

# IAB *Werkstattbericht*

---

Diskussionsbeiträge des Instituts für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit  
Ausgabe Nr. 10 / 21.10.2003

## **Von Verwaltungsdaten zu Forschungsdaten**

Die Individualdaten für die Evaluation des ESF-BA-Programms  
2000 bis 2006

*Thomas Kruppe, Martina Oertel*



**Bundesanstalt für Arbeit**

---

## Die letzten Ausgaben des **IAB** *Werkstattbericht* im Überblick

---

- Nr. 15      **Steigerung von Effizienz und Reputation in der Arbeitsvermittlung**  
21.10.2002      Fragen der Privatisierung oder Modernisierung im Spiegel internationaler Ansätze und Erfahrungen
- Nr. 16      **Neuordnung der Arbeitslosenhilfe im Rahmen eines dreistufigen Systems**  
25.11.2002
- Nr. 17      **Einkommen von Männern und Frauen beim Berufseintritt**  
17.12.2002      Betriebliche Ausbildung und geschlechtsspezifische berufliche Segregation in den 90er Jahren
- Nr. 18      **Dreifache Heterogenität von ABM und SAM und der Arbeitslosigkeitsstatus der Teilnehmer sechs Monate nach Programm-Ende – Erste deskriptive Befunde**  
18.12.2002
- Nr. 1      **Beschäftigungsfähigkeit als Evaluationsmaßstab?**  
10.2.2003      Inhaltliche und methodische Aspekte der Wirkungsanalyse beruflicher Weiterbildung im Rahmen des ESF-BA-Programms
- Nr. 2      **Evaluation der individuellen Netto-Effekte von ABM in Deutschland**  
10.3.2003      Ein Matching-Ansatz mit Berücksichtigung von regionalen und individuellen Unterschieden
- Nr. 3      **Substitution von bezahlten Überstunden durch Arbeitslose**  
14.3.2003      Das BAFF-Modell im Arbeitsamtsbezirk Neuwied – eine vertiefende Analyse
- Nr. 4      **Systeme des Leistungsbezugs bei Arbeitslosigkeit**  
31.3.2003      - Ein zwischenstaatlicher Vergleich -
- Nr. 5      **Ältere ab 55 Jahren: Erwerbstätigkeit, Arbeitslosigkeit und Leistungen der Bundesanstalt für Arbeit**  
16.4.2003
- Nr. 6      **Licht am Ende des Tunnels?**  
17.4.2003      Eine aktuelle Analyse der Situation schwerbehinderter Menschen am Arbeitsmarkt
- Nr. 7      **Green Card, IT-Krise und Arbeitslosigkeit**  
20.5.2003      Mit einer Kündigung verlieren ausländische IT-Fachkräfte oft mehr als nur ihren Job – Eine Fallstudie in München
- Nr. 8      **Trainingsmaßnahmen - Was verbirgt sich dahinter?**  
6.6.2003
- Nr. 9      **Förderung auf hohem Niveau**  
7.8.2003      Das Jugendsofortprogramm zum Abbau der Jugendarbeitslosigkeit – 1999 bis 2002

*Die Reihe "IAB Werkstattbericht" gibt es seit 1991. Eine vollständige Themenübersicht finden Sie in den „Veröffentlichungen“ des Instituts für Arbeitsmarkt- und Berufsforschung (Tel. 0911/179-3025).*

---

### **IAB** *Werkstattbericht*

Nr. 10 / 21.10.2003

#### **Redaktion**

Ulrich Möller, Elfriede Sonntag

#### **Graphik & Gestaltung**

Monika Pickel, Elisabeth Strauß

#### **Technische Herstellung**

Hausdruckerei der BA

#### **Rechte**

Nachdruck - auch auszugsweise - nur mit Genehmigung des IAB gestattet

#### **Bezugsmöglichkeit**

Institut für Arbeitsmarkt- und Berufsforschung,  
Regensburger Str. 104,  
D-90327 Nürnberg  
Tel.: 0911/179-3025

#### **IAB im Internet:** <http://www.iab.de>

Dort finden Sie u.a. ausgewählte **IAB** *Werkstattberichte* im Volltext zum Download

#### **Rückfragen zum Inhalt an**

Thomas Kruppe, Tel. 0911/179-5649  
oder e-Mail: [thomas.kruppe@iab.de](mailto:thomas.kruppe@iab.de)

**ISSN** 0942-1688

# Von Verwaltungsdaten zu Forschungsdaten

## Die Individualdaten für die Evaluation des ESF-BA-Programms 2000 bis 2006

### Inhaltsverzeichnis

Kurzfassung .....	4
<b>1. Einleitung</b> .....	<b>5</b>
<b>2. Rahmenvorgaben für eine ESF-BA-Individualdatenbank</b> .....	<b>6</b>
<b>3. ESF-BA-Geschäftsstatistik</b> .....	<b>10</b>
3.1 Gemeinsamkeiten und Unterschiede zwischen alter und neuer Geschäftsstatistik .....	10
3.2 St 38 und BA-Statistik zu den regulären Förderinstrumenten .....	11
<b>4. Der Aufbau einer Individualdatenbank</b> .....	<b>11</b>
4.1 Die Ausgangslage: Fallbezogene Einzeldatensätze .....	12
4.2 Erstes Ziel: Teilnahmebezogene Informationen und Historisierung .....	14
4.3 Zweites Ziel: Personenbezogene Informationen .....	17
<b>5. Probleme der Datengense und Lösungswege bei der Datenaufbereitung</b> .....	<b>18</b>
5.1 Analyse der zeitlichen Differenz zwischen Eintritt und Datenziehung .....	19
5.2 Datenanreicherung für Evaluationszwecke .....	21
5.3 Datenbereinigung .....	22
<b>6. Erste Zusammenführung personenbezogener Einzeldatensätze</b> .....	<b>23</b>
<b>7. Schlussfolgerungen</b> .....	<b>26</b>
Literatur .....	27
Quellen .....	28

## Kurzfassung

Gegenstand dieses Berichts ist die Darstellung des Aufbaus einer Individualdatenbank für erweitertes Monitoring und Evaluation im Rahmen der Begleitforschung zum ESF-BA-Programm. Da es sich um Daten handelt, die in der Bundesanstalt für Arbeit im Rahmen des Geschäftsprozesses anfallen (prozessproduziert), sind die dabei auftretenden Fragestellungen und Probleme durchaus von genereller Bedeutung. Denn die Nutzung von prozessproduzierten Daten zur Evaluation von arbeitsmarktpolitischen Maßnahmen spielt eine immer größere Rolle. So fallen z. B. bestimmte Probleme bei der Nutzung von Prozessdaten weg, die sich bei Auswertungen von Befragungsdaten ergeben, (u.a. geringe Fallzahlen, die eine tiefere Gliederung nicht zulassen). Welche Besonderheiten im Gegenzug aber gerade bei Daten existieren, die im Rahmen des Geschäftsprozesses erhoben werden und deshalb in erster Linie nicht auf wissenschaftliche Evaluationen zugeschnitten sind, soll hier thematisiert werden. Insbesondere spielt der Umbau der Daten von fallbezogenen Einzeldatensätzen hin zu personenbezogenen historisierten Verläufen eine Rolle. Es werden aber auch Probleme und deren Auswirkungen bzw. Lösungsmöglichkeiten vorgestellt, die speziell durch den Prozess der dezentralen Eingabe der Daten und deren monatlicher zentraler Zusammenführung entstehen.

Insgesamt geht es hier nicht um eine technische Darstellung des Datenbankaufbaus. Vielmehr soll beispielhaft für mögliche Probleme bei anderen prozessproduzierten Daten sensibilisiert werden und insofern z. B. notwendige Datentransformationsprozesse eher aus inhaltlicher Sicht verständlich gemacht werden. Gerade auch wegen der weiterhin anstehenden Umstrukturierung der Datenproduktion in der Bundesanstalt für Arbeit („Virtueller Arbeitsmarkt“ etc.) erscheint dies unabdingbar.

Die Daten, auf die sich dieser Bericht bezieht, decken den Zeitraum von 3 Jahren – Januar 2000 bis Dezember 2002 – ab und berücksichtigen Dateneingänge bis Februar 2003, die sich auf diesen Zeitraum beziehen („Nachläufer“). Dies entspricht der Hälfte der Laufzeit des ESF-BA-Programms. Die Aufbereitung der Daten erfolgte auch mit Blick auf die Generierung von Vergleichsgruppen für eine im zweiten Halbjahr 2003 anstehende Befragung von Teilnehmer/innen und Nichtteilnehmer/innen an Maßnahmen zur Förderung der beruflichen Weiterbildung<sup>1</sup> und bilden den Grundstock für weitergehende Kausalanalysen.

---

<sup>1</sup> Vgl. hierzu Deeke/Kruppe (2003).

# 1. Einleitung

Die Evaluation arbeitsmarktpolitischer Maßnahmen mit mikro-ökonomischen Methoden hat in den letzten Jahren immer mehr an Bedeutung gewonnen. Eines der Probleme ist dabei oftmals der Rückgriff auf Datenquellen mit nur geringem Stichprobenumfang.<sup>2</sup> Evaluationen können aber auch direkt auf die im Geschäftsprozess der Bundesanstalt für Arbeit anfallenden, prozessproduzierten Daten aufsetzen, so auch ein Teil der Evaluation der die Arbeitsmarktpolitik des Bundes (Regelförderung nach SGB III) ergänzenden Förderung durch den Europäischen Sozialfond (ESF-BA-Programm).<sup>3</sup> Solche Prozessdaten weisen gegenüber anderen Quellen neben Vorteilen auch spezifische Probleme auf. So sind sie in ihrer Rohform z. B. weder historisch abgelegt – also entlang einer Zeitachse – noch ist daraus der Gesamtverlauf einer individuellen Maßnahmeteilnahme direkt abzulesen. Deshalb spielt der Umbau der Daten von fallbezogenen Einzeldatensätzen über maßnahmebezogene Kombinationen hin zu personenbezogenen historisierten Verläufen im Folgenden eine besondere Rolle (zu den Definitionen vgl. Kapitel 4).<sup>4</sup>

Des Weiteren werden Probleme, deren Auswirkungen und entsprechende, beispielhafte Lösungsmöglichkeiten vorgestellt, die speziell durch den Prozess der dezentralen Eingabe der Daten (in den Arbeitsämtern vor Ort) und deren monatlicher zentralen Zusammenführung entstehen. Da diese Spezifika – zumindest in abgewandelter Form – auch für die Nutzung von Einzeldaten für die Evaluation anderer arbeitsmarktpolitischer Maßnahmen relevant sind, sollen sie beispielhaft in diesem Bericht vorgestellt werden – auch, um andere für einen entsprechenden Umgang mit solchen Daten zu sensibilisieren.

Die Begleitforschung zum sogenannten ESF-BA-Programm,<sup>5</sup> bei dem der Förderung von Personen durch die Bundesanstalt für Arbeit (BA) mit Instrumenten des SGB III zusätzlich Mittel aus dem Europäischen Sozialfond (ESF) zur Seite gestellt werden, begann im Herbst 2000.<sup>6</sup> Die Erschließung der individuellen Förderdaten und der Aufbau einer „Individualdatenbank“ startete nach konzeptionellen Vorarbeiten im Frühsommer 2001, nachdem die fallbezogenen Einzeldatensätze zu Zugängen, Beständen, Abgängen und Verbleib der ESF-BA-Förderfälle dem IAB zugänglich gemacht worden waren. Die Aufbereitung und Validierung dieser Daten erfolgt monatlich fortlaufend mit den jeweils aktuellen Daten aus der Förderung. Nach der bereits erwähnten Transformation der fallbezogenen Datensätze zu teilnahme- und anschließend zu personenbezogenen Datensätzen wurden bzw. werden diese mit weiteren Merkmalen aus der zentralen Bewerberangebotsdatei der Bundesanstalt für Arbeit komplettiert. Weiterhin wurden die Arbeiten an einer Verknüpfung mit den Einzeldatensätzen der BA zur gesetzlichen Förderung beruflicher Weiterbildung, von Trainingsmaßnahmen und zum Überbrückungsgeld für Existenzgründer<sup>7</sup> durchgeführt. Erst damit wurden Grundlagen für vergleichende Analysen des Einsatzes der arbeitsmarktpolitischen Instrumente mit und ohne ESF-Ergänzung geschaffen und die Ausgangsbasis für mikroanalytische Wirkungsanalysen hergestellt.

---

<sup>2</sup> Vgl. hierzu die Übersicht in Fitzenberger/Speckesser (2002)

<sup>3</sup> Weitere Projekte, die auf prozessproduzierte Daten zugreifen, sind z. B. die Evaluation der Netto-Effekte von ABM in Deutschland von Caliendo/Hujer/Thomsen (2003) und die Evaluation von Lohnkostenzuschüssen von Jaenichen (2003).

<sup>4</sup> An dieser Stelle gilt besonderer Dank den Mitarbeiter/innen des IAB-Arbeitsbereichs 9, ohne deren tatkräftige Unterstützung der Aufbau der Individualdatenbank nicht möglich gewesen wäre.

<sup>5</sup> Vgl. hierzu auch Kapitel 2. Zur weitergehenden Erläuterung des ESF-BA-Programms vgl. Deeke/Kruppe/Müller/Schuler (2002) und Deeke/Kruppe/Kurtz/Müller/Schuler (2003).

<sup>6</sup> Zum Konzept der Begleitforschung vgl. Deeke/Kruppe (2002). Auch zum Vorläuferprogramm „AFG-Plus“ (1995 – 1999) führte das IAB die Begleitforschung durch – „notgedrungen“ auf einer damals weitaus schlechteren Datengrundlage als heute (vgl. Deeke, Schuler 2003).

<sup>7</sup> Zu den Daten und ersten Ergebnissen der Analyse einer Zusammenführung von ESF-BA-Daten mit Daten zur Teilnahme an Trainingsmaßnahmen vgl. Kurtz (2003) und zur Existenzgründungsförderung vgl. u.a. Kurtz (2002).

Zusammengefasst dient der Aufbau der Individualdatenbank folgenden Zielen:

- Differenzierte Zielgruppenanalyse
- Regional differenzierte Analysen
- Verbleibsanalysen
- Vergleich der Teilnahme an Maßnahmen zur Förderung beruflicher Weiterbildung und Trainingsmaßnahmen jeweils mit und ohne ESF-Ergänzung
- Bildung von Vergleichsgruppen für mikroanalytische Wirkungsanalysen – auch als Grundlage für ergänzende repräsentative Befragungen von Teilnehmer/innen und Nichtteilnehmer/innen

Da die prozessproduzierten Datensätze zur gesetzlichen Arbeitsförderung nicht unmittelbar für Forschungszwecke nutzbar sind, müssen diese ebenso wie die Einzeldaten zum ESF-BA-Programm mit erheblichem Aufwand erst erschlossen und für die Analysezwecke aufbereitet werden. In dieser Hinsicht hat die Begleitforschung zum ESF-BA-Programm ähnlich wie andere Projekte des IAB den Charakter eines Pilotprojekts für die Erschließung der Einzeldaten der gesetzlichen Arbeitsförderung als zentrale Voraussetzung für deren wissenschaftliche Evaluation.<sup>8</sup>

Nach einer Beschreibung der Rahmenvorgaben und der ESF-BA-Geschäftsstatistik werden im Folgenden zunächst der konzeptionelle Aufbau der Individualdatenbank und die entsprechenden Umsetzungsarbeiten vorgestellt.

## 2. Rahmenvorgaben für eine ESF-BA-Individualdatenbank

Der Europäische Sozialfond (ESF) soll als Teil der Europäischen Strukturfonds als qualitative Ergänzung zur nationalen aktiven Arbeitsmarktpolitik (die in Deutschland sowohl vom Bund als auch von den Bundesländern umgesetzt wird) eingesetzt werden. Inhaltlich lehnt sich die von der Europäischen Kommission erwünschte zusätzliche Förderung im Rahmen der ESF-Ziele an die jeweilig aktuellen Beschäftigungspolitischen Leitlinien im Rahmen der Europäischen Beschäftigungsstrategie an, deren Umsetzung in Nationalen Aktionsplänen (NAP) jährlich geplant wird. Die ESF-Förderung soll dabei der Umsetzung der Beschäftigungspolitischen Leitlinien und NAP dienen. Bund und Länder haben dazu übergreifende „Dachprogramme“ formuliert, in denen neben qualitativen Zielen auch quantitative Inputziele vereinbart sind. In diesem Rahmen sind wiederum die Einzelprogramme des Bundes und der Länder eingebettet.<sup>9</sup>

Die Zielsteuerung liegt während der Laufzeit von 2000 bis 2006 in Form von Entscheidungen über die konkreten Programme und deren Umsetzung dezentral auf der Ebene der Mitgliedstaaten und deren Regionen, in Deutschland also bei Bund und Bundesländern. Dies geht gleichzeitig einher mit einer – gegenüber dem Vorläuferprogramm – erheblichen Ausweitung des Monitoring (einschließlich Finanz-

---

<sup>8</sup> Vgl. hierzu auch: Deeke/Wiedemann (2002)

<sup>9</sup> Für Ostdeutschland als Ziel 1-Gebiet gibt es das Gemeinschaftlichen Förderkonzept (GFK) von Bund und Ländern (bezogen auf alle EU-Strukturfonds), wobei die entsprechende Planung der Umsetzung je Bundesland und gesondert für den Bund in einem jeweiligem Operationellem Programm (OP) geregelt ist. Das OP des Bundes wird ausschließlich aus dem ESF kofinanziert. Für die ESF-Förderung in Westdeutschland als Ziel 3-Gebiet wurde ein Einheitliches Programmplanungsdokument (EPPD) von Bund und Ländern beschlossen. Konkrete Umsetzungsregelungen und Ziele zu diesen „Dachprogrammen“ sind wiederum in jeweiligen Programmergänzungsdokumenten ausgeführt. Das bundesweite ESF-BA-Programm ist darin als Einzelprogramm des Bundes enthalten. Vgl. zu diesem auf mehreren Ebenen „verflochtenen“ Programmzusammenhängen und daraus resultierenden Umsetzungsaspekten: Deeke (2003).

kontrolle) mit verstärkter Berichtspflicht, um sowohl den Beteiligten bei der Umsetzung eine effektivere Steuerung und der Europäischen Kommission ein systematisches Zielcontrolling zu ermöglichen. Gleichzeitig wird mit dem Monitoring eine Datenbasis für die Bewertung (Evaluierung) erzeugt.<sup>10</sup>

Zur Begleitung und Beobachtung der Umsetzung aller ESF-Programme wurde ein Indikatorensystem entwickelt, das im Rahmen des Programmergänzungsdokuments bzw. im jeweiligen OP der Länder beschrieben ist. So soll mit Hilfe dieser Indikatoren die Leistung des Programms hinsichtlich der ESF-Zielvorgaben beurteilt werden können. Auch sollen sie als gemeinsame Basis von Monitoring und Evaluation in Bund und Ländern umgesetzt werden, um eine möglichst weitgehende Vergleichbarkeit gerade auch auf aggregierter Ebene zu erreichen.<sup>11</sup>

Diese Indikatoren sind in sechs Gruppen gegliedert.<sup>12</sup> Im einzelnen handelt es sich um Kontext-, Input-Verlaufs- oder Output-Indikatoren und Ergebnisindikatoren im Rahmen der Datenbeschaffung und Berichterstattung (Monitoring), während sich die Evaluation auf die Wirkungs- und Effizienz-Indikatoren konzentriert.

Mit Hilfe der Kontextindikatoren soll innerhalb des Umsetzungsrahmens der ESF-kofinanzierten Programme die Entwicklung des Arbeitsmarktes und der Beschäftigung beobachtet und die (weitere) Planung entsprechend ausgerichtet werden.

Die Input-Indikatoren sollen die Planungszahlen für

- Finanzmittel,
- Teilnehmer/-innen insgesamt und
- von Beratungsstellen und Projekten

erfassen und damit die einzubringenden Ressourcen nach Schwerpunkten/Politikzielen und Maßnahmen abbilden,

Die Verlaufs- bzw. Output-Indikatoren (Indikatoren zum finanziellen und materiellen Verlauf, Kontingentindikatoren) sollen durch den Vergleich mit den Input-Indikatoren die Möglichkeit bieten, recht zeitnah Fehlentwicklungen in Finanzströmen oder Teilnehmerstrukturen zu erkennen und falls erforderlich korrigierend einzugreifen.

Für die Ergebnis-Indikatoren ist das Monitoring innerhalb des Förderzeitraumes auch verpflichtet, entsprechende Individualdaten zu erheben, und zwar

- zum einen für Abbruchs- und Erfolgs-Indikatoren, die individuelle Ereignisse im Verlauf bis zum Ende der Maßnahmeteilnahme aufzeichnen,
- zum anderen für Übergangs- und Verbleibs-Indikatoren, die nach Beendigung der Maßnahme erhoben werden und den Verbleib bzw. die Beschäftigungssituation nach einem bzw. sechs Monat(en) abbilden (Stichtagskonzept).

---

<sup>10</sup> Zum Zusammenhang zwischen Steuerungsansatz und Monitoringanforderungen sowie der Rolle von Evaluation bei arbeitsmarktpolitischen Programmen vor dem Hintergrund auch der Entwicklung der EU-Strukturfonds vgl. Deeke/Kruppe (2002), zur Entwicklung von Monitoring im Zusammenspiel mit Evaluation in Europa vgl. Auer/Kruppe (1996).

<sup>11</sup> Vgl. zum Feedback-System im Rahmen des Konzepts der reflexiven Steuerung des ESF: Deeke (2003).

<sup>12</sup> Quelle: Programmergänzungsdokument, Stand vom 26.03.2001, Seite 21 f.

Eine zusätzliche Bedeutung kommt den Indikatoren in Hinsicht auf die Verteilung der sogenannten leistungsgebundenen Reserve<sup>13</sup> zu. Zu Beginn der Programmlaufzeit wird von jedem geförderten Mitgliedsland der Europäischen Union ein Betrag von 4 % der ESF-Mittel zurückbehalten. Diese sollen im Rahmen der Halbzeitbewertung in diesem Jahr, also nach Ablauf der halben Planungsperiode, durch die Kommission in enger Zusammenarbeit mit dem jeweiligen Mitgliedsstaat auf die leistungsstarken Programme aufgeteilt werden. Dazu wurden in den Programmergänzungen zur ESF-Umsetzung in West- und Ostdeutschland spezifische Indikatoren vorgeschlagen.

Bestimmte Indikatoren, namentlich die Wirkungs- und Effizienz-Indikatoren, werden ausschließlich für die Evaluation benötigt und auch von dieser erhoben, entsprechend der Vorgabe, dass *„durch das Monitoring die materiellen und finanziellen Verlaufsdaten sowie die individualisierten Daten zu den Ergebnisindikatoren (...) erfasst (werden), während die Evaluation für die Erfassung der Wirkungen und die Analyse der Effizienz der Förderung zuständig ist“* (EPPD Ziel 3: 204).

Zu dieser – im Unterschied zu früheren Programmen deutlich verstärkten – Schaffung einer Datengrundlage für Begleitung und Bewertung in Form eines sogenannten „Stammblattverfahrens“ heißt es im EPPD:

*„Eine wesentliche Erweiterung gegenüber den bisherigen Begleitungs- und Bewertungssystemen betrifft die Erfassung von individuellen Teilnehmerdaten bei personenbezogenen Maßnahmen; diese Erweiterung ist erforderlich, um eine Aggregation und Auswertung von Daten bis auf EPPD- bzw. GFK-Ebene zu ermöglichen bzw. zu verbessern. Für jeden Maßnahmeteilnehmer wird ein individueller Datensatz angelegt und ereignisbezogen (= Eintritt, Austritt und Austrittsgründe, Verbleib nach Maßnahmeende) fortgeschrieben. Dieser Individualdatensatz muss so aufgebaut sein, dass für die Auswertung eine freie Kombinierbarkeit aller Angaben gewährleistet ist und zugleich die Datenschutzbelange berücksichtigt werden. Der konkrete Umfang der zu erfassenden Teilnehmer- und anderen Daten wird für jeden Maßnahmetyp gesondert definiert.“* (EPPD Ziel 3: 206).

Dieser im Vergleich zur Regelförderung geforderte hohe Aufwand an Datenerfassung und Aufbereitung soll hier nicht weiter diskutiert werden.<sup>14</sup> Zu berücksichtigen ist aber, dass sowohl zu einem ausgereiften Monitoring als auch zur Evaluation eine entsprechende Datengrundlage wichtige Voraussetzung ist. Die Anforderungen an diese Datengrundlage kann in Teilbereichen jedoch durchaus unterschiedlich zwischen Monitoring und Evaluation sein. Deshalb kann es durch eine einheitliche Vorgabe u. U. zu Dysfunktionalitäten in beide Richtungen kommen: Einem (zu) hohen Verwaltungsaufwand stehen (zu) wenig Informationen auf der anderen Seite gegenüber.

Für das ESF-BA- Programm (vgl. **Kasten**) hat die Bundesanstalt für Arbeit als programmumsetzende Organisation entsprechend den ESF-Regelungen zur Berichtserstattung auch die Aufgabe des spezifischen Monitoring. Die Begleitforschung durch das IAB unterstützt dabei die zuständigen Referate in der Hauptstelle und erarbeitet für ihre eigenen Forschungszwecke eine Individualdatenbank, die aufbauend auf dem sogenannten „Stammblattverfahren“ des Monitoring strukturiert wird. Im Folgenden wird der tatsächliche Aufbau der Geschäftsstatistik der BA zum ESF-BA-Programm erläutert.

<sup>13</sup> Vgl. hierzu Verordnung (EG) 1260/1999 des Rates vom 21. Juni 1999.

<sup>14</sup> Vgl. dazu Deeke (2000).

## Das ESF-BA-Programm 2000-2006

*Die aktive Arbeitsförderung des dritten Sozialgesetzbuches (SGB III) wird im Rahmen des ESF-BA-Programms zusätzlich aus Mitteln des Europäischen Sozialfonds (ESF) ergänzt.\* Mit diesem Programm wird eine individuelle Teilnahme an Maßnahmen gefördert, bei der in jedem einzelnen Fall Mittel des SGB III und des ESF zusammen kommen müssen. Zielgruppen sind insbesondere Arbeitslose (darunter besonders von Langzeitarbeitslosigkeit Bedrohte und Ältere), sowie Arbeitslose mit Migrationshintergrund und Berufsrückkehrerinnen. Es sollen dabei nicht nur, aber insbesondere Personen ohne gesetzlichen Anspruch auf Leistungen zum Lebensunterhalt nach dem SGB III einbezogen werden.*

*Ein Schwerpunkt des Programms lag im Zeitraum 2000 bis 2002 bei der Förderung beruflicher Weiterbildung (FbW), mit dem Ziel der Erhaltung bzw. der Erhöhung der individuellen Beschäftigungsfähigkeit und deren Realisierung durch einen anschließenden Übergang in (nicht geförderte) Beschäftigung.\*\* Die Regelförderung der beruflichen Weiterbildung wurde durch das ESF-BA-Programm in zweierlei Hinsicht ergänzt:*

- *Erstens durch die Förderung sogenannte „Nichtleistungsbezieher/innen“ in Weiterbildungsmaßnahmen - also von Personen, die aufgrund fehlender Vorbeschäftigungszeiten kein Unterhaltsgeld nach dem SGB III beziehen konnten. Für sie konnte während der Teilnahme an einer Weiterbildungsmaßnahme ein ESF-Unterhaltsgeld gezahlt.*
- *Zweitens durch die Finanzierung der Teilnahme an zusätzlicher ESF-Modulen in Verbindung mit einer Weiterbildungsmaßnahme, d.h. von Maßnahmeabschnitten, die in der Regelförderung nicht möglich sind, weil sie der Vermittlung berufsbezogenen Allgemeinwissens dienen (z.B. Deutschkenntnisse, Mathematik, Fremdsprachen) oder ein Auslandspraktikum beinhalteten.*

*Ebenfalls in die Förderung einbezogen waren u.a. die Teilnahme von Nichtleistungsbezieher/innen an Trainingsmaßnahmen, von (potentiellen) Überbrückungsgeldbezieher/innen an einem vorbereitenden Existenzgründerseminar und ein Coaching nach erfolgter Gründung, eine Qualifizierung während struktureller Kurzarbeit. Auch wurden zusätzliche Leistungen wie sozialpädagogische Betreuung gewährt.*

*Seit 2003 ist die Förderung von Nichtleistungsbezieher/innen bei der Teilnahme an einer beruflichen Weiterbildung ohne Modul und an Trainingsmaßnahmen sowie ein Coaching für Existenzgründer/innen und die Qualifizierung bei Kurzarbeit möglich.*

---

\* Zum ESF-BA-Programm und seiner bisherigen Umsetzung vgl. Deeke (2003) und Müller/Schuler (2003).

\*\* Vgl. hierzu Deeke/Kruppe (2003)

### 3. ESF-BA-Geschäftsstatistik

Die Bundesanstalt für Arbeit führt im Rahmen des Monitoring des ESF-BA-Programms eine entsprechende Geschäftsstatistik.<sup>15</sup> Diese wird auf der Grundlage von ereignisbezogenen Einzeldatensätzen erstellt und wird intern als St 38 bezeichnet, wie bereits die Statistik des Vorläuferprogramms. Im Vergleich zu dieser vorherigen Version ist die neue St 38 erheblich differenzierter. An der Vorbereitung dieser Statistik war die Begleitforschung beratend beteiligt.

#### 3.1 Gemeinsamkeiten und Unterschiede zwischen alter und neuer Geschäftsstatistik

Betrachtet man die Fassung der alten St 38, die mit Ende der Förderperiode 1995 - 1999 eingestellt wurde, und vergleicht diese mit der neuen Geschäftsstatistik gleichen Namens, welche die Anforderungen an das Monitoring in der laufenden Programmperiode abdecken soll, so spiegeln sich die gestiegenen Anforderungen in folgender Form wider: Es werden nicht mehr nur Zugänge, sondern nun auch Bestände, Abgänge und Verbleib (am Stichtag nach einem Monat und nach sechs Monaten) nach Maßnahmeart unterschieden berichtet. Nun wird auch die Anzahl der Fälle sozialpädagogischer Betreuung ausgewiesen. Generell wird davon ausgegangen, dass es sich bei allen berichteten Fällen auch um tatsächliche Eintritte und Bestände in bzw. Austritte aus Maßnahmen handelt, da im Falle von Nichtantritt bzw. vorzeitigem Abbruch der Maßnahme der Fall gesondert gekennzeichnet wird.

Wesentliche Merkmale, die vorher nicht erfasst wurden, die aber im Rahmen des sogenannten „Stammblattverfahrens“ gefordert sind, werden nun in die Berichterstattung einbezogen, darunter neben Personenmerkmalen (z.B. Alter nach Gruppen, Berufsausbildung) u. a. auch Abgangsgründe. Allerdings fehlen auch einige Merkmale. So ist keine direkte Identifikation von Nicht-Leistungsbeziehern (keine Lohnersatzleistung nach SGB III) möglich, da das entscheidende Merkmal „Leistungsbezug vor Eintritt in die Maßnahme“ nicht erfasst wird. Hilfsweise kann der Anteil der geförderten Nichtleistungsbezieher nur über die Zahl der Bezieher des für diese Gruppe als Kann-Leistung vorgesehenen ESF-Unterhaltsgeldes indiziert werden.

Auch werden zum Förderbereich Qualifizierung bei Kurzarbeit die Angaben zu Wirtschaftsklassen und Betriebsnummern der kurzarbeitenden Betriebe erst seit 2001 erhoben. Die Erfassung der Betriebsgröße erfolgt erst seit Mai 2003.

Spezifische Erfassungsregeln existieren für die Leistungsart sozialpädagogische Betreuung: Hier soll aufgrund unterschiedlicher Häufigkeit und Diskontinuität der Betreuung nur die Erstbetreuung erfasst werden. Es sind deshalb für diese Fälle verschiedene statistische Merkmale nicht ausweisbar (z. B. Dauer und zeitliche Lage vor, während und/oder nach der Teilnahme an einer Qualifizierungsmaßnahme). Zudem ist es auf der bisher in der aggregierten Statistik berichteten Fallebene unklar, ob sozialpädagogische Betreuung als eine ESF-Leistung z. B. begleitend zur Teilnahme an regulärer Förderung beruflicher Weiterbildung (FbW) bzw. Trainingsmaßnahmen nach dem Sozialgesetzbuch III ohne sonstige ESF-Ergänzung oder als weitere Leistung begleitend zu ESF-gestützter beruflicher Weiterbildung bzw. Trainingsmaßnahmen gewährt wird.

---

<sup>15</sup> Die Geschäftsstatistik (St 38) ist zur Zeit vom Statistikreferat der BA nicht freigegeben, d. h., dass sie nur intern und für die Berichterstattung an das BMWA und die Europäischen Kommission ausgewertet wird. Seit 2002 wird lediglich die Gesamtzahl der Zugangsfälle (einschließlich der Fälle sozialpädagogischer Betreuung) in West- und Ostdeutschland veröffentlicht.

## 3.2 St 38 und BA-Statistik zu den regulären Förderinstrumenten

Innerhalb der BA-Statistik zu den regulären Förderinstrumenten kann (bislang) die Gruppe derjenigen Teilnehmer und Teilnehmerinnen nicht identifiziert werden, die neben einer gesetzlichen Leistung zusätzlich auch eine ergänzende ESF-Leistung beziehen, also Geförderte des ESF-BA-Programms sind. So kann beispielsweise nicht ausgewiesen werden, welcher Anteil der bei ihrer beruflichen Weiterbildung Geförderten (nur) mit Hilfe des ESF teilnehmen konnte bzw. zusätzliche Leistungen im Rahmen des ESF-BA-Programms erhielt. Dies gilt auch für die Eingliederungsbilanz des SGB III. Voraussetzung wäre die generelle Einführung eines entsprechenden ESF-Merkmals in den einschlägigen Fachverfahren. Weil diese ebenso wie auch die Eingliederungsbilanz zur gesetzlichen Förderung berichten sollen, hat die BA bisher auf die Aufnahme eines ESF-Identifizierers verzichtet.

Diese „Merkmalslücke“ in der BA-Statistik hat folgenreiche Konsequenzen sowohl für das Monitoring als auch für die Evaluation. Denn es ist so nicht möglich, Auskunft über die Merkmalstruktur der im Rahmen des ESF-BA-Programms geförderten Personengruppen im Vergleich zu den in der gesetzlichen Regelförderung ohne ergänzende ESF-Förderung geförderten Personengruppen zu geben.

## 4. Der Aufbau einer Individualdatenbank

Ausgehend von der im vorherigen Kapitel beschriebenen Datengrundlage zum ESF-BA-Programm, bei der es sich um fallbezogene Einzeldatensätze getrennt für Zugang, Bestand und Abgang, Verbleib nach einem Monat und Verbleib nach 6 Monaten handelt, muss zunächst eine teilnahmebezogene und darauf basierend dann eine personenbezogene Individualdatenbank aufgebaut werden.<sup>16</sup> Dies ist für die wirkungsanalytische Evaluation und vorgeschaltete deskriptive Analysen (die auch das Monitoring der BA ergänzen können) unumgänglich. Für diese Zwecke müssen die ESF-BA-Datensätze zusätzlich mit anderen Datenquellen verknüpft werden. Dabei handelt es sich u. a. um die Einzeldaten

aus den Fachverfahren der SGB III-Förderung<sup>17</sup> von

- beruflicher Weiterbildung und Trainingsmaßnahmen,
- Existenzgründungen,

sowie aus der

- Beschäftigtenstatistik,
- Bewerberangebotsdatei (BewA) zu Arbeitslosen und auch nicht arbeitslos gemeldeten Arbeitssuchenden

und zusätzlich auch

- zum Leistungsbezug (Arbeitslosengeld bzw. Arbeitslosenhilfe).

Ziel ist es, Maßnahme-, Erwerbs-, und Arbeitslosigkeitsverläufe verfolgen und analysieren zu können. Weil der Geschäftsbereich Statistik der Bundesanstalt für Arbeit zur Zeit derartige Verknüpfungen noch nicht herstellen kann<sup>18</sup>, arbeitet die Begleitforschung am Aufbau einer entsprechenden Individualdatenbank für ihre eigenen Forschungszwecke. Diese ist mittlerweile weitgehend fertiggestellt. Im Folgenden werden die Ausgangslage und die einzelnen Arbeitsschritte auf dem Weg zu dieser Individualdatenbank beschrieben, die prototypisch auch für die Bearbeitung von Daten aus anderen Fachverfahren stehen können und deshalb über den spezifischen Projektzusammenhang hinausweisen.

<sup>16</sup> Die Unterschiede zwischen fall-, teilnehmer- und personenbezogenen Datensätzen werden noch genauer erläutert.

<sup>17</sup> Zur Kurzarbeit bzw. dem Kurzarbeitergeld (Kug) hat die BA kein vergleichbares Fachverfahren, weil Kug nach herkömmlichem Verständnis als „passive“ Leistung gilt.

<sup>18</sup> Dies wird erst im Rahmen des BA-Data-Warehouse ermöglicht, welches sich zur Zeit noch im Aufbau befindet. Vgl. hierzu auch: Schwarzfärber (2002).

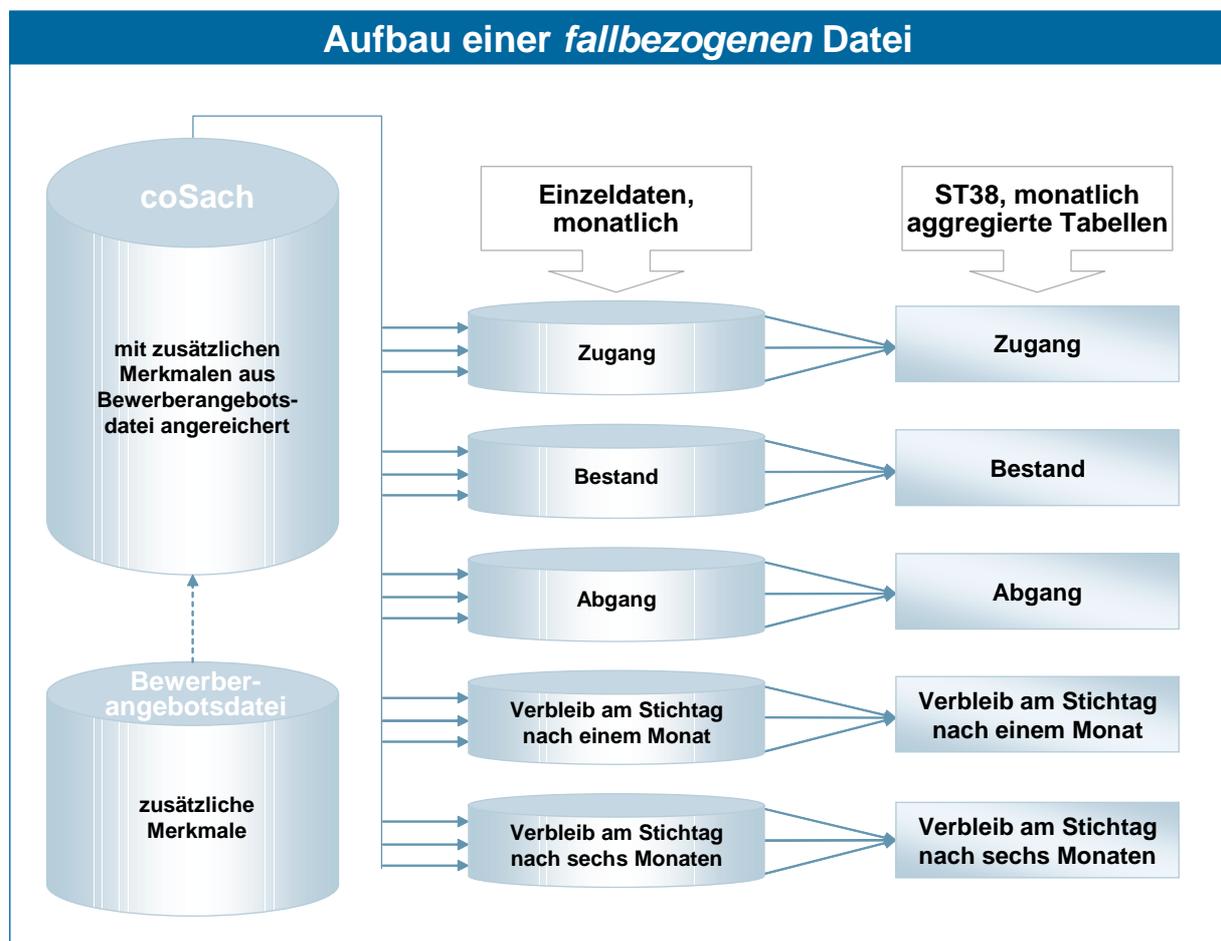
#### 4.1 Die Ausgangslage: Fallbezogene Einzeldatensätze

Die Daten zur ESF-BA-Statistik (St 38-Einzeldaten) werden dezentral in den Arbeitsämtern in die Förderdatenbank coSach,<sup>19</sup> Teilverfahren FbW, eingegeben. Diese Eingabe erfolgt nach einer Bewilligung dem Grunde nach, die der jeweilige Sachbearbeiter/Vermittler entschieden hat. Dieser gibt jedoch die Daten nicht selber ein, sondern reicht die Papierversion weiter, die anschließend gesammelt an anderer Stelle im Amt elektronisch erfasst wird.

Die Daten werden monatlich an bestimmten Stichtagen als fallbezogene Daten auf zentraler Ebene zusammengeführt (vgl. auch *Schaubild 1*). Der Berichtszeitraum der St 38 umfasst so immer den Zeitraum zwischen zwei Stichtagen, die beispielsweise im Jahr 2002 zwischen dem 19. und 27. des jeweiligen Monats variierten. Der jeweilige Monat wird also nicht exakt abgebildet, da er sich mit notwendigem Vorlauf nach den monatlichen Berichtsterminen der Pressekonferenz der Bundesanstalt für Arbeit richtet.

Bei der zentralen Zusammenführung („Datenziehung“) der fallbezogenen Datensätze werden aus der Bewerberangebotsdatei (BewA) auf Einzeldatenbasis ausgewählte Merkmale hinzugespielt. Dazu wird eine Kundennummer als Identifikator genutzt, die ab 1.1.2000 bundesweit zur Verfügung steht, so dass alle zu erfassenden Datensätze mit dieser Kundennummer versehen sein sollten (und im Falle der Daten zur Teilnahme am ESF-BA-Programm in der aktuellen Förderperiode auch tatsächlich sind).

Schaubild 1



<sup>19</sup> coSach: Computergestützte Sachbearbeitung.

**Fallbezogenheit der Datensätze** bedeutet, dass ihre Ziehung/Erstellung in zeitlicher Abhängigkeit eines Ereignisses stattfindet. Dem entsprechend werden die Datensätze getrennt als Zugangs-, Bestands-, Abgangs- und Verbleibsdatsätze (nach einem bzw. sechs Monaten) nach folgenden Regeln am Stichtag generiert:

Als **Zugang** wird ein Datensatz gezählt, wenn das Datum des geplanten Eintritts in die Maßnahme in den Berichtszeitraum fällt, also nach dem letzten Stichtag und nicht in der Zukunft liegt. Ob der Eintritt tatsächlich erfolgte, kann dabei aus Ablaufgründen nicht mit völliger Sicherheit gefolgert werden. So werden die Teilnahmedaten bereits eingegeben, wenn eine Bewilligung „dem Grunde nach“ durch den Arbeitsberater stattgefunden hat – unabhängig von einer tatsächlichen Bewilligung. Auch handelt es sich bei dem eingegebenen Datum zunächst um den geplanten Eintritt, der sich durchaus vom Datum des tatsächlichen Eintritts unterscheiden kann. Bei über den Stichtag hinaus verzögerter bzw. verspäteter Eingabe von Änderungen (z.B. bei Nichtantritt etc.) wird der Zugangsdatensatz auf Basis der noch nicht geänderten Daten erstellt und so u.U. fälschlicherweise ein Nichtantritt als Eintritt gewertet.

Bei verzögerter bzw. verspäteter Eingabe der Teilnahmedaten liegen diese erst nach dem Stichtag vor, so dass sie auch erst zu einem späteren Zeitpunkt in die Statistik einfließen können. Insofern handelt es sich streng genommen um Zugangsdatensätze (Zugang im Sinne von Fall-Erfassung mit Bezug auf den geplanten Eintritt nach einer Bewilligung dem Grunde nach) im Unterschied zu echten Eintrittsdatsätzen (Bezug tatsächlich realisierter Eintritt).<sup>20</sup> Als Austrittsdatum wird in den Zugangsdatensätzen zunächst der geplante Austritt aus der Maßnahme eingetragen.<sup>21</sup>

Um den **Bestand** zu erfassen werden am Stichtag alle Datensätze gezogen, deren Eintrittsdatum in die Maßnahme vor und das geplante Austrittsdatum nach dem aktuellen Stichtagsdatum liegt. Dieses führt dazu, dass ein Datensatz, wenn das Eintrittsdatum innerhalb des Berichtszeitraums liegt und das geplante Austrittsdatum größer als das Stichtagsdatum ist, im gleichen Monat sowohl als Zugang als auch bereits als Bestand gezählt wird. Für eine Maßnahmeteilnahme können im gesamten Maßnahmezeitraum – abhängig von der Laufzeit der Maßnahme – mehrere Bestandsdatensätze entstehen, und zwar jeweils ein Bestandsdatensatz je Stichtag, der in den Maßnahmezeitraum fällt.

**Abgangsdatensätze** werden mit einem Monat Verspätung gezählt, um zu gewährleisten, dass eine tatsächliche Aktualisierung des Datenbestands (z.B. in Bezug auf das individuelle Ergebnis der Maßnahme) auch wirklich stattgefunden hat. Ein Datensatz wird folglich dann als Abgang gezählt, wenn das Austrittsdatum im vorherigen Berichtsmonat liegt. Bei einem Austritt wird das geplante mit dem tatsächlichen Austrittsdatum überschrieben, falls zwischen beiden eine Abweichung besteht. Gleichzeitig soll ein „negativer“ Erfolg bzw. Abbruch der Maßnahme erfasst werden (Prüfung nicht bestanden, Arbeitsaufnahme, mangelnde Leistung, längere Fehlzeiten, sonstige Gründe, Nichtantritt).

In Zugangs- und Bestandsdatensätzen steht also das geplante, in Abgangsdaten das tatsächliche Austrittsdatum. Allerdings kann es auch hier zu Fehlern kommen, wenn ein vorzeitiger Austritt aus der Maßnahme erst nach der automatisierten Datenziehung eingegeben wird. Durch die um einen Monat verzögerte Erfassung soll dies aber eine geringere Relevanz haben.

**Verbleibsdatsätze** werden zweimal erfasst: **einen Monat** nach Austritt aus der Maßnahme und **sechs Monate** nach Austritt aus der Maßnahme. Diese Datensätze enthalten vor allem ein Merkmal, welches angibt, ob der Maßnahmeteilnehmer ein bzw. sechs Monate nach Maßnahmeaustritt am Stichtag arbeitslos gemeldet ist oder nicht.

<sup>20</sup> Im Bewusstsein dieser Problematik werden dennoch im folgenden die Begriffe Zugänge und Eintritte synonym verwandt, da eine Prüfung, ob die berichteten Zugänge mit den jeweiligen geplanten Eintritten auch tatsächlich Eintritte waren, hier nicht geleistet werden kann (vgl. aber Kapitel 5.1).

<sup>21</sup> Ausnahme: sozialpädagogische Betreuung, bei der das Erstbetreuungsdatum als Ein- und Austrittsdatum erfasst wird.

## 4.2 Erstes Ziel: Teilnahmebezogene Informationen und Historisierung

Wie im vorherigen Kapitel dargestellt, enthalten die Einzeldatensätze, aus denen die Geschäftsstatistik zum ESF-BA-Programm erstellt wird, nur getrennte Informationen zu Zugang, Abgang, Bestand und Verbleib. Die Beschreibung und Analyse eines Teilnahmeverlaufs inklusive Verbleib ist so jedoch noch nicht möglich, da hierzu diese fallbezogenen Informationen verknüpft werden müssen, um die Teilnahme an einer Maßnahme bzw. den Bezug einer Leistung im Rahmen des ESF-BA-Programms analysieren zu können.

**Teilnahmebezogenen Daten** unterscheiden sich von fallbezogenen Daten also darin, dass es sich nicht mehr nur um eine ereignisorientierte<sup>22</sup> getrennte Sammlung von Daten handelt. Durch das Zusammenführen dieser Datensätze anhand eines teilnahmebezogenen Identifikators können Informationen zum Verlauf einer Maßnahmeteilnahme und dem anschließenden Verbleib nach einem bzw. sechs Monat(en) bereit gestellt werden. So können z.B. Abweichungen zwischen geplantem und realisiertem Austrittsdatum erst durch die teilnahmebezogene Aufbereitung der Daten erfasst werden.

Durch die fallbezogene Konstruktion der St38-Einzeldaten entstehen starke Redundanzen, da für eine Maßnahmeteilnahme mehrere einzelne Datensätze mit großteils gleichen Informationen existieren. Dies gilt insbesondere für die monatlichen Bestandsdatensätze, die unabhängig davon erstellt werden, ob sich eine Änderung der Daten ergeben hat oder nicht. Aber auch zwischen Daten zum Zugang, Bestand und Abgang, werden viele redundante Informationen geliefert.

Aber nicht alle Informationen in den einzelnen zu einer Teilnehmernummer gehörenden Datensätzen sind redundant, da sich verschiedene Merkmale während der Laufzeit der Maßnahme faktisch ändern können. Neben der Korrektur von fehlenden oder fehlerhaft erfassten Informationen können dies zum Beispiel veränderte Informationen zum geplanten Austrittsdatum etc. sein. Diese Veränderungen sollen aus inhaltlichen Gründen nachvollziehbar bleiben.

Der zusammengeführte Maßnahmeteilnahme-Datensatz besteht folglich aus zwei verschiedenen Kategorien von Feldern: zum einen aus konstanten, nicht veränderlichen und zum anderen aus Feldern, deren Inhalt sich verändern kann. Im folgenden wird der Weg des Aufbaus einer teilnahmebezogenen Datenbank beschrieben. Dieser Weg leitet sich ab vom Konzept einer Maßnahme-Teilnahme-Grunddatei (MTG), wie sie zur Zeit vom IAB für die Förderung von Arbeitsbeschaffungsmaßnahmen, Strukturpassungsmaßnahmen, Maßnahmen der beruflichen Weiterbildung, Trainingsmaßnahmen und Maßnahmen der Freien Förderung nach § 10 SGB III aufgebaut wird.<sup>23</sup>

Der Teilnahmedatensatz wird sukzessive aus den Informationen aus Zugang, ggf. Bestand / Beständen, Abgang und Verbleib aufgebaut (vgl. **Schaubild 2, Seite 15**).<sup>24</sup>

➤ **Die Neuanlage eines Teilnahmedatensatzes** wird durch einen fallbezogenen **Zugangssatz** verursacht. Dabei werden die entsprechenden Datenfelder inklusive dem geplanten Austrittsdatum gefüllt, das Feld „tatsächlicher Austritt“ bleibt noch leer.

<sup>22</sup> Ereignis meint hier: Zugang, Bestand, Abgang, Verbleib am Stichtag nach einem Monat, Verbleib am Stichtag nach sechs Monaten (vgl. Kapitel 4.1).

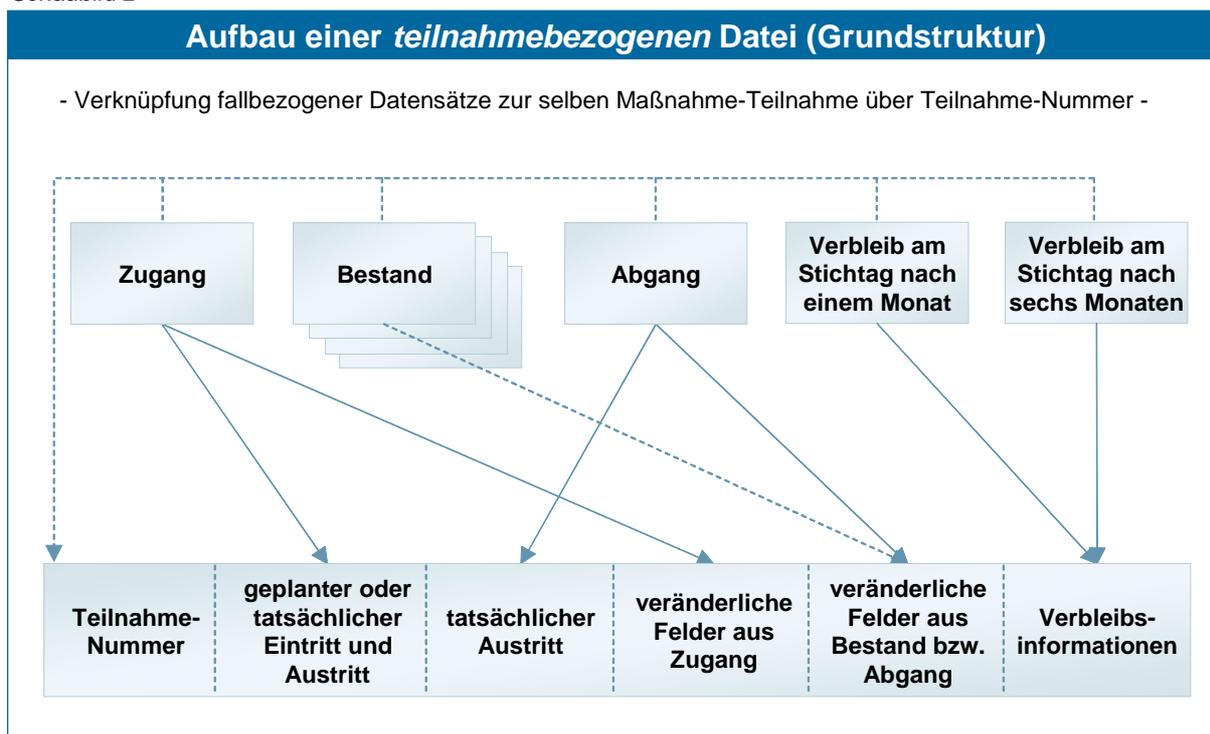
<sup>23</sup> Siehe dazu: Kellner, Elmar (2002): MTG - Maßnahme-Teilnehmer-Grunddatei. Machbarkeitsstudie zum Aufbau einer alle Maßnahmentearten umfassenden „Maßnahme-Teilnehmer-Grunddatei“ – Abschlussbericht.

<sup>24</sup> In dem im folgenden beschriebenen Verfahren werden dazu vornehmlich die Zugangs- und Abgangsdatensätze herangezogen und die Informationen aus den Bestandsdatensätzen nur bei zwischenzeitlich relevanten Änderungen genutzt.

- Ein **Bestandsatz** kann Eintragungen und/oder Änderungen im Bereich der veränderlichen Datenfelder zur Folge haben. Die vorher erfassten Zugangsdaten bleiben aber unverändert. Die Änderungen werden zusätzlich mit dem Datum der Änderung gespeichert. Ein besonderes Problem besteht darin, eine Korrektur, eine Änderung oder eine Falscheingabe voneinander zu unterscheiden. Abgesehen von bestimmten Plausibilitätsprüfungen kann die Entscheidung darüber nur bei einer späteren Datenauswertung getroffen werden. Alle Datenfelder des Bestandsdatensatzes ohne Änderungen werden verworfen.
- Ein **Abgangssatz** liefert u.a. das Datum des tatsächlichen Austritts und Informationen über eine (nicht) erfolgreiche Teilnahme. Änderungen in den veränderlichen Datenfelder gegenüber den letzten Bestandsdatensatz werden im Bereich der Abgangsdaten gespeichert. Alle unveränderten Datenfelder werden verworfen.
- Ein **Verbleibssatz** enthält Information zum Verbleib am Stichtag nach einem bzw. nach sechs Monat(en). Nur die verbleibsrelevanten Daten werden im entsprechenden Bereich abgelegt, alle anderen Felder verworfen.

Durch den Abgleich der Bestandsdaten und der Anbindung von Bestands-, Abgangs- und Verbleibsdaten an den Zugangssatz wird die in den Ursprungsdaten bestehende Redundanz erheblich verringert und so eine wesentlich geringere Menge an Daten gespeichert, ohne dass Informationen verloren gehen. Allerdings kommen durch die Hinzufügung des jeweiligen Gültigkeitszeitraums (s. u.) bei veränderlichen Informationen auch wieder zusätzliche Daten hinzu. Denn, wie bereits beschrieben, können solche Änderungen innerhalb eines Merkmales in den Daten häufiger auftreten. Um eine stichtagsgenaue Historisierung durchzuführen, genügt es deshalb nicht, nur die Veränderungen zwischen Zugangs- und Abgangsdatensatz zu betrachten, sondern es müssen auch die Änderungen innerhalb der Bestandsdaten festgehalten werden.

Schaubild 2



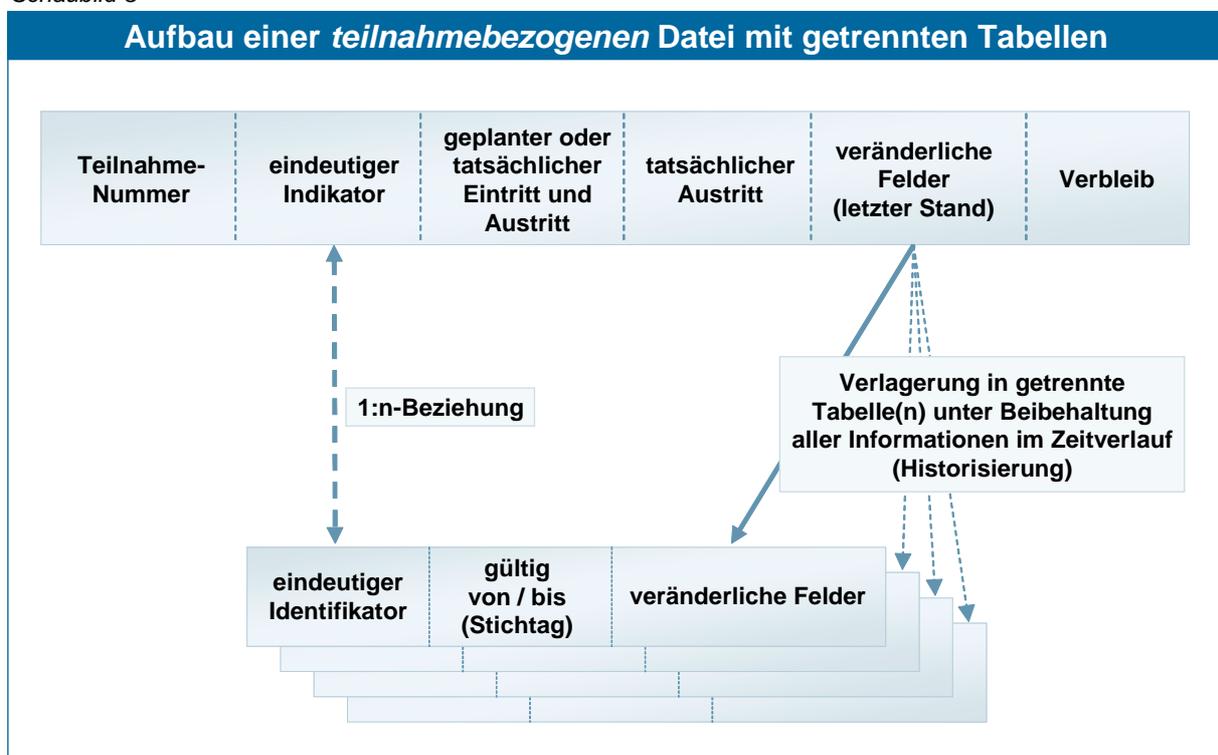
Dabei muss ein Kompromiss zwischen Redundanz und Performanz gefunden werden, d. h., dass als zweiter Gesichtspunkt auch die Verarbeitungsgeschwindigkeit zu berücksichtigen ist. Für die veränderlichen Merkmale können entweder für alle Merkmale eine einzige Datenbanktabelle, oder für eine Gruppe von Merkmalen oder sogar für jedes einzelne Merkmal jeweils eine separate Tabelle angelegt werden. Innerhalb dieser Tabelle(n) wird bei jeder Veränderung eines Merkmals ein neuer Datensatz mit Identifizierer und Veränderungsdatum erzeugt (vgl. auch **Schaubild 3**).<sup>25</sup>

Dabei ist zu bedenken:

- Wird eine einzige Tabelle für alle Merkmale verwendet, so werden bei der Veränderung einer einzigen Merkmalsausprägung erneut Redundanzen aufgebaut, da auch alle nicht-veränderten Merkmale erneut abgespeichert werden.
- Wird eine Tabelle pro Merkmal verwendet, müssen bei einer Auswertung viele Tabellen miteinander verknüpft werden. Dieses wirkt sich jedoch negativ auf die Verarbeitungsgeschwindigkeit (Performanz) aus. Weiterhin wird durch Schlüsselbegriffe und verwaltungstechnische Felder der Datenbestand aufgebläht.

Welchem Modell der Vorzug gegeben wird, hängt letztlich von der Performanz ab, die durch die technischen Gegebenheiten jeweils vorgegeben ist. Umgesetzt wird nun ein Mittelweg, der sich zur Zeit als beste Lösung anbietet. Hierbei wird der Aufbau der ESF-BA-Individualdatenbank in Abstimmung mit dem Aufbau der Maßnahme-Teilnehmer-Grunddatei (MTG) durchgeführt, um eine weitere Verknüpfung der ESF-BA-Daten mit anderen Datenquellen zu optimieren. Aus dem gleichen Grund findet ebenfalls bereits in der Aufbauphase eine Abstimmung mit den Aufbauarbeiten des BA-Data-Warehouse statt.

Schaubild 3



<sup>25</sup> Auf eine Tabelle mit den unveränderlichen Merkmalen des Teilnahmedatensatzes kommt eine unterschiedliche Anzahl von Tabellen für veränderliche Merkmale („1:n Beziehung“).

### 4.3 Zweites Ziel: Personenbezogene Informationen

Wie im vorherigen Kapitel dargestellt, kann nach einer entsprechenden Zusammenführung der Daten die Teilnahme an einer Maßnahme im Zeitverlauf – also historisch – verfolgt werden.<sup>26</sup> Die Beschreibung und Analyse des Erwerbsverlaufs einer Person (Personenhistorie) ist jedoch so noch nicht möglich, da sich die verknüpften Informationen bisher nur auf die Teilnahme an einer einzelnen Maßnahme oder den Bezug einer Leistung im Rahmen des ESF-BA-Programms beziehen. Möglich ist z. B., dass eine Person im Rahmen der Gesamtlaufzeit des Programms entweder gezielt – im Sinne einer „Förderkette“, wegen des Wechsel von einer Maßnahme in eine andere (z. B. bei Maßnahmeabbruch) oder aus sonstigen Gründen – in größerem zeitlichen Abstand an zwei oder mehr Maßnahmen teilnimmt.

**Personenbezogene Datensätze** unterscheiden sich von teilnahmebezogenen Datensätzen dadurch, dass es sich nicht mehr um eine Sammlung von Daten handelt, die sich nur auf genau eine Teilnahme an einer einzigen Maßnahme beziehen, sondern in einem weiteren Schritt durch das Zusammenführen jeweils teilnahmebezogener Datensätze und deren Verknüpfung mit weiteren Informationen zu Arbeitslosigkeit und/oder Beschäftigung etc. anhand eines personenbezogenen Identifikators Informationen zum individuellen Erwerbsverläufen bereit stellen – auch zu sogenannten „Maßnahmekarrieren“.

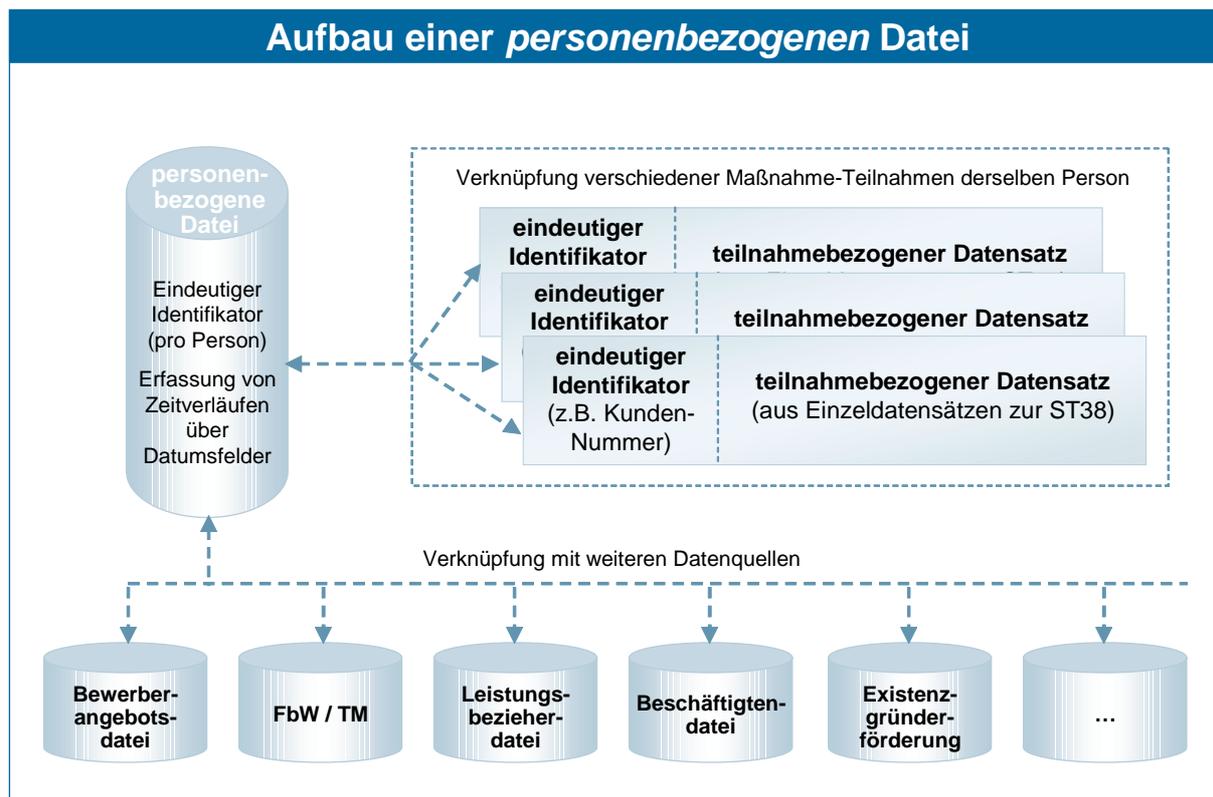
Geplant ist, im Rahmen des BA-Data-Warehouse alle prozess erzeugten Daten (Maßnahmedaten, Leistungsbezug, Beschäftigungsdaten) personenbezogen zu verknüpfen. Hierzu sollen sowohl alle Personenmerkmale als auch alle „Betätigungen“ (Beschäftigung, Maßnahmeteilnahme, Leistungsbezug, Arbeitssuche) historisch – also mit allen Veränderungen im Zeitverlauf – abgespeichert werden. Zum einen wird das BA-Data-Warehouse jedoch deutlich später realisiert als es für die Begleitforschung zum ESF-BA-Programm und weitere IAB-Evaluierungsprojekte<sup>27</sup> erforderlich ist. Zum anderen aber sind teilweise die Erfordernisse an Datenstrukturen unterschiedlicher Natur (Geschäftsprozess- Controlling, Regel-Statistik, Wissenschaft). Auch ist eine langfristige Haltung der Daten, u. a. zum Aufbau längerer Zeitreihen, gesetzlich bisher nur dem IAB erlaubt. Für die Individualdatenbank zum ESF-BA-Programm bedeutet dies: Nach dem Aufbau einer teilnahmeorientierten Datenbasis aus den fallbezogenen Einzeldatensätzen der St 38 zum ESF-BA-Programm wird diese Basis mit den Einzeldaten der Bewerberangebotsdatei, mit Einzeldaten zur Förderung beruflicher Weiterbildung und Trainingsmaßnahmen sowie mit weiteren Datenquellen verknüpft (vgl. auch **Schaubild 4**, Seite 18).

---

<sup>26</sup> Beim Aufbau der Datenstruktur ist zusätzlich zu beachten, dass verschiedene Auswertungstechniken bis hin zu Ereignisanalysen mittels Episodensplitting möglich sein sollen. Zu den spezifischen Anforderungen und Problemen einer Aufbereitung von Prozessdaten zu Forschungsdaten für Episodensplitting vgl. z.B. Bender/Fitzenberger/Lechner (2002).

<sup>27</sup> Zu Planung und Stand der Entwicklung des Data Warehouse der BA vgl. Deeke/Wiedemann (2002).

Schaubild 4



## 5. Probleme der Datengenerierung und Lösungswege bei der Datenaufbereitung

Beim Aufbau von Individualdatenbanken, wie hier zur Begleitforschung des ESF-BA-Programms, betritt das IAB Neuland. Neu ist dabei nicht nur, dass der zuvor beschriebene Aufbau der Datenbank mit der Zusammenführung verschiedener Datenquellen in einem kontinuierlich fortzuführenden Prozess stattfinden muss, neu sind auch die durch die Bundesanstalt bereitgestellten Daten. Daraus folgt ein hoher Erschließungsaufwand, da es sich zum einen um sehr große Datenmengen handelt<sup>28</sup> und zum anderen die Generierung von Einzeldaten im Verwaltungsprozess einer anderen als der Logik der Begleitforschung folgen. Deshalb treten zum Teil Probleme auf, die weit vor der Analyse solcher Daten zu aller erst erkannt und weit möglichst gelöst werden müssen.

Wie bereits beschrieben, werden die Einzeldaten zur ESF-BA-Geschäftsstatistik (St 38) dezentral in den Arbeitsämtern eingegeben. Die Eingabe erfolgt aber nicht direkt durch den jeweilige Sachbearbeiter/Vermittler, sondern diese reichen die Papierversionen weiter, die dann gesammelt an anderer Stelle im Amt elektronisch erfasst werden. Dabei kommt es zum Teil zu Verzögerungen bei der elektronischen Erfassung (Zugangsbuchung). Anschließend werden die Daten monatlich an bestimmten Stichtagen auf zentraler Ebene zusammengeführt. Unmittelbar bei dieser zentralen Zusammenführung („Datenziehung“) werden aus der jeweiligen dezentralen Bewerberangebotsdatei (BewA) auf Einzeldatenbasis ausgewählte Merkmale mit gezogen und zu den Einzeldatensätzen hinzugespielt. Die erste Ziehung der Daten – rückwirkend bis einschließlich Januar 2000 – fand im März 2001 statt, da das Ziehungsverfahren erst zu diesem Zeitpunkt implementiert wurde.

<sup>28</sup> Das Volumen der Rohdaten allein der ST38 (in ASCII-Format, ohne Beschreibung etc.) betrug 2000 ca. 95 MB, 2001 ca. 240 MB und 2002 ca. 300 MB. In einer SPSS-Datei mit Variablen- und Wertebeschreibung entspricht dies insgesamt ca. einem Gigabyte benötigtem Speicherplatz.

Dieses Verfahren beinhaltet für die Nutzbarkeit der Daten für die Zwecke der Begleitforschung spezifische Probleme: Sowohl eine verzögerte Erfassung der Daten als auch das Zusammenspielen auf zentraler Ebene mit der gleichzeitigen Hinzufügung von Merkmalen aus der Bewerberangebotsdatei jeweils am Stichtag hat zur Folge, dass nicht sichergestellt ist, dass der aufgefüllte Datenstand tatsächlich dem jeweiligen Zustand bei Eintritt, Austritt bzw. Verbleib entspricht. Das Risiko, dass die Daten nicht mehr dem angezielten Stand entsprechen, ist bei Zugangsdatensätzen am größten, da besonders bei kurzfristigen Maßnahmen bzw. Abbrüchen die Daten nach Beendigung der Maßnahme bereits entsprechend aktualisiert sein können (Dauer der vorherigen Arbeitslosigkeit, letzter Berufsabschluss, Berufswunsch, etc.) bevor eine Ziehung des Zugangs überhaupt stattgefunden hat.

Ganz besonders trifft dies natürlich für die im März 2001 nachträglich gesammelt gezogenen Daten zu. Aus Sicht der Begleitforschung ist dies deshalb besonders kritisch, weil es nicht nur bei der aggregierten Auswertung zu Verzerrungen kommen kann (die im Saldo betrachtet ggf. statistisch nur von relativ geringer Bedeutung sein mögen), sondern vor allem weil die Bildung von Vergleichsgruppen zur Ermittlung von Nettoeffekten (Arbeitslose ohne Förderung der beruflichen Weiterbildung, Nicht-ESF-BA-geförderte Teilnahme an beruflichen Weiterbildung) auf Basis individueller Merkmale entsprechend der Struktur der ESF-BA-Geförderten misslingen könnte.

Ein weiterer „produktionsbedingter“ Effekt ist, dass durch das beschriebene verzögerte Einlaufen der Datensätze eine exakte Zahl z. B. der Zugänge für einen bestimmten Zeitraum erst nach längerer Zeit festgestellt werden kann. Dieses Problem wird in der ESF-Statistik der Bundesanstalt für Arbeit (St 38) durch die Berichterstattung nach Stichtagen umgangen, da z. B. Zugänge generell (erst) dem Monat zugerechnet werden, in dem sie statistisch zentral erfasst werden (statistisches Berichtsdatum). Im Hinblick auf die Widerspiegelung eines exakteren Bildes geht die Begleitforschung – im Unterschied zum Vorgehen der Statistik – dazu bei ihren Betrachtungen immer vom Datum des (geplanten) Eintritts aus.

## 5.1 Analyse der zeitlichen Differenz zwischen Eintritt und Datenziehung

Die Größenordnung des Problems wurde in einem ersten Schritt durch einen Abgleich zwischen (geplantem) Eintrittsdatum und statistischem Berichtsdatum (Stichtagsdatum) für die Ziehungen zwischen März 2001<sup>29</sup> und Dezember 2002 abgeschätzt. Wie aus der *Tabelle 1* (Seite 20) zu ersehen ist, fanden die Hälfte aller Ziehungen innerhalb von 4 Wochen, ca. 8 % allerdings später als 90 Tage nach Eintritt in die Maßnahme statt (ohne Maßnahmeart „Qualifizierung bei Kurzarbeit“ und ohne Leistungsart „sozialpädagogische Betreuung“).<sup>30</sup> Die letzten 3 % der Fälle wurden erst nach einem halben Jahr oder später erfasst.

Diese zeitliche Differenz divergiert allerdings nach Maßnahmeart (vgl. *Tabelle 2*, Seite 20). Während Trainingsmaßnahmen am zeitnächsten erfasst werden (im Mittel liegen zwischen Eintrittsdatum und Ziehungsstichtag 35 Tage), liegt dieser Wert bei Coaching bei 88 Tagen, also bei fast 3 Monaten. Die Maximaldauer zwischen (geplantem) Eintritts- und Ziehungsdatum beträgt bei allen Maßnahmearten mehr als 9 Monate, zum Teil bis zu knapp unter 3 Jahren.

<sup>29</sup> Wie bereits beschrieben, fand die erste Ziehung der Daten – rückwirkend bis einschließlich Januar 2000 – im März 2001 statt, da das Ziehungsverfahren erst zu diesem Zeitpunkt implementiert wurde. Eine Analyse der „regelmäßigen“ Verzögerung zwischen (geplantem Eintrittsdatum und zeitlicher statistischer Zuordnung (Stichtag) kann deshalb erst ab diesem Zeitpunkt aufsetzen.

<sup>30</sup> Dabei handelt es sich um gesonderte Problembereiche, insbesondere bei der Leistungsart „sozialpädagogische Betreuung“, bei der von vorn herein weniger Merkmale erfasst werden. Hier waren die verspäteten Eingaben im betrachteten Zeitraum allerdings am gravierendsten. Bei Qualifizierung bei Kurzarbeit werden Personen gefördert, die sich noch in Beschäftigung befinden. Ab hängig vom Ablauf der Kurzarbeit und der Umsetzung der Qualifizierungsmaßnahmen findet die Eingabe hier nicht schon bezogen auf den (geplanten) Eintritt, sondern erst später statt.

Tabelle 1

Differenz zwischen Eintrittsdatum und Stichtag		
Differenz in Tagen (gruppiert)	Prozent	Kumulierte Prozente
0 – 7	5,8	5,8
8 – 14	7,7	13,5
15 – 28	36,6	50,1
28 – 59	28,5	78,7
60 – 89	9,2	87,9
90 – 119	4,3	92,2
120 – 149	2,4	94,6
150 – 179	1,5	96,1
180 – 209	1,0	97,1
210 – 239	0,8	97,9
240 – 269	0,5	98,4
270 – 299	0,5	98,9
300 – 329	0,4	99,3
330 – 359	0,3	99,5
360 und mehr	0,5	100,0

Eintritte zwischen März 2001 und Dezember 2003.  
Dateneingänge: März 2001 bis Februar 2003.  
Alle Maßnahmentearten (ohne Qual. w. Kurzarbeit und sozialpädagogische Betreuung).  
Quelle: Einzeldaten zur St38 (Zugangsdatensätze), eigene Berechnungen.

Der Eindruck, als würde bei allen Maßnahmentearten die statistische Erfassung generell mindestens einen Monat verspätet erfolgen (alle Mittelwerte sind größer als 30 Tage), entsteht dadurch, dass die Mittelwerte der Arbeitsämter ungewichtet in die Berechnung eingegangen sind. So beeinflusst ein Amt mit nur einem Fall und dessen schneller Erfassung die Mittelwerte mit dem gleichen Gewicht wie ein Amt mit vielen Fällen und generell verspäteter Eingabe (vgl. auch **Tabelle 3**, Seite 21). So fiel zwischen März 2001 und Dezember 2002 z.B. der Eintritt in ESF-geförderte Maßnahmen zur beruflichen Weiterbildung mit Modul in einem Amt jeweils unmittelbar mit dem Stichtag zusammen (Mittelwert Minimum = 0 Tage), wobei die Erfassung vor Ort entweder vor oder gleichzeitig zu Eintrittsdatum und Stichtag stattfand. In diesem Amt war die Anzahl der Fälle insgesamt sehr gering. In einem anderen Amt dagegen vergingen im Mittel über 7 Monate (Max = 228 Tage) zwischen Eintrittsdatum und statistischer Zählung. Dies lag nur zum Teil an Einführungsproblemen zu Beginn der Programmlaufzeit, da sich die Dauer im weiteren Zeitverlauf nur leicht verringert hat.

Da, wie zuvor beschrieben, bestimmte Variablen erst bei der Datenziehung zugespielt werden, besteht durch die Verzögerungen zwischen Eintrittsdatum und Ziehung das Risiko, möglicherweise nicht (mehr) dem Stand bei Eintritt entsprechende Informationen zu erhalten. Die tatsächliche Größenordnung ist zwar hier noch nicht abschätzbar. Die Möglichkeit einer größeren Verzerrung ließ aber eine Korrektur der gelieferten Datensätze notwendig erscheinen, um vor allem bei der Bildung von Vergleichsgruppen zur Ermittlung von Nettoeffekten im Rahmen der Wirkungsanalyse (Arbeitslose ohne Förderung der beruflichen Weiterbildung, Nicht-ESF-BA-geförderte Teilnahme an beruflichen Weiterbildung) die tatsächlichen Strukturmerkmale der ESF-BA-Geförderten entsprechend berücksichtigen zu können.

Der Eindruck, als würde bei allen Maßnahmentearten die statistische Erfassung generell mindestens einen Monat verspätet erfolgen (alle Mittelwerte sind größer als 30 Tage), entsteht dadurch, dass die Mittelwerte der Arbeitsämter ungewichtet in die Berechnung eingegangen sind. So beeinflusst ein Amt mit nur einem Fall und dessen schneller Erfassung die Mittelwerte mit dem gleichen Gewicht wie ein Amt mit vielen Fällen und generell verspäteter Eingabe (vgl. auch **Tabelle 3**, Seite 21). So fiel zwischen März 2001 und Dezember 2002 z.B. der Eintritt in ESF-geförderte Maßnahmen zur beruflichen Weiterbildung mit Modul in einem Amt jeweils unmittelbar mit dem Stichtag zusammen (Mittelwert Minimum = 0 Tage), wobei die Erfassung vor Ort entweder vor oder gleichzeitig zu Eintrittsdatum und Stichtag stattfand. In diesem Amt war die Anzahl der Fälle insgesamt sehr gering. In einem anderen Amt dagegen vergingen im Mittel über 7 Monate (Max = 228 Tage) zwischen Eintrittsdatum und statistischer Zählung. Dies lag nur zum Teil an Einführungsproblemen zu Beginn der Programmlaufzeit, da sich die Dauer im weiteren Zeitverlauf nur leicht verringert hat.

Da, wie zuvor beschrieben, bestimmte Variablen erst bei der Datenziehung zugespielt werden, besteht durch die Verzögerungen zwischen Eintrittsdatum und Ziehung das Risiko, möglicherweise nicht (mehr) dem Stand bei Eintritt entsprechende Informationen zu erhalten.

Tabelle 2

Differenz zwischen Eintrittsdatum und Stichtag nach Maßnahmenteart (in Tagen)		
	Mittelwert (ungewichtet)	Maximum
FbW ohne Modul	40	717
FbW mit Modul	42	659
Trainingsmaßnahmen	35	296
Existenzgründungsseminare	52	555
Coaching	88	716
<b>Insgesamt</b>	<b>49</b>	<b>717</b>

Eintritte zwischen März 2001 und Dezember 2003.  
Dateneingänge: März 2001 bis Februar 2003.  
Quelle: Einzeldaten zur St38 (Zugangsdatensätze), eigene Berechnungen.

## 5.2 Datenanreicherung für Evaluationszwecke

Wie zuvor beschrieben, besteht durch die technische Verzögerung zwischen Eintrittsdatum und statistischer Erfassung/Datenziehung die Gefahr der Verzerrung bestimmter Variablen mit zeitkritischem Inhalt, da sich der zugespielte Status möglicherweise nicht mehr auf den Zustand bei bzw. vor Eintritt in die Maßnahmen bezieht, sondern bereits mit aktualisierten Werten überschrieben wurde. Um die tatsächlichen Strukturmerkmale der ESF-BA-Geförderten bei der Bildung von Vergleichsgruppen zur Ermittlung von Nettoeffekten im Rahmen der Wirkungsanalyse (Arbeitslose ohne Förderung der beruflichen Weiterbildung, Nicht-ESF-BA-geförderte Teilnahme an beruflichen Weiterbildung) entsprechend

Tabelle 3

Differenz zwischen Eingabe und Datenziehung nach Maßnahmentypen und Arbeitsämtern		
	Mittelwert	
	Minimum	Maximum
FbW ohne Modul	16	180
FbW mit Modul	0	228
Trainingsmaßnahmen	8	150
Existenzgründungsseminare	3	226
Coaching	23	280

Eintritte zwischen März 2001 und Dezember 2003.

Dateneingänge: März 2001 bis Februar 2003.

Quelle: Einzeldaten zur St38 (Zugangsdatensätze), eigene Berechnungen.

Merkmale und den Merkmalen aus dem Bewerberangebot eben nicht durch den Stichtag, sondern durch das jeweilige Datum (z.B. geplantes Eintrittsdatum) tagesgenau durchzuführen.

Das Ergebnis der erneuten Zuspiegelung aus der Bewerberangebotsdatei (BewA-Konten) wird im folgenden am Beispiel der Variable „Arbeitslos vor Eintritt in die Maßnahme (ja/nein)“ im Vergleich zu den ursprünglichen Datensätzen vorgestellt, wobei die Erfassung dieser Variablen einer Besonderheit unterliegt: Grundlage der Berichterstattung in der Geschäftsstatistik St 38 ist die Eingabe dieses Merkmals im Verfahren coSach durch die jeweiligen Bearbeiter in den Ämtern. Dementsprechend fußen die Angaben in den Einzeldatensätzen der St 38 auch auf dieser Erfassung.

Anders ist es bei den Einzeldaten zur regulären Förderung der beruflichen Weiterbildung und der Teilnahme an Trainingsmaßnahmen. Hier werden die Angaben zum Arbeitslosenstatus und zur Dauer der vorherigen Arbeitslosigkeit nicht durch Eingaben in coSach generiert, sondern durch eine unmittelbare Abfrage aus der Bewerberangebotsdatei (BewA/BewE) zugespielt. Damit bestünde für einen teilnahme- oder gar personenbezogenen Vergleich zwischen der ESF-BA-Förderung und der Regelförderung ohne ESF-Ergänzung das Problem, den Arbeitslosenstatus und die Arbeitslosigkeitsdauer aus zwei unterschiedlichen – und damit potentiell widersprüchlichen – Quellen zu vergleichen. Darüber hinaus ist für die mikroanalytischen Wirkungsuntersuchungen die Bildung von Vergleichsgruppen mit nicht geförderten Arbeitslosen erforderlich, deren Daten ebenfalls aus der Bewerberangebotsdatei gezogen werden. Auch deshalb ist eine entsprechende Berechnung der Angaben zu Arbeitslosigkeitsstatus und -dauer auf Basis der Angaben aus der BewA notwendig.

Vergleicht man diese nachträglich zugespielte, korrigierte Datenerfassung der Variablen aus der Bewerberangebotsdatei mit dem in coSach erfassten Merkmal (**Tabelle 4**), so zeigt sich für die Gruppe der vor Eintritt in die Maßnahme Arbeitslosen eine Abweichung von ca. 7 Prozentpunkten, wobei hier die korrigierte Variable aus der Bewerberangebotsdatei mit 82,6 % den geringeren Anteil zuvor Arbeitsloser ausweist.

berücksichtigen zu können, wurde anfangs in einem komplexen Verfahren eine erneute Zuspiegelung der Variablen aus zentral verfügbaren Bewerberangebotsdateien<sup>31</sup> zu Beständen und Bewegungen durchgeführt. Die nun deutlich verbesserte Datenhaltung in der Bundesanstalt für Arbeit machte es mittlerweile möglich, auf die Informationen aus der Bewerberangebotsdatei in Form von „Konten“ zurückzugreifen. Hierbei ist es möglich, die Zuordnung zwischen den in coSach speziell zur ESF-BA-Förderung erfassten

Tabelle 4

Arbeitslosigkeit vor Maßnahmebeginn		
Vergleich zwischen der Variablen aus coSach und der BewA-Fassung für Teilnehmende an Einzelinstrumenten der individuellen Hilfe zur Qualifizierung		
Status vor Förderung	CoSach	BewA
	in Prozent	
Nicht arbeitslos	10,1	17,4
Arbeitslos	89,9	82,6
<b>Gesamt</b>	<b>100</b>	<b>100</b>

Quelle: Bereinigte Einzeldaten zur St38 (Zugänge in Maßnahmen), eigene Berechnungen.

<sup>31</sup> Bewerberangebotsdateien: monatliche Bestandsdaten (BewA) und tagesgenaue Bewegungsdaten (BewE) zu An-, Ab- und Ummeldungen.

Zur Überprüfung der Validität dieser Datenkorrektur und gleichzeitiger Überprüfung ihrer tatsächlichen Notwendigkeit wurde noch einmal auf ebenfalls in den St 38 Einzeldaten enthaltene BewA-Angaben zum Arbeitslosenstatus und zur vorherigen Dauer zurückgegriffen. Diese St 38 Merkmale sind ausschließlich für die Beobachtung des Verbleibs (arbeitslos ja/nein) vorgesehen. Sie werden zwar auch in den Zugangsdatensätzen automatisch gefüllt, sind dort aber nicht verwertbar, weil z. B. das BewA-Merkmal wegen des zwischenzeitlichen Eintritts in die Maßnahme überschrieben sein kann und ein eigentlich zuvor Arbeitsloser bei der anschließenden Stichtagsziehung dann fälschlich als „nicht arbeitslos“ erfasst würde. Ein Vergleich dieser Daten mit einerseits den coSach-Angaben und andererseits mit den Ergebnissen der oben beschriebenen BewA-Korrektur für die Zeit unmittelbar vor Maßnahmeintritt bestätigte diese Einschätzung.

### 5.3 Datenbereinigung

Zusätzlich wurden verschiedene Bereinigungen der Daten vorgenommen. So wurden beispielweise Datensätze mit Eintrittsdatum vor Programmstart, also vor 2000, entfernt. Da die für einen Abgleich mit anderen Datensätzen notwendige Zuordnung einzig auf Basis der Kundennummer zu keinem befriedigenden Ergebnis führte,<sup>32</sup> wurde ein eindeutiger Identifikator neu generiert. Da es bei den Datensätzen gelegentlich – z. B. technisch bedingt – zu Mehrfacherfassungen kommen kann, wurden diese in einem mehrstufigen Verfahren identifiziert und bereinigt. Die Identifizierung wurde schrittweise anhand der erweiterten Kundennummer, der Maßnahmennummer, Teilnehmernummer, Eintrittsdatum und Maßnahmeart vorgenommen. Stimmt bei zwei Zugangsdatensätzen alle überprüften Merkmale überein, wurde von einer Doppelerfassung des Zugangsdatensatzes ausgegangen (0,03 %). Stimmt alle Ausprägungen außer dem (geplanten) Eintrittsdatum überein, so wurde von einer Verschiebung des Eintritts ausgegangen und ebenfalls nur einer der Datensätze beibehalten (0,02 %). Stimmt immerhin noch Kundennummer und Maßnahmennummer überein, so wurde auch hier eine Modifikation vermutet und ebenfalls nur einer der Datensätze beibehalten (3,4 %).

Nach der Durchführung der Bereinigungsschritte, die im übrigen analog zu den ESF-BA-Einzeldatensätzen auch für andere Einzeldaten durchgeführt werden – so z. B. für die zu verknüpfenden Einzeldatensätze zur Regelförderung nach SGB III, genauer zur Förderung beruflicher Weiterbildung, Trainingsmaßnahmen, Existenzgründungsförderung und Freier Förderung – konnte eine erste Analyse von Datensätzen auf Personen-Ebene durchgeführt werden.

---

<sup>32</sup> In einigen Fällen ist die Kundennummer nicht „eindeutig“, da sie z. B. versehentlich mehrfach vergeben wurde. Außerdem kommt es – wie generell bei prozessproduzierten Daten nicht hundertprozentig vermeidbar – zu fehlerhaften Eingaben. In anderen Fällen wurden zum Teil „künstliche“ Kundennummern vergeben (z.B. nach dem Schema 111111 ff), damit einem Anfügen von BewA-Daten technisch nichts entgegensteht, da nur bei „erfolgreicher“ Zusammenspielung die Daten in die zentrale Statistik eingehen können. Wenn die Kundennummer fehlt oder in der Bewerberangebotsdatei nicht gefunden wird, wird der Datensatz zurückgewiesen und ein Fehlerprotokoll im Arbeitsamt vor Ort erstellt. Hier sollen dann diese Datensätze korrigiert werden, wenn sie statistisch noch relevant sind, d.h. das geplante Austrittsdatum aus der Maßnahme noch nicht überschritten ist. Für den Zeitraum bis Ende 2001 gab es insgesamt lediglich 320 Fälle mit erkennbaren Dummy-Kundennummern sowie weitere 81 Fälle mit nicht eindeutigen Kundennummern, zusammen genommen waren davon also weniger als 0,2 % der Datensätze betroffen.

## 6. Erste Zusammenführung personenbezogener Einzeldatensätze

In diesem Kapitel werden erste Auswertungen der im Rahmen der Individualdatenbank zur Evaluation des ESF-BA-Programms sowohl auf Teilnahme- wie auch auf Personenebene zusammengeführten Einzeldatensätze vorgestellt, um das hierdurch neu erschlossene Informationspotential zu demonstrieren. Die Dateien der im Ausbau befindlichen integrierten Datenbasis des IAB, erlaubt analoge Analysen z.B. für die Förderung von Arbeitsbeschaffungsmaßnahmen, Strukturanpassungsmaßnahmen, Maßnahmen der beruflichen Weiterbildung, Trainingsmaßnahmen und Maßnahmen der Freien Förderung nach § 10 SGB III.

In **Tabelle 5** werden im Überblick die teilnahmebezogenen Zugänge in die ESF-BA-Förderung nach Maßnahmenteilen ausgewiesen. Ein Vergleich mit der Geschäftsstatistik St 38 ist nicht ohne weiteres möglich, da die Zugänge in **Tabelle 5** (und auch im weiteren) nach (geplantem) Eintrittsdatum im entsprechenden Jahr aufgeführt sind (und nicht nach Stichtag der Datenziehung). Zu beachten ist außerdem, dass es sich um bereinigten Daten handelt (vgl. Kapitel 5). Eingegangen in die Auswertungen sind letztlich alle Zugänge, die von Januar 2000 bis zum Stichtag Ende Februar 2003 zentral als Einzeldaten zur St 38 zusammengespield wurden und deren Eintrittsdatum im Zeitraum von Januar 2000 bis Dezember 2002 liegt. Die aufgeführten Werte sind deshalb als vorläufig zu betrachten, weil sich die Werte im Zuge später eingehender, weiterer Datensätze sehr wahrscheinlich noch erhöhen werden.<sup>33</sup> Nicht in alle Auswertungen eingegangen sind die Zugänge zu „sozialpädagogischer Betreuung“, da es sich hierbei nicht um einen Zugang bzw. Eintritt in eine Maßnahme handelt, sondern um die Erfassung einer flankierenden ESF-Leistung. Dieses Fehlen wird jedoch ausdrücklich an den jeweiligen Tabellen angemerkt.

Tabelle 5

		Zugänge in ESF-BA-Förderung nach Maßnahmenteilen (teilnahmebezogene Auswertung)								
		2000			2001			2002		
		Ost	West	Gesamt	Ost	West	Gesamt	Ost	West	Gesamt
BW ohne Modul*	abs	10.806	11.978	22.784	9.136	10.864	20.000	7.727	9.149	16.876
	in %	41,7	51,5	46,3	35,0	43,9	39,3	32,3	32,9	32,6
BW mit Modul	abs	6.500	3.218	9.718	7.187	4.269	11.456	5.287	4.255	9.542
	in %	25,1	13,8	19,8	27,5	17,2	22,5	22,1	15,3	18,5
Trainingsmaßnahmen	abs	3.456	777	4.233	3.636	781	4.417	2.791	807	3.598
	in %	13,3	3,3	8,6	13,9	3,2	8,7	11,7	2,9	7,0
Existenzgründungsseminare	abs	1.852	1.998	3.850	1.600	3.163	4.763	1.459	4.665	6.124
	in %	7,2	8,6	7,8	6,1	12,8	9,4	6,1	16,8	11,8
Coaching	abs	2.255	1.281	3.536	3.721	1.876	5.597	5.714	3.188	8.902
	in %	8,7	5,5	7,2	14,2	7,6	11,0	23,9	11,5	17,2
Qual. w. KuG	abs	1.022	4.023	5.045	846	3.811	4.657	924	5.733	6.657
	in %	3,9	17,3	10,3	3,2	15,4	9,2	3,9	20,6	12,9
<b>Gesamt</b>	abs	<b>25.891</b>	<b>23.275</b>	<b>49.166</b>	<b>26.126</b>	<b>24.764</b>	<b>50.890</b>	<b>23.902</b>	<b>27.797</b>	<b>51.699</b>
	in %	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>

\* inklusive betrieblicher Einzelmaßnahmen

Quelle: Bereinigte Einzeldaten zur St38 (Zugänge in Maßnahmen ohne Leistungsfälle „sozialpädagogische Betreuung“), eigene Berechnungen.

<sup>33</sup> Die erste Abschätzung aus einem Zeitraum im Jahr 2000 ergab einen Eingang von 8 % der Zugangsdatensätze erst nach mehr als drei Monaten (vgl. Kapitel 5.2). Eine Veränderung der Werte insbesondere für das Jahr 2002 in dieser Höhe ist deshalb durchaus möglich.

Ohne an dieser Stelle die in **Tabelle 5** dargestellten Ergebnisse inhaltlich weiter zu interpretieren<sup>34</sup>, soll im Folgenden dargestellt werden, welche Ergebnisse die Zusammenfassung der Daten zu personenbezogenen Konten liefert.

Dazu wurde auf Basis der bereinigten Einzeldaten auf einen Identifikator zurück gegriffen, der beim Aufbau der Datenbank zu diesem Zweck angelegt wurde. Soweit technisch möglich, wird dabei versucht, eine eindeutige Zuordnung auf Personenebene durchzuführen.<sup>35</sup>

Darauf aufsetzend wurde jeder gültige Zugangsdatensatz zu einer ESF-kofinanzierten Maßnahme sowie ggf. der ESF geförderte Leistungsfall „sozialpädagogische Betreuung“ der Person analysiert. Die Zusammenführung führt zu den in **Tabelle 6** dargestellten Ergebnissen, wobei hier auch die Leistungsfälle „sozialpädagogischer Betreuung“ enthalten sind, da es auch Personen gibt, die im Rahmen der ergänzenden ESF-Förderung nur diese Leistung aus dem ESF-BA-Programm erhielten.

Die Zahl von 407.293 teilnahmebezogenen Zugängen der bereinigten Einzeldaten lassen sich also auf 360.514 Personen zurückführen, von denen 40.196 bis zu 15 weitere Zugangsdatensätze aufweisen, die sich auf insgesamt weitere 46.779 Zugänge addieren (vgl. **Tabelle 6**). Es haben somit 11,1 % an mehr als einer Maßnahme/Leistung im betrachteten Zeitraum teilgenommen/beansprucht.

Diese mehrfachen Maßnahmeteilnahmen derselben Person können sich auf Teilnahmen an unterschiedlichen Maßnahmen beziehen, so z.B. als Abfolge von einer Trainingsmaßnahme zur Vorbereitung auf eine sich anschließende Maßnahme zur beruflichen Weiterbildung oder parallel zu einer Maßnahme eine sozialpädagogische Betreuung erhalten. Analysiert man die Daten auf Personenebene

**Tabelle 7**

Mehrfache Zugänge in gleiche Maßnahmen / Leistungen (personenbezogene Auswertung)	
Anzahl der Personen mit	
2 Teilnahmen	26.745
3 Teilnahmen	2.371
4 Teilnahmen	338
5 Teilnahmen	85
6 Teilnahmen	42
7 Teilnahmen	27
8 Teilnahmen	17
9 Teilnahmen	16
10 Teilnahmen	6
11 Teilnahmen	4
12 Teilnahmen	3
13 Teilnahmen	2
14 Teilnahmen	0
15 Teilnahmen	0
16 Teilnahmen	1

Quelle: Bereinigte Einzeldaten zur St38 (Zugänge in Maßnahmen einschl. Leistungsfälle „sozialpädagogische Betreuung“), eigene Berechnungen.

**Tabelle 6**

Zugänge in ESF-Förderung (personenbezogene Auswertung)	
<b>Teilnahmebezogene Datensätze insgesamt</b>	<b>407.293</b>
<b>Personen insgesamt</b>	<b>360.514</b>
Weitere Zugangsdatensätze	46.779
Personen mit mehreren Zugängen	40.196
darunter Personen mit	
2 Zugangsdatensätzen	35.101
3 Zugangsdatensätzen	4.197
4 Zugangsdatensätzen	635
5 Zugangsdatensätzen	132
6 Zugangsdatensätzen	54
7 Zugangsdatensätzen	28
8 Zugangsdatensätzen	17
9 Zugangsdatensätzen	16
10 Zugangsdatensätzen	6
11 Zugangsdatensätzen	4
12 Zugangsdatensätzen	3
13 Zugangsdatensätzen	2
14 Zugangsdatensätzen	0
15 Zugangsdatensätzen	0
16 Zugangsdatensätzen	1

Quelle: Bereinigte Einzeldaten zur St38 (Zugänge in Maßnahmen einschl. Leistungsfälle „sozialpädagogische Betreuung“), eigene Berechnungen.

<sup>34</sup> Vgl. dazu den Bericht der Begleitforschung zur Halbzeitbewertung: Deeke/Kruppe/Kurtz/Müller/Schuler (2003). Die Ergebnisse in der **Tabelle 5** weichen in der Summe der Jahre und nach Maßnahmeart nur geringfügig von den aggregierten Daten der Geschäftsstatistik St 38 ab (Deeke 2003).

<sup>35</sup> Vergleiche hierzu Kapitel 5.3. Da es nicht völlig auszuschließen ist, dass eine Person z. B. zwei unterschiedliche Kundennummern erhält (etwa wenn sie in einen anderen Arbeitsamtsbezirk umzieht), und ebenfalls eine Person unterschiedliche BewA-Eingabenummern erhalten kann und gleichzeitig auch im Zeitverlauf für dieselbe Person unterschiedliche Sozialversicherungsnummern existieren können, ist eine wirkliche Eindeutigkeit nicht gegeben. Weitergehende Abgleiche sind jedoch unter Berücksichtigung des zu erwartenden Ertrags zu aufwändig und wurden daher nicht durchgeführt.

daraufhin, so ergibt dies, dass im betrachteten Zeitraum 12.268 Personen an zwei, 423 Personen an drei und 13 Personen sogar an vier ESF-geförderten Maßnahmen unterschiedlichen Typs teilnahmen.

Es kann sich aber auch um mehrfache Teilnahmen an der gleichen Maßnahmeart handeln, so z. B. mehrere Trainingsmaßnahmen im Zeitraum zwischen Januar 2000 und Dezember 2002 (vgl. **Tabelle 7**, Seite 24).

**Tabelle 8**

<b>Zugänge in ESF-geförderte Maßnahmen (personenbezogene Auswertung)</b>	
<b>Teilnahmebezogene Datensätze insgesamt</b>	<b>155.829</b>
<b>Personen insgesamt</b>	<b>137.875</b>
Weitere Zugangsdatensätze	17.954
Personen mit mehreren Zugängen	15.680
darunter Personen mit	
2 Zugangsdatensätzen	13.882
3 Zugangsdatensätzen	1.515
4 Zugangsdatensätzen	212
5 Zugangsdatensätzen	27
6 Zugangsdatensätzen	16
7 Zugangsdatensätzen	8
8 Zugangsdatensätzen	7
9 Zugangsdatensätzen	5
10 Zugangsdatensätzen	3
11 Zugangsdatensätzen	2
12 Zugangsdatensätzen	2
13 Zugangsdatensätzen	1

Quelle: Bereinigte Einzeldaten zur St38 (Zugänge in Maßnahmen ohne Leistungsfälle „sozialpädagogische Betreuung“), eigene Berechnungen.

Bei dieser Betrachtung muss jedoch berücksichtigt werden, dass bei einer Person z. B. sowohl eine mehrfache Teilnahme in einer Maßnahmeart als auch eine mehrfache Teilnahme in einer anderen Maßnahmeart auftreten kann.

Lässt man die Zugänge zu „sozialpädagogischer Betreuung“ bei den Analysen außer Betracht, da es sich ja nicht um einen Zugang bzw. Eintritt in eine Maßnahme handelt, sondern um die Erfassung einer flankierenden ESF-Leistung, ergibt sich das in **Tabelle 8** dargestellte Bild. Die Zahl von 155.829 teilnahmebezogenen Zugängen der bereinigten Einzeldaten lassen sich also auf 137.875 Personen zurückführen, von denen 15.680 bis zu 12 weitere Zugangsdatensätze aufweisen, die sich auf insgesamt weitere 17.954 Zugänge addieren (vgl. **Tabelle 8**). Es haben somit 11,4 % an mehr als einer Maßnahme im betrachteten Zeitraum teilgenommen.

Bei der Datenerfassung der Leistung „sozialpädagogische Betreuung“ wurde vor allem in der Startphase des ESF-BA-Programms möglicherweise (in geringem Umfang) von der Vorgabe abgewichen, nur den Erst-Betreuungstermin als Eintrittsdatum zu erfassen und für jeden weiteren Betreuungstermin ein neuer Datenzugang erzeugt. Darauf lassen zumindest z. B. teilweise monatlich wiederkehrende Zugänge schließen. Auch deshalb wird hier die Berechnungen der mehrfachen Teilnahme an gleichen Maßnahmen unter Ausschluss der sozialpädagogischen Betreuung dargestellt (vgl. **Tabelle 9**).

Ob die gefundenen Mehrfachteilnahmen tatsächlich stattgefunden haben oder ob es sich hierbei um Fehlbuchungen handelt, ist aus den Daten nicht ersichtlich. Da aber eine Mehrfachförderung in bestimmten Konstellationen durchaus möglich ist, erscheint die Anzahl der möglichen Fehler als durchaus tolerabel.

**Tabelle 9**

<b>Mehrfache Zugänge in gleiche Maßnahmen (personenbezogene Auswertung)</b>	
Anzahl der Personen mit	
2 Teilnahmen	9.803
3 Teilnahmen	786
4 Teilnahmen	101
5 Teilnahmen	18
6 Teilnahmen	13
7 Teilnahmen	8
8 Teilnahmen	7
9 Teilnahmen	5
10 Teilnahmen	3
11 Teilnahmen	2
12 Teilnahmen	2
13 Teilnahmen	1

Quelle: Bereinigte Einzeldaten zur St38 (Zugänge in Maßnahmen ohne Leistungsfälle „sozialpädagogische Betreuung“), eigene Berechnungen.

## 7. Schlussfolgerungen

Gegenstand dieses Berichts war die Darstellung des Aufbaus einer Individualdatenbank für erweitertes Monitoring und Evaluation im Rahmen der Begleitforschung zum ESF-BA-Programm auf Basis der in der Bundesanstalt für Arbeit im Rahmen des Geschäftsprozesses anfallenden prozessproduzierten Daten. Da die Nutzung von prozessproduzierten Daten zur Evaluation von arbeitsmarktpolitischen Maßnahmen eine immer größere Rolle spielt, wurden hier Besonderheiten aufgezeigt, die gerade bei solchen Daten existieren, die im Rahmen des Geschäftsprozesses erhoben werden und deshalb in erster Linie nicht auf eine wissenschaftliche Evaluationen zugeschnitten sind. Von besonderer Bedeutung ist daher der in diesem Bericht vorgestellte Umbau der Daten von fallbezogenen Einzeldatensätzen hin zu personenbezogenen historisierten Verläufen. In Zukunft stehen der Forschung zwar im Rahmen der integrierten Datenbasis am IAB oftmals bereits historisierte Datensätze mit Identifikatoren auf Personenebene zur Verfügung. Das Wissen um die Genese dieser Datensätze und die spezifischen Probleme bei ihrer Erzeugung ist bei der Interpretation der Daten hilfreich. So ist z. B. eine Kenntnis der hier vorgestellten Datenberechnungen nach Eintrittsdatum als Erklärung für eine Abweichung zu den durch die Geschäftsstatistik publizierten Werten notwendig, die auf der Datenziehung nach Stichtagskonzept beruhen. Ebenso ist es hilfreich, die Veränderlichkeit des Datenbestandes im Zeitverlauf z. B. durch verspätete Erfassung bzw. Ziehung („Nachläufer“) zu kennen, um so Rückschlüsse auf eine mögliche Unterschätzung am aktuellen Rand ziehen zu können. Eine Aufbereitung der prozessproduzierten Daten in der hier beschriebenen Form eröffnet zudem wesentliche weitere Möglichkeiten, so z.B. eine bessere Generierung von Vergleichsgruppen bei Analysen zu individuellen Nettoeffekten und können den Grundstock für weitergehende Kausalanalysen bilden.

## Literatur

- Auer P./ Kruppe T. (1996):* Monitoring of Labour Market Policy in EU Member States. In: Schmid, Günther and Jacqueline O'Reilly, Klaus Schömann (eds), International Handbook of Labour Market Policy and Evaluation, Edward Elgar, Cheltenham.
- Bender, S./ Fitzenberger, B./ Lechner, M. (2002):* Pber die Wirksamkeit von FuU-Maßnahmen – ein Evaluationsversuch mit prozessproduzierten Daten aus dem IAB, in: Deeke/ Wiedemann (2002): Anhang 4
- Caliendo, M./ Hujer, R./ Thomsen, S. L. (2003):* Evaluation individueller Netto-Effekte von ABM in Deutschland – Ein Matching-Ansatz mit Berücksichtigung von regionalen und individuellen Unterschieden. IAB-Werkstattbericht Nr. 02/2003, Nürnberg,
- Deeke, A. (2003):* Die Umsetzung des ESF-BA-Programms in der Zeit von Anfang 2000 bis Ende 2002, in: Deeke/ Kruppe/ Kurtz/ Müller/ Schuler (2003).
- Deeke, A. (2000):* Wedelt der Schwanz mit dem Hund? Anforderungen an das Monitoring und die Evaluation der Arbeitsförderung mit den Mitteln des Europäischen Sozialfonds, in: Ministerium für Arbeit, Soziales, Gesundheit und Frauen des Landes Brandenburg (Hrsg.): Evaluierung der Arbeitsmarktpolitik aus Ländersicht. Konzepte, Möglichkeiten und Grenzen politikberatender Wirkungsforschung, Potsdam, S. 29 - 33.
- Deeke, A. (1999):* Vier Jahre ESF-BA-Programm. Die Umsetzung der ergänzenden Förderung zum AFG und SGB III aus dem Europäischen Sozialfonds, IAB-Werkstattbericht 17/1999, Nürnberg.
- Deeke, A./ Kruppe, T. (2003):* Beschäftigungsfähigkeit als Evaluationsmaßstab? – Inhaltliche und methodische Aspekte der Wirkungsanalyse beruflicher Weiterbildung im Rahmen des ESF-BA-Programms. IAB-Werkstattbericht Nr. 01/2003, Nürnberg.
- Deeke, A./ Kruppe, T. (2002):* Prozess- und kontextorientierte Begleitforschung zu arbeitsmarktpolitischen Programmen, in: Kleinhenz, G. (Hrsg.): IAB-Kompodium Arbeitsmarkt- und Berufsforschung, Beiträge aus der Arbeitsmarkt- und Berufsforschung (BeitrAB) 250, Nürnberg, S. 399 – 414.
- Deeke, A./ Kruppe, T./ Kurtz, B./ Müller, P./ Schuler, W. (2003):* Begleitforschung zum „ESF-BA-Programm 2000 - 2006“ – Bericht zur Halbzeitbewertung, Nürnberg
- Deeke, A./ Kruppe, T./ Müller, P./ Schuler, W. (2002):* Das „ESF-BA-Programm 2000 - 2006“ und seine Umsetzung im ersten Jahr, IAB-Werkstattbericht 5/2002, Nürnberg.
- Deeke, A./ Schuler, W. (2003):* Fünf Jahre „AFG-Plus“: Arbeitsförderung aus dem Europäischen Sozialfonds, Beiträge aus der Arbeitsmarkt- und Berufsforschung (BeitrAB) Bd. 265, Nürnberg
- Deeke, A./ Wiedemann, E. (2002):* Evaluierung aktiver Arbeitsmarktpolitik und Datengrundlagen – Bericht von einem Workshop in der Bundesanstalt für Arbeit am 9. November 2001, IAB-Werkstattbericht 2/2002, Nürnberg.
- Fitzenberger B./ Speckesser, S. (2002):* Weiterbildungsmaßnahmen in Ostdeutschland – Ein Misserfolg der Arbeitsmarktpolitik? ZEW-Discussion Paper No. 02-16, Zentrum für Europäische Wirtschaftsforschung, Mannheim.
- Jaenichen, U. (2002):* Lohnkostenzuschüsse und individuelle Arbeitslosigkeit \* Analysen auf der Grundlage kombinierter Erhebungs- und Prozessdaten unter Anwendung von Propensity Score Matching. In: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, Jg. 35, H. 3. S. 327 - 351.

- Kellner, E. (2002):* MTG – Maßnahme-Teilnehmer-Grunddatei. Machbarkeitsstudie zum Aufbau einer alle Maßnahmenteilnehmer umfassenden „Maßnahme-Teilnehmer-Grunddatei“ – Abschlussbericht, Nürnberg.
- Kurtz, B. (2003):* Trainingsmaßnahmen im ESF-BA-Programm und in der gesetzlichen Regelförderung nach dem SGB III, in: Deeke/ Kruppe/ Kurtz/ Müller/ Schuler (2003).
- Kurtz, B. (2002):* Förderung von Existenzgründungen: Das ESF-BA-Programm im Zusammenspiel mit der Regelförderung des SGB III, IAB-Werkstattbericht Nr. 8/2002, Nürnberg.
- Müller, P./Kurtz, B. (2002):* Aktive Arbeitsmarktpolitik und Gender Mainstreaming: Strukturen von Teilnahme und Verbleib bei ausgewählten Instrumenten der Bundesanstalt für Arbeit, in: Engelbrech, G. (Hg.): Arbeitsmarktchancen für Frauen, Beiträge aus der Arbeitsmarkt- und Berufsforschung (BeitrAB) 258, Nürnberg.
- Müller, P./Schuler, W. (2003):* Ergebnisse zur dezentralen Umsetzung des ESF-BA-Programms, in: Deeke/ Kruppe/ Kurtz/ Müller/ Schuler (2003).
- Müller, P./Schuler, W. (2001):* Erste Ergebnisse zur Umsetzung des ESF-BA-Programms auf dezentraler Ebene. In: Deeke, A./ Kruppe, T./ Müller, P./ Schuler, W. (2001): Begleitforschung zum „ESF-BA-Programm 2000 - 2006“ – Erster Zwischenbericht. IAB-Projektbericht, Nürnberg.
- Schwarzfärber, H. P. (2002):* Das Data-Warehouse der BA. In: Deeke, A./ Wiedemann, E (2002): Anhang 1.

## Quellen

- Durchführungsanweisungen (DA) zu den Richtlinien des Bundesministeriums für Arbeit und Sozialordnung für aus Mitteln des Europäischen Sozialfonds (ESF) mitfinanzierte zusätzliche arbeitsmarktpolitische Maßnahmen im Bereich des Bundes. In: Bundesanstalt für Arbeit, Dienstblatt-Runderlass 8/2000 vom 09.03.2000, Nürnberg.
- EPPD Ziel 3 Stand 2000: Einheitliches Programmplanungsdokument zur Entwicklung des Arbeitsmarktes und der Humanressourcen für die Interventionen des Ziels 3 in Deutschland vom 30.08.2000. Struktur-fondsperiode 2000 - 2006, Bonn.
- GFK Ziel 1. Gemeinschaftliches Förderkonzept Ziel 1 und Ziel 1-Übergangsregion in Deutschland 2000 - 2006, Bonn.
- OP Ziel 1, Stand Dezember 2000: Regionalübergreifendes Operationelles Programm des Bundes zur Entwicklung des Arbeitsmarktes und der Humanressourcen für die Interventionen des Ziels 1 in Deutschland. Struktur-fondsperiode 2000 - 2006, Bonn.
- Programmergänzungsdokument für die Interventionen des Ziels 3 in Deutschland vom 28.03.2001. Struktur-fondsperiode 2000 - 2006, Bonn.
- Programmergänzungsdokument für das regionalübergreifende Operationelle Programm des Bundes zu Ziel 1 in Deutschland vom 21.05.2001. Struktur-fondsperiode 2000 - 2006, Bonn.
- Richtlinien (RL) des Bundesministeriums für Arbeit und Sozialordnung für aus Mitteln des Europäischen Sozialfonds (ESF) mitfinanzierte zusätzliche arbeitsmarktpolitische Maßnahmen im Bereich des Bundes vom 20.01.2000. Bonn.
- Sozialgesetzbuch (SGB) Drittes Buch (III) – Arbeitsförderung. 4. Ausgabe, Stand: 15. Februar 2001. Textausgabe. Nürnberg (Bundesanstalt für Arbeit).