidende LeH-Spells werden für die IEB entsprechend der folgenden Regeln bereinigt:

- Alle Überschneidungen bei denen in beiden Spells **dieselbe Leistungsart** und **dieselbe Leistungshöhe** steht, werden wie folgt bereinigt:

- Wenn ein Spell vollständig innerhalb eines anderen liegt, wird der kürzere Spell gelöscht. Bei gleicher Länge wird ein (zufällig ausgewählter) Satz gelöscht.

- Bei Teilüberschneidungen wird das Enddatum des früher beginnenden Spells korrigiert.

---



- Überschneidungen zwischen Spells mit **unterschiedlicher Leistungsart** und/oder unterschiedlicher Leistungshöhe werden nur dann bereinigt, wenn es sich um eine Überschneidung von genau einem Tag handelt und die Originaldauer beider Spells größer als ein Tag ist. In diesen Fällen wird das Enddatum des früher beginnenden Spells um einen Tag zurückgesetzt.

### 3.3.3 MTH

Für die IEBS werden die MTH-Daten aus der ISAAK-Anwenderschicht, bestehend aus Auswertefiles für die Teilverfahren ABM/SAM, FbW, FF, ESF-BA, PSA, JG, AFL//JP, P37, AGH, ESG und SWL verwendet. Dabei werden die folgenden Bereinigungs-schritte und Filterbedingungen ausgeführt:

- Die aus der MTH verwendeten Daten umfassen einen Zeitraum von 01/1995 bis 09/2007. Datensätze mit einem Beginndatum vor 1995 wurden gelöscht. Vollständig enthalten sind die Maßnahmeteilnahmen jedoch erst ab 01/2000 (Beginndatum der Maßnahme).
  - Die Versicherungsnummer wird soweit die Kundennummer vorhanden ist, aus der Relation KNR-VSNR geholt. Ist die Kundennummer nicht vorhanden, wird sie den BA-Kunden-Daten (HIST-ISAAK-Person) entnommen.
  - Stillgelegte, sowie geänderte Versicherungsnummern werden aus der VDR-Datei zum Stand vom 30.06.2006 aktualisiert. Sätze mit stillgelegter Versicherungsnummer bleiben erhalten, wenn eine Kundennummer vorhanden ist; die Versicherungsnummer wird auf 'missing' gesetzt.
  - Geschlecht und Geburtsdatum werden – soweit möglich – aus der Versicherungsnummer übernommen.
  - Eine Gebietsstands-bereinigung wird nicht durchgeführt.
  - Der Regionaltyp06 wird über die aktuelle Relation AA-Regiotyp des DWH generiert.
  - Für die Bereitstellung der ESF-Daten aus MTH für IEB werden alle sozialpädagogischen Maßnahmen ausgefiltert, da es sich hierbei nicht um Maßnahmen, sondern um Leistungen handelt und in den Daten nur ein Beginndatum, jedoch kein Zeitraum enthalten ist.
  - Bei den Daten aus dem Verfahren ABM werden alle einmaligen Leistungen (MOBI und UBV) ausgeschlossen.
  - Datensätze werden ausgefiltert, wenn Versicherungsnummer und Kundennummer leer sind, denn damit ist keine Identifikation und Konto-zuordnung möglich.
  - Sätze, deren Enddatum kleiner als das Beginndatum ist, werden ausgefiltert.
  - Konten mit mehr als 99 Spells werden nicht in die IEB übernommen.
-

### 3.3.4 BewA

Auf die Daten aus dem ASU-Auswertefile der ISAAK-Anwenderschicht werden für die IEB die folgenden Bereinigungs-schritte und Filtersetzungen ausgeführt:

- Die aus ASU verwendeten Daten umfassen den Zeitraum von 01/1997 bis 09/2007. Ab 1999 sind die Daten vollständig enthalten.
- Die Versicherungsnummer wird, soweit die Kundennummer vorhanden ist, aus der Relation KNR–VSNR geholt. Ist die Kundennummer nicht vorhanden, wird sie aus den BA-Kunden-Daten (HIST-ISAAK-Person) entnommen.
- Stillgelegte, sowie geänderte Versicherungsnummern werden aus der VDR-Datei zum Stand vom 30.06.2006 bereinigt. Sätze mit stillgelegter Versicherungsnummer bleiben erhalten, wenn die Kundennummer vorhanden ist. Die Versicherungsnummer wird in diesen Fällen auf 'missing' gesetzt.
- Geschlecht und Geburtsdatum werden – soweit möglich – aus der Versicherungsnummer übernommen.
- Eine Gebietsstandsbereinigung wird nicht durchgeführt.
- Bildung des Erwerbsstatus: In der Datenbasis gibt es nur die Unterscheidung zwischen Sätzen mit Status "arbeitslos" und "arbeitsuchend". In der IEB wird unterschieden zwischen "arbeitslos", "krank" und "arbeitsuchend". Sätze mit Erwerbsstatus "krank" sind Datensätze, bei denen (1) ein Alo-Vorgängersatz existiert, der lückenlos anschließt und den Abgangsgrund "krank" (sna = 6004) hat, (2) ein Alo-Nachfolgesatz existiert, der ebenfalls lückenlos anschließt und (3) der Satz selbst nicht den Status "arbeitslos" sondern den Status "arbeitsuchend" hat.
- Der Regionaltyp06 wird über die aktuelle Relation AA-Regiotyp des DWH generiert.
- Datensätze werden ausgefiltert, wenn Versicherungsnummer und Kundennummer leer sind, weil damit keine Identifikation und Kontobildung möglich ist.
- Sätze mit einem Enddatum, das kleiner als das Beginndatum ist, werden ausgefiltert.
- Die Kontenbildung erfolgt nach Versicherungsnummer und Kundennummer (VSNR hat Vorrang).

### 3.3.5 Gesamtdatei

Die Gesamtdatei IEB wird aus den Datenquellen für Beschäftigte (BeH), Leistungsempfänger SGB III (LeH), Leistungsempfänger SGB II (LHG), Maßnahmeteilnehmer (MTH) und Arbeitssuchende (ASU/BewA) erstellt. Dabei werden folgende Aufbereitungen durchgeführt:

- Kontenbildung über Versicherungsnummer bzw. Kundennummer, wobei vorrangig immer die Versicherungsnummer verwendet wird.

- Vergabe einer technischen Konto-ID als Identifikator. Dabei bleiben ab der IEB V5.00 die in der Vorversion vergebenen Konto-IDs mit ihrer Zuordnung zu Versicherungsnummer bzw. Kundennummer als kontenbildende Schlüssel erhalten.
- Datensätze ohne Kundennummer und ohne Versicherungsnummer werden gelöscht.
- Datensätze, bei denen das Alter kleiner 13 oder größer 75 ist, werden gelöscht.
- Datensätze, bei denen das Enddatum kleiner als das Beginndatum ist, werden gelöscht.
- Konten mit einem Wechsel von Geschlecht oder Geburtsdatum werden gelöscht.

### 3.4 Rechtlich unzulässige zeitgleiche Zustände

In der IEB werden Daten aus verschiedenen operativen Systemen zusammengeführt, die untereinander nicht abgeglichen werden. Die Bereinigung rechtlich unzulässiger paralleler Zustände erfordert umfassende Kenntnisse der rechtlichen Regelungen. Eine sehr hilfreiche Übersicht zulässiger und unzulässiger Zustände bieten Fitzenberger et al. 2005b. Die Ergebnisse eines Projektes zur Überprüfung inkonsistenter Zustände ist in Jaenichen et al. 2005 beschrieben. Hier sei eine Auswahl einführender Hinweise gegeben:

- Es können Unstimmigkeiten durch Falscheingaben vorkommen. Wenn Kundennummern falsch eingegeben wurden, kann es auch vorkommen, dass Konten nicht korrekt zusammengeführt wurden und deshalb eigentlich zu erwartende parallelliegende Sätze fehlen.
  - Darüber hinaus kann es vorkommen, dass für Maßnahmeteilnahmen bei Abbruch einer Maßnahme kein aktualisierter Satz erfasst wird, sodass in den Daten bereits während der Maßnahmeteilnahme wieder parallel ALG-Bezug auftaucht.
  - Es kann vorkommen, dass die Daten in den verschiedenen operativen Systemen nicht am selben Tag, sondern mit einigen Tagen Unterschied erfasst werden, sodass z.B. eine Abmeldung aus Arbeitslosigkeit erst einige Tage nach Maßnahmebeginn erfolgt.
  - Zu ALG-Sätzen aus der LeH gibt es ab 1999 parallele ASU-Sätze.
  - Zu BewA-Sätzen muss es keine parallelen LeH-Sätze geben, wenn der Arbeitsuchende keinen Anspruch auf Leistungen hat.
  - Auch zu BewA-Sätzen mit Arbeitslosigkeits-Status oder LeH-Sätzen kann es parallel BeH-Sätze geben.
  - Bei reinen FbW-Maßnahmen sollen die Teilnehmer nicht gleichzeitig arbeitsuchend gemeldet sein. Es muss also keine parallelen BewA-Sätze für den gesamten Zeitraum geben. Erst kurz vor Ende der Maßnahme melden sich die Teilnehmer wieder arbeitsuchend. Bis zum Ende der Maßnahme kann es dann parallele BewA-Sätze geben.
  - FbW, TM und DSL haben in der Regel parallele Spells in LeH mit Unterhaltsgeld (UHG). Aufgrund einer Gesetzesänderung werden ab dem 01.01.2004 die Teilnehmer an Trai-
-

ningsmaßnahmen nicht mehr als arbeitslos gezählt. Ob und wie sich das in den Daten niederschlägt ist noch nicht geprüft. Ausnahme: Parallel zur FbW-Maßnahme hat der Teilnehmer eine ESF-BA-Maßnahme. In diesem Fall darf es keinen UHG-Spell in der LeH geben.

- Bei TM und DSL sind die Teilnehmer gleichzeitig arbeitsuchend und eventuell arbeitslos gemeldet.
- Bei ABM und SAM sollten die Teilnehmer arbeitsuchend aber nicht arbeitslos gemeldet sein.
- Bei ABM und SAM sollte es außerdem parallele Beschäftigtensätze geben.

### 3.5 Datenqualität und Probleme

Die Darstellung der Verknüpfungsproblematik, der bereits vorgenommenen Bereinigungen, sowie der Hinweis auf das Vorkommen rechtlich unzulässiger zeitgleicher Zustände in den vorangegangenen Abschnitten vermitteln ein Gefühl für die Datenqualität dieser prozessproduzierten Daten. Von Seiten des IAB werden nur Bereinigungen vorgenommen, deren allgemeiner Nutzen sicher scheint. Weitere bekannte Qualitätsprobleme sind:

- Unstimmigkeiten entstehen durch Falscheingaben. Wenn Kundennummern falsch eingegeben wurden, werden Konten nicht korrekt zusammengeführt und deshalb fehlen zu erwartende parallelliegende Spells.
- Es kommt vor, dass bei Abbruch einer Maßnahme kein aktualisierter Spell erfasst wird, der den Abbruch beschreibt, sodass in den Daten bereits während der Maßnahmeteilnahme wieder parallel Leistungsbezug auftaucht.

Die Nutzer der Daten sind aufgerufen, die Liste der Inkonsistenzen und Bereinigungsansätze zu ergänzen.

Im Gegensatz zu den von der BA veröffentlichten Statistikdaten und dem BA-Beschäftigtenpanel sind in der IEBS verspätete Erfassungen und nachträgliche Korrekturen von Spells berücksichtigt, die bei den zeitnah veröffentlichten Statistikdaten nicht berücksichtigt werden können. Aus diesem Grund ist es nicht möglich, mit Hilfe der IEBS die Zahlen aus den Statistikdaten der BA zu reproduzieren. Insbesondere bei der Neueinführung von Maßnahmen ist aufgrund von verspäteten Erfassungen mit erheblichen Diskrepanzen im Vergleich zu den von der Statistik der BA veröffentlichten Zahlen zu rechnen.

### 3.5.1 BeH

#### Geringfügig Beschäftigte

Bei den BeH-Datensätzen besteht eine Besonderheit bezüglich der geringfügig Beschäftigten. Die Meldepflicht für geringfügig Beschäftigte besteht seit 01.04.1999 (genauer: die Übermittlung dieser Daten an die BA). Vorher sind keine Daten über geringfügig Beschäftigte enthalten. In den Daten sieht das dann so aus, als ob die Beschäftigungen am 01.04.1999 beginnen würden (`ieb_beg_orig = 1.4.1999`). Vermutlich handelt es sich bei einem nicht geringen Anteil der Fälle, die als Beginndatum den 01.04.1999 haben, um Beschäftigungen, die schon länger (auch schon im Vorjahr) bestanden und als Beginndatum eigentlich den 01.01.1999 haben müssten. Durch die Meldepflicht für geringfügige Beschäftigung steigt die Gesamtzahl der BeH-Datensätze 1999 erheblich an. Außerdem gibt es ab dem 01.04.1999 mehr Überschneidungen – sowohl innerhalb von BeH-Sätzen, als auch mit anderen Datenquellen –, da z. B. geringfügige Beschäftigung und Leistungsbezug gleichzeitig möglich sind.

#### Füllgrad der Bildungsvariable

Auswertungen zeigen, dass die Anzahl der fehlenden Werte bei der Bildungsvariable mit der Zeit zunimmt. Deutlich wird dies bei Auszählungen der Erstmeldungen nach Jahren. Am aktuellen Rand erreicht der Anteil der fehlenden Werte für die Quelle BeH über 30%. Weitere Auswertungen lassen vermuten, dass auch Betriebsgrößeneffekte eine Rolle spielen (höherer Anteil fehlender Werte bei Kleinbetrieben).

### 3.5.2 LeH

#### Einführung des Fachverfahrens Colibri

In den Jahren 2004 (Pilotagentur Gotha) und 2005 wurde schrittweise ein neues Fachverfahren zur Erfassung der Leistungsempfangszeiten eingeführt. Ein besonderes Problem, das sich in den Daten wiederfinden kann, liegt in der Übergangsphase, d. h. in den Zeiträumen bei denen Datensätze in der LeH und im DWH vorhanden sind (01.01.2003 bis 31.12.2005). Aufgrund der Umstellung vom Alt- auf das Neuverfahren und der sich überlappenden Erfassungszeiträume in alter LeH und DWH kann die Datenübernahme in die neue LeH nicht nach festen Gültigkeitszeiträumen erfolgen. Dies kann zu einer Unterschätzung in der Anzahl der Leistungsempfangsmeldungen während des Übergangszeitraums führen.

### 3.5.3 MTH

#### Umstellung in der Datengenese (MTG / MTH)

Das Fachverfahren coSach wurde 2004 durch ein neues Verfahren abgelöst, das sich insbesondere durch eine verbesserte Anbindung an das DWH der BA kennzeichnet. Die Regeln, mit denen die MTG als Datengrundlage für die Maßnahmeteilnahmen aufgebaut wurde, können mit dieser Verfahrensumstellung nicht mehr beibehalten werden. Die MTG wird durch die MTH abgelöst. Die dort verwendeten Regeln und Prüfverfahren sorgen für eine Erhöhung der Teilnehmerzahlen in den Maßnahmen der aktiven Arbeitsmarktpolitik (vgl. auch Engelhardt et al., 2008).

#### Beginn- und Enddatum

Bei MTH-Datensätzen sind Eintritts- und Austrittsdatum ursprünglich von Sachbearbeitern erfasst und im Altverfahren coSach nicht per Programm auf Plausibilität geprüft worden. Es ist davon auszugehen, dass diese Datumsangaben bei einem geringen Prozentsatz der Daten fehlerhaft sind. Da hier in einzelnen Datensätzen unplausible Angaben, wie z. B. mit Beginn- und Enddatum am 01.01.1900 oder Beginndatum in 1920 auftreten, wurden alle Datensätze mit einem Eintrittsdatum vor dem 01.01.1995 generell ausgeschlossen.

Im Verfahren coSach NT, das seit 21.12.2004 für alle coSach Teilverfahren im Einsatz ist, werden, wo es möglich ist, bereits bei der Eingabe Plausibilitätsprüfungen durchgeführt.

#### Bekanntes Problem im Fall des Überbrückungsgeldes

Bei Maßnahmeteilnahmedatensätzen zu Überbrückungsgeld (ÜG) gibt es eine nicht geringe Anzahl von Fällen, bei denen die Dauer 2 Jahre beträgt, obwohl maximal ein halbes Jahr gewährt werden darf. Dies kommt dadurch zustande, dass in coSach ABM in der Eingabemaske der 01.01. des jeweiligen Jahres als Eintrittsdatum und der 31.12. des Folgejahres als Austrittsdatum voreingestellt sind.

### 3.5.4 ASU / BewA

#### Einführung des Fachverfahrens Verbis

Von Mitte 2005 bis Mitte 2006 wurde das Verfahren coArb, aus dem die Arbeitssuchenden- und Bewerber-Angebotsdaten stammen, durch das Verfahren VerBIS abgelöst. Ab Juli 2005 wurde das operative System coArb erstmals in der Agentur Wiesbaden als Pilotprojekt, ab Dezember 2005 dann sukzessive in mehreren Tranchen in allen AA durch VerBIS ersetzt. Für diese Umstellung findet in VerBIS für jede AA zum Zeitpunkt der Umstellung eine Umladung der bisherigen Daten aus coArb und deren Migration nach VerBIS statt. Die aus VerBIS stammenden Daten werden für das DWH aber (bis auf weiteres) wieder in das für coArb entwickelte Datenmodell „zurück-

migriert“. Dieses Mapping der VerBIS-Daten auf die bisherigen Merkmale im DWH führt zu unvollständigen ASU-Spells, die durch eine hohe Anzahl an Missings gekennzeichnet sind. Zeitweise kommt es sogar zu einem kompletten Ausfall von ASU-Spells.

#### Unvollständigkeit der ASU-Spells vor 2000

Die Daten liegen für den Zeitraum von 01/2000 bis 09/2007 vollständig vor, sofern Arbeitssuchende vor dem Jahr 2000 unter derselben Kundennummer geführt wurden. Die Vollständigkeit von Konten, die sich über die Zeit vor 2000 erstrecken, ist allerdings problematisch.

Die IEB-Datenquelle ASU wurde mit Bestandsdaten ab 1997 aufgebaut. Für diesen Zeitraum werden unterschiedliche Identifikatoren verwendet, sodass wegen Zuordnungsproblemen Lücken in den ASU-Meldungen je Person existieren können:

Konten, die ausschließlich ASU-Spells vor 2000 enthalten, sind hinsichtlich der ASU-Spells unvollständig, da diesen u. U. keine KNR zugewiesen wurde (ca. 1% aller Konten in der IEB mit ASU-Spells).

Konten, die Spells mit ASU-Zeiten vor 2000 beinhalten und zusätzlich Spells besitzen, die in den Zeitraum 01/2000 bis 08/2002 hineinragen oder darin liegen, sind hinsichtlich der ASU-Spells jedoch vollständig (ca. 1/3 aller Konten in der IEB mit ASU-Spells).

Konten, die ausschließlich ASU-Spells beinhalten, die nach 01/2000 beginnen, sind hinsichtlich der ASU-Spells ebenso vollständig (ca. 2/3 aller Konten in der IEB mit ASU-Spells).

Konten, die ASU-Spells vor 2000 und nach 08/2002, jedoch keine Spells im Zeitraum 01/2000 bis 08/2002 enthalten, sind für den Zeitraum vor 2000 nicht vollständig, da die Möglichkeit zur Verknüpfung mit Spells vor 2000 fehlt (weniger als 1 % aller Konten mit ASU-Spells).

Für Konten, die nur ASU-Spells nach 8/2002 enthalten, ist nicht bekannt, ob vor 2000 ASU-Spells vorlagen (ca. 30 % aller Konten mit ASU-Spells).

#### Regional unterschiedliche Erhöhung des Suchendenaufkommens

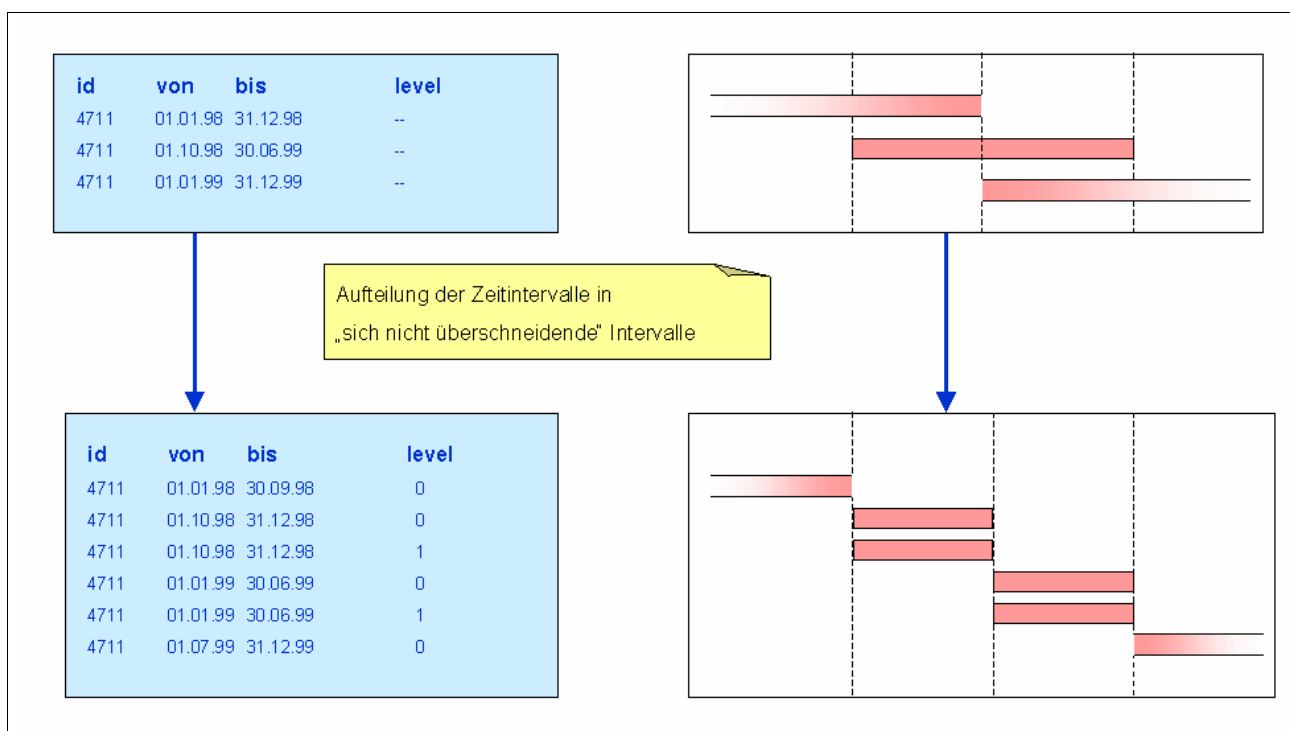
Die Einführung des Rechtskreises SGB II in die aktive Arbeitsmarktpolitik der BA war mit einer unterschiedlichen organisatorischen Umsetzung der Verantwortlichkeit (Arbeitsgemeinschaften, getrennte Trägerschaft, optimierende Kommunen) und unterschiedlichen Erfassungssystemen verbunden. In den dispositiven Systemen der BA und den aufsetzenden Forschungsdatensätzen des IAB sind sowohl die Trägerschaftswchsel eines Kunden, wie auch die Zuordnung von Arbeitssuchenden zu Trägern in Längsschnitten nur eingeschränkt abgebildet und nur bedingt statistisch auswertbar. Zum Teil zeigt sich dies in regionalen Unterschieden im Arbeitsuchendenaufkommen im Wechsel von 2004 auf 2005. Ein Beispiel: Personen, die Ende 2004 Arbeitsuchendenmeldungen für den Rechtskreis SGB III mit Gültigkeit bis Mitte 2005 aufweisen, können ggf. ab 01. Feb-



ruar 2005 im Rechtskreis SGB II geführt werden. Dies ist aber nicht in den Daten erkennbar. Die tatsächliche Gültigkeit der SGB III Arbeitssuchendenmeldung wäre dann tatsächlich nur bis Ende Januar 2005.

### 3.6 Episodensplitting

Bei der Erstellung der Gesamtdatei wird ein Episodensplitting durchgeführt. Dabei werden bei Überschneidungen von Zeiträumen innerhalb eines Kontos diese Spells so geschnitten und dupliziert, dass vollständig parallele Zeiträume und Zeiträume ohne Überschneidung entstehen. Die Anzahl der Spells erhöht sich dadurch.



Durch das Episodensplitting werden Spells verdoppelt und der Gültigkeitszeitraum wird geändert. Die ursprünglichen Datumsvariablen Beginn und Ende des Originalspells (*begorig* und *endorig*) bleiben erhalten, die Variablen Beginn und Ende der Episode (*begepi* und *endepi*) kennzeichnen Anfang und Ende der gesplitteten Episoden. Ob Spells gesplittet wurden, kann an einem Vergleich des ursprünglichen Zeitraums (*begorig* und *endorig*) mit dem Episodenzeitraum (*begepi* und *endepi*) festgestellt werden. Gesplittete Spells, die ursprünglich in einem Spell zusammengefasst waren, haben zudem dieselbe Satznummer.

Um die Originaldaten ohne Episodensplitting wieder herzustellen bzw. die durch das Episodensplitting künstlich erzeugten Episoden zu löschen, gibt es verschiedene Möglichkeiten. Man selektiert:



- alle Spells, für die gilt: Beginn des Originalspells = Beginn der gesplitteten Episode,
- alle Spells, für die gilt: Ende des Originalspells = Ende der gesplitteten Episode oder
- einen beliebigen Spell pro Satznummer.

Bei Merkmalen, die aufgrund ihres Bezugs zum Originalspell innerhalb der Satznummer konstant sind (siehe die Zeile Zeitbezug in den Beschreibungen der einzelnen Merkmale), macht es keinen Unterschied, welche Episode pro Satznummer gewählt wird. Bei den im Anschluss an das Episodensplitting erzeugten Merkmalen, die innerhalb der Satznummer variieren, ist das sehr wohl von Bedeutung. Soll dieses Merkmal nach dem Löschen der gesplitteten Episoden verwendet werden, muss das Kriterium Beginn des Originalspells = Beginn der gesplitteten Episode verwendet werden.

Löscht man die durch das Episodensplitting erzeugten Spells nicht, ist Folgendes zu beachten: Drei Variablen beziehen sich auf den Zustand vor Beginn des Originalspells und treffen daher nur für den Beginn der ersten von mehreren gesplitteten Episoden zu:

- Erwerbsstatus vor Arbeitsuche (estatvor)
- Art der Beendigung der letzten Tätigkeit (art\_kuend)
- Geplantes Beginndatum (begplan)

Vier Variablen beziehen sich auf den Zustand nach dem Ende des Originalspells und treffen daher nur für das Ende der letzten, von mehreren gesplitteten Episoden zu:

- Grund des Spellendes/Beendigungsgrund (grund),
- Status nach Abgang/Ausschlussgrund (sna),
- Restanspruchsdauer Arbeitslosengeld (restanspruch)
- Geplantes Enddatum (endplan)

### **3.7 Datenanonymisierung und sensible Merkmale**

Nach §75 SGB X ist eine Übermittlung von Sozialdaten einschließlich persönlicher Identifikatoren möglich, wenn dies zur Erreichung des Forschungsziels notwendig ist. Der Grundsatz der Datensparsamkeit verlangt, dass nicht mehr Informationen zur Verfügung gestellt werden dürfen, als zur Erreichung des Forschungsziels benötigt werden. In jedem Fall gilt es Datenmissbrauch (Deanonymisierung, vom Gesetzgeber nicht vorgesehene Analysen, etc.) zu verhindern. Bei dieser schwach anonymisierten (pseudo-anonymisierten) IEBS erfolgt dies zum einen durch die Beschränkung des Zugangswegs auf Gastaufenthalte (und mittelfristig Datenfernverarbeitung) und den damit verbundenen organisatorischen und technischen Maßnahmen.

---

Weiterhin werden folgende, geringe Modifikationen der Daten vorgenommen: Identifikatoren wie Sozialversicherungs-, Kunden- und Betriebsnummern werden durch systemfreie Nummerierungen ersetzt. An Stelle des Geburtsdatums ist das Geburtsjahr enthalten.

Bestimmte Merkmale, die eine Deanonymisierung von Personen erleichtern, werden nur dann im Original weitergegeben, wenn es für das Analyseziel notwendig ist und dies im Antrag auf den Datenzugang<sup>17</sup> explizit begründet wird. Diese aus datenschutzrechtlicher Sicht besonders sensiblen Merkmale sind in der IEBS:

die Staatsangehörigkeit

die Arbeits- und Wohnorte auf Kreisebene

die Arbeits- und Wohnorte auf Ebene der Arbeitsagenturen

Wohnort im Ausland

Gesundheitszustand

### **3.8 Testdaten und Hinweise für Datenauswertungen**

Damit sich Datennutzer auf ihren Gastaufenthalt vorbereiten und die Zeit am FDZ effizient nutzen können, werden Testdaten bereitgestellt. Es handelt sich dabei um Zufallsvariablen, die die univariaten Verteilungen der Variablen und ausgewählte Zusammenhänge in den Daten abbilden. Sie dienen zur Anfertigung und Erprobung von Auswertungssyntaxen aber keinesfalls für eine inhaltliche Interpretation der Ergebnisse.

## **4 Variablenbeschreibung**

Die Variablenbeschreibung gliedert sich in einen allgemeinen Teil mit einer Merkmalsübersicht und einer Beschreibung der Typisierung der in den Daten vorkommenden fehlenden Werte. Es folgen Beschreibungen der einzelnen Merkmale, die inhaltlich sortiert sind. Dadurch werden ähnliche Merkmale in nahe beieinander liegenden Abschnitten beschrieben, teilweise auch im selben Abschnitt, mit dem Ziel, dass Redundanzen vermieden und Zusammenhänge der Variablen deutlicher werden.

---

<sup>17</sup> Aktuelle Ausführungen zum Datenzugang finden Sie unter <http://fdz.iab.de>.

---

## 4.1 Allgemeiner Teil

### 4.1.1 Merkmalsübersicht

Die Merkmalsübersicht listet die Variablennamen und die längeren Merkmalsbezeichnung auf. Weiterhin bietet sie eine Übersicht darüber, welche Merkmale für die Spells der vier Datenquellen vorkommen bzw. gefüllt sind:

in den Spells dieser Quelle vorhanden,

in den Spells dieser Quelle nicht vorhanden, die Variable enthält hier ausschließlich Systemmissings.

Beispiel: Das geplante Beginndatum einer Maßnahme ist lediglich für MTH-Sätze gefüllt, die Spells der anderen Datenquellen enthalten in dieser Variablen Systemmissings. Eine weitere Besonderheit der IEBS liegt darin, dass einige Variablen je nach Datenquelle unterschiedliche Inhalte haben. So enthält der Erwerbsstatus für die BeH-Spells die Personengruppe des Meldeverfahrens, für LeH-Spells die Leistungsart, für MTH-Spells die Maßnahmeart und für BewA-Spells den Arbeitsuche-Status. Nicht bei jeder Variablen sind diese Unterschiede direkt im Variablennamen ersichtlich. Eine nach Variablennamen alphabetisch sortierte Merkmalsübersicht findet sich in Anhang 6.2.

Variablen-name	Seite	Bezeichnung	BeH	LEH	MTH	BewA
<b>Identifikationsnummern</b>						
persnr	29	Personennummer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
satznr	30	Satznummer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
betnr	30	Betriebsnummer	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<b>Spellanfang und -ende</b>						
begorig	31	Beginndatum Originalspell	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
endorig	31	Enddatum Originalspell	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
begepi	31	Beginndatum der gesplitteten Episode	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
endepi	31	Beginn- und Enddatum der gesplitteten Episode	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>Generierte technische Merkmale</b>						
quelle	32	Quelle der Information	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
kom_quel	32	Kombination der Quellen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
spell	33	Spellzähler pro Konto	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
nspell	33	Anzahl der Spells pro Konto	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
level2	34	Spellzähler pro Episode	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
nlevel2	34	Anzahl der zeitgleichen Spells pro Episode	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
level1	34	Spellzähler pro Episode und Quelle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
nlevel1	34	Anzahl der zeitgleichen Spells pro Episode und Quelle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fortsetzung der Tabelle auf der nächsten Seite

Variablen-Name	Seite	Bezeichnung	BeH	LEH	MTH	BewA
<b>Personenstatus vor, während und nach dem aktuellen Spell</b>						
estatvor	35	Erwerbsstatus vor Arbeitsuche				
erwstat	35	Erwerbsstatus: Personengruppe, Leistungsart, Maßnahmeart, Arbeitsuche-Status	ab 99			
grund	37	Grund des Spellendes/Beendigungsgrund				
sna	38	Status nach Abgang				
<b>Persönliche Merkmale</b>						
gebjahr	38	Geburtsjahr				
sex	39	Geschlecht				
nation	39	Staatsangehörigkeit				
schweb	40	Schwerbehindertenstatus				
schbild	40	Schulabschluss				
bild	40	Ausbildung				
quali	41	Qualifikation				
famstand	42	Familienstand				
kind	42	Anzahl der Kinder im Haushalt				
gesund_ein	43	Gesundheitliche Einschränkungen				
<b>Angaben zu Beschäftigungsverhältnis und Arbeitsuche</b>						
stib	43	Stellung im Beruf und Arbeitszeit				
beruf	44	Beruf				
tentgelt	44	Tagesentgelt / täglicher Leistungssatz				
gleitz	45	Gleitzone	ab 03			
w73	45	Wirtschaftszweig 73				
w93	45	Wirtschaftszweig 93	ab 99			
w03	45	Wirtschaftszweig 03				
mobil	46	Bereitschaft zu bundesweiter Stellensuche				
kunden_gr	47	BA-Kundengruppe				
art_kuend	47	Art der Beendigung der letzten Tätigkeit				
arbzeit	47	Arbeitszeit des Stellengesuchs				
restanspruch	48	Restanspruchsdauer Arbeitslosengeld				
begplan	48	Geplantes Beginndatum				
endplan	48	Geplantes Enddatum				
<b>Angaben zum Beschäftigungsbetrieb</b>						
az_hpt	49	Anzahl der hauptbeschäftigten Mitarbeiter				
grd_jahr	49	Gründungsjahr des Betriebes				

Fortsetzung der Tabelle auf der nächsten Seite

Variablen-Name	Seite	Bezeichnung	BeH	LEH	MTH	BewA
<b>Ortsangaben</b>						
ao_bula	49	Arbeitsort Bundesland (Kreis)				
wo_bula	50	Wohnort Bundesland (Kreis)	ab 99			
ao_rd	51	Arbeitsort Regionaldirektion (Agenturbezirk)				
wo_rd	51	Wohnort Regionaldirektion (Agenturbezirk)	ab 99			
wo_kreis	50	Wohnort Kreis				
ao_kreis	49	Arbeitsort Kreis				
ao_aa	51	Arbeitsort Agenturbezirk				
wo_aa	52	Wohnort Agenturbezirk				
wo_aatyp06	52	Wohnort Regionaltyp der Arbeitsagentur 06	ab99			
wo_ausl	53	Wohnort im Ausland				

#### 4.1.2 Fehlende Werte

Die BA-Daten enthalten keine einheitliche Codierung der fehlenden Werte. In der IEBS wurde folgende Standardisierung durchgeführt:

Bezeichnung	Wert	Beschreibung	Beispiel
Keine Angabe und keine Zuordnung möglich	-7	Häufig wird bei den ursprünglichen Schlüsseln der Datenquellen bereits ein Wert für das Fehlen der Angabe im Anschluss an die Datenerhebung vorgesehen. Dem Sinn entsprechend ist diese Kategorie auch zutreffend, wenn Felder in Spells leer sind.  Ebenfalls mit der -7 kodiert sind nunmehr die in der Vorversion mit -9 markierten Ausprägungen. Hierunter fallen Werte, die nicht zu den vorgegebenen Kategorien passen und vermutlich auch durch Eingabefehler oder fehlende Angaben entstanden sind.	Arbeitskräfte ohne nähere Tätigkeitsangabe in der Klassifikation der Berufe.
Systemmissing	.	Wenn ein Merkmal für eine Quelle nicht gefüllt ist (graue Felder der Merkmalsübersicht im vorhergehenden Abschnitt).  Bei Datumsangaben, da negative Werte als Datum angezeigt würden.	

## 4.2 Identifikationsnummern

### 4.2.1 Personennummer

Variablenname	persnr
Variablenlabel	Personennummer
Gefüllt für	alle Quellen
Herkunft	generiert

Zeitbezug	personenbezogen
Anonymisierung	keine
Detailbeschreibung	Die Personennummer stellt einen systemfreien Identifikator für eine Person dar. Der ursprüngliche Identifikator wird vorrangig aus der Sozialversicherungsnummer gebildet. Ist keine Versicherungsnummer vorhanden, wird stattdessen die Kundennummer der BA verwendet. Sätze, die weder eine gültige Versicherungsnummer noch eine gültige BA-Kundennummer haben, werden verworfen, da sie nicht verknüpft werden können. Dieser Identifikator wird für jede Person durch eine generierte Zufallszahl ersetzt, die eindeutig ist.
Besonderheiten	Die Nichtweitergabe der Sozialversicherungs- und BA-Kundennummern erfolgt aus Datenschutzgründen. Damit erfüllt dieser Datensatz die Vorgabe der schwachen Anonymisierung.  Zur Eindeutigkeit von Sozialversicherungs- und BA-Kundennummern und resultierenden Probleme bei der Verknüpfung der Spells aus unterschiedlichen Quellen siehe Abschnitt 3.1.

#### 4.2.2 Satznummer

Variablenname	satznr
Variablenlabel	Satznummer
Gefüllt für	alle Quellen
Herkunft	generiert
Zeitbezug	innerhalb der Personennummer variabel
Anonymisierung	keine
Detailbeschreibung	Die Satznummer ist über alle Originalsätze der IEB-Quellen – also auch über die ungesplitteten Sätze der Gesamt-IEB – eindeutig. Das Merkmal satznr weist jedem Originalspell eine eindeutige Zahl zu. Dadurch wird das Zuspielen weiterer Merkmale aus anderen IAB-Datenquellen möglich. Durch das Episodensplitting neu gebildete Spells behalten die gleiche satznr wie der Originalspell.
Besonderheiten	keine

#### 4.2.3 Betriebsnummer

Variablenname	betnr
Variablenlabel	Betriebsnummer
Gefüllt für	BeH
Herkunft	BeH
Anonymisierung	systemfreier Identifikator; nur gültig für die IEBS
Zeitbezug	innerhalb der Personennummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Betriebsnummern werden von den Arbeitsagenturen an Betriebe mit sozialversicherungspflichtig Beschäftigten vergeben. Die Originale wurden in der IEBS durch einen systemfreien Betriebsidentifikator ersetzt. Jeder Betrieb ist mit einem eindeutigen Identifikator geführt.
Besonderheiten	Da die Betriebsmerkmale Arbeitsort und Wirtschaftszweig über die Betriebsnummer zugespielt werden, fehlen diese Informationen in Meldungen ohne Betriebsnummer.

### 4.3 Spellanfang und -ende

#### 4.3.1 Beginndatum Originalspell

Siehe Abschnitt 4.3.4

#### 4.3.2 Enddatum Originalspell

Siehe Abschnitt 4.3.4

#### 4.3.3 Beginndatum der gesplitteten Episode

Siehe Abschnitt 4.3.4

#### 4.3.4 Beginn- und Enddatum

Variablennamen	Originalspells: begorig, endorig gesplittete Episoden: begepi, endepe
Variablenlabel	Beginndatum des Originalspells (begepi: der gesplitteten Episode) Enddatum des Originalspells (endepe: der gesplitteten Episode)
Gefüllt für	alle Quellen
Herkunft	BeH, LeH, MTH, BewA
Anonymisierung	keine
Zeitbezug	begorig, endorig: innerhalb der Personennummer variabel, innerhalb der Satznummer konstant begepi, endepe: innerhalb der Personennummer und innerhalb der Satznummer variabel
Detailbeschreibung	begorig und endorig geben tagesgenau den Beginn und das Ende des Originalspells an, begepi und endepe den Beginn und das Ende der überschneidungsfreien, gesplitteten Episoden. Bei MTH-Spells ist das Enddatum des Originalspells dem Abgangssatz entnommen, sofern ein solcher vorhanden ist. Ansonsten stammt es aus dem Zugangssatz. Beginn- und Enddatum sind wie alle Datumsangaben der IEBS ganzzahlig gespeichert, wobei der 31.12.1959 den Wert -1 annimmt, der 01.01.1960 den Wert 0, der 02.01.1960 den Wert 1 u.s.w.
Besonderheiten	Wenn eine Meldung aufgrund von Überlappungen durch das Episodensplitting auf mehrere Episoden aufgeteilt wurde, dann stimmt der Beginn des Originalspells mit dem Beginn der <u>ersten</u> gesplitteten Episode überein und das Ende des Originalspells entspricht dem Ende der <u>letzten</u> gesplitteten Episode. Alle anderen Beginn- und Enddaten der Originalspells und der gesplitteten Episoden weichen voneinander ab. Wurde ein Originalspell nicht gesplittet, so stimmt das Episodenanfangs- bzw. -enddatum mit dem Anfangs- bzw. Enddatum des Originalspells überein. Sollen sich Auswertungen auf die ungesplitteten Spells beziehen, lässt sich ein Filter beispielsweise durch die Bedingung begorig = begepi bilden. Gesplittete Spells, die ursprünglich in einem Spell zusammengefasst waren, haben dieselbe Satznummer (satznr). Eine ausführliche Erläuterung des Episodensplittings findet sich in Abschnitt 3.6.

	<p>Achtung: Manche Merkmale gelten nur für die Originalspells, nicht für die durch das Episodensplitting hinzugefügten Spells (siehe Abschnitt 3.6).</p> <p>Durch das Episodensplitting ergeben sich Episoden, die schon vor dem Beobachtungszeitraum einer Quelle beendet sind oder erst nach dem Beobachtungszeitraum einer Quelle beginnen.</p> <p>Bei den Datumsangaben ist bei Überlappung zu berücksichtigen, dass die Validität der Daten mit der Quelle variiert, z. B. tauchen in der BewA-Quelle u.a. Spells vor 1990 auf. Hierbei dürfte es sich vermutlich um Tippfehler handeln. Derartig frühe (oder auch späte) Spells, auch in den anderen Quellen außerhalb des eigentlichen Gültigkeitszeitraums, sind nicht in den Daten bereinigt.</p>
--	--

## 4.4 Generierte technische Merkmale

### 4.4.1 Quelle der Information

Variablenname	quelle
Variablenlabel	Quelle der Information
Gefüllt für	alle Quellen
Herkunft	generiert
Zeitbezug	innerhalb der Personnummer variabel, innerhalb der Satznummer konstant
Anonymisierung	keine
Detailbeschreibung	Enthält die Information, aus welcher Datenquelle die Informationen des Spells kommen. Bei MTH wird nach der Herkunft aus der spezifischen Maßnahmeteilnahme-Biografie unterschieden, was sich teilweise, aber nicht immer mit der Unterscheidung der Teilverfahren in coSach deckt.
Besonderheiten	keine

### 4.4.2 Kombination der Quellen

Variablenname	kom_quel							
Variablenlabel	Kombination der Quellen							
Gefüllt für	alle Quellen							
Herkunft	generiert							
Anonymisierung	keine							
Zeitbezug	innerhalb der Personnummer und innerhalb der Satznummer variabel							
Detailbeschreibung	<p>Das Merkmal wird nach dem Episodensplitting erzeugt und zeigt an, ob und welche Überschneidungen zwischen verschiedenen Quellen vorliegen. Die Variable ist als Bitmuster angelegt, sodass jede mögliche Kombination abgebildet werden kann. Dazu werden einfach die Werte der Variablen quelle (siehe Vorseite) aufaddiert, allerdings jeder Wert pro Quelle nur einmal. Zeitliche Überschneidungen innerhalb einer Quelle werden somit nicht angezeigt (siehe dazu die Merkmale level1 und nlevel1, Abschnitt 4.4.8), ausgenommen Überschneidungen der acht Teilverfahren innerhalb der Quelle MTH.</p> <p>Um festzustellen, welche Quellen gleichzeitig vorliegen, muss von dem Wert von kom_quel solange der jeweils höchste enthaltene Wert abgezogen werden, bis das Ergebnis 0 ist.</p> <p>Beispiel: Für eine Episode hat kom_quel den Wert 59:</p> <table border="1" style="margin-left: 20px;"> <tr> <td>Der höchste in</td> <td>59</td> <td>enthaltene Wert ist</td> <td>32:</td> <td>59 – 32</td> <td>=</td> <td>27</td> </tr> </table>	Der höchste in	59	enthaltene Wert ist	32:	59 – 32	=	27
Der höchste in	59	enthaltene Wert ist	32:	59 – 32	=	27		



	<p>Der höchste in 27 enthaltene Wert ist 16: <math>27 - 16 = 11</math>                  Der höchste in 11 enthaltene Wert ist 8: <math>11 - 8 = 3</math>                  Der höchste in 3 enthaltene Wert ist 2: <math>3 - 2 = 1</math>                  Der höchste in 1 enthaltene Wert ist 1: <math>1 - 1 = 0</math></p> <p>In 59 waren also 32, 16, 8, 2 und 1 enthalten, d. h. in diesem fiktiven Beispiel wären für die gleiche Episode für eine Person Spells zu Beschäftigung, Leistungsempfang, FbW/TM/DSL, FF und ESF-BA vorhanden.</p> <p>Die Ablage als Bitmuster minimiert den Speicherbedarf. Es lässt sich jedoch leicht in Indikatorvariablen umwandeln, die anzeigen, ob es in der betreffenden Episode einen Spell aus einer bestimmten Quelle gibt.</p> <p>Beispielsyntax in Stata:</p> <pre>generate q1 = mod(kom_quel, 2)~=0 generate q2 = mod(int(kom_quel/ 2), 2)~=0 generate q4 = mod(int(kom_quel/ 4), 2)~=0 generate q8 = mod(int(kom_quel/ 8), 2)~=0 generate q16 = mod(int(kom_quel/ 16), 2)~=0 generate q32 = mod(int(kom_quel/ 32), 2)~=0 generate q64 = mod(int(kom_quel/ 64), 2)~=0 generate q128 = mod(int(kom_quel/ 128), 2)~=0 generate q256 = mod(int(kom_quel/ 256), 2)~=0 generate q512 = mod(int(kom_quel/ 512), 2)~=0 generate q1024 = mod(int(kom_quel/1024), 2)~=0</pre> <p><math>\text{mod}(x,y) = x - y \cdot \text{int}(x/y)</math> (Modulo-Funktion)                  int: schneidet Nachkommastellen ab</p>
	1 bis 2047 (mögliche Werte) / 1 bis 1856 (tatsächliche Werte)

#### 4.4.3 Spellzähler pro Konto

Siehe folgender Abschnitt.

#### 4.4.4 Anzahl der Spells pro Konto

Variablennamen	spell, nspell
Variablenlabel	Spellzähler pro Konto (spell) ; Anzahl der Spells pro Konto (nspell)
Gefüllt für	alle Quellen
Herkunft	generiert
Anonymisierung	keine
Zeitbezug	spell: innerhalb der Personnummer und innerhalb der Satznr variabel nspell: innerhalb der Personnummer (und somit auch der Satznummer) konstant
Detailbeschreibung	<p>Der Spellzähler <b>spell</b> zählt alle Spells pro Konto, beginnend mit 1. Sowohl die Originalspells, als auch die durch das Episodensplitting hinzugefügten, werden gezählt. Die Daten wurden vorher in der folgenden Reihenfolge sortiert:</p> <ul style="list-style-type: none"> <li>• Personnummer (Abschnitt 4.2.1)</li> <li>• Beginndatum der gesplitteten Episode (Abschnitt 4.3.3)</li> <li>• Quelle (Abschnitt 4.4.1)</li> </ul>

	<p>BeH-Spells sind zudem sortiert nach:</p> <ul style="list-style-type: none"> <li>• Geringfügigkeitskennzeichen (aufsteigend, nicht in der IEBS enthalten)</li> <li>• Tagesentgelt / täglicher Leistungssatz (absteigend, Abschnitt 4.7.3)</li> </ul> <p>Das Merkmal <b>nspell</b> enthält die Gesamtzahl der Spells für das aktuelle Konto, d. h. es gilt: <math>nspell = \max(spell)</math> pro Konto.</p>
Besonderheiten	keine

#### 4.4.5 Spellzähler pro Episode

Siehe folgenden Abschnitt.

#### 4.4.6 Anzahl der zeitgleichen Spells

Siehe folgenden Abschnitt.

#### 4.4.7 Spellzähler pro Episode und Quelle

Siehe folgenden Abschnitt.

#### 4.4.8 Anzahl der zeitgleichen Spells pro Quelle

Variablennamen/ Variablenlabel:	<p>Spellzähler pro Episode: level2</p> <p>Anzahl der Spells pro Episode: nlevel2</p> <p>Spellzähler pro Episode und Quelle: level1</p> <p>Anzahl der Spells pro Episode und Quelle: nlevel1</p>
Gefüllt für	alle Quellen
Herkunft	generiert
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer und innerhalb der Satznummer variabel
Detailbeschreibung	<p>Der Spellzähler pro Episode <b>level2</b> nummeriert parallele Spells einer Episode von Null bis n. Gibt es in einer anderen Episode des Kontos erneut parallele Spells, wird wieder mit Null beginnend hoch gezählt.</p> <p>Der Spellzähler <b>level1</b> nummeriert parallele Sätze einer Episode <u>innerhalb einer Quelle</u> von Null bis n. Die Teilverfahren der MTH-Spells (quelle = 4, 8, 16, 32, 128, 256, 512, 1024) werden als separate Quellen betrachtet. Gibt es in derselben Episode parallele Spells einer anderen Quelle oder in einer anderen Episode des Kontos erneut parallele Sätze innerhalb dieser Quelle, wird jeweils wieder mit Null beginnend hoch gezählt. Die Spells eines Kontos werden vorher nach Quelle und innerhalb der Quellen in der folgenden Reihenfolge sortiert:</p> <p>BeH</p> <ul style="list-style-type: none"> <li>• Versicherungsnummer (aufsteigend, nicht in der IEBS enthalten)</li> <li>• Geringfügigkeitskennzeichen (aufsteigend, nicht in der IEBS enthalten)</li> <li>• Tagesentgelt (absteigend)</li> </ul> <p>LeH</p> <ul style="list-style-type: none"> <li>• Versicherungsnummer bzw. Kundennummer (aufsteigend, nicht in der IEBS)</li> <li>• Leistungsart (aufsteigend)</li> </ul>

	<ul style="list-style-type: none"> <li>• Satznummer (aufsteigend)</li> </ul> <p>MTH</p> <ul style="list-style-type: none"> <li>• Versicherungsnummer bzw. Kundennummer (aufsteigend, nicht in der IEBS enthalten)</li> <li>• Quelle (aufsteigend)</li> <li>• Satznummer (aufsteigend)</li> </ul> <p>BewA</p> <ul style="list-style-type: none"> <li>• Versicherungsnummer bzw. Kundennummer (aufsteigend, nicht in der IEBS enthalten)</li> <li>• Satznummer (aufsteigend)</li> </ul> <p>Das Geringfügigkeitskennzeichen ist folgendermaßen definiert: Geringfügig und kurzfristig Beschäftigte (= Personengruppen 109 und 209) erhalten den Wert 2, alle anderen Personengruppen erhalten den Wert 1. Diese Sortierung führt dazu, dass bei Mehrfachbeschäftigung immer die sozialversicherungspflichtige Beschäftigung – sofern eine solche vorliegt – den Wert 0 für <b>level1</b> erhält. Außerdem wird innerhalb dieser Sortierung immer die Beschäftigung mit dem höheren Entgelt mit dem Wert 0 gekennzeichnet. Der BeH-Spell mit dem höchsten Tagesentgelt hat also i.d.R. den Wert 0 und kennzeichnet die Hauptbeschäftigung. Die Variable Anzahl der Spells pro Episode <b>nlevel2</b> gibt an, wie viele zeitgleiche Spells es innerhalb der Episode gibt, d.h. es gilt: <math>nlevel2 = \max(level2) + 1</math> pro Episode. Die Variable Anzahl der Spells pro Episode und Quelle <b>nlevel1</b> gibt an, wie viele zeitgleiche Spells es innerhalb der Episode pro Quelle gibt, d.h. es gilt: <math>nlevel1 = \max(level1) + 1</math> pro Episode und Quelle.</p>
Besonderheiten	Konten mit mehr als 2000 Spells werden systembedingt gelöscht (betrifft 4 Konten) .

## 4.5 Personenstatus vor, während und nach dem aktuellen Spell

### 4.5.1 Erwerbsstatus vor Arbeitsuche

Variablenname	estatvor
Variablenlabel	Erwerbsstatus vor Arbeitssuche
Gefüllt für	BewA
Herkunft	BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personennummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Das Attribut zeigt für BewA-Spells den Erwerbsstatus vor der Arbeitsuche an (zum Status der Arbeitsuche siehe erwstat, Abschnitt 4.5.2).
Besonderheiten	Ab 12/2002 erfolgte eine deutliche Reduzierung der Ausprägungen. Die Werte älterer Sätze wurden auf die aktuell gültigen Werte umgeschlüsselt, die somit für den gesamten Zeitraum gültig sind.  Die Angabe ist für den ungesplitteten Originalspell und für die erste gesplittete Episode gültig.

### 4.5.2 Erwerbsstatus: Personengruppe, Leistungsart, Maßnahmeart, Arbeitsuche-Status

Variablenname	erwstat
Variablenlabel	Erwerbsstatus: Personengruppe, Leistungsart, Maßnahmeart, Arbeitsuche-Status

Gefüllt für	alle Quellen
Herkunft	BeH (ab 1999), LeH, MTH, BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>Das Merkmal enthält den Status der Personen während des aktuellen Spells, also von Beginndatum bis Enddatum des Originalspells (siehe begorig und endorig, Abschnitt 4.3.1). Sie nimmt für jede Datenquelle unterschiedliche Werte mit unterschiedlichen Bedeutungen an.</p> <p>Für <b>BeH-Spells</b> ist die Personengruppe der Beschäftigtenmeldung enthalten, die mit dem neuen Meldeverfahren zum 01.01.1999 eingeführt wurde. Sie weist beitrags- oder leistungsrechtliche Besonderheiten des Beschäftigungsverhältnisses auf, die mindestens einen Zweig der Sozialversicherung betreffen, unterschieden nach meldenden Institutionen.</p> <p>Sozialversicherungspflichtige Arbeitnehmer, deren Beschäftigungsverhältnis keine Besonderheiten aufweist, erhalten den Schlüssel 101. Treffen mehrere Schlüssel zu, ist der kleinste anzugeben. Geringfügig Beschäftigte lassen sich als die Personengruppen 109, 110, 202, und 209 abgrenzen.</p> <p>Das Merkmal Personengruppe ist erst seit der Einführung des neuen Meldeverfahrens zum 01.01.1999 Bestandteil der Meldungen zur Sozialversicherung. Die Personengruppe kann in Meldungen, die sich auf die Jahre vor 1999 beziehen, aber erst 1999 oder später eingegangen sind, enthalten sein. Für Meldungen, die vor 1999 eingegangen sind, wird versucht, anhand bestimmter Regeln mit Hilfe der Merkmale „Ausbildung“, „Stellung im Beruf und Arbeitszeit“ und „Beruf“ (siehe Abschnitt 4.6.6, 4.7.1 und 4.7.2) sowie weiterer Informationen eine Zuordnung zu den Personengruppen durchzuführen. In vielen Fällen ist das jedoch nicht eindeutig möglich.</p> <p>Für <b>LeH-Spells</b> ist die gruppierte Leistungsart angegeben, d.h. es ist angegeben, ob Arbeitslosengeld, Arbeitslosenhilfe oder Unterhaltsgeld bezogen wird.</p> <p>Für <b>MTH-Spells</b> ist eine Zusammenfassung der Merkmale Maßnahmebereich und Leistungsart enthalten. Eine Aufstellung aller möglichen Werte findet sich im Auszahlungshandbuch (siehe <a href="http://fdz.iab.de">http://fdz.iab.de</a>).</p> <p>Bei <b>BewA-Spells</b> ist der Arbeitsuche-Status enthalten. In der BewA gibt es nur die Unterscheidung zwischen Spells mit Status "arbeitslos" und "arbeitsuchend". In den IEB wird zusätzlich der Status „krank“ generiert. Spells erhalten den Arbeitsuche-Status "krank" wenn</p> <ul style="list-style-type: none"> <li>• ein Vorgängerspell „arbeitslos“ existiert, der lückenlos anschließt und Abganggrund „krank“ (sna = 6004, siehe Abschnitt 4.5.4) hat und</li> <li>• der Spell selbst nicht den Status "arbeitslos" sondern den Status "arbeitsuchend" hat.</li> </ul>
Besonderheiten	<p>Die DEÜV-Personengruppen 301 bis 303 werden nicht voll erhoben.</p> <p>Mitunter finden sich Personen zeitgleich in mehreren Erwerbszuständen, die rechtlich nicht zulässig sind. Aufgrund der Vielzahl von Maßnahmen, deren Teilnahmevoraussetzungen sich im Zeitablauf häufig ändern, gestaltet sich die Bereinigung dieser Inkonsistenzen sehr aufwändig. Eine erste Aufstellung zulässiger und unzulässiger Zustände findet sich in Fitzenberger et al. 2005b.</p> <p>Achtung: im Vergleich zur Version 2005 haben sich z.T. die Ausprägungen geändert.</p> <p>Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4)</p>

## 4.5.3 Grund des Spellendes/Beendigungsgrund

Variablenname	grund
Variablenlabel	Grund des Spellendes / Beendigungsgrund
Gefüllt für	BeH, LeH, BewA
Herkunft	BeH, LeH, BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>Der Grund des Spellendes hat für die Datenquellen unterschiedliche Bedeutungen:          BeH: Grund der Abgabe der Meldung          LeH: Grund der Beendigung des Leistungsbezugs          BewA: Grund des Abgangs aus Arbeitsuche</p> <p>In <b>BeH-Spells</b> kennzeichnet der Abgabegrund den Anlass, aus dem der Arbeitgeber den Sozialversicherungsträgern die betreffende Beschäftigungsmeldung erstattet hat. Die Abgabegründe für Beschäftigungsmeldungen sind nach den Regeln des seit 01.01.1999 geltenden Meldeverfahrens (gemäß DEÜV<sup>18</sup>) verschlüsselt. Im Vergleich zum alten Meldeverfahren (nach den DEVO-/DÜVO-Regeln<sup>19</sup>) werden die Abgabegründe differenzierter erfasst. Die bisherigen Abgabegründe wurden auf die neuen Abgabegründe umgeschlüsselt.</p> <p>In der BeH und damit auch in der IEBS treten nicht alle, im Rahmen des Meldeverfahrens möglichen, Abgabegründe auf. So beinhaltet die BeH nur Meldungen mit Entgeltangaben (also Jahres-, Unterbrechungs- und Abmeldungen), während Anmeldungen wegen fehlender Entgeltinformationen nicht enthalten sind. Ein Informationsverlust ist damit jedoch nicht verbunden, da die Angaben aus einer Anmeldung mit der darauf folgenden Jahres-, Unterbrechungs- oder Abmeldung erneut übermittelt werden.</p> <p>Für <b>LeH-Spells</b> zeigt das Merkmal den Grund für das Ende des Bezugs von Arbeitslosengeld, -hilfe oder Unterhaltsgeld an. Ab dem 01.01.2003 haben einige Ausprägungen eine veränderte Bedeutung. Maßgeblich dafür, welche Bedeutung zutrifft, ist, ob das Enddatum des Originalspells (endorig, Abschnitt 4.3.2) kleiner oder größer/gleich dem 01.01.2003 ist. Die hiervon betroffenen Ausprägungen sind anhand des Auszahlungshandbuchs (siehe <a href="http://fdz.iab.de">http://fdz.iab.de</a>) ersichtlich.</p> <p>Gründe für den Beginn des Leistungsbezugs finden sich in der IEBS nicht, da sie sich aus den Meldungen der Arbeitsagenturen an die Krankenkassen über den Abschluss einer Leistungsbezugsperiode speist.</p> <p>Bei <b>BewA-Spells</b> ist der Abmelde- bzw. Abgangsgrund enthalten. Hierbei ist zu beachten, dass ab dem 26.04.2003 eine Reduktion der Merkmalsausprägungen erfolgte. Die alten Werte sind, soweit es möglich war, auf die neuen Werte umgeschlüsselt.</p>
Besonderheiten	<p>Bei LeH- und BewA-Spells ist besonders auf die Änderungen der Merkmalsbedeutungen im Zeitablauf zu achten.</p> <p>Die Angabe ist für den ungesplitteten Originalspell und für die letzte gesplittete Episode gültig.</p> <p>Achtung: im Vergleich zur Version 2005 haben sich z.T. die Ausprägungen geändert.</p> <p>Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4).</p>

<sup>18</sup> DEÜV: Datenerfassungs- und -übermittlungsverordnung, in Kraft getreten am 1. Januar 1999.

<sup>19</sup> DEVO: Datenerfassungsverordnung; DÜVO: Datenübermittlungsverordnung; die 2. DEVO und die 2. DÜVO wurden jeweils am 1. Januar 1999 durch die Datenerfassungs- und -übermittlungsverordnung abgelöst.

#### 4.5.4 Status nach Abgang

Variablenname	sna
Variablenlabel	Status nach Abgang
Gefüllt für	BewA, MTH
Herkunft	generiert aus den BewA-Merkmalen Abgangsgrund und „geplante Förderung“ bei Abgang
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>Dieses Merkmal gibt an, in welchen Status Arbeitsuchende nach Ende des BewA-Spells wechseln. Er gibt u.a. Auskunft darüber, ob nach Abgang aus Arbeitslosigkeit eine geförderte oder ungeförderte Beschäftigung aufgenommen wurde. Insbesondere am aktuellen Rand, wo noch keine parallelen Beschäftigendaten vorliegen, kann dies eine wichtige Information über den Verbleib darstellen. Das Merkmal wird aus dem BewA-Abgangsgrund (siehe Kapitel 4.5.3) und dem Merkmal Berufspraxis/Förder-ID (nicht in der IEBS) gebildet.</p> <p>Hintergrund:</p> <p>Durch Änderungen in dem den BewA-Daten zugrunde liegenden operativen System coArb, werden bestimmte Informationen, die früher im Merkmal Abmeldegrund erfasst wurden, seit 26.04.2003 nicht mehr dort, sondern im Merkmal "Geplante Förderung" (ehemals Merkmal "Förder-ID") bei Abgang abgebildet. Die „geplante Förderung“ ist jedoch im zugrunde liegenden Datenmodell nicht der Tabelle zum Meldestatus (ASU-Status, Identifikator = Kundennummer) zugeordnet, sondern ein Merkmal aus dem Bewerberangebot (Identifikator = BewA-Eingabenummer) mit eigener Gültigkeitsdauer. Um weiterhin zu jedem ASU-Datensatz den zugehörigen Beendigungsgrund zuordnen zu können, wurde in der Forschungsdatenbank ISAAK die "Geplante Förderung" bei Abgang den entsprechenden ASU-Status-Sätzen zugeordnet und ein neues Merkmal als Kombination aus dem bisherigen Abmeldegrund und dem Merkmal "Geplante Förderung" bei Abgang gebildet, das vor und ab 2003 dieselben Ausprägungen hat, sodass es keinen Bruch über die Zeit gibt.</p>
Besonderheiten	<p>Einige Ausprägungen sind nicht über den gesamten Zeitraum gültig. Ein Teil der Ausprägungen entfällt in 2003, da die Merkmalsausprägung beim Abgangsgrund nicht mehr erhoben wird, siehe Variable Grund des Spellendes (grund). Der Wechsel der Merkmalsausprägungen erfolgt bei einigen Spells auch schon vor oder erst nach 2003.</p> <p>Die Angabe ist für den ungesplitteten Originalspell und für die letzte gesplittete Episode gültig.</p> <p>Für die MTH gibt es jedoch nur einen Wert: „Beschäftigung selbst gesucht“.</p> <p>Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4). Das Merkmal wird mit guter Datenqualität gefüllt.</p>

## 4.6 Persönliche Merkmale

### 4.6.1 Geburtsjahr

Variablenname	gebjahr
Variablenlabel	Geburtsjahr
Gefüllt für	alle Quellen
Herkunft	BeH, BA-Kundenhistorik
Anonymisierung	keine

Zeitbezug	innerhalb der Personennummer (und somit auch der Satznummer) konstant
Detailbeschreibung	Das Geburtsjahr wird aus der Versicherungsnummer gebildet, sofern diese vorhanden ist. Bei Konten ohne Versicherungsnummer wird auf Basis der BA-Kundennummer das aktuellste Geburtsjahr aus der Kundenhistorik der BA zugespielt.
Besonderheiten	Da das Jahr in der Sozialversicherungsnummer 2-stellig angegeben ist, wird das Jahrhundert '19' ergänzt. Ein Wechsel des Geburtsjahres innerhalb eines Kontos kann nicht auftreten. Die Variable wird auf Basis des Geburtsdatums generiert.

#### 4.6.2 Geschlecht

Variablenname	sex
Variablenlabel	Geschlecht
Gefüllt für	alle Quellen
Herkunft	BeH, BA-Kundenhistorik
Anonymisierung	keine
Zeitbezug	innerhalb der Personennummer (und somit auch der Satznummer) konstant
Detailbeschreibung	Das Geschlecht wird aus der Versicherungsnummer gebildet, sofern diese vorhanden ist. Bei Konten ohne Versicherungsnummer wird auf Basis der BA-Kundennummer die aktuellste Geschlechtsangabe aus der Kundenhistorik der BA zugespielt.
Besonderheiten	Ein Wechsel des Geschlechts innerhalb eines Kontos kann nicht auftreten.

#### 4.6.3 Staatsangehörigkeit

Variablenname	nation
Variablenlabel	Staatsangehörigkeit
Gefüllt für	alle Quellen
Herkunft	BeH, LeH, BA-Kundenhistorik
Anonymisierung	keine
Zeitbezug	innerhalb der Personennummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Die Variable Staatsangehörigkeit liegt in den Daten leicht aggregiert vor.
Besonderheiten	Das Merkmal wird nicht bereinigt, d.h. in einem Konto können für einen Zeitpunkt oder im Zeitablauf unterschiedliche Staatsangehörigkeiten auftreten, die auf unterschiedlichen Angaben in den Quelldaten beruhen und nicht unbedingt mit einem tatsächlichen Wechsel der Staatsangehörigkeit verbunden sind. Die Zuordnung bezieht sich auf den Zeitpunkt der Erfassung, daher sind Staaten enthalten, die nicht mehr existieren (z.B. die Sowjetunion oder Jugoslawien). Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4).



#### 4.6.4 Schwerbehindertenstatus

Variablenname	schweb
Variablenlabel	Schwerbehindertenstatus
Gefüllt für	BewA, MTH
Herkunft	BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Das Merkmal stammt aus dem BewA. Es wird der Schwerbehindertenstatus aus dem BewA zugespielt, der am Tag vor Beginn der Maßnahmeteilnahme ( <b>MTH-Spells</b> ) bzw. der Arbeitsuche ( <b>BewA-Spells</b> ) gültig ist.
Besonderheiten	Die Information wird benötigt, um festzustellen, ob jemand eine Behinderung nach dem Schwerbehindertengesetz hat. Eine Schwerbehinderung liegt ab dem Grad der Behinderung von 50% vor.  Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4). Das Merkmal wird mit guter Datenqualität gefüllt.

#### 4.6.5 Schulabschluss

Variablenname	schbild
Variablenlabel	Schulabschluss
Gefüllt für	BewA, MTH
Herkunft	BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Das Merkmal stammt aus der Quelle BewA und enthält für die Quellen MTH und BewA den Schulabschluss. ( <b>MTH-Spells</b> : gültig ab dem Tag vor Beginn der Maßnahme, <b>BewA-Spells</b> : gültig ab Beginndatum Arbeitslosigkeit bzw. Arbeitssuchendenstatus)
Besonderheiten	Angaben zum Schulabschluss von Beschäftigten sind im Merkmal Ausbildung enthalten. Siehe hierzu auch Abschnitt 4.6.6.  Die Angaben der MTH- und BewA-Sätze scheinen zuverlässiger zu sein als die der BeH. Hintergrund ist, dass die BeH-Angaben aus dem Meldeverfahren stammen, der Schulabschluss aber keinen sozialversicherungsrechtlichen Anspruch begründet und daher vermutlich mit weniger Sorgfalt angegeben wird.  Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4). Das Merkmal wird mit guter Datenqualität gefüllt.

#### 4.6.6 Ausbildung

Variablenname	bild
Variablenlabel	Ausbildung
Gefüllt für	BeH, MTH, BewA
Herkunft	BeH, BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant



Detailbeschreibung	<p>Für <b>BeH-Spells</b> ist der Ausbildungsschlüssel (B2-Schlüssel) der Beschäftigtenstatistik enthalten, d.h. die Angaben stammen aus dem Meldeverfahren und dort aus den „Angaben zur Tätigkeit“. Bei dem Signierschlüssel „Ausbildung“ handelt es sich um einen kombinierten Schlüssel, durch den sowohl die erreichte Schulbildung als auch die abgeschlossene Berufsausbildung (jeweils höchster Abschluss) der/des Beschäftigten erhoben wird. Dabei ist grundsätzlich zunächst die Schulbildung festzustellen, danach die Berufsausbildung. In Bezug auf den Schulabschluss ist das Merkmal nur dichotom interpretierbar: Die Allgemeine Hochschulreife ist vorhanden oder nicht vorhanden. Wenn sie nicht vorhanden ist, kann aus den Angaben nicht mit Sicherheit abgelesen werden, ob an anderen Schulen tatsächlich ein Abschluss erlangt wurde, oder lediglich der Besuch einer solchen Schule vermerkt ist. In Bezug auf die Berufsausbildung wird lediglich bei den Werten 5 (Abschluss einer Fachhochschule) und 6 (Hochschul-/Universitätsabschluss) auf die Feststellung einer eventuell sonstigen Berufsausbildung verzichtet. Nicht als Ausbildung zählen hierbei berufliche Fortbildung (z.B. Kurse an Sprach- und Dolmetscherschulen) sowie Fachkenntnisse, die durch praktische Tätigkeiten erworben wurden.</p> <p>Für <b>MTH-</b> und <b>BewA-Spells</b> stammt die abgeschlossene Berufsausbildung aus dem BewA. Für MTH-Spells wird sie zum Beginn der Maßnahme zugespielt, für BewA-Spells zum Beginn der Arbeitsuche.</p>
Besonderheiten	<p>In den BeH-Spells gibt es für bestimmte Teilgruppen einen hohen Anteil fehlender Angaben, weil das Merkmal versicherungsrechtlich keine Bedeutung hat. Zusammen mit einem Betriebswechsel treten häufig auch Änderungen im Ausbildungsstatus auf. Das liegt daran, dass beim neuen Betrieb die Meldedaten neu zusammengestellt werden. Wenn ein Beschäftigter z.B. durch eine berufsbegleitende Weiterbildung einen höheren Abschluss erlangt hat, dann wird dieser Statuswechsel wahrscheinlich erst von einem neuen Beschäftigungsbetrieb erfasst. Es ist in der Regel davon auszugehen, dass bei längeren Beschäftigungszeiten innerhalb eines Betriebes die einmal erhobenen personenbezogenen Daten fortgeschrieben werden.</p> <p>Konsistenzchecks der Bildungsvariablen (Schulabschluss, Abschnitt 4.6.5 und Ausbildung, Abschnitt 4.6.6) von Fitzenberger et al. (2005b, S. 36 ff.) ergeben eine nicht geringe Anzahl von Inkonsistenzen, sowohl zwischen den Quellen als auch im Zeitablauf. Die Angaben der MTH und BewA-Sätze scheinen zuverlässiger zu sein als die der BeH.</p> <p>Fitzenberger et al. (2005a) machen konkrete Korrekturvorschläge für die Bildungsvariable BILD der IABS, die der IEBS-Variablen der BeH-Spells entspricht. Bei der IEBS empfiehlt es sich, zusätzlich die Informationen der MTH- und BewA-Spells zu verwenden.</p> <p>Neue Auszählungen zeigen eine erhebliche Zunahme der fehlenden Werte für jüngere Einstiegskohorten (bis zu 30%).</p> <p>Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4). Das Merkmal ist derzeit für Daten aus VerBIS nicht ausreichend gefüllt und deshalb nicht auswertbar.</p>

#### 4.6.7 Qualifikation

Variablenname	quali
Variablenlabel	Qualifikation des Bewerbers
Gefüllt für	BewA
Herkunft	BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer und innerhalb der Satznummer variabel
Detailbeschreibung	Bei diesem Merkmal wird nicht nur die Ausbildung berücksichtigt, sondern auch die Berufserfahrung. Bei gleichwertiger Befähigung aufgrund von Vorbildung, Leistung oder Erfahrung kann auch ohne einen entsprechenden Abschluss eine Qualifikationsstufe eingetragen werden

Besonderheiten	Das Merkmal enthält die Qualifikation von Bewerber/Innen, wie sie in coArb erfasst wird. Das Merkmal wird aus VerBIS nicht mehr geliefert. Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4).
----------------	---

#### 4.6.8 Familienstand

Variablenname	famstand
Variablenlabel	Familienstand
Gefüllt für	BEH und MTH
Herkunft	BEH und MTH
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer und innerhalb der Satznummer variabel
Detailbeschreibung	Das Merkmal beschreibt den Familienstand von Bewerber/innen (BewA). Der Familienstand schließt Kategorien der häuslichen Gemeinschaft mit Kindern oder Lebenspartnern ein. Allerdings ist nicht trennscharf, ob <ul style="list-style-type: none"> <li>• eine verheiratete / in häuslicher Gemeinschaft lebende Person Kinder erzieht,</li> <li>• eine verheiratete / in häuslicher Gemeinschaft / alleinlebende Person überhaupt Kinder hat,</li> <li>• eine alleinerziehende / in häuslicher Gemeinschaft / alleinlebende Person verheiratet ist</li> </ul> Werteausprägungen kleiner 3 sind Werte aus der Quelle LeH und umfassen nur 2 Ausprägungen. Ein Abgleich der Informationen aus beiden Quellen wurde nicht vorgenommen.
Besonderheiten	Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4). Das Merkmal wird mit guter Datenqualität gefüllt. Die Informationen aus der LeH sind von geringerer Qualität.

#### 4.6.9 Anzahl der Kinder im Haushalt

Variablenname	kind
Variablenlabel	Anzahl der Kinder im Haushalt
Gefüllt für	BewA
Anonymisierung	keine
Herkunft	BewA
Zeitbezug	innerhalb der Personnummer und innerhalb der Satznummer variabel
Detailbeschreibung	Es werden die im Haushalt lebenden Kinder gezählt, die jünger als 15 Jahre sind. Der Grund für die Angabe liegt darin, dass die Zumutbarkeit von Vermittlungsvorschlägen im Zusammenhang mit betreuungspflichtigen Kindern eine wichtige Rolle spielt. Hilfebedürftigen mit Kindern unter drei Jahren ist die Ausübung einer Arbeit nicht zumutbar, vgl. §10 (1) Nr. 3 SGB II.
Besonderheiten	Es ist fraglich, ob tatsächlich nur Kinder unter 15 Jahren gezählt werden Die mögliche Anzahl der Kinder hat sich mit der VerBIS Umstellung geändert: <ul style="list-style-type: none"> <li>• CoArb: Kinderanzahl &gt; 10 konnte nicht erfasst werden</li> <li>• VerBIS: Kinderanzahl bis 99 möglich</li> </ul> Im IEBS-Datensatz wurden Fälle mit mehr als drei Kindern zu einer Klasse zusammengefasst. Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4). Das

	Merkmal wird mit guter Datenqualität gefüllt.
--	---

#### 4.6.10 Gesundheitliche Einschränkungen

Variablenname	gesund_ein
Variablenlabel	Gesundheitliche Einschränkungen
Gefüllt für	BewA
Herkunft	BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer und innerhalb der Satznummer variabel
Detailbeschreibung	Bestehende gesundheitliche Einschränkungen von Bewerber/innen
Besonderheiten	Angaben zu den Auswirkungen auf die Vermittlung der gesundheitlichen Einschränkungen gibt es nur in der Quelle BewA. Die Variable informiert nur darüber, ob (keine) Einschränkungen vorliegen.  Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4). Das Merkmal wird mit guter Datenqualität gefüllt.

### 4.7 Angaben zu Beschäftigungsverhältnis und Arbeitsuche

#### 4.7.1 Stellung im Beruf und Arbeitszeit

Variablenname	stib
Variablenlabel	Stellung im Beruf und Arbeitszeit
Gefüllt für	BeH
Herkunft	BeH
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Das Merkmal Stellung im Beruf und Arbeitszeit stammt aus den im Rahmen des Meldeverfahrens vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“.  Neben den Auszubildenden, Praktikanten und Volontären zählen auch Anlernlinge, Schüler an Schulen des Gesundheitswesens und Teilnehmer an geförderten Maßnahmen zur beruflichen Fortbildung, Umschulung und betrieblichen Einarbeitung zu der Kategorie „Beschäftigte in beruflicher Ausbildung“ (Wert 0).  Die Differenzierung nach Voll- und Teilzeitbeschäftigten richtet sich nach dem Verhältnis zwischen vertraglich vereinbarter und betriebsüblicher Arbeitszeit. Die Teilzeitbeschäftigten (Werte 8 und 9) sind in zwei Gruppen unterteilt, wobei die Grenze entlang der Hälfte der üblichen Arbeitsstunden eines Vollzeitbeschäftigten verläuft. Sowohl für diese, als auch für Meister und Poliere kann die Zuordnung des Status als Arbeiter oder Angestellter nur über die jeweiligen Rententräger erfolgen.
Besonderheiten	Infolge der Erfassung im Meldeverfahren liegt also nur für Vollzeitbeschäftigte die Stellung im Beruf vor. Für Teilzeitbeschäftigte wiederum ist lediglich der Teilzeitstatus bekannt, aber nicht die Stellung im Beruf.

## 4.7.2 Beruf

Variablenname	beruf
Variablenlabel	Beruf
Gefüllt für	alle Quellen
Herkunft	BeH, BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznr konstant
Detailbeschreibung	<p>Die Ausprägungen der Berufe entsprechen der Berufsklassifikation der Bundesanstalt für Arbeit: „Klassifizierung der Berufe. Systematisches und alphabetisches Verzeichnis der Berufsbenennungen“. Sie enthält ca. 25.000 Berufsbezeichnungen, die Aggregationsebene Berufsordnung besteht aus einer 3-stelligen Kennziffer und umfasst rund 330 Ausprägungen.</p> <p>Bei <b>BeH-Spells</b> stammt die Berufsbezeichnung der aktuell vom Beschäftigten ausgeübten Tätigkeit aus den vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“. Der Arbeitgeber verschlüsselt die Tätigkeit des Beschäftigten hierbei auf Basis der Berufsordnung (3-Steller). Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet.</p> <p>Wie bei allen Merkmalen der BeH gilt, dass bestimmte Gruppen von Erwerbstätigen nicht sozialversicherungspflichtig sind. Folglich ist die Verteilung der Berufe nicht repräsentativ für die Gesamtheit der Erwerbstätigen.</p> <p><b>MTH-</b> und <b>BewA-Spells</b> enthalten den Beruf der letzten Erwerbstätigkeit aus dem BewA. <b>MTH-Spells</b> enthalten diese Information aus dem BewA bezogen auf den Tag vor Beginn der Maßnahme, in den <b>BewA-Spells</b> bezieht sich die Information auf den Stand zu Beginn des Arbeitssuche-Spells.</p>
Besonderheiten	<p>Der hohe Anteil fehlender Werte bei den BewA-Spells erklärt sich z.T. dadurch, dass viele Arbeitsuchende vormals nie oder längere Zeit nicht erwerbstätig waren (siehe Variable Erwerbsstatus vor Arbeitssuche, Abschnitt 4.5.1).</p> <p>Diese Variable ist von der Verbis-Umstellung betroffen (siehe Abschnitt 4.5.4). Das Merkmal wird mit guter Datenqualität gefüllt.</p>

## 4.7.3 Tagesentgelt / täglicher Leistungssatz

Variablenname	tentgelt,
Variablenlabel	Tagesentgelt, täglicher Leistungssatz
Gefüllt für	BeH, LeH
Herkunft	BeH, LeH
Anonymisierung	Keine (tentgelt_d: Vergrößerung)
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>In <b>BeH-Spells</b> weist das Merkmal jeweils das Brutto-Tagesentgelt des Beschäftigten aus. Es wurde aus dem vom Arbeitgeber gemeldeten Zeitraumentgelt und der Dauer des Zeitraums der ungesplitteten Originalmeldung in Kalendertagen berechnet.</p> <p>Im Fall eines Fehlers beim Entgelt wird das Tagesentgelt gelöscht. Der Spell bleibt aber erhalten.</p> <p>In den Jahren 1975 bis 1998 meldeten Arbeitgeber grundsätzlich nur das sozialversicherungspflichtige Arbeitsentgelt, also Entgelt oberhalb der Geringfügigkeitsgrenze und bis zur Beitragsbemessungsgrenze in der gesetzlichen Rentenversicherung. Mit der Einbeziehung der geringfügig Beschäftigten in das Meldeverfahren ab</p>

	<p>01.04.1999 werden auch Entgelte unterhalb der Geringfügigkeitsgrenze ausgewiesen. Die Beitragsbemessungsgrenze als obere Kappungsgrenze gilt jedoch weiterhin. Ausnahmsweise können Entgeltangaben die Beitragsbemessungsgrenze allerdings doch übersteigen. Dies dürfte in der Regel auf Jahressonderzahlungen zurückzuführen sein, die der Arbeitgeber bei den Jahres-, Unterbrechungs- oder Abmeldungen zum normalen Entgelt hinzurechnen kann. Dabei ist es dann unerheblich, ob durch diese Addition die für den Meldezeitraum maßgebliche Beitragsbemessungsgrenze in der Rentenversicherung überschritten wird. Es könnte sich allerdings auch um fehlerhafte Angaben bei Beschäftigungszeit oder Entgelthöhe handeln, was wegen der Versicherungsrelevanz dieser Informationen jedoch eher selten der Fall sein dürfte. Die Geringfügigkeits- und Beitragsbemessungsgrenzen unterscheiden sich von Jahr zu Jahr, nach Rentenversicherungsträger sowie zwischen alten und neuen Bundesländern. Maßgeblich ist dabei jeweils der Standort des Betriebs. Eine Übersicht über diese Grenzen ist am FDZ erhältlich.</p> <p>Bei <b>LeH-Spells</b> handelt es sich um den täglichen Leistungssatz an Arbeitslosengeld, Arbeitslosenhilfe oder Unterhaltsgeld. Dabei ist zu berücksichtigen, dass der tägliche Leistungssatz bei Spells mit einem Original-Beginndatum vor dem 01.01.1998 für Werktage gilt und dass er bei Spells mit einem Original-Beginndatum ab dem 01.01.1998 für Kalendertage gilt.</p>
Besonderheiten	<p>Tagesentgelte und Leistungssätze sind in Euro ausgewiesen.</p> <p>Tagesentgelte gleich oder nahe 0 können dadurch entstehen, dass Unterbrechungsmeldungen vorliegen, bzw. rechtlich bestehende Beschäftigungsverhältnisse ohne Entgeltzahlungen – sog. Ruhende Beschäftigungsverhältnisse (möglich wären Tagesentgelte größer als 0 aber nahe 0, die dadurch entstehen, dass z.B. Sonderausschüttungen einkommenswirksam werden). Ein typisches Beispiel für ein ruhendes Beschäftigungsverhältnis wären Frauen im Erziehungsurlaub.</p>

#### 4.7.4 Gleitzone

Variablenname	gleitz
Variablenlabel	Gleitzone
Gefüllt für	BeH, seit 01.04.2003
Herkunft	BeH
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Das Merkmal zeigt an, ob es sich bei der Beschäftigungsmeldung um eine Beschäftigung im Niedriglohnbereich innerhalb der so genannten Gleitzone mit einem Bruttoarbeitsentgelt von 400 bis 800 Euro handelt, für die der Arbeitnehmer nur einen reduzierten Beitrag zum Gesamtsozialversicherungsbeitrag leisten muss. Die entsprechende gesetzliche Regelung besteht seit 01.04.2003

#### 4.7.5 Wirtschaftszweige (73, 93, 03)

Variablenname	w73, w93, w03
Variablenlabel	Wirtschaftszweig 93; Wirtschaftszweig 73, Wirtschaftszweig 03
Gefüllt für	BeH, seit 1999
Herkunft	BeH
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZXX als 3-Steller aus, also die Wirtschaftsgruppe. WZXX steht für die „Klassifikation der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1993, 1973

	<p>bzw. 2003<sup>4</sup>. Der Wirtschaftszweig nach der 93er Klassifikation ist erst seit 1999 (seit Einführung des neuen Meldeverfahrens) enthalten und löst die Zuordnungssystematik WZ73 ab. 2003 wurde eine erneute Systematik (WZ03) eingeführt.</p> <p>Die Klassifikation der Wirtschaftszweige WZXX, dient dazu, die wirtschaftlichen Tätigkeiten von statistischen bzw. örtlichen Einheiten, "Betrieben", zu erfassen<sup>20</sup>. Unter einer örtlichen Einheit ist ein an einem räumlich festgelegten Ort<sup>21</sup> gelegenes Unternehmen oder Teil eines Unternehmens<sup>22</sup> zu verstehen. Für eine genauere Darstellung siehe u.a. <a href="http://www.statistik-portal.de/Statistik-Portal/klassiWZ93.asp">http://www.statistik-portal.de/Statistik-Portal/klassiWZ93.asp</a> sowie <a href="http://www.statistik-portal.de/Statistik-Portal/de_klassiWZ03.asp">http://www.statistik-portal.de/Statistik-Portal/de_klassiWZ03.asp</a>.</p> <p>Exemplarisch zur WZ93: Diese baut auf der statistischen Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft NACE Rev.1<sup>23</sup> auf, die vier Gliederungsebenen vorsieht und bezüglich der ersten beiden wiederum auf dem internationalen Standard ISIC Rev.3<sup>24</sup> basiert. Um eine Vergleichbarkeit der statistischen Daten zwischen den Mitgliedsstaaten zu gewährleisten, sollte diese grundsätzliche Gliederung vollständig übernommen werden. Die Aufnahme aller Systematiken kann ggf. dazu genutzt werden, zeitkonsistente Zuordnungen für die Auswertungen zu generieren. Alternativ können alte Zuordnungen aber auch durch den Betriebsindikator zeitlich fortgeschrieben werden.</p> <p>Jedem Betrieb darf nur eine Kennziffer zugeordnet werden. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, muss die wirtschaftliche Haupttätigkeit ermittelt werden. Diese orientiert an der Bruttowertschöpfung zu Faktorkosten oder anhand einer geeigneten Ersatzgröße (z.B. Beschäftigtenzahl). Bei der praktischen Arbeit muss sich die BA bei der Zuordnung mit der Selbsteinschätzung der Betriebe begnügen.</p>
Besonderheiten	Der Wirtschaftszweig wird über die Betriebsnummer zugespielt und fehlt, wenn diese ungültig ist.

4.7.6 Bereitschaft zu bundesweiter Stellensuche

Variablenname	mobil
Variablenlabel	Bereitschaft zur bundesweiter Stellensuche
Gefüllt für	BewA
Herkunft	BewA
Anonymisierung	Keine
Zeitbezug	innerhalb der Personennummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>Die Variable beschreibt die regionale Ausgleichsfähigkeit einer Person.</p> <p>Bei Meldungen zur Arbeitsuche muss der/die Arbeitssuchende angeben, ob bundesweit oder regional eingeschränkt Arbeit gesucht wird. Wenn nur ausgewählte Regionen angegeben werden, erfolgt die Nachfrage, ob trotzdem auch andere Regionen in Betracht gezogen werden können. Wenn dies der Fall ist, gilt eine Person als "regional ausgleichsfähig", andernfalls nicht. Bei bundesweiter Suche gilt eine Person dementsprechend automatisch als "ausgleichsfähig".</p>
Besonderheiten	keine

<sup>20</sup> Die in der BA verwendete statistische bzw. örtliche Einheit ist grundsätzlich der in dem „Betriebsdatenverfahren – co-Ber-coStat“ definierte DEVO/DÜVO-Betrieb sowie der von diesem abgeleitete Fachbetrieb.

<sup>21</sup> Der Begriff des räumlich festgelegten Orts ist eng begrenzt: zwei Produktionseinheiten desselben Unternehmens, die sich an verschiedenen Orten befinden, sind selbst wenn beide Orte in derselben Gemeinde liegen, als zwei örtliche Einheiten zu betrachten.

<sup>22</sup> Werkstätte, Werk, Verkaufsladen, Büro, Grube, Lagerhaus, etc.

<sup>23</sup> NACE Rev.1: „Nomenclature générale des activités économiques dans les communautés européennes“, erste überarbeitete Fassung (1990); die erste Version stammt aus dem Jahr 1970.

<sup>24</sup> ISIC Rev.3: „International Standard Industrial Classification of All Economic Activities“, dritte überarbeitete Fassung (1990).

## 4.7.7 BA-Kundengruppe

Variablenname	kunden_gr
Variablenlabel	BA-Kundengruppe
Gefüllt für	BewA
Herkunft	BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Das Merkmal enthält die Eingruppierung der BA-Kunden in verschiedene Typen wie Marktkunde, Beratungskunde (aktivieren), Beratungskunde (fördern) sowie Betreuungskunde. Diese Unterscheidung berücksichtigt die Handlungsbedarfe mit Blick auf die individuellen Integrationshemmnisse.
Besonderheiten	keine

## 4.7.8 Art der Kündigung

Variablenname	art_kuend
Variablenlabel	Art der Beendigung der letzten Tätigkeit
Gefüllt für	BewA
Herkunft	BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Dieses Merkmal beschreibt die Art der Beendigung des letzten Arbeits- oder Ausbildungsverhältnisses vor einer erneuten Arbeitsuche. Es kann daher zur Identifikation von Job-to-Job-Vermittlungen verwendet werden.
Besonderheiten	keine

## 4.7.9 Arbeitszeit des Stellengesuchs

Variablenname	arbzeit
Variablenlabel	Arbeitszeit des Stellengesuchs
Gefüllt für	BewA
Herkunft	BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Arbeitssuchende geben beim Vermittlungsvorgang an, welchen zeitlichen Umfang die gesuchte Stelle annehmen sollte.
Besonderheiten	Bei diesem Merkmal sind VerBIS-Effekte erkennbar, da bei VerBIS die Eingabe nicht obligatorisch ist. Der Füllgrad soll jedoch durch die entsprechenden Arbeitsanweisungen verbessert werden.



## 4.7.10 Restanspruchsdauer Arbeitslosengeld

Variablenname	restanspruch
Variablenlabel	Restanspruchsdauer Arbeitslosengeld
Gefüllt für	LeH
Herkunft	LeH
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>Die Restanspruchsdauer für Arbeitslosengeldzahlungen wird in der LeH ausgewiesen. Sie bezieht sich auf das Enddatum einer Leistungsempfangszeit.</p> <p>Der Restanspruch ist der Anspruch, der nach Ende des aktuellen Leistungsbezugs übrig bleibt. Der Restanspruch bleibt 7 Jahre bestehen.</p> <p>Es ist davon auszugehen, dass die Angabe "Missing" eine Restanspruchsdauer von 0 Tagen ausdrückt.</p>
Besonderheiten	Liegt das Enddatum vor dem 01.01.1998, wird der Restanspruch in Werktagen, danach in Kalendertagen angegeben. Dieser Anspruch zählt maximal bis auf die Höchstdauer für das jeweilige Lebensalter.

## 4.7.11 Geplantes Beginndatum

Variablenname	planbeg
Variablenlabel	Geplantes Beginndatum
Gefüllt für	MTH
Herkunft	MTH
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Geplantes Beginndatum einer Förderperiode. Das Merkmal ist nur für Maßnahmedatensätze gefüllt. Das geplante Beginndatum kann, muss aber nicht, identisch mit dem Beginndatum des Original-Gültigkeitszeitraums sein.
Besonderheiten	keine

## 4.7.12 Geplantes Enddatum

Variablenname	planend
Variablenlabel	Geplantes Enddatum
Gefüllt für	MTH
Herkunft	MTH
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Geplantes Enddatum einer Förderperiode. Das Merkmal ist nur für Maßnahmedatensätze gefüllt. Das geplante Enddatum kann, muss aber nicht, identisch mit dem Enddatum des Original-Gültigkeitszeitraums sein.
Besonderheiten	keine



## 4.8 Angaben zum Beschäftigungsbetrieb

### 4.8.1 Anzahl der hauptbeschäftigten Mitarbeiter

Variablenname	az_hpt
Variablenlabel	Anzahl Hauptbeschäftigte im Betrieb
Gefüllt für	BeH
Herkunft	BHP, Betriebshistorik-Panel
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Die Variable gibt die Anzahl der Mitarbeiter an, die in einem Betrieb nach dem Haupterwerbskonzept tätig waren. Die Hauptbeschäftigung ist dabei definiert als diejenige mit dem höchsten Tagesentgeltsatz. Gibt es für eine Person mehrere Beschäftigungsmeldungen mit dem gleichen Entgeltsatz zum 30. Juni eines Jahres, zählt diejenige mit der längsten Beschäftigungsdauer. Eine geringfügige Beschäftigung wird nur dann als Hauptbeschäftigung gezählt, wenn keine sozialversicherungspflichtige Beschäftigung vorliegt.
Besonderheiten	Diese Information wird generiert. Es handelt sich dabei nicht um ein Datum aus dem Meldeverfahren.

### 4.8.2 Gründungsjahr des Betriebes

Variablenname	<b>grd_dat</b>
Variablenlabel	Gründungsjahr des Betriebes
Gefüllt für	BeH
Herkunft	BHP, Betriebshistorik-Panel
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	Basiert auf dem ersten BLH-Datum des Betriebes.
Besonderheiten	Diese Information wird generiert. Es handelt sich dabei nicht um ein Datum aus dem Meldeverfahren.

## 4.9 Ortsangaben

### 4.9.1 Arbeitsort Bundesland (Kreis)

Variablenname	ao_bula (ao_kreis)
Variablenlabel	Arbeitsort Bundesland; auch Arbeitsort Kreis
Gefüllt für	BeH
Herkunft	BeH
Anonymisierung	keine
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant

Detailbeschreibung	<p>Das Merkmal enthält den Standort des Betriebes, in dem die Person beschäftigt ist. Die Kodierung basiert auf einem 8-stelligen Gemeindegeschlüssel des Statistischen Bundesamtes. Dieser enthält in den Stellen 1-2 das Bundesland, in den Stellen 1-3 den Regierungsbezirk, in den Stellen 1-5 den Kreis und in den Stellen 1-8 die Gemeinde. Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null. Bis 1999 wurde der Betriebsort auf Gemeindeebene verschlüsselt.</p> <p>Da sich die Grenzen der Kreise im Zeitverlauf regelmäßig ändern, treten ohne Gebietsstandsaktualisierungen Fälle auf, bei denen die Gemeindegeschlüsselnummer wechselt, ohne dass der Betrieb seinen Standort verlagert hat. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Angaben der BeH-Spells auf den Gebietsstand vom 31.12.2005 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Betriebsortes zu einer Gemeinde nach den Grenzen zum 31.12.2005.</p>
Besonderheiten	<p>Der Arbeitsort auf Kreisebene (ao_kreis) wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt (→ sensible Merkmale, siehe Abschnitt 3.7). Andernfalls ist lediglich das Bundesland (ao_bula) in den Daten enthalten.</p> <p>Im Jahr 1991 häufen sich fehlenden Werte. Dies ist darauf zurückzuführen, dass in der BeH für 1991 in den neuen Bundesländern fehlerhafte Gemeindegeschlüssel enthalten sind, die auf einen fehlenden Wert umkodiert wurden.</p> <p>Der Arbeitsort wird über die Betriebsnummer zugespielt und fehlt, wenn diese ungültig ist.</p>

#### 4.9.2 Wohnort Bundesland (Kreis)

Variablenname	wo_bula (wo_kreis)
Variablenlabel	Wohnort Bundesland; (auch Wohnort Kreis)
Gefüllt für	BeH (ab 1999), MTH, BewA
Herkunft	BeH, MTH, BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>Das Merkmal enthält den Wohnort der Person. Die Kodierung basiert auf einem 8-stelligen Gemeindegeschlüssel des Statistischen Bundesamtes. Dieser enthält in den Stellen 1-2 das Bundesland, in den Stellen 1-3 den Regierungsbezirk, in den Stellen 1-5 den Kreis und in den Stellen 1-8 die Gemeinde. Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null.</p> <p>Da sich die Grenzen der Kreise im Zeitverlauf regelmäßig ändern, treten ohne Gebietsstandsaktualisierungen Fälle auf, bei denen die Gemeindegeschlüsselnummer wechselt, ohne dass die Person umgezogen ist. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Angaben der <b>BeH-Spells</b> auf den Gebietsstand vom 31.12.2005 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Wohnortes zu einer Gemeinde nach den Grenzen zum 31.12.2005.</p> <p>Für <b>MTH-</b> und <b>BewA-Spells</b> erfolgte keine Gebietsstandsaktualisierung. In 2007 führt dies wegen Kreisgebietsreformen in Sachsen-Anhalt und Thüringen zu neu ausgewiesenen Kreisen und kreisfreien Städten (Kodierungsnummern 14000 bis 14099 sowie 15001 bis 15091). Die Zuordnung zum Agenturbezirk und Bundesland bleibt davon unberührt.</p>
Besonderheiten	<p>Der Wohnort auf Kreisebene (wo_kreis) wird wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt (→ sensible Merkmale, siehe Abschnitt 3.7). Andernfalls ist lediglich das Bundesland (wo_bula) in den Daten enthalten.</p>

## 4.9.3 Arbeitsort Regionaldirektion (Agenturbezirk)

Variablenname	ao_rd (ao_aa)
Variablenlabel	Arbeitsort Regionaldirektion (Agenturbezirk)
Gefüllt für	BeH, MTH, BewA
Herkunft	BeH, MTH, BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>Die Bundesagentur für Arbeit gliedert sich in:</p> <ul style="list-style-type: none"> <li>• Regionaldirektionen (bis 31.12.2003 Landesarbeitsämter),</li> <li>• Agenturen für Arbeit (bis 31.12. 2003 Arbeitsämter),</li> <li>• Geschäftsstellen.</li> </ul> <p>Inhalt dieser Variable ist</p> <ul style="list-style-type: none"> <li>• bei <b>BeH-Spells</b> die Geschäftsstelle des Arbeitsortes,</li> <li>• bei <b>MTH-Spells</b> die Geschäftsstelle des Maßnahmeortes und</li> <li>• bei <b>BewA-Spells</b> ist die statistische Geschäftsstelle, also diejenige, die das Bewerberangebot führt.</li> </ul> <p>Da sich die Grenzen der Geschäftsstellen im Zeitverlauf geändert haben, treten ohne Gebietsstandsaktualisierungen Fälle auf, bei denen die Geschäftsstellenkennung wechselt, ohne dass ein tatsächlicher Ortswechsel vorliegt. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Angaben der <b>BeH-Spells</b> auf den Gebietsstand vom 31.12.2007 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung zu einer Geschäftsstelle nach den Grenzen zum 31.12.2007.</p> <p>Für <b>MTH-</b> und <b>BewA-Spells</b> erfolgte <b>keine</b> Gebietsstandsberichtigung, d.h. die Angaben über die Geschäftsstellen entsprechen dem Stand zu Beginn des Originalspells.</p>
Besonderheiten	<p>Die Ortsangaben auf Agenturebene (ao_aa) werden wegen ihrer besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt (→ sensible Merkmale, siehe Abschnitt 3.7). Andernfalls ist lediglich die Regionaldirektion (ao_rd) in den Daten enthalten.</p> <p>Der Arbeitsort der <b>BeH-Spells</b> wird über die Betriebsnummer zugespielt und fehlt, wenn diese ungültig ist.</p>

## 4.9.4 Wohnort Regionaldirektion (Agenturbezirk)

Variablenname	wo_rd (wo_aa)
Variablenlabel	Wohnort Regionaldirektion (Agenturbezirk)
Gefüllt für	BeH (ab 1999), LeH, MTH, BewA
Herkunft	BeH, LeH, BewA
Anonymisierung	keine
Zeitbezug	innerhalb der Personnummer variabel, innerhalb der Satznummer konstant

Detailbeschreibung	<p>Die Bundesagentur für Arbeit gliedert sich derzeit in:</p> <ul style="list-style-type: none"> <li>• Regionaldirektionen (bis 31.12.2003 Landesarbeitsämter),</li> <li>• Agenturen für Arbeit (bis 31.12. 2003 Arbeitsämter) und</li> <li>• Geschäftsstellen.</li> </ul> <p>Inhalt dieser Variable ist</p> <ul style="list-style-type: none"> <li>• bei <b>BeH-Spells</b> die Geschäftsstelle des Wohnortes (ab 1999 gefüllt),</li> <li>• bei <b>LeH-Spells</b> die für den Leistungsbezug zuständige <b>Arbeitsagentur</b>,</li> <li>• bei <b>MTH-Spells</b> die Geschäftsstelle des Wohnortes aus dem BewA am Tag vor Beginn der Maßnahme und</li> <li>• bei <b>BewA-Spells</b> die Geschäftsstelle des Wohnortes zu Beginn der Arbeitsuche.</li> </ul> <p>Da sich die Grenzen der Geschäftsstellen im Zeitverlauf geändert haben, treten ohne Gebietsstandsaktualisierungen Fälle auf, bei denen die Geschäftsstellenkennung wechselt, ohne dass ein tatsächlicher Ortswechsel vorliegt. Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die Angaben der <b>BeH-Spells</b> auf den Gebietsstand vom 31.12.2007 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung zu einer Geschäftsstelle nach den Grenzen zum 31.12.2007.</p> <p>Für <b>MTH-</b> und <b>BewA-Spells</b> erfolgte <b>keine</b> Gebietsstandsberichtigung, d.h. die Angaben über die Geschäftsstellen entsprechen dem Stand zu Beginn des Orginalspells.</p>
Besonderheiten	<p>Die Ortsangaben auf Agenturebene (wo_aa) werden wegen ihrer besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt (→ sensible Merkmale, siehe Abschnitt 3.7). Andernfalls ist lediglich die Regionaldirektion (wo_rd) in den Daten enthalten.</p>

#### 4.9.5 Wohnort Regionaltyp der Arbeitsagentur

Variablenname	wo_aatyp06
Variablenlabel	Wohnort Regionaltyp der Arbeitsagentur
Gefüllt für	BeH (ab 1999), LeH, MTH, BewA
Anonymisierung	keine
Herkunft	BeH, LeH, BewA
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Satznummer konstant
Detailbeschreibung	<p>Die Arbeitsagenturen werden einem bestimmten Arbeitsmarkttyp zugeordnet. Das Regionalkennzeichen dient zur Typisierung der Regionen bezüglich der dort vorherrschenden Arbeitsmarktlage im Jahr 2006.</p> <p>Das Merkmal enthält die von Blien et al. (2004) entwickelte Typisierung von Arbeitsagenturbezirken bezüglich der dort vorherrschenden Arbeitsmarktlage. Die 12 Vergleichstypen lassen sich zu 5 Strategietypen aggregieren. Die Zuordnung bezieht sich auf den Wohnort (siehe Wohnort Regionaldirektion (Agenturbezirk), Abschnitt 4.9.4).</p>
Besonderheiten	keine

## 4.9.6 Wohnort im Ausland

Variablenname	wo_ausl
Variablenlabel	Wohnort im Ausland
Gefüllt für	BeH (ab 1999), LeH, MTH, BewA
Anonymisierung	keine
Herkunft	BeH, LeH, BewA
Zeitbezug	innerhalb der Personenummer variabel, innerhalb der Fall-ID konstant
Detailbeschreibung	Grundlage ist die Auslandsinformation in der Variablen Wohnort (wo_kreis). Eine Umschlüsselung lässt es zu, Wohnorte im Ausland zu identifizieren. Unterschieden wird in: Europa (differenziert), Afrika, Amerika, Asien, Ozeanien, sonstiges Ausland.
Besonderheiten	Der Wohnort im Ausland ist wegen seiner besonderen datenschutzrechtlichen Sensibilität nur auf Antrag und nur in begründeten Fällen bereitgestellt (→ sensible Merkmale, siehe Abschnitt 3.7).

## 5 Literatur

Blien, Uwe, Franziska Hirschenauer, Manfred Arendt, Hans Jürgen Braun, Dieter-Michael Gunst, Sibel Kilcioglu, Helmut Kleinschmidt, Martina Musati, Hermann Roß, Dieter Vollkommer und Jochen Wein (2004), Typisierung von Bezirken der Agenturen für Arbeit, Zeitschrift für Arbeitsmarktforschung, Jg. 37, H. 2. S. 146-175.

Blien, Uwe; Hirschenauer, Franziska; Arendt, Manfred; Braun, Hans Jürgen; Gunst, Dieter-Michael; Kilcioglu, Sibel; Kleinschmidt, Helmut; Musati, Martina; Roß, Hermann; Vollkommer, Dieter; Wein, Jochen (2004): Typisierung von Bezirken der Agenturen für Arbeit. In: Zeitschrift für ArbeitsmarktForschung, Jg. 37, H. 2, S. 146-175

Engelhard, Astrid, Oberschachtsiek, Dirk und Scioch, Patrycja (2008), Datengeneese zweier Datenkonzepte – MTG (Maßnahme-Teilnahme-Grunddatei) und ISAAK (Instrumente Aktiver Arbeitsmarktpolitik). Eine Betrachtung ausgewählter Fälle am Beispiel der Förderung im Rahmen des ESF-BA-Programms. FDZ Methodenreport 08/2008

Fitzenberger, Bernd und Ralf Wilke (2004), Unemployment Durations in West-Germany Before and After the Reform of the Unemployment Compensation System During the 1980s , ZEW Discussion Paper No. 04-24, Mannheim.

Fitzenberger, Bernd, Aderonke Osikominu und Robert Völter (2005a), Imputation Rules to Improve the Education Variable in the IAB Employment Subsample, ZEW Discussion Paper No. 05-10.

Fitzenberger, Bernd, Michael Lechner, Ruth Miquel, Stephan Wiehler, Conny Wunsch, Martin Biewen, Sonja Lischke, Aderonke Osikominu, Tobias Wenzel (2005b), Die Beschäftigungswirkungen der FbW-Maßnahmen 2000-2002 auf individueller Ebene: Eine Evaluation auf Basis der prozessproduzierten Daten des IAB: Eine Evaluation auf Basis der prozessproduzierten Daten des IAB –vorläufiger, unveröffentlichter Zwischenbericht-.

Jaenichen, Ursula, Thomas Kruppe, Gesine Stephan, Britta Ullrich und Frank Wießner (2005), You can split it if you really want: Korrekturvorschläge für ausgewählte Inkonsistenzen in IEB und MTH, FDZ-Datenreport 4/2005.

Lee, Sokbae und Ralf Wilke (2005), Reform of Unemployment Compensation in Germany: A Non-parametric Bounds Analysis Using Register Data, ZEW Discussion Paper No. 05-29, Mannheim.

Ratgeber zur Sozialversicherung 2005, Broschüre des Deutsche BKK Arbeitgeberservice, [http://www1.deutschebkk.de/Service/Arbeitgeberservice/Wissen\\_von\\_A-Z/5B2D7F649D116FBDC1256DC900510B3E/RatgeberSV\\_2009\\_Internet.pdf](http://www1.deutschebkk.de/Service/Arbeitgeberservice/Wissen_von_A-Z/5B2D7F649D116FBDC1256DC900510B3E/RatgeberSV_2009_Internet.pdf) , S. 32 (Link vom 04.05.2009)

## 6 Anhang

### 6.1 Abkürzungsverzeichnis

AA	Agentur für Arbeit/ Arbeitsamt
AB	Arbeitsberatung
ABM	Arbeitsbeschaffungsmaßnahmen
AEZ	Arbeitsentgeltzuschuss
AG	Arbeitgeber
ALG	Arbeitslosengeld
ALHI	Arbeitslosenhilfe
ASU	Arbeitsuche
AU	Arbeitsunfähigkeit
AV	Arbeitsvermittlung
BA	Bundesagentur für Arbeit
BeH	Beschäftigten-Historik
BewA	Bewerberangebot
BHD	Beschäftigten-Historik-Datei
BHI	Beschäftigungshilfe für Langzeitarbeitslose
BLH	Beschäftigten- und Leistungsempfänger-Historik
BSI	Beschäftigung schaffende Infrastrukturförderung
coArb	Computerunterstützte Arbeitsvermittlung (operatives Verfahren zur Verwaltung der Vermittlung (Altverfahren))
coLei	Computerunterstützte Leistungsgewährung (operatives Verfahren zur Leistungsgewährung (Altverfahren))
coLibri	Operatives Verfahren zur Leistungsgewährung (Neuverfahren)
coSach	Operatives Verfahren zur Verwaltung von Maßnahmen und Teilnahmen
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und –übermittlungsverordnung -
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung –
DSL	Deutschsprachlehrgang
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung –
DWH	Data Warehouse
EGHI	Eingliederungshilfe
EGN	Eingliederungszuschuss bei Neugründung
EGZ	Eingliederungszuschuss
ESF	Europäischer Sozialfonds
EV	Eingliederungsvertrag
ExGZ	Existenzgründerzuschuss
EZV	Einstellungszuschüsse bei Vertretung
FbW	Förderung beruflicher Weiterbildung
FDZ	Forschungsdatenzentrum
FELEG	Gesetz zur Förderung der Einstellung der landwirtschaftlichen Erwerbstätigkeit

---

FF	Freie Förderung
HOGA	Hotel- und Gaststättenvermittlung
IAB	Institut für Arbeitsmarkt- und Berufsforschung
IABS	IAB-Beschäftigtenstichprobe
KuG	Kurzarbeitergeld
LeH	Leistungsempfänger-Historik
LIAB	Linked-Employer-Employee-Daten des IAB
LKZ	Lohnkostenzuschuss
LZA	Langzeit-Arbeitslosigkeit
MTH	Maßnahme-Teilnehmer-Gesamtdatenbank
NACE	Allgemeine Systematik der Wirtschaftszweige in den Europäischen Gemeinschaften des Statistischen Amtes der Europäischen Gemeinschaften
PSA	Personal-Service-Agentur
SAM	Strukturanpassungsmaßnahmen
SchwBg	Gesetz zur Sicherung der Eingliederung Schwerbehinderter in Arbeit, Beruf und Gesellschaft – Schwerbehindertengesetz-
SGB	Sozialgesetzbuch
TM	Trainingsmaßnahme
UEG	Überbrückungsgeld
UHG	Unterhaltsgeld
VDR	Verband deutscher Rentenversicherungsträger
ZND	Zentrale Nachweisdatei
zPDV	Zentrale Personen-Daten-Verwaltung

## 6.2 Merkmalsliste alphabetisch

Variablen-name	Seite	Bezeichnung
persnr		Personennummer
satznr		Satznummer
betnr		Betriebsnummer
begorig		Beginndatum Originalspell
endorig		Enddatum Originalspell
begepi		Beginndatum der gesplitteten Episode
endepi		Enddatum
quelle		Quelle der Information
kom_quel		Kombination der Quellen
spell		Spellzähler pro Konto
nspell		Anzahl der Spells pro Konto
level2		Spellzähler pro Episode
nlevel2		Anzahl der Spells pro Episode
level1		Spellzähler pro Episode und Quelle
nlevel1		Anzahl der Spells pro Episode und Quelle
estatvor		Erwerbsstatus vor Arbeitsuche



erwstat		Erwerbsstatus: Personengruppe, Leistungsart, Maßnahmeart, Arbeitsuche-Status
grund		Grund des Spellendes
sna		Status nach Abgang
gebjahr		Geburtsjahr
sex		Geschlecht
nation		Staatsangehörigkeit
schweb		Schwerbehindertenstatus
schbild		Schulabschluss
bild		Ausbildung
quali		Qualifikation des Bewerbers
famstand		Familienstand
kind		Anzahl Kinder im Haushalt
gesund_ein		gesundheitliche Einschränkungen
stib		Stellung im Beruf und Arbeitszeit
beruf		Beruf
tentgelt		Tagesentgelt / täglicher Leistungssatz
gleitz		Gleitzone
w73		Wirtschaftszweig 73
w93		Wirtschaftszweig 93
w03		Wirtschaftszweig 03
mobil		Bereitschaft zu bundesweiter Stellensuche
kunden_gr		BA-Kundengruppe
art_kuend		Art der Beendigung der letzten Tätigkeit
arbzeit		Arbeitszeit des Stellengesuchs
restanspruch		Restanspruchsdauer Arbeitslosengeld
planend		Geplantes Enddatum einer Förderperiode
planbeg		Geplantes Beginndatum einer Förderperiode
az_hpt		Anzahl Hauptbeschäftigte im Betrieb
grd_jahr		Gründungsjahr des Beschäftigungsbetriebes
ao_bula		Arbeitsort Bundesland (Kreis)
wo_bula		Wohnort Bundesland (Kreis)
ao_rd		Arbeitsort Regionaldirektion (Arbeitsagentur)
wo_rd		Wohnort Regionaldirektion (Arbeitsagentur)
wo_kreis		Wohnort Kreis
ao_kreis		Arbeitsort Kreis
ao_aa		Arbeitsort Agenturbezirk
wo_aa		Wohnort Agenturbezirk
wo_aatyp06		Wohnort Regionaltyp der Arbeitsagentur 06
wo_ausl		Wohnort im Ausland

**Imprint****FDZ *Datenreport***

No. 3/2009

**Publisher**

The Research Data Centre (FDZ)  
of the Federal Employment Service  
in the Institute for Employment Research  
Regensburger Str. 104  
D-90478 Nuremberg

**Editorial staff**

Stefan Bender, Dagmar Herrlinger

**Technical production**

Dagmar Herrlinger

**Copyright**

Reproduction – also in parts – only with permission of the FDZ

**Download**[http://doku.iab.de/fdz/reporte/2009/DR\\_03-09.pdf](http://doku.iab.de/fdz/reporte/2009/DR_03-09.pdf)**Internet**<http://fdz.iab.de/>**Corresponding author**

Jörg Heining, Institute for Employment  
Research, Regensburger Str. 104,  
D-90478 Nuremberg  
Phone: +49-(0)911/179-5392  
E-Mail: [Joerg.Heining@iab.de](mailto:Joerg.Heining@iab.de)