

Fachkonzept Datenpaket G – Grobsegmentierung

Stand: 13.08.2026

Dieses Dokument ist streng vertraulich und ausschließlich für Kunden der VR Data GmbH bestimmt. Eine Weitergabe an Dritte, insbesondere an Wettbewerber der VR Data, ist strengstens untersagt. Für den Fall der Zuwiderhandlung behält sich die VR Data vor, Schadensersatzansprüche geltend zu machen.

Inhaltsverzeichnis

A. Hintergrund.....	3
B. Prozess der Erstellung des Datenpakets Grobsegmentierung	3
C. Zuordnung der Kunden der Bank im Datenbestand der VR Data GmbH.....	4
D. Die Umsatzdaten der VR Data.....	5
E. Beschreibung der Datei, die an die Bank zurückgeliefert wird, und Verwendung der Ergebnisse in der Bank.....	7
Appendix 1: Beschreibung der Datenquellen.....	8

A. Hintergrund

Die Einteilung der Firmenkunden in Segmente ermöglicht eine effizientere und zielgerechtere Marktbearbeitung. Dabei spielt die Umsatzgröße eines Firmenkunden eine zentrale Rolle. Firmen mit einem größeren Jahresumsatz weisen aus Sicht der Bank idR auch ein größeres Umsatzpotenzial auf. Leider liegen der Bank idR jedoch nicht für alle ihre Firmenkunden Umsatzangaben vor.

Die VR Data bezieht von einer großen deutschen Wirtschaftsauskunftei Daten über alle deutschen Firmen: vom kleinen Gewerbetreibenden bis zum weltweit tätigen Konzern. In dieser Datenbank liegt die Abdeckung mit Umsatzdaten bei rund 98%. Mit dem Datenpaket G - Grobsegmentierung stellt die VR Data diese Daten den Nutzern des BankingWorkspace (im Folgenden als BWS bezeichnet) zur Verfügung, um eine möglichst flächendeckende Grobsegmentierung durchführen zu können. Fragen zur Anwendung Grobsegmentierung der BMS Corporate Solutions GmbH (im Folgenden als BMS CS bezeichnet) richten Sie bitte direkt an die BMS CS. Hierüber wissen wir zu wenig, um Ihnen Auskunft geben zu können.

Bitte beachten Sie, dass die Umsatzdaten der VR Data, die Ihnen mit der Anwendung Grobsegmentierung zur Verfügung gestellt werden, drei mögliche Quellen haben: Sie stammen entweder aus einer im Bundesanzeiger veröffentlichten Bilanz (bilanziert), sie wurden von der Auskunftei, bei der wir diese Daten einkaufen, im Internet, in Pressemeldungen oder direkt bei den Unternehmen recherchiert (recherchiert), oder sie wurden auf der Basis von anderen Daten geschätzt bzw. modelliert (modelliert). Weitere Informationen zu den Umsatzdaten entnehmen Sie bitte dem Abschnitt D.

Der Prozess der Datenanreicherung für die Grobsegmentierung wurde mit der BMS CS entwickelt und abgestimmt, um die Bank bei einer Grobsegmentierung mit fehlenden Umsatzdaten zu versorgen bzw. mit aktuelleren Umsatzdaten zu versorgen, wenn die Umsatzdaten der Bank schon relativ alt sind. Die Regeln zum Umgang mit veralteten Daten der Bank werden in Abschnitt D beschrieben. Der Prozess wird im nachfolgenden Abschnitt B beschrieben.

B. Prozess der Erstellung des Datenpakets Grobsegmentierung

Das Datenpaket G Grobsegmentierung wurde gemeinsam mit der BMS CS entwickelt. Der Prozess ist weitestgehend automatisiert. Die Bank kann die Umsatzdaten direkt im BWS in der Anwendung Grobsegmentierung bestellen. Im Folgenden wird der Prozess beschrieben:

1. Bestellung im BWS in der Anwendung Grobsegmentierung
2. Notwendige Verträge mit der VR Data
3. Lieferung von Kundendaten an die VR Data
4. Anreicherung der an die VR Data gelieferten Kundendaten

5. Auslieferung der Daten an die Bank und Verwendung im Rahmen der Grobsegmentierung im BWS

Nach der Bestellung der Bank im BWS per Knopfdruck kümmert sich die VR Data zunächst darum, dass die erforderlichen Verträge geschlossen werden: einen Servicevertrag Datenpakete und bei Bedarf auch einen Vertrag zur Auftragsverarbeitung (Datenschutz / DSGVO) sowie eine DORA-Vertragsergänzung, wenn diese noch nicht zwischen der Bank und der VR Data geschlossen wurden. Da geprüft werden muss, ob diese Verträge bereits vorliegen, ist dieser Schritt ein manueller Prozess, den die VR Data nach Eingang der Bestellung über den BWS startet. Die VR Data kommt bei Bedarf auf die Bestellerbank zu.

Wenn alle Verträge geschlossen sind, startet die VR Data den Erstellungsprozess. Aus dem BWS heraus werden automatisiert bestimmte Kundendaten an die VR Data geliefert, um die Firmenkunden der Bank eindeutig identifizieren zu können (siehe Abschnitt C). Welche Kundendaten an die VR Data geliefert werden, können Sie dem Fachkonzept „Datenabzug bei der Bank“ entnehmen. Anders als im Fachkonzept „Datenabzug bei der Bank“ beschrieben, werden die Kundendaten, die an die VR Data geliefert werden, zwar nicht per IDA-Bericht erzeugt, die Datenfelder sind aber mehr oder weniger die gleichen. Der Datenaustausch wurde mit der BMS CS abgestimmt und erfordert kein manuelles Eingreifen seitens der Bank.

Wenn die Kunden eindeutig identifiziert sind, wird geprüft, ob der Bank bereits eine Umsatzangabe vorliegt und wie alt diese ist. Wenn der Umsatz fehlt oder wenn der bei der Bank vorhandene Umsatz zu alt ist, wird ein neuer Umsatzwert aus dem Datenbestand der VR Data geliefert. Die erzeugte Datei wird automatisch an den BWS zurückgeliefert und dort im Rahmen der Grobsegmentierung verarbeitet.

C. Zuordnung der Kunden der Bank im Datenbestand der VR Data GmbH

Um den Firmenkunden der Bank Daten zuordnen zu können, muss die Firma zunächst im Datenbestand der VR Data eindeutig identifiziert bzw. zugeordnet werden. Die Zuordnung der Kunden der Bank im Datenbestand der VR Data erfolgt in einem zweistufigen Prozess:

- i. automatisierte Abgleichregeln (sogenannte Matchingverfahren)
- ii. manuelle Nachbearbeitung (dort, wo erforderlich)

i.) Automatisierte Matchingverfahren werden dann angewendet, wenn die Eindeutigkeit mit hinreichender Sicherheit gegeben ist. Dies ist bei in einem Register geführten Firmen zum Beispiel dann der Fall, wenn der Firmenname, die Anschrift und die Registernummer in den Stammdaten der Bank genau mit dem Firmennamen, der Anschrift und der Registernummer im Datenbestand der VR Data übereinstimmt. Die Eindeutigkeit der Zuordnungen, die aus dem automatisierten Matchingverfahren resultieren, wurden im Rahmen des Realisierungsprozesses intensiv validiert.

ii.) Eine manuelle Nachbearbeitung ist in den Fällen erforderlich, in denen die Zuordnung nicht eindeutig ist. In diesen Fällen werden die Zuordnungen mithilfe einer eigens hierfür programmierten Bedienoberfläche manuell überprüft. Kann die Firma vom Sachbearbeiter eindeutig identifiziert werden, erfolgt eine Zuordnung. Falls nicht, können für die betreffende Firma im Rahmen des Datenpaketes keine Daten geliefert werden.

Beim Matching unterscheiden wir grundsätzlich zwischen in einem öffentlichen Register geführten Firmen (intern nennen wir dies die Kategorie FK) und nicht in einem Register geführten Firmen (diese nennen wir intern Kategorie GuF, da sie insbesondere die große Gruppe der Gewerbetreibenden und Freiberufler umfasst. Diese Kategorie umfasst aber auch kommunale und kirchliche Unternehmen). Die Zuordnungsquoten zwischen diesen beiden Kategorien unterscheiden sich deutlich. Für in einem Register geführte Firmen erreichen wir je nach Qualität der von der Bank gelieferten Daten regelmäßig Quoten von rund 97%. In der Kategorie der Gewerbetreibenden und Freiberufler liegen die Quoten i. d. R. bei rund 55%. Dies liegt zum einen daran, dass es hier keine Registernummern gibt, die wir für die Zuordnung nutzen können. Zum anderen liegt im Datenbestand der Bank i. d. R. die Privatanschrift vor und im Datenbestand der VR Data die Anschrift der Firma des Gewerbetreibenden. Diese beiden Anschriften weichen häufig voneinander ab. Daher werden in der Kategorie GuF zusätzlich weitere Daten für das Matching verwendet, so z.B. weitere, in agree21 hinterlegte Anschriften, aber auch Telefonnummern und Mailadressen usw.

Im Datenpaket G kann die Bank auswählen, ob sie nur Daten über in einem Register geführte Firmen (Kategorie FK) oder nur Daten über nicht in einem Register geführte Firmen (Kategorie GuF) oder ob sie für beide Kategorien Daten beziehen möchte. Dies legt die Bank bei der Bestellung des Datenpakets G im BankingWorkspace fest. Sollte sich die Bank nach der Bestellung noch einmal umentscheiden, können Sie dies der VR Data mitteilen. Die VR Data hat die Möglichkeit, den Bestellprozess zu stornieren. Daraufhin muss die Bank den Bestellvorgang nochmals neu im BWS mit der letztendlich gewünschten Kategorie bzw. den letztendlich gewünschten Kategorien starten.

D. Die Umsatzdaten der VR Data

Die Umsatzdaten, die die VR Data von einer großen deutschen Wirtschaftsauskunftei bezieht, haben drei mögliche Quellen: Sie stammen entweder aus einer im Bundesanzeiger veröffentlichten Bilanz (bilanziert), sie wurden von der Auskunftei recherchiert (recherchiert), oder sie wurden auf der Basis von anderen Daten geschätzt bzw. modelliert (modelliert).

- Bilanziert: Der Umsatz stammt von Firmen, die ihre Umsatzzahlen im Rahmen der Bilanzpublizität veröffentlichen. Die Auskunfteien sammeln die veröffentlichten Jahresabschlussdaten und pflegen sie in ihre Wirtschaftsdatenbanken ein. Der Jahresumsatz ist jedoch leider nur für große Kapitalgesellschaften nach § 267 HGB publizitätspflichtig, daher sind bilanzierte Umsatzdaten nur für wenige deutsche Firmen verfügbar. Die Zahl liegt im mittleren fünfstelligen Bereich. Der Umsatz ist meistens größer als 50 Mio. Euro.

- **Recherchiert:** Recherchierte Umsatzdaten sind Umsatzdaten von nicht offenlegungspflichtigen Firmen. Die Daten werden vom Datenlieferanten der VR Data auf Internetseiten, in Pressemeldungen oder direkt beim jeweiligen Unternehmen recherchiert. Die Zahl liegt im mittleren sechsstelligen Bereich. Da große Unternehmen ihren Umsatz publizieren müssen, liegen recherchierte Umsatzdaten v.a. für kleine und mittlere Unternehmen vor.
- **Modelliert:** Für alle anderen Firmen wird der Jahresumsatz auf Basis weiterer externer Daten hochgerechnet (beispielsweise auf Basis der Branche, des Unternehmensalters, der Region und – sofern verfügbar – der Bilanzsumme und der Anzahl der Beschäftigten). Die Zahl liegt im mittleren siebenstelligen Bereich. Modellierte Daten werden von den Auskunftsteilen hauptsächlich für kleine Firmen bis 5 Mio. Euro Umsatz ausgewiesen. Rund 99% der modellierten Umsatzdaten sind unter der Schwelle von 5 Mio. Euro.

Wir haben die modellierten Daten stichprobenartig geprüft. Da die Auskunftsteil die tatsächlichen Werte in diesen Fällen nicht kennt, ist unseres Erachtens nachvollziehbar, dass die modellierten Werte häufig von den tatsächlichen Werten abweichen. In den Fällen, in denen wir die tatsächlichen Werte kannten, weil uns Pilotbanken diese im Rahmen einer Pilotierung zur Verfügung gestellt haben, war die Abweichung aber nicht so groß, dass diese Angaben nicht trotzdem im Rahmen einer Grobsegmentierung sehr wertvoll waren. Das gleiche gilt für recherchierte Daten. Auch hier stimmt der Wert verständlicherweise nicht immer mit dem tatsächlichen Wert überein. Schließlich kann die Firma beispielsweise bei einer Selbstauskunft oder der Veröffentlichung im Internet auch einen Wert angeben, der vom Wert im Jahresabschluss abweicht, um den Umsatz in der Außenwirkung größer oder kleiner darzustellen. Einige Firmen im Datenbestand der VR Data haben einen recherchierten Umsatz von 0 EUR. Dies ist i.d.R. vor allem dann der Fall, wenn die betreffende Firma nach unseren Informationen ihre Geschäftstätigkeit eingestellt hat. Wenn die Bank den Firmenkunden trotzdem noch in ihrem Bestand hat, weil vielleicht noch ein Kredit bedient werden muss, wird dieser Firma im Datenpaket Grobsegmentierung als Umsatz 0 EUR zugeordnet.

Für eine Grobsegmentierung sind modellierte und recherchierte Werte unseres Erachtens trotzdem wertvoll, weil sie eine erste Indikation darüber geben, welches Umsatzpotenzial ungefähr hinter diesen Firmenkunden steckt. Natürlich sind direkt von der Bank beim Firmenkunden erhobene Umsatzdaten besser. Der Aufwand einer massenhaften Erhebung von Umsatzdaten steht aber i.d.R. in keinem Verhältnis zum Nutzen.

Da der Bank zu einem gewissen Prozentsatz bereits Umsatzangaben vorliegen, muss definiert werden, wann Umsatzdaten an die Bank zurückgeliefert werden und wann nicht. Diese Definition erfolgt folgendermaßen:

- Wenn der von der Bank im Rahmen der Kundendaten gelieferte Umsatzwert leer ist, dann wird der Umsatzwert der VR Data zurückgeliefert.
- Wenn der von der Bank gelieferte Umsatzwert nicht leer ist, dann wird dann ein Umsatzwert zurückgeliefert, wenn der Umsatzwert der VR Data neuer (jünger) ist als der von der Bank gelieferte Umsatzwert.

Der VR Data liegen sowohl Umsatzdaten für in einem deutschen Register geführte Firmen als auch für nicht in einem Register geführte Firmen vor. Im Wesentlichen hat die VR Data Daten über jede deutsche Firma. Mit Sicherheit werden einige Firmen wie z.B. Nebenerwerbslandwirte oder der eine oder andere Gewerbetreibende oder Kleinstgewerbetreibende fehlen. Wir schätzen die Fehlquote aber auf einen Wert von unter 2%. Insgesamt stehen der VR Data Daten von rd. 5,2 Mio. wirtschaftlich aktiven Firmen zur Verfügung. Davon liegen bei 5,13 Mio. Firmen Umsatzdaten vor.

Die Bank kann wählen, ob sie das Datenpaket Grobsegmentierung nur für in einem Register geführte Firmen oder nur für nicht in einem Register geführte Firmen oder für beide Kategorien bestellen möchte. Die VR Data fragt dies nach der eingegangenen Bestellung nochmals bei der Bank ab, um mögliche Missverständnisse zu vermeiden.

E. Beschreibung der Datei, die an die Bank zurückgeliefert wird, und Verwendung der Ergebnisse in der Bank

In der Ergebnislieferung über den BWS werden im Wesentlichen folgende Daten zurückgeliefert:

Personennummer	Personennummer des Firmenkunden der Bank
GENO-UID	Eindeutige ID für den Firmenkunden der Bank im Datenbestand der VR Data (GENO-Unique Identifier)
Umsatz	Umsatz des Firmenkunden der Bank in Euro
Quelle	Quelle der Umsatzangabe, also bilanziert, recherchiert oder modelliert
Umsatztyp	EINZEL oder KONZERN. Spezifikation, ob es sich bei dem von der VR Data gelieferten Umsatzwert um den Umsatz der Einzelfirma oder um den Umsatz des Konzerns handelt. In der Regel hat die VR Data nur die Einzelwerte. Konzernwerte liegen eher selten vor. Wenn beide Werte vorliegen, wird in Absprache mit der BMS CS der Konzernwert zurückgeliefert.
Jahr des Umsatzes	Jahr, aus dem der Umsatzwert stammt. Bei bilanzierten Werten handelt es sich um das Jahr des veröffentlichten Jahresabschlusses. Bei recherchierten und modellierten Werten um das Jahr, in dem der Wert recherchiert bzw. modelliert wurde.

Diese Daten werden automatisiert an den BWS geliefert und dort im Rahmen der Grobsegmentierung verarbeitet. Ein manuelles Verarbeiten der Daten ist nicht erforderlich.

Appendix 1: Beschreibung der Datenquellen

Die Daten der VR Data stammen aus einer kommerziellen Wirtschaftsdatenbank einer großen deutschen Wirtschaftsauskunftei. Mit der Auskunftei besteht ein entsprechender Kooperations- bzw. Lizenzvertrag.