



Digitalisierung in der amtlichen Statistik – Nutzung von Verwaltungsdaten

Integration von Daten des elektronischen Handelsregisters
zur Qualitätssicherung im statistischen Unternehmensregister

Volker Schneider



Dipl.-Volkswirt Volker Schneider ist Referent im Referat „Unternehmensregister, Gewerbeanzeigen, Verkehr“ des Statistischen Landesamtes Baden-Württemberg.

Erklärtes Ziel der amtlichen Statistik ist es, die Statistikproduktion möglichst effizient sowie belastungsarm für die Auskunftspflichtigen zu gestalten. Eine seit langem eingeübte Praxis stellt in diesem Zusammenhang die Nutzung von digitalisierten Verwaltungsdaten dar. Die Ausweitung der Nutzung von digitalen Verwaltungsdaten und Registern gehört unter anderem zu den Maßnahmen, die die Statistischen Ämter des Bundes und der Länder im Rahmen ihrer „Digitalen Agenda“ im November 2018 beschlossen haben. Damit stärkt der statistische Verbund diesen Ansatz. Eine bislang unzureichend genutzte Quelle zur Qualitätssicherung des statistischen Unternehmensregisters ist das elektronische Handelsregister. Der Beitrag geht konkret auf dessen Nutzbarmachung und die damit verbundenen Potenziale ein. Mit der im Beitrag beschriebenen technischen Umsetzung kann die Qualität des statistischen Unternehmensregisters deutlich verbessert werden.

Verwaltungsdaten – das Statistische Unternehmensregister und das elektronische Handelsregister

Verwaltungsdaten sind eine wichtige Datenquelle der amtlichen Statistik. Es handelt sich dabei um Informationen, die zunächst rein aus verwaltungstechnischen Gründen anfallen, bei Eignung und nach Nutzbarmachung aber zur effizienten und belastungsarmen Statistikproduktion und zur Qualitätssicherung in der amtlichen Statistik genutzt werden können. Ein Beispiel für die Nutzung von Verwaltungsdaten in der amtlichen Statistik ist das statistische Unternehmensregister.

Das statistische Unternehmensregister (im weiteren Unternehmensregister) ist eine Datenbank, die Unternehmen und Betriebe über nahezu alle Wirtschaftsbereiche führt und ausweist.¹ Es speist sich im Wesentlichen aus Daten der Bundesagentur für Arbeit und der Finanzverwaltung. Das Unternehmensregister ist ein wichtiges Instrument der amtlichen Statistik zur Unterstützung der Erhebungen. Es dient den Unternehmensstatistiken zur Berichtskreispflege und stellt bei Stichprobenerhebungen

die Auswahlgrundlage dar. Gleichzeitig werden aus dem Unternehmensregister aber auch Statistiken erstellt, für die keine weiteren Befragungen notwendig sind. Exemplarisch stehen hierfür die Handwerksstatistiken (Vierteljährliche Handwerksberichterstattung und Jährliche Handwerkszählung). Nach der Klassifizierung von im Unternehmensregister geführten Unternehmen als Handwerksunternehmen² werden diese Statistiken direkt aus dem Unternehmensregister erzeugt. Ein weiteres Beispiel ist die Statistik der Unternehmensdemografie, die ebenfalls aus dem Unternehmensregister erzeugt wird.³

Die Nutzung von Verwaltungsdaten für die amtliche Statistik ist nur zulässig, wenn dies explizit in einem Gesetz geregelt ist. Für das statistische Unternehmensregister sind das Statistikregistergesetz⁴ und das Verwaltungsdatenverwendungsgesetz⁵ maßgebend. In diesen Gesetzen wird geregelt, welche Daten zu welchen Zwecken genutzt werden dürfen.

Mit der Nutzung der oben genannten Verwaltungsdaten der Bundesagentur für Arbeit und der Finanzverwaltung konnten auskunftspflichtige Unternehmen entlastet und die Effizienz der amtlichen Statistik gesteigert werden. Wichtig bei der Nutzung von Registerdaten für die Erstellung von Statistiken ist auch die Qualitätssicherung dieser Datenquellen. Ein unzureichend genutzter Ansatz ist die Verwendung der digitalisierten Handels-, Genossenschafts- und Partnerschaftsregister (im weiteren elektronisches Handelsregister). Nach Umstellung des Unternehmensregisters auf eine neue Version im Jahr 2014 stand die Integration der Handelsregisterdaten in die teilautomatisierten Verarbeitungsschritte des zentralen Unternehmensregisters weit oben auf der Agenda. Das Statistische Landesamt Baden-Württemberg hat die in der Verarbeitung des elektronischen Handelsregisters steckenden Möglichkeiten zur Reduzierung der Arbeitsbelastungen in den Landesämtern und die mit einer integrierten Verarbeitung verbundenen Qualitätspotenziale immer hoch eingeschätzt und die Umsetzung des Projektes stets befürwortet. Aufgrund anderer Prioritätensetzungen wurde dieses, als EHUG-Verarbeitung bezeichnete Projekt, im

1 Nicht ausgewiesen werden derzeit die Abschnitte A (Land- und Forstwirtschaft, Fischerei), O (Öffentliche Verwaltung, Verteidigung), T (Private Haushalte mit Hauspersonal) und U (Exterritoriale Organisationen und Körperschaften) der Klassifikation der Wirtschaftszweige 2008 (WZ 2008).

2 Für die Klassifizierung werden Verwaltungsdaten der Handwerkskammern verwendet.

3 Siehe hierzu bspw. Volker Schneider: Unternehmensgründungen und ihre wirtschaftliche Bedeutung – Methodik und erste Ergebnisse – Vorstellung der Statistik der Unternehmensdemografie, in Statistisches Monatsheft Baden-Württemberg 1/2019, S. 3 ff.

4 Statistikregistergesetz vom 16. Juni 1998 (BGBl. I S. 1300), das zuletzt durch Artikel 2 des Gesetzes vom 21. Juli 2016 (BGBl. I S. 1768) geändert worden ist.

5 Verwaltungsdatenverwendungsgesetz vom 4. November 2010 (BGBl. I S. 1480), das durch Artikel 1 des Gesetzes vom 18. Dezember 2018 (BGBl. I S. 2637) geändert worden ist.

Statistischen Verbund bislang allerdings nicht umgesetzt. Aus diesen Gründen ergriff das Statistische Landesamt Baden-Württemberg die Initiative, die Potenziale für die Pflege des Unternehmensregisters in Baden-Württemberg zu heben. Vergleichbare Initiativen betrieben auch IT.NRW und das Bayerische Landesamt für Statistik.

Datenzugang

§ 4a des Statistikregistergesetzes gewährt den Statistischen Landesämtern den direkten Zugriff auf Daten des elektronischen Handelsregisters. Nach Abstimmung und mit Unterstützung der Landesjustizverwaltung erhielt das Statistische Landesamt Baden-Württemberg schließlich über eine Web-Schnittstelle Zugriff auf die sogenannten Bekanntmachungen des elektronisch geführten Handelsregisters. Die Bekanntmachungen liefern Informationen über Änderungen von zentraler Bedeutung von im Handelsregister eingetragenen Einheiten. Konkret enthalten die Bekanntmachungen Hinweise auf Insolvenzverfahren, Umfirmierungen oder Änderungen von Geschäftsanschriften. Es handelt sich dabei um frei zugängliche Informationen, denn die Einsichtnahme in das Handelsregister ist nach Handelsgesetzbuch § 9 jedem zu Informationszwecken gestattet.

Datenabruf

Ausgestattet mit den grundlegenden Informationen zur Web-Schnittstelle des elektronisch geführten Handelsregisters konnte ein Client entwickelt werden, der es ermöglicht, die monatlichen Bekanntmachungen der baden-württembergischen Registergerichte Freiburg, Mannheim, Stuttgart und Ulm abzurufen.

Für das Geschäftsjahr 2018 konnte der Fachbereich so rund 60 000 Bekanntmachungen abrufen. Bei den Bekanntmachungen kann es sich um einfache Umfirmierungen bis hin zu komplexen Vorgängen, beispielsweise eine Umfirmierung plus Änderung der Geschäftsanschrift und der Übernahme weiterer Gesellschaften, handeln. Die monatlichen Abrufe liefern als Merkmale

- das Registergericht,
- den Registerort,
- die Registernummer,
- das Bekanntmachungsdatum,
- den Gegenstand,
- den Bekanntmachungstext,
- den Firmennamen,
- den Firmensitz,
- die Straße,

- die Hausnummer,
 - die Postleitzahl und
 - den Ort
- der eingetragenen rechtlichen Einheiten.

Die Merkmale Registergericht, Registerort und Registernummer, sowie der Firmenname, der Firmensitz, die Straße, die Hausnummer, die Postleitzahl und der Ort werden zur Identifikation der zugehörigen Unternehmensregister-Einheiten benötigt. Die für den Fachbereich-zentralen Informationen verbergen sich in den Bekanntmachungstexten und müssen analysiert werden.

Maschinelle Textanalyse

Die Analyse der Bekanntmachungstexte erfolgt mittels eines speziell entwickelten Algorithmus. Zunächst führt der Algorithmus vorbereitende Arbeiten durch, das heißt er liest die relevanten Daten ein, benennt Merkmale um und berechnet neue Merkmale. Danach beginnt der eigentliche Kernprozess, die Analyse und Klassifizierung der in den Bekanntmachungstexten dokumentierten Ereignisse. Dazu wird der Bekanntmachungstext auf bestimmte Signalwörter untersucht, über die sich die derzeit zehn in der *Übersicht* beschriebenen Ereignisse klassifizieren lassen. Existiert beispielsweise im Bekanntmachungstext eines untersuchten Datensatzes die Zeichenfolge „Firma geändert“, so wird diesem Datensatz das Ereignis Umfirmierung zugeschrieben. In der Praxis kommt es natürlich vor, dass eine Umfirmierung und eine Änderung der Geschäftsanschrift Hand in Hand gehen. Der Algorithmus stellt solche multiplen Ereignisse ebenfalls über Signalwörter fest und klassifiziert die Bekanntmachung entsprechend als Umfirmierung und Änderung der Geschäftsanschrift.

Im Rahmen der Entwicklung des Algorithmus wurde untersucht, welche Signalwörter welche Ereignisse treffsicher beschreiben und abgrenzen, welche Ereignisse für die Unternehmensregister-Pflege von Relevanz sind. Das *Schaubild* gibt die absoluten und relativen Häufigkeiten der, für das Geschäftsjahr 2018, erkannten Ereignisse wieder.

Verknüpfung mit dem Unternehmensregister

Um die gewonnenen Informationen im Rahmen der Pflege und der Qualitätssicherung des Unternehmensregisters nutzen zu können, müssen diese schließlich den passenden Einheiten im Unternehmensregister zugeordnet werden. Dabei sind drei Teilmengen zu unterscheiden.

Ü

Abgrenzung der für die Unternehmensregister-Pflege relevanten Ereignisse

Ereignis	Erläuterung
Neueintragung	Es wurde eine neue rechtliche Einheit im Handelsregister eingetragen.
Geschäftsanschrift geändert	Die Geschäftsanschrift einer rechtlichen Einheit wurde geändert.
Registerblatt geschlossen	Eine rechtliche Einheit wurde geschlossen.
Sitzverlegung	Der Satzungssitz einer rechtlichen Einheit wurde verlegt.
Firma geändert	Die Firmierung einer rechtlichen Einheit wurde geändert.
Gegenstand geändert	Der Gegenstand einer rechtlichen Einheit wurde geändert.
Löschungsankündigung	Eine rechtliche Einheit wird/wurde wegen Vermögenslosigkeit oder aus sonstigen Gründen gelöscht.
Verschmelzung	Die rechtliche Einheit wurde mit einer anderen rechtlichen Einheit zu einer neuen rechtlichen Einheit zusammengeführt.
Ausgliederung	Die Vermögenswerte einer rechtlichen Einheit wurden auf eine andere rechtliche Einheit übertragen.
Registernummer geändert	Die als Identifikator dienende Registerkennung einer rechtlichen Einheit wurde geändert.

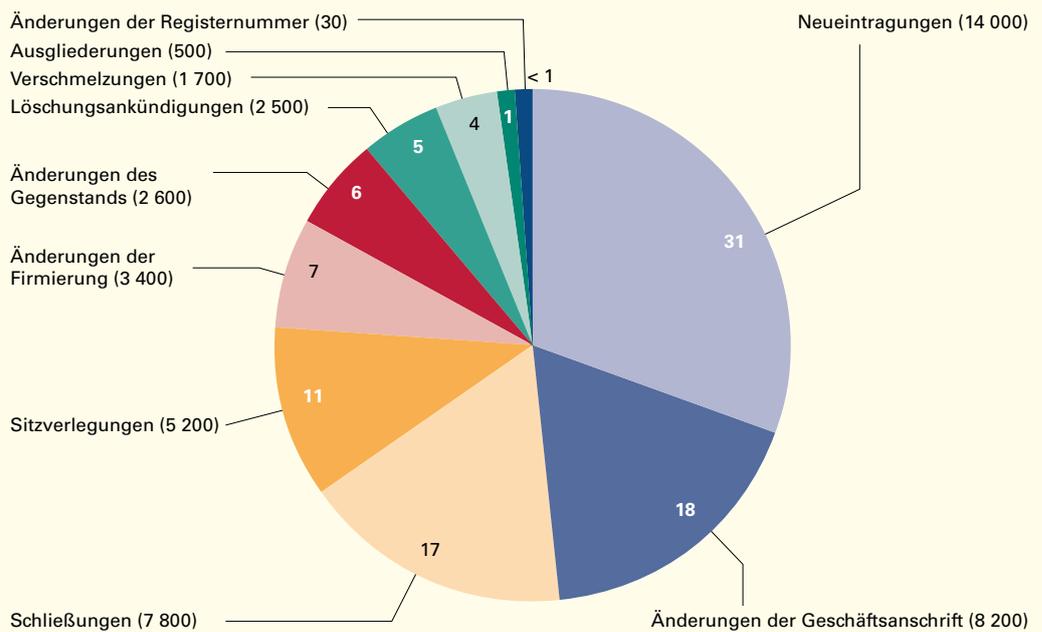
Die erste Teilmenge bilden die rechtlichen Einheiten aus dem Handelsregister, die sich eindeutig mit im Unternehmensregister gepflegten rechtlichen Einheiten über die Kombination aus Registergericht, Registerart und Registernummer verknüpfen lassen. Diese Menge wird nach Zuordnung an weitere Verarbeitungsstufen übergeben.

Die zweite Teilmenge umfasst Einheiten, die sich nur über Namensabgleiche mit rechtlichen Einheiten des Unternehmensregisters verknüpfen lassen. Dies kommt vor, wenn für eine im Handelsregister eingetragene rechtliche Einheit, die ebenfalls bereits im Unternehmensregister vorliegt, die oben genannten Identifikatoren nicht im Unternehmensregister einge-

S

Handelsregister-Bekanntmachungen in Baden-Württemberg im Geschäftsjahr 2018

Anzahl in (), Anteile in %



Datenquelle: Elektronisches Handelsregister; Statistisches Unternehmensregister.

tragen sind. Ursache dafür kann beispielsweise sein, dass diese Merkmale in den administrativen Daten nicht besetzt waren.

Die dritte Teilmenge ist schließlich die Menge an im Handelsregister eingetragenen rechtlichen Einheiten, die sich nicht im Unternehmensregister wiederfinden. In der Regel handelt es sich dabei um Neueintragungen, die für das Unternehmensregister relevant sind, sobald die mit den rechtlichen Einheiten verbundenen Unternehmen die Relevanzschwelle⁶ überschreiten.

Herausforderung Namensabgleich

Eine besondere Herausforderung stellen die nicht direkt zuordenbaren rechtlichen Einheiten der zweiten Teilmenge dar. Wegen fehlender eindeutiger Identifikatoren müssen sie über Namensabgleiche den paarigen Unternehmensregister-Einheiten zugeordnet werden. Hierzu wurde ein separater Algorithmus entwickelt, dem die von *Wladimir Lewenstein*⁷ entwickelte sogenannte Levenshtein-Distanz zugrunde liegt. Der implementierte Algorithmus gleicht die Firmenbezeichnung und die Adressangaben zwischen Einheiten im Unternehmensregister und im Handelsregister ab und ermittelt daraus eine Zuordnungswahrscheinlichkeit. Liegt die Übereinstimmung in diesen Merkmalen zwischen zwei Einheiten über einem voreingestellten Wert von beispielsweise 90 %, wird dieser und werden alle weiteren Übereinstimmungen größer 90 % als Zuordnungsvorschläge ausgegeben. Die abschließende Bewertung der Vorschläge unterliegt der manuellen Prüfung.

Nutzung der Daten im Unternehmensregister

Nach Zuordnung der erweiterten Datensätze zu den rechtlichen Einheiten des Kernregisters folgt der für das Unternehmensregister gewinnbringende letzte Arbeitsschritt, die Übernahme der den Handelsregister-Bekanntmachungen entnommenen Ereignisse.

Alle verknüpften Datensätze werden dazu in die manuelle Bearbeitung gegeben und die Änderungen von dort im Unternehmensregister umgesetzt. Die Datensätze enthalten Informationen über singuläre oder multiple Ereignisse, von denen die rechtlichen Einheiten im Unternehmensregister betroffen sind. Ein multiples Ereignis liegt beispielsweise vor, wenn eine Ausgliederung mit einer Umfirmierung zusammenfällt. Dabei ist die Ausgliederung ein besonders komplexes Ereignis. Bei einer

Ausgliederung ist nicht nur zu klären, welche Niederlassungen des betroffenen Unternehmens gegebenenfalls auszugliedern sind, sondern auch, welches Unternehmen diese Niederlassungen aufnimmt. Neben den ausführlichen Bekanntmachungstexten werden in der Bearbeitung weitere Quellen (Internet, administrative Daten der Bundesagentur für Arbeit und der Finanzverwaltung) hinzugezogen. Mittelfristig besteht auch die Möglichkeit, Änderungen maschinell umzusetzen – dies erfordert allerdings noch konzeptionelle Entwicklungsarbeit. Mit der Verarbeitung der Daten des elektronischen Handelsregisters können relevante Ereignisse und damit verbundene Veränderungen umfassend und in hoher Qualität im Unternehmensregister abgebildet werden. Die Bekanntmachungen des Handelsregisters liefern hierfür wertvolle Informationen. Neben den konkret betroffenen Unternehmen gibt es auch Hinweise zu indirekt betroffenen Unternehmen. Zudem erhöhen die Angaben und strukturellen Informationen die Qualität der Unternehmensinformationen und die Verknüpfungsqualität im Unternehmensregister. Gerade die Verknüpfungsqualität spielt im Unternehmensregister eine zentrale Rolle. Mit der Kombination aus Registerort, Registerart und Registerkennung ist ein wesentlicher Anteil der im Unternehmensregister gepflegten Einheiten eindeutig identifizierbar. Damit werden automatische Verknüpfungen mit Quellen, die ebenfalls diesen Schlüssel führen, direkt möglich. Dies gilt insbesondere für die Verarbeitung von Unternehmensgruppendaten, auf deren Basis im Unternehmensregister Unternehmenszusammenhänge abgebildet werden.

Fazit

Die Bekanntmachungen des elektronisch geführten Handelsregisters liefern zuverlässige Informationen über Unternehmensmerkmale und demografische Ereignisse. Die Kombination aus Registerort, Registerart und Registerkennung ermöglicht es zudem diese Informationen eindeutig mit bereits im Unternehmensregister gepflegten Einheiten zu verknüpfen. Mit dem entwickelten Algorithmus hat sich das Statistische Landesamt Baden-Württemberg bereits heute die Möglichkeit geschaffen, diese Informationen im Sinne der Verarbeitungseffizienz und der Qualitätssteigerung des Unternehmensregisters zu nutzen. Neben der Etablierung der Verarbeitung und Fortentwicklung des Algorithmus setzt sich Baden-Württemberg weiterhin für den Aufbau einer zentralen bundesweiten und im Unternehmensregister integrierten Verarbeitungsmöglichkeit von Handelsregisterdaten ein. ■

⁶ Ein Unternehmen wird im Unternehmensregister relevant, wenn es im Berichtsjahr einen Umsatz von mehr als 17 500 Euro erzielte und/oder über Beschäftigte verfügte.

⁷ Wladimir Iossifowitsch Lewenstein (1935–2017), engl. Levenshtein.