

IW-Trends

Data Sharing in Deutschland

Jan Büchel / Barbara Engels

IW-Trends 2/2023

Vierteljahresschrift zur
empirischen Wirtschaftsforschung
Jahrgang 50



Herausgeber

Institut der deutschen Wirtschaft Köln e.V.

Postfach 10 19 42
50459 Köln
www.iwkoeln.de

Das Institut der deutschen Wirtschaft (IW) ist ein privates Wirtschaftsforschungsinstitut, das sich für eine freiheitliche Wirtschafts- und Gesellschaftsordnung einsetzt. Unsere Aufgabe ist es, das Verständnis wirtschaftlicher und gesellschaftlicher Zusammenhänge zu verbessern.

Das IW in den sozialen Medien

Twitter
[@iw_koeln](https://twitter.com/iw_koeln)

LinkedIn
[@Institut der deutschen Wirtschaft](https://www.linkedin.com/company/institut-der-deutschen-wirtschaft)

Facebook
[@IWKoeln](https://www.facebook.com/IWKoeln)

Instagram
[@IW_Koeln](https://www.instagram.com/IW_Koeln)

Verantwortliche Redakteure

Prof. Dr. Michael Grömling

Senior Economist
groemling@iwkoeln.de
0221 4981-776

Holger Schäfer

Senior Economist
schaefer.holger@iwkoeln.de
030 27877-124

**Alle Studien finden Sie unter
www.iwkoeln.de**

Rechte für den Nachdruck oder die elektronische Verwertung erhalten Sie über lizenzen@iwkoeln.de.

In dieser Publikation wird aus Gründen der besseren Lesbarkeit regelmäßig das grammatische Geschlecht (Genus) verwendet. Damit sind hier ausdrücklich alle Geschlechteridentitäten gemeint.

ISSN 1864-810X (Onlineversion)

© 2023
Institut der deutschen Wirtschaft Köln Medien GmbH
Postfach 10 18 63, 50458 Köln
Konrad-Adenauer-Ufer 21, 50668 Köln
Telefon: 0221 4981-450
iwmedien@iwkoeln.de
iwmedien.de

Data Sharing in Deutschland: Verbreitung, Intensität und Potenziale

Jan Büchel / Barbara Engels, Mai 2023

Zusammenfassung

Digitale Daten sind ein zentrales Element von Prozessen, Produkten und Geschäftsmodellen. Dementsprechend wichtig ist das Teilen dieser Daten. In diesem Beitrag wird untersucht, wie intensiv Unternehmen in Deutschland Data Sharing betreiben. Dazu wurden 1.051 Industrieunternehmen und industriennahe Dienstleister befragt. Die Antworten zur Qualität und Quantität des Data Sharing werden in ein Modell eingespeist, das den Unternehmen Intensität-Scores für ihre Datenabgabe und ihren Datenempfang zuweist. Den Ergebnissen zufolge beteiligen sich 58 Prozent der Unternehmen in Deutschland nicht am Data Sharing – weder geben sie Daten ab noch empfangen sie welche. 38 Prozent der befragten Unternehmen empfangen Daten und nur 21 Prozent der Unternehmen stellen Daten zur Verfügung. Wenn Unternehmen Daten abgeben, dann eher mit höherer (14 Prozent) als mit niedriger Intensität (7 Prozent). Der Datenempfang hingegen erfolgt bei 18 Prozent der Unternehmen mit hoher Intensität und bei 20 Prozent mit niedriger. Besonderes Entwicklungspotenzial zeigt sich daher bei der Datenabgabe in Bezug auf die Verbreitung und beim Datenempfang bezüglich der Intensität. Generell liegt bei der Intensität der Datenabgabe und des Datenempfangs ein positiver statistischer Zusammenhang vor. Vor allem Datengeber empfangen in den meisten Fällen auch Daten von anderen Unternehmen. Einen Kooperationspartner zu finden, Rechtssicherheit und Datensouveränität zu gewährleisten, sind Herausforderungen beim Data Sharing, die durch Unternehmen oft nicht allein zu meistern sind.

Stichwörter: Digitalisierung, Daten, Datenwirtschaft, Unternehmen

JEL-Klassifikation: D29, L29, O12

DOI: 10.2373/1864-810X.23-02-02

Data Sharing als Basis für Wettbewerbsfähigkeit

Daten sind ein begehrter Produktionsfaktor in der digitalisierten Wirtschaft. Die Digitalisierung von Prozessen, Produkten und Geschäftsmodellen basiert auf Daten. Vor allem ist Künstliche Intelligenz nur durch die Nutzung von Daten möglich (Rusche et al., 2022, 19 f.). Um die Wettbewerbs- und Zukunftsfähigkeit der deutschen Wirtschaft zu beurteilen, ist Wissen darüber erforderlich, inwiefern Unternehmen Daten nutzen und insbesondere auch mit anderen Unternehmen teilen. Zum Beispiel steigt die Wahrscheinlichkeit, dass Unternehmen innovative Produkte entwickeln, die sich am Markt etablieren, wenn sie große Datenmengen analysieren (Niebel et al., 2019). Ein internationaler Vergleich für das Jahr 2022 zeigt, dass die Datenökonomie gemessen an der Wertschöpfung, die durch Daten gewonnen wird, in der EU nach den USA am zweitgrößten ist (IDC/Lisbon Council, 2023, 157 f.). Allerdings wächst sie im Vergleich zum Jahr 2021 sowohl in den USA als auch in China und Japan schneller als in der EU. Data Sharing beinhaltet, Daten an andere Unternehmen abzugeben oder von diesen Daten zu empfangen. Denn die größten Potenziale der Datennutzung können realisiert werden, wenn Daten aus unterschiedlichen Quellen kombiniert werden. Dies gilt etwa für die Optimierung von Prozessen: Wenn Produktionsdaten von verschiedenen Akteuren der Wertschöpfungskette geteilt werden, können Lieferketten transparenter werden (Büchel/Engels, 2022). Kommt es zu Engpässen auf vorgelagerten Stufen der Wertschöpfungskette, sind Akteure auf nachgelagerten Stufen bestenfalls in Echtzeit darüber informiert und können entsprechend automatisiert reagieren. Neue Geschäftsmodelle und Produkte sind oft nur möglich, wenn Unternehmen nicht nur ihre eigenen Daten nutzen, sondern auch Daten von anderen Unternehmen und Institutionen einbeziehen können. Zudem kann die Abgabe von Daten an andere Teil eines neuen Geschäftsmodells sein. So wie Daten ein integraler Bestandteil von Prozessen, Produkten und Geschäftsmodellen geworden sind, ist Data Sharing nunmehr zentral für die Zukunftsfähigkeit von Unternehmen, da es Zugang zu mehr Daten und somit Möglichkeiten der Datenanalyse ermöglicht.

Im Folgenden wird ermittelt, inwiefern sich Unternehmen in Deutschland am Data Sharing beteiligen, also Daten teilen oder empfangen. Dafür wird die Intensität des Data Sharing von Unternehmen in Deutschland bestimmt. Basis für diese Analyse ist eine repräsentative CATI-Befragung (Computer Assisted Telephone Interview) unter

1.051 Unternehmen aus Industrie und industrienahen Dienstleistern in Deutschland, die im Herbst 2022 von der IW Consult durchgeführt wurde (Primärdaten). In einem Data-Sharing-Modell wird jedes der befragten Unternehmen verschiedenen Intensitätsstufen zugeordnet: Je mehr Unternehmen einer höheren Intensitätsstufe zuzuordnen sind, umso positiver kann die Zukunfts- und Wettbewerbsfähigkeit der deutschen Wirtschaft eingeschätzt werden.

Data-Sharing-Modell

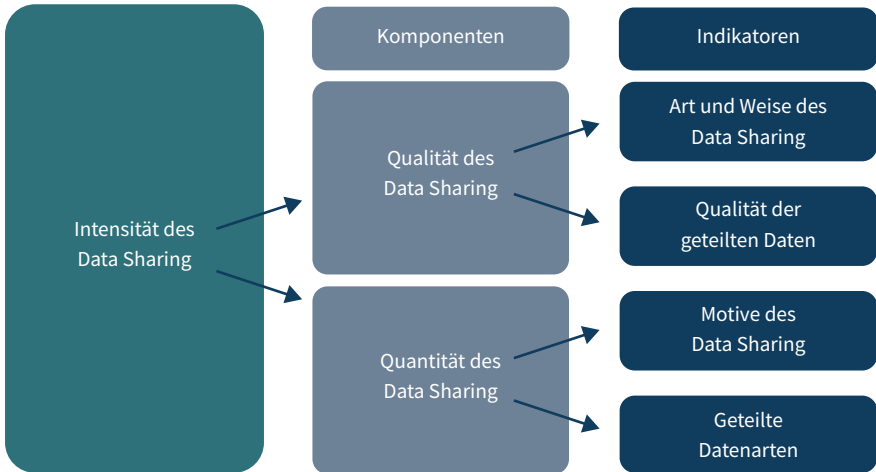
Es gibt Unternehmen, die nur wenige, ausgewählte Daten teilen. Beispiele hierfür sind Unternehmen, die lediglich Finanzdaten an Behörden weitergeben, weil sie rechtlich dazu verpflichtet sind. Andere Unternehmen teilen mehr und unterschiedliche Daten aus verschiedenen Gründen und zu diversen Zwecken. Zum Beispiel gibt es Unternehmen, die ihre Daten an Unternehmen abgeben, die damit datengetriebene Dienstleistungen aus dem Bereich der vorausschauenden Wartung anbieten. So werden Ausfallwahrscheinlichkeiten von Maschinen präziser ermittelt und Produktionsstopps vermieden. Die Dienstleistung ist für Kunden attraktiv, da sie so Kosten sparen können, erfordert von ihnen aber eine kontinuierliche Abgabe von Maschinen- und Produktionsdaten an den Dienstleister. Die geteilten Daten müssen den Anforderungen des Datenempfängers entsprechen, etwa wenn dieser Daten von verschiedenen Kunden in einem gemeinsamen Datensatz vereint.

Die vorstehenden Beispiele geben einen Eindruck davon, wie unterschiedlich Data Sharing gestaltet sein kann. Um zu beurteilen, inwiefern sich ein Unternehmen am Data Sharing beteiligt und wie hoch die Intensität des Data Sharing also ist, werden in dem hier verwendeten Data-Sharing-Modell die Qualität und die Quantität des Data Sharing betrachtet. Beide Komponenten werden anhand von jeweils zwei Indikatoren approximiert, die objektiv messbar und auf die Geschäftspraxis der befragten Unternehmen anwendbar sind (Übersicht). Alle vier Indikatoren werden für den Datenempfang und die Datenabgabe getrennt analysiert.

Die Komponente der Qualität des Data Sharing wird über die Indikatoren Art und Weise des Data-Sharing und die Qualität der geteilten Daten gemessen:

Komponenten und Indikatoren des Data-Sharing-Modells

Übersicht



Quelle: Institut der deutschen Wirtschaft

- **Art und Weise des Data Sharing:** Data Sharing kann papierbasiert, mit händischen Zwischenschritten (etwa als Dateianhang einer E-Mail), automatisiert oder automatisiert und in Echtzeit erfolgen.
- **Qualität der geteilten Daten:** Die Qualität der abgegebenen oder empfangenen Daten wird durch zahlreiche Faktoren determiniert. Eine Herausforderung ist, dass verschiedene Unternehmen – und oft sogar verschiedene Unternehmensbereiche – unterschiedliche Anforderungen an die Datenqualität haben (Engels/Schäfer, 2020). Im Data-Sharing-Modell werden deshalb Merkmale von Daten benutzt, die möglichst objektiv sind. Für eine hohe Datenqualität spricht etwa, dass sie maschinenlesbar sind, ein standardisiertes Format haben, in sich widerspruchsfrei und qualitätsgeprüft sind sowie über einen Identifier verfügen. Dieser macht jeden Datenpunkt eindeutig identifizierbar und somit mit anderen Datensätzen verknüpfbar.

Die Komponente der Quantität des Data Sharing wird mittels der Indikatoren Motive des Data Sharing und Anteil an geteilten Datenarten approximiert:

- **Motive des Data Sharing:** Unternehmen geben Daten aus unterschiedlichen Gründen und zu verschiedenen Zwecken ab. Bei mehreren Motiven ist davon auszugehen, dass Unternehmen Daten mit verschiedenen Unternehmen teilen und mehrere Datenquellen betroffen sind, was tendenziell die geteilte Datenmenge erhöht. Zu den abgefragten Gründen gehören unter anderem vertragliche Verpflichtungen, rechtliche Vorgaben, die Verbesserung von eigenen Prozessen, Produkten und Geschäftsmodellen sowie die Stärkung von Geschäftsbeziehungen.
- **Geteilte Datenarten:** Da es kaum möglich ist, die Menge an geteilten Daten zu quantifizieren und unternehmensübergreifend vergleichbar zu machen, dient der Anteil der geteilten Datenarten an den im Unternehmen digital gespeicherten Datenarten als weiteres Indiz für die geteilte Datenmenge. Zu den abgefragten Datenarten gehören unter anderem Produktdaten, Produktionsdaten, Nutzungsdaten von Kunden, Lieferantendaten, Finanzdaten sowie Forschungs- und Entwicklungsdaten.

Basierend auf dem Antwortverhalten zu den jeweils zwei Fragen zur Qualität und Quantität des Data Sharing werden die befragten Unternehmen in dem zugrunde liegenden Data-Sharing-Modell getrennt für Datenabgabe und Datenempfang drei verschiedenen Data-Sharing-Intensitätsstufen zugeordnet:

- **Stufe 0: keine** Datenabgabe oder **kein** Datenempfang,
- **Stufe 1:** Datenabgabe oder Datenempfang mit **niedriger** Intensität,
- **Stufe 2:** Datenabgabe oder Datenempfang mit **hoher** Intensität.

Eine Annahme des Modells ist, dass Qualität und Quantität in gleichem Umfang auf die Intensität des Data Sharing einzahlen. Beispielsweise ordnet das Modell ein Unternehmen, das nur wenige Daten in hoher Qualität teilt, der gleichen Stufe zu wie ein Unternehmen, das viele Daten in niedriger Qualität teilt. Gleiches gilt für ein Unternehmen, das Daten in mittelmäßiger Qualität teilt und deren Quantität exakt dem Mittelwert aus den ersten beiden Szenarien entspricht. Methodisch gelingt dies, indem das Antwortverhalten zunächst pro Frage oder Indikator auf einen Wert zwischen 0 und 1 skaliert wird. Die vier Indikatoren-Scores werden für jedes Unternehmen gleichgewichtet zum Geber- oder Empfänger-Score zusammengefasst. Aufgrund der

Gleichgewichtung und der identischen Anzahl an Fragen beeinflussen Qualität und Quantität das Endergebnis in identischem Umfang.

Das Modell stuft Unternehmen, deren Geber- oder Empfänger-Score den Wert 0,5 übersteigt, auf Stufe 2 (hohe Intensität) ein. Unternehmen, deren Score unterhalb von 0,5 liegt, werden auf Stufe 1 (niedrige Intensität) eingeordnet. Die Grenze von 0,5 wird gewählt, da ein Unternehmen mit diesem Score im Durchschnitt exakt die Hälfte der relevanten Aspekte aus Qualität und Quantität erfüllt. Unternehmen, die angeben, keine Daten an andere Unternehmen abzugeben oder von ihnen zu empfangen, werden auf Stufe 0 (keine Datenabgabe oder kein Datenempfang) eingestuft.

Im Detail werden die einzelnen Indikatoren-Scores wie folgt gebildet:

- Qualitätsindikator **Art und Weise**: Jedes Unternehmen gibt an, wie viel Prozent seiner Daten digital mit händischen Zwischenschritten, automatisiert oder automatisiert und in Echtzeit abgegeben oder empfangen werden. Die Antwortmöglichkeiten sind nach Fortschrittlichkeit sortiert. Sie enthalten teilweise Eigenschaften aus vorgelagerten Optionen und ergänzen diese um weitere Merkmale. Daher fließen fortschrittlichere Aspekte mit höherer Gewichtung in die Score-Bildung ein als grundlegende.
- Qualitätsindikator **Qualität der geteilten Daten**: Bei diesem Indikator wird geprüft, wie viele der Antwortitems, also der Datenqualitätskriterien, vom Unternehmen als gegeben genannt werden. Die verschiedenen Datenqualitätskriterien bauen im Vergleich zum ersten Indikator nicht inhaltlich aufeinander auf. Zum Beispiel sind die Eigenschaften, dass Daten vollständig sind sowie über cloudbasierte Plattformen abgerufen werden können, nicht eindeutig zueinander in Bezug zu setzen. Somit ist unklar, wie wichtig die einzelnen Kriterien für die Qualität der geteilten Daten sind. Aus diesem Grund werden die einzelnen Frageitems über eine Principal-Component-Analyse (PCA) gewichtet. Sie untersucht das Antwortverhalten aller Unternehmen auf wesentliche Muster und bildet sogenannte Hauptkomponenten. Auf sie können einzelne Frageitems in unterschiedlicher Stärke einwirken. Anhand der Stärke können Gewichte für die Frageitems abgeleitet werden. Folglich erhalten Qualitätskriterien eine höhere Gewichtung, wenn sie im Muster des Antwortverhal-

tens vieler Unternehmen eine bedeutende Rolle einnehmen. Stark korrelierende Frageitems wirken auf gemeinsame Hauptkomponenten ein. Dies drückt aus, dass ein gemeinsamer inhaltlicher Aspekt die Frageitems lenkt. Die PCA berücksichtigt diesen impliziten Zusammenhang, wodurch allgemeingültige Tendenzen hervorgehoben werden und wenig informatives statistisches Rauschen reduziert wird.

- Quantitätsindikator **Motive des Data Sharing**: Ähnlich zum Indikator der Datenqualität sind die Antwortitems zur Frage nach den Motiven des Data Sharing nicht eindeutig zueinander in Bezug zu setzen. Unternehmen können beispielsweise mithilfe von geteilten Daten Produkte und Geschäftsmodelle weiterentwickeln oder den datengebenden Unternehmen im Gegenzug datenbasierte Dienstleistungen anbieten. Deshalb wird auch für diesen Indikator eine PCA angewendet, um die Motive des Datenteilens zu gewichten.
- Quantitätsindikator **Datenarten**: Bei diesem Indikator geben Unternehmen an, wie viele der bei ihnen digital gespeicherten Datenarten sie mit anderen Unternehmen teilen. Der Anteil der geteilten Datenarten an allen im Unternehmen digital gespeicherten Datenarten wird berechnet, um Unternehmen aus verschiedenen Branchen besser miteinander vergleichen zu können. Ein Unternehmen, das beispielsweise keine Forschungsdaten digital speichert, kann sie auch nicht teilen – und sollte dafür folglich nicht im Modell „bestraft“ werden. Aufgrund der jeweils unterschiedlichen Anzahl an digital gespeicherten Datenarten pro Unternehmen werden die Anteile der geteilten Datenarten gleichgewichtet gebildet.

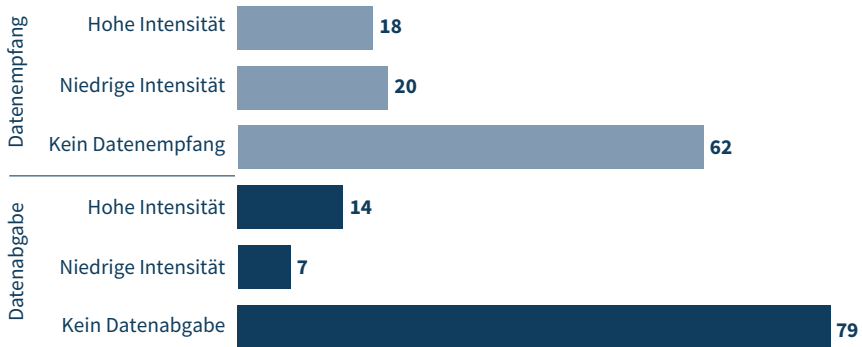
Intensität des Data Sharing in Deutschland

Gemäß dem Data-Sharing-Modell empfangen 20 Prozent der 1.051 befragten Unternehmen Daten mit niedriger Intensität (Stufe 1), 18 Prozent mit hoher Intensität (Stufe 2). Abbildung 1 zeigt, dass 62 Prozent keine Daten empfangen (Stufe 0). Noch weniger verbreitet ist die Datenabgabe. 79 Prozent der befragten Unternehmen geben keine Daten an andere Unternehmen ab. Wenn die Unternehmen Daten abgeben, dann eher mit hoher Intensität als mit niedriger: 14 Prozent der Unternehmen sind der Intensitätsstufe 2 zuzuordnen. 7 Prozent der Unternehmen geben Daten mit niedriger Intensität ab. Insgesamt ist der Datenempfang also deutlich verbreiteter unter den Unternehmen in Deutschland als die Datenabgabe, auch wenn im Verhältnis zur niedrigen Intensi-

Intensität des Data Sharing

Abbildung 1

Anteil der Unternehmen in Deutschland, die Daten mit hoher oder niedriger Intensität an andere Unternehmen abgeben oder von anderen Unternehmen empfangen, in Prozent aller Unternehmen¹⁾



1) Befragung von 1.051 Industrieunternehmen und industrienahen Dienstleistern im Herbst 2022.
Quelle: Institut der deutschen Wirtschaft

Abbildung 1: <http://dl.iwkoeln.de/index.php/s/iGZXAcAj4GN3xFD>

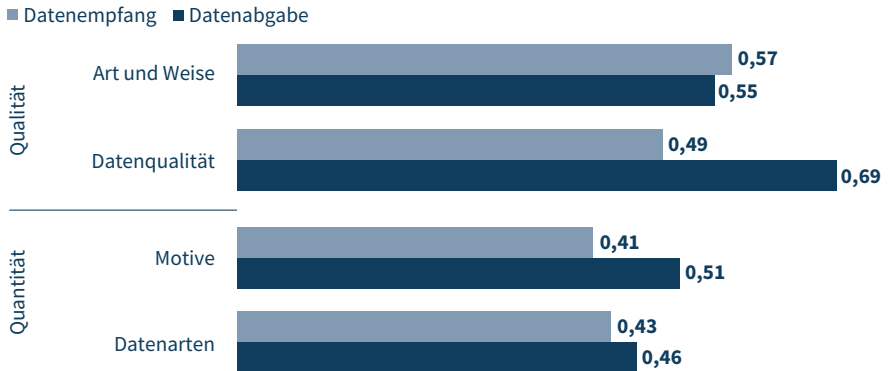
tät mehr Unternehmen Daten mit hoher Intensität abgeben als Unternehmen Daten mit hoher Intensität empfangen.

Damit die Wertschöpfungspotenziale des Data Sharing gehoben werden können, sollten mehr Unternehmen Daten abgeben. Dazu müssen individuelle Vorteile der Datenabgabe vorliegen, von den Unternehmen erkannt und genutzt werden. Es sollte lukrativ sein, Daten abzugeben, beispielsweise indem Unternehmen ein Entgelt für ihre bereitgestellten Daten erhalten. Die Ergebnisse des Data-Sharing-Modells deuten darauf hin, dass derzeit wenige Unternehmen Daten mit hoher Intensität abgeben. Wünschenswert wäre, dass viele Unternehmen Daten mit hoher Intensität abgeben, um alle möglichen Synergien zu heben. Potenziale des Data Sharing, etwa um Prozesse, Produkte und Geschäftsmodelle zu optimieren oder neu zu gestalten, könnten so besser eingelöst werden. Eine Analyse der einzelnen Indikatoren des Modells verdeutlicht im Folgenden, wo die Unternehmen vor allem Entwicklungspotenzial zeigen. Bei der Betrachtung der durchschnittlichen Indikatoren-Scores ergeben sich zum Teil deutliche Unterschiede zwischen Datenempfang und Datenabgabe (Abbildung 2).

Indikatoren-Scores

Abbildung 2

Arithmetisches Mittel der vier Indikatoren-Scores zur Qualität und Quantität des Data Sharing auf einer Skala von 0 (kein relevanter Aspekt erfüllt) bis 1 (alle relevanten Aspekte erfüllt)



Basis: nur Unternehmen, die Daten an andere Unternehmen abgeben oder Daten von anderen Unternehmen empfangen. Befragung von 1.051 Industrieunternehmen und industrienahen Dienstleistern im Herbst 2022. Quelle: Institut der deutschen Wirtschaft

Abbildung 2: <http://dl.iwkoeln.de/index.php/s/nt8j525PoRo8yxL>

Die **Art und Weise** des Data Sharing ist sowohl beim Datenempfang als auch bei der Datenabgabe mit durchschnittlichen Scores von 0,57 und 0,55 tendenziell als fortschrittlich einzuschätzen. Empfangen Unternehmen Daten von anderen Unternehmen, erfolgt dies im Durchschnitt zu 33 Prozent digital, automatisiert und in Echtzeit, zu 18 Prozent digital, automatisiert, aber nicht in Echtzeit sowie zu 34 Prozent digital mit händischen Zwischenschritten und zu 12 Prozent papierbasiert. Die Art und Weise der Datenabgabe ist leicht weniger fortschrittlich als der Datenempfang. Geben Unternehmen Daten ab, erfolgt dies zu 30 Prozent digital, automatisiert und in Echtzeit, zu 18 Prozent digital, automatisiert, aber nicht in Echtzeit und zu 39 Prozent digital mit händischen Zwischenschritten. 12 Prozent der Daten werden papierbasiert abgegeben.

Bei der **Qualität der geteilten Daten** zeigen sich erhebliche Unterschiede zwischen Datenempfang und Datenabgabe. Die Unternehmen bewerten die Qualität der empfangenen Daten mit einem durchschnittlichen Score von 0,49. Die Qualität der abgegebenen Daten hingegen wird mit 0,69 deutlich höher bewertet. Das entspricht dem höchsten Wert unter allen vier Indikatoren-Scores. Auf die abgegebenen Daten

treffen laut den befragten Unternehmen fast alle Qualitätskriterien deutlich häufiger zu als auf die empfangenen Daten. Besonders groß sind die Unterschiede hinsichtlich der Qualitätsprüfung, der Widerspruchsfreiheit und dem Vorhandensein eines Identifiers. Während 78 Prozent der Unternehmen angeben, ihre abgegebenen Daten seien qualitätsgeprüft, gehen davon nur 60 Prozent der Datenempfänger aus. 80 Prozent der Datengeber behaupten, ihre abgegebenen Daten seien in sich widerspruchsfrei – dies bestätigen allerdings nur 66 Prozent der Datenempfänger. Die Verknüpfung über einen Identifier ist laut 50 Prozent der Datengeber möglich, aber nur laut 41 Prozent der Datenempfänger. Die objektiven Qualitätskriterien bieten also offensichtlich Interpretationsspielraum. Es ist möglich, dass Datengeber die Qualität ihrer abgegebenen Daten subjektiv überschätzen. Eine andere Möglichkeit ist, dass Data Sharing mit einem Informationsverlust hinsichtlich der die Daten beschreibenden Metadaten einhergeht, weshalb die Unternehmen, die Daten empfangen, deren Qualität nicht so hoch einschätzen. Zudem könnten Unternehmen Daten nicht nur von anderen Unternehmen empfangen, sondern beispielsweise auch von öffentlichen Institutionen, die in der Unternehmensbefragung nicht als Datengeber auftauchen und deren Qualität der abgegebenen Daten unterschiedlich sein könnte.

Ebenfalls deutlich sind die Unterschiede zwischen Datenempfang und Datenabgabe bei den **Motiven des Data Sharing**. Für die Datenabgabe erzielen die Unternehmen im Durchschnitt einen höheren Score (0,51) als für den Datenempfang (0,41). Insgesamt zeigt sich bei den Motiven noch Entwicklungspotenzial. Die Datenabgabe erfolgt vor allem wegen vertraglicher Verpflichtungen: 81 Prozent der Unternehmen nennen dies als Grund. Auch beim Datenempfang sind vertragliche Verpflichtungen der am häufigsten genannte Grund (72 Prozent). Auf Rang zwei liegt sowohl beim Datenempfang als auch bei der Datenabgabe das Motiv, dass die geteilten Daten im Produktionsprozess benötigt werden – entweder im eigenen (Empfang) oder im anderen Unternehmen (Abgabe). Die Datenabgabe erfolgt außerdem häufig wegen rechtlicher Vorgaben (Rang 3 mit 57 Prozent). Unternehmen empfangen oft Daten, weil sie im Gegenzug Dienstleistungen anbieten (Rang 3 mit 63 Prozent). Aus diesem Grund geben allerdings nur 51 Prozent der Unternehmen Daten ab. Deutlich häufiger als die Datenabgabe geschieht der Datenempfang zur Verbesserung der eigenen Prozesse

(58 versus 43 Prozent). Lediglich für 26 und 27 Prozent der datenteilenden Unternehmen ist der Datenempfang oder die Datenabgabe Kern des Geschäftsmodells.

Die Unterschiede zwischen **Datenarten**, die empfangen oder abgegeben werden, fallen geringer aus. Für die Datenabgabe liegt der Score der geteilten Datenarten mit durchschnittlich 0,46 etwas über dem für den Datenempfang (0,43). Vor allem Produktdaten werden von Unternehmen empfangen: Unternehmen, die Daten empfangen und bei denen Produktdaten digital gespeichert werden, erhalten im Durchschnitt 52 Prozent ihrer genutzten Produktdatenmenge von anderen Unternehmen. Andersherum geben Datengeber, die Produktdaten digital speichern, im Durchschnitt 47 Prozent der anfallenden Produktdatenmenge an andere Unternehmen ab. Der Anteil der abgegebenen Datenmenge ist für Finanzdaten mit 58 Prozent am höchsten. Auch relativ viele Personaldaten werden abgegeben (55 Prozent). Es liegt nahe, dass diese Daten etwa zu buchhalterischen und steuerlichen Zwecken an andere Institutionen oder Dienstleister wie Wirtschaftsprüfer weitergegeben werden. Am wenigsten abgegeben und empfangen werden Forschungs- und Entwicklungsdaten (33 Prozent). Diese werden offenbar als besonders schützenswert empfunden und verbleiben meist im eigenen Unternehmen, wo die vermeintlich bestmögliche Kontrolle über die Daten gewährleistet werden kann. 40 Prozent der Nutzungsdaten von Kunden und Lieferanten werden empfangen. Potenzial zeigt sich auch beim Teilen von Produktions- und Prozessdaten: Ebenfalls etwa 40 Prozent dieser Daten werden empfangen oder abgegeben. Es ist zu erwarten, dass sich Produktionsprozesse verbessern ließen, wenn mehr Daten darüber auch mit anderen Unternehmen geteilt würden. Beispielsweise ist die vorausschauende Wartung nur möglich, wenn Daten über die zu wartende Maschine – und damit über die Produktion – mit dem Maschinenhersteller geteilt werden.

Zusammenfassend kann zunächst festgehalten werden, dass die befragten Unternehmen die Datenabgabe eher mit höherer Intensität als den Datenempfang betreiben, vor allem in Bezug auf die Datenqualität. Der Anteil an Datengebern ist dennoch deutlich geringer als der Anteil an Datenempfängern. Dementsprechend zeigt sich beim Datenempfang Entwicklungspotenzial in Bezug auf dessen Intensität. Die Anzahl der Motive und empfangenen Datenarten ist – gerade auch relativ zur Datenabgabe – ausbaubar.

Die Datenqualität der empfangenen Daten ist vom Datenempfänger hingegen direkt kaum zu beeinflussen. Hier scheint es zu Informationsverlusten zwischen Datenempfängern und Datengebern zu kommen, denn Letztere bewerten ihre abgegebenen Daten deutlich besser als die Empfänger. Möglicherweise muss das Verständnis von hoher Datenqualität zwischen Datenempfänger und Datengeber besser abgestimmt werden. Ebenfalls Entwicklungspotenzial gibt es bei der Datenabgabe: Es wäre wünschenswert, wenn sich die Anteile der Datengeber erhöhten. Zudem zeigt sich gerade bei den abgegebenen Datenarten und den Motiven zur Datenabgabe noch Potenzial.

Zusammenhang zwischen Datenabgabe und Datenempfang

Im Folgenden werden die bislang separat betrachteten Intensität-Scores der Unternehmen für die Datenabgabe und den Datenempfang zusammengeführt, um einen möglichen Zusammenhang zu untersuchen. Konkret wird die Schnittmenge der jeweiligen Intensitätsniveaus analysiert (Tabelle 1).

Schnittmengen aus Datenempfängern und Datengebern

Tabelle 1

Anteil der Unternehmen in Deutschland, die Daten mit hoher oder niedriger Intensität an andere Unternehmen abgeben oder von anderen Unternehmen empfangen¹⁾, in Prozent

		Datenabgabe			Zeilensumme
		Keine Datenabgabe	Niedrige Intensität	Hohe Intensität	
Datenempfang	Kein Datenempfang	58	2	2	62
	Niedrige Intensität	12	4	5	20
	Hohe Intensität	10	1	8	18
Spaltensumme		79	7	14	100

1) Befragung von 1.051 Industrieunternehmen und industrienahen Dienstleistern im Herbst 2022. Teilweise Rundungsdifferenzen.

Quelle: Institut der deutschen Wirtschaft

Tabelle 1: <http://dl.iwkoeln.de/index.php/s/g9rYMYkppffz4nB>

58 Prozent der befragten Unternehmen geben weder Daten ab noch empfangen sie welche. Entsprechend sind 42 Prozent der Unternehmen als Datengeber und/oder Datenempfänger aktiv. Im Vorjahr waren es lediglich 27 Prozent (Büchel/Engels, 2022). Dass der Anteil der Unternehmen, die als Datengeber und/oder Datenempfänger aktiv sind, nur etwas höher ist als der Anteil der Datenempfänger (38 Prozent; s. Abbildung 1), zeigt, dass die meisten der Datengeber ebenfalls Daten empfangen. Nur 4 Prozent der Unternehmen, die Daten abgeben, empfangen keine Daten. Gibt ein Unternehmen also Daten ab, ist es sehr wahrscheinlich, dass es auch Daten empfängt.

Etwa 17 Prozent der Unternehmen sind sowohl Datengeber als auch Datenempfänger. Der größte Anteil unter ihnen (8 Prozentpunkte) entfällt auf Unternehmen, die mit hoher Intensität Daten empfangen und abgeben. Dies könnte ein Hinweis auf gegenseitige Effekte sein. Beispielsweise könnten Unternehmen hochqualitative Daten empfangen und daraufhin ähnlich hohe Qualitätsstandards im Unternehmen adaptieren. Dies kann dazu führen, dass sie eigene Daten ebenfalls in hoher Qualität abgeben. Gleichzeitig können Unternehmen, die Daten automatisiert und in Echtzeit abgeben, erwarten, dass dies auch auf die von ihnen empfangenen Daten zutrifft. Empfangen Unternehmen viele verschiedene Datenarten, können sie es als vorteilhaft empfinden, diese Vielzahl an Datenarten selbst an andere Unternehmen abzugeben. Um den Zusammenhang zwischen der Intensität der Datenabgabe und des Datenempfangs weiter zu untersuchen, werden die Empfänger- und Geber-Scores der 234 Unternehmen, die Daten empfangen und Daten abgeben, gegenübergestellt (Abbildung 3).

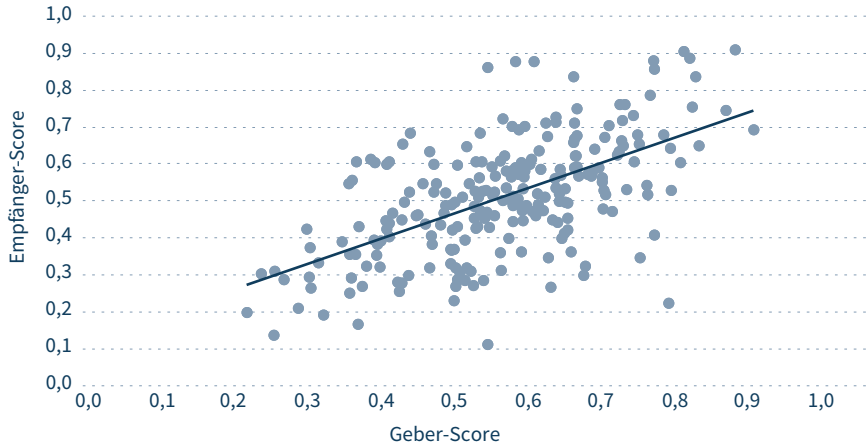
Der Vergleich der erzielten Empfänger- und Geber-Scores zeigt einen positiven statistischen Zusammenhang zwischen der Intensität der Datenabgabe und des Datenempfangs: Eine Korrelationsanalyse ergibt einen Koeffizienten von 0,6. Unternehmen mit hohem Geber-Score haben tendenziell auch einen hohen Empfänger-Score. Interessant ist, dass keines der Unternehmen, das Geber und Empfänger ist, einen Geber-Score unter 0,2 hat. Die Empfänger-Scores unterschreiten nur in einzelnen Fällen den Wert von 0,2, nicht aber von 0,1.

Die Analyse kann empirisch allerdings nicht nachweisen, ob die Datenabgabe dabei den Datenempfang beeinflusst oder umgekehrt und wie stark ein möglicher Effekt ist.

Zusammenhang zwischen Empfänger- und Geber-Score

Abbildung 3

Empfänger- und Geber-Score der Unternehmen in Deutschland, die sowohl Daten an andere Unternehmen abgeben als auch Daten von anderen Unternehmen empfangen¹⁾, auf einer Skala von 0 bis 1, Werte von 0 bis 0,5 (niedrige Intensität) und 0,5 bis 1 (hohe Intensität)



1) Basis: 234 Unternehmen von insgesamt 1.051 befragten Industrieunternehmen und industrienahen Dienstleistern im Herbst 2022.

Quelle: Institut der deutschen Wirtschaft

Abbildung 3: <http://dl.iwkoeln.de/index.php/s/rqRDHayxcJwZLrA>

Allerdings verdeutlichen die Argumente aus dem vorherigen Abschnitt, dass in beide Wirkungsrichtungen positive Effekte möglich sind. Liegen sie in beiden Richtungen vor, könnten diese sich gegenseitig verstärken und Synergien bilden.

Chi-Quadrat-Tests zeigen auch, dass Unternehmen, die Daten mit hoher Intensität abgeben, mit statistischer Signifikanz eher auch Daten mit hoher Intensität empfangen (Tabelle 2). Dies unterstützt die vorstehenden Analysen zum Zusammenhang zwischen Datenempfang und Datenabgabe. Darüber hinaus wurde mittels Chi-Quadrat-Tests geprüft, ob Unternehmen, die bestimmte Charakteristika erfüllen, eher Daten mit hoher Intensität empfangen oder abgeben als Unternehmen, die diese Charakteristika nicht erfüllen. Die getesteten Charakteristika sind, ob ein Unternehmen die Voraussetzungen dafür erfüllt, Daten effizient zu bewirtschaften („data economy ready“), ob

Intensität des Datenteilens nach Unternehmenscharakteristika

Tabelle 2

Ergebnisse auf Basis von Chi-Quadrat-Tests

	Datenempfang mit hoher Intensität	Datenabgabe mit hoher Intensität
„data economy ready“	+*** (0,37)	+*** (0,33)
Viele Beschäftigte	= (0,02)	= (0,003)
Hoher Umsatz	= (0,14)	= (0,06)
Datenabgabe mit hoher Intensität	+*** (0,36)	

*** / ** / *: signifikant auf dem 1- / 5- / 10-Prozent-Niveau; Chi-Quadrat-Tests bezüglich Differenz der Mittelwerte verschiedener Subsamples; + bedeutet: Unternehmen, für die die jeweilige Charakteristik zutrifft, teilen Daten signifikant häufiger mit hoher Intensität im Vergleich zu Unternehmen, für die die jeweilige Charakteristik nicht zutrifft. – bedeutet entsprechend, dass ein negativer Zusammenhang vorliegt; = bedeutet, dass es keine statistisch signifikanten Unterschiede in der Häufigkeit gibt, dass ein Unternehmen Datengeber oder Datenempfänger mit hoher Intensität ist, je nachdem, ob die Unternehmenseigenschaft zutrifft oder nicht; Cramér's V-Wert in Klammern: Maßzahl für die Stärke des Zusammenhangs auf einer Skala von 0 bis 1.

Quelle: Institut der deutschen Wirtschaft

Tabelle 2: <http://dl.iwkoeln.de/index.php/s/tGwbFTeSDPip6BQ>

ein Unternehmen klein oder mittelgroß bis groß ist und ob es einen niedrigeren oder mittleren bis hohen Umsatz hat.

Unternehmen, die „data economy ready“ sind, geben eher Daten mit hoher Intensität (versus niedrige Intensität) ab als Unternehmen, die dies nicht sind. Das trifft auch auf den Datenempfang mit hoher Intensität zu. Dieses Ergebnis passt zu dem Antwortverhalten der Unternehmen hinsichtlich der Frage, was die Wahrscheinlichkeit erhöhen würde, dass sie mehr Daten teilten: 21 Prozent gaben an, eine bessere Datenverwaltung in ihrem Unternehmen würde die Wahrscheinlichkeit erhöhen, Daten zu teilen. Eben diese bessere Datenverwaltung wird unter anderem mit der Data Economy Readiness beschrieben.

Die Höhe des Umsatzes und die Anzahl der Beschäftigten führten nicht zu signifikanten Unterschieden in der Wahrscheinlichkeit, dass Unternehmen den Datenempfang oder die Datenabgabe mit hoher statt niedriger Intensität betreiben. Dies deutet darauf hin, dass intensives Data Sharing nicht notwendigerweise mit einem hohen personellen

oder finanziellen Ressourceneinsatz verbunden ist. Insofern ist intensives Data Sharing für ganz unterschiedliche Unternehmenstypen eine Option. Das Vorhandensein von entsprechenden Kompetenzen im Unternehmen fördert dabei das Data Sharing: 17 Prozent der befragten Unternehmen geben an, dass mehr Fachpersonal im Bereich IT die Wahrscheinlichkeit erhöhen würde, dass sie Daten teilen. Positiv für die Entwicklung des Data Sharing könnte daher sein, dass in den vergangenen Jahren zunehmend Unternehmen in Deutschland nach neuen Beschäftigten mit Datenkompetenzen suchen (Büchel et al., 2023a; 2023b).

Anreize für intensives Data Sharing

Vielen Unternehmen in Deutschland ist der Nutzen der Datenabgabe (58 Prozent) und des Datenempfangs (47 Prozent) noch unklar. 61 Prozent der Unternehmen sind sich sicher, dass die Risiken der Datenabgabe größer als ihr Nutzen sind. 74 Prozent der Unternehmen meinen, alle relevanten Daten würden in ihrem Unternehmen selbst generiert, weshalb sie keine oder nicht noch mehr Daten empfangen müssten.

Gleichzeitig wird aus den Umfrageergebnissen ein Matching-Problem deutlich. 43 Prozent der Unternehmen empfangen nicht mehr Daten, weil für ihr Unternehmen interessante Daten nicht vorhanden seien. 40 Prozent meinen, solche Daten würden nicht von anderen Unternehmen abgegeben. Andersherum geben Unternehmen nicht mehr Daten ab, weil sie keinen Nutzen für andere Unternehmen böten (36 Prozent) oder weil es keine Interessenten für diese Daten gäbe (30 Prozent). Möglicherweise liegt an dieser Stelle ein Informationsdefizit vor: Potenzielle Datengeber wissen nicht, welche Daten für potenzielle Datenempfänger interessant sind. Potenzielle Datenempfänger wissen nicht, über welche Daten potenzielle Datengeber verfügen. Könnten diese Informationsasymmetrien gelöst werden, würde mehr wertschöpfungssteigerndes Data Sharing ermöglicht. In der Tat geben 37 Prozent der befragten Unternehmen an, sie würden eher Daten mit anderen Unternehmen teilen, wenn sie passende Kooperationspartner dafür hätten.

Für ein Viertel der Unternehmen steigt die Wahrscheinlichkeit, dass sie Daten teilen, wenn es eine Instanz gäbe, die das Datenteilen regelt und prüft, ob die rechtlichen Vorschriften und vertraglichen Abmachungen eingehalten werden. Das passt zu dem

Befund, dass 64 Prozent (51 Prozent) der Unternehmen nicht mehr Daten abgeben (empfangen), weil sie rechtliche Bedenken hinsichtlich des Datenschutzes, Urheberrechts oder Kartellrechts haben. Gut ein Viertel der Unternehmen wünscht sich eine bessere Verfügbarkeit von Plattformen und Technologien zum Datenteilen. Data-Sharing-Plattformen können eine Transaktionsinfrastruktur bieten oder Datenservices wie Datenanalyse, Datenanreicherung oder Datenbereinigung übernehmen (Meisel/Spiekermann, 2019; Lindner et al., 2020). Plattformen könnten auch die technischen Bedenken hinsichtlich der Datensicherheit lindern und für mehr Datensouveränität sorgen. Jene Bedenken sind für 46 Prozent (38 Prozent) der Unternehmen ein Grund, nicht mehr Daten abzugeben (zu empfangen).

Aus der Auswertung der Befragung zur Intensität des Data Sharing entsteht der Eindruck, dass Unternehmen in Deutschland noch deutliches Entwicklungspotenzial beim Data Sharing haben und auf unternehmensexterne Unterstützung angewiesen sind. Einen Kooperationspartner zu finden, Rechtssicherheit und Datensouveränität zu gewährleisten, sind Herausforderungen beim Data Sharing, die durch Unternehmen oft nicht allein zu meistern sind. Es ist denkbar, dass Data-Sharing-Pioniere in einzelnen Branchen eine Zugwirkung auf andere Unternehmen der Branche haben. Aber auch durch die Politik unterstützte Plattformen und Informationsangebote könnten dazu beitragen, die Verbreitung und Intensität von Data Sharing in Deutschland zu steigern. Ein Umfrageexperiment zeigt jedenfalls, dass Unternehmen, denen konkrete Potenziale der Datennutzung aufgezeigt werden, signifikant häufiger planen, Daten in Zukunft vielfältiger zu nutzen als Unternehmen ohne entsprechende Information zu den Potenzialen (Erdsiek/Rost, 2022). Eine zentrale Herausforderung für Unternehmen ist es auch, sich des eigenen Datenbestands und der eigenen Datenflüsse bewusst zu werden sowie eine Data Governance zu etablieren. Dies ist die Voraussetzung dafür, um definieren zu können, welche Daten von anderen Unternehmen zum eigenen Wertschöpfungsprozess beitragen könnten – und welche Daten das Unternehmen selbst anderen Unternehmen anbieten kann.

Wir danken Lennart Bolwin und Christian Kestermann für die Unterstützung. Diese Studie ist Teil des vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Projekts „Incentives and Economics of Data Sharing – IEDS“.

Literatur

Büchel, Jan / Engels, Barbara, 2022, Branchentrends beim Data Sharing. Status Quo und Use Cases in Deutschland, IW-Report, Nr. 53, Köln

Büchel, Jan / Engler, Jan Felix / Mertens, Armin, 2023a, The Demand for Data Skills in German Companies. Evidence from Online Job Advertisements, in: EconPol Forum, 24. Jg., Nr. 2, CESifo, München

Büchel, Jan / Engler, Jan Felix / Mertens, Armin, 2023b, Gesuchte Datenkompetenzen in Deutschland – Ergebnisse einer Big-Data-Analyse, in: IW-Trends, 50. Jg., Nr. 2, S. 3–17

Engels, Barbara / Schäfer, Christin, 2020, Data Governance in deutschen Unternehmen, Gutachten im Rahmen des BMWi-Verbundprojekts DEMAND – DATA ECONOMICS AND MANAGEMENT OF DATA DRIVEN BUSINESS, Köln

Erdsiek, Daniel / Rost, Vincent, 2022, Datenbewirtschaftung in deutschen Unternehmen: Umfrageergebnisse zu Status-quo und mittelfristigem Ausblick, ZEW-Kurzexpertise, Nr. 9, Mannheim

IDC / Lisbon Council, 2023, European Data Market Study 2021–2023. First Report on Facts and Figures, Studie im Auftrag der Europäischen Kommission, Luxemburg

Lindner, Maximilian / Straub, Sebastian / Kühne, Bettina, 2020, How to Share Data? Data-Sharing-Plattformen für Unternehmen, https://www.iit-berlin.de/wp-content/uploads/2021/04/SDW_Studie_DataSharing_210407.pdf [3.4.2023]

Meisel, Lukas / Spiekermann, Markus, 2019, Datenmarktplätze – Plattformen für Datenaustausch und Datenmonetarisierung in der Data Economy, ISST-Bericht, hrsg. v. Boris Otto und Jakob Rehof, Dortmund

Niebel, Thomas / Rasel, Fabienne / Viète, Steffen, 2019, BIG data – BIG gains? Understanding the link between big data analytics and innovation, in: Economics of Innovation and New Technology, 28. Jg., Nr. 3, S. 296–316

Rusche, Christian et al., 2022, KI-Monitor 2022. Künstliche Intelligenz in Deutschland, Gutachten im Auftrag des Bundesverbands Digitale Wirtschaft (BVDW) e.V., Köln

Data Sharing in Germany: Extent, Intensity and Future Potential

Digital data is a key element in processes, products and business models. Thus, data sharing is a matter of great importance. The present analysis examines how intensively companies in Germany share data. A survey was conducted among 1,051 industrial companies and industry-related service providers. Their responses regarding the quality and quantity of their data sharing were incorporated into a model that assigns intensity scores to companies for their data provision and data reception. The results show that 58 per cent of companies in Germany are not involved in data sharing at all – neither providing data nor receiving it. 38 per cent receive data and only 21 per cent provide it to others. When companies provide data, they are more likely to do so at a high intensity (14 per cent) than at a low intensity (7 per cent). The opposite is true for enterprises that receive data, only 18 per cent doing so at a high, but 20 per cent at a low, intensity. These results suggest a clear potential to make data provision more widespread and data reception more intensive. There is, in general, a positive statistical correlation between the intensity of data provision and data reception. In particular, most data providers also receive data from other companies. Finding a cooperation partner, ensuring legal certainty and data sovereignty are challenges in data sharing that many companies are not able to master on their own.