Open Data. The Benefits

Das volkswirtschaftliche Potential für Deutschland

Marcus M. Dapp, Dian Balta,
Walter Palmetshofer, Helmut Krcmar

Pencho Kuzev (Hrsg.)

Konrad
Adenauer
Stiftung

Open Data. The Benefits

Das volkswirtschaftliche Potential für Deutschland

Marcus M. Dapp, Dian Balta,
Walter Palmetshofer, Helmut Krcmar

Pencho Kuzev (Hrsg.)

Eine Studie im Auftrag der Konrad-Adenauer-Stiftung e.V.

Herausgegeben im Auftrag der Konrad-Adenauer-Stiftung e.V.
von Dr. Pencho Kuzev

Diese Studie wurde vom Institute for Public Information Management (ipima) der fortiss GmbH, An-Institut der TU München durchgeführt.

Autoren:
Dr. Marcus M. Dapp
Dian Balta
Mag. Walter Palmetshofer
Prof. Dr. Helmut Krcmar

Lektorat: Jenny Kahlert, Konrad-Adenauer-Stiftung e.V.
Gestaltung: SWITSCH KommunikationsDesign, Köln
Druck: Bonifatius GmbH, Paderborn

Dieses Werk ist lizenziert unter den Bedingungen von „Creative Commons Namensnennung-Weitergabe unter gleichen Bedingungen 3.0 Deutschland", CC BY-SA 3.0 DE (abrufbar unter: http://creativecommons.org/licenses/by-sa/3.0/de/)

© 2016, Konrad-Adenauer-Stiftung e.V., Sankt Augustin/Berlin

Printed in Germany.
Gedruckt mit finanzieller Unterstützung der Bundesrepublik Deutschland.
ISBN 978-3-95721-202-3

ClimatePartner klimaneutral
Druck | ID: 53323-1603-1093

Inhalt

7 | VORWORT

9 | ZUSAMMENFASSUNG

13 | SUMMARY

16 | 1. EINLEITUNG

19 | 2. KURZEINFÜHRUNG

2.1 Die kurze und dynamische Geschichte von Open Data 19

2.2 Begriffsdefinition und Abgrenzung 21

25 | 3. WERTSCHÖPFUNG DURCH OPEN DATA

3.1 Wie Open Data Wirkung entfaltet 25

3.2 Entstehung und Publikation von Open Data 27

3.3 Akteure und Aktivitäten in der Nachnutzung 29

3.4 Ökonomische Betrachtung 31

37 | 4. METHODIK DER POTENTIALSCHÄTZUNG

4.1 Zwei Ansätze: nationale Sekundär- und kommunale Primärdaten . . 37

4.2 Potentialschätzung durch internationale Sekundärdaten 38

4.3 Hochrechnung kommunaler Open-Data-Kosten 45

55 | 5. PROJEKTION DES POTENTIALS FÜR DEUTSCHLAND

5.1 Die Quantifizierung des Potentials 55

5.2 Konservativ, ambitioniert oder optimistisch? 55

5.3 Kommunale Open-Data-Kosten als Berechnungsbasis 60

62 | KRITISCHE REFLEKTION UND FAZIT

64 | ANMERKUNGEN

66 | VERZEICHNIS

Vorwort

Digitales Gold, Rohstoff des 21. Jahrhunderts, neues Öl – mit diesen Begriffen wird das wirtschaftliche Potential von Daten gegenwärtig verglichen. Sollten die Datenbestände im Besitz des Staates nicht ebenso wertvoll sein und als offene Verwaltungsdaten zur Verfügung stehen, um einen volkswirtschaftlichen Mehrwert zu erzeugen? Um das volkswirtschaftliche Potential offener Verwaltungsdaten für Deutschland zu ermitteln, hat die Konrad-Adenauer-Stiftung das Institute for Public Information Management (ipima) der fortiss GmbH mit der vorliegenden Studie beauftragt.

Die Studie nutzt dafür zwei Ansätze, die eine methodisch nachvollziehbare Schätzung gestatten sollen. Der top-down-Ansatz erlaubt eine langfristige Potentialabschätzung. Die bottom-up-Herangehensweise liefert Erkenntnisse, wo Deutschland stünde, wenn die Durchdringung in den Kommunen jener der heute in Deutschland führenden „Open-Data-Kommunen" entspräche. Die stark auseinanderliegenden Potentialwerte der zwei Ansätze weisen auf die Fragen hin, ob die aktuelle Umsetzung der G8-Open-Data-Charta zufriedenstellend ist und die erhofften Open-Data-Wertschöpfungspotentiale realisiert werden oder bessere Rahmenbedingungen erforderlich sind.

Die Studie belegt, dass offene Verwaltungsdaten einen beachtlichen volkswirtschaftlichen Mehrwert erzeugen können. Sie liefert damit ein starkes Argument, die systematische Bereitstellung offener Daten zu forcieren. Mit den Ergebnissen der Studie möchten wir einen Beitrag dazu leisten, Open Data als gesellschaftliches Thema besser zu bewerten, auf der politischen Agenda adäquat zu positionieren und breiter in die Verwaltungsprozesse hineinzubringen.

Wir wünschen Ihnen eine anregende Lektüre.

Dr. Pencho Kuzev

Zusammenfassung

Die Digitalisierung geht mit einem rasanten Anwachsen der Datenbestände einher und eröffnet in allen gesellschaftlichen Bereichen völlig neue Möglichkeiten. Daten stellen dadurch eine neue, wichtige Ressource des 21. Jahrhunderts dar und bergen ein in Deutschland derzeit noch wenig genutztes Potential für die öffentliche Hand, die Wirtschaft und die Zivilgesellschaft. Neben großen Kundendatenbeständen („big data") und Sensordaten („Internet der Dinge") ist ein weiterer Ansatz zur Erschließung dieses Potentials die Öffnung von Daten.

Offene Daten („Open Data") sind Daten, die von jedermann ohne Einschränkungen verwendet werden können. Nach Francoli (2011) unterstützen offene Daten wichtige Ziele der Verwaltung: Neben der Stärkung der Demokratie durch Bürgerbeteiligung und anderem sozialen Nutzen, bergen offene Daten auch wirtschaftlichen Nutzen. Gurin (2014) führt beispielhaft einige Arten auf, wie wirtschaftlicher Nutzen entstehen kann: Firmengründer bauen neue, zuvor nicht mögliche, Geschäftsmodelle auf; Verwaltungen bieten zentrale Datenbestände an, die wirtschaftlich verwertbar sind; Investoren nutzen Daten zur Risikoanalyse und -minimierung; Verwaltungen und Unternehmen werden transparenter und erhöhen so das Vertrauen von Bürgern und Kunden; wissenschaftliche Arbeit kann beschleunigt werden; aufbereitete Daten verbessern Entscheidungsfindungsprozesse bei Kunden, in Firmen und Verwaltungen.

Deutschland befindet sich im internationalen Vergleich im Mittelfeld, was den Aufbau und die Entwicklung eines Open-Data-Ecosystems angeht. Zum Beispiel ist Deutschland im Open-Data-Index zwischen 2013 und 2015 vom 11. auf den 27. Platz abgerutscht. Das liegt nicht daran, dass die Open-Data-Aktivitäten in Deutschland weniger wurden, sondern dass sie in den letzten zwei Jahren in 16 anderen Ländern deutlich gesteigert wurden. Deutschland nimmt vor diesem Hintergrund das Risiko in Kauf, den Anschluss an die Entwicklungen im Bereich Open Data zu verlieren. Als Baustein und Beitrag zu dieser Diskussion beantwortet diese Studie die Frage: Was ist das volkswirtschaftliche Potential von Open Data in Deutschland?

Die vorgenommene Schätzung erfolgt unter Berücksichtigung aller bis dato bekannten Potential-Studien zu Open Data bzw. Public Sector Information (PSI).

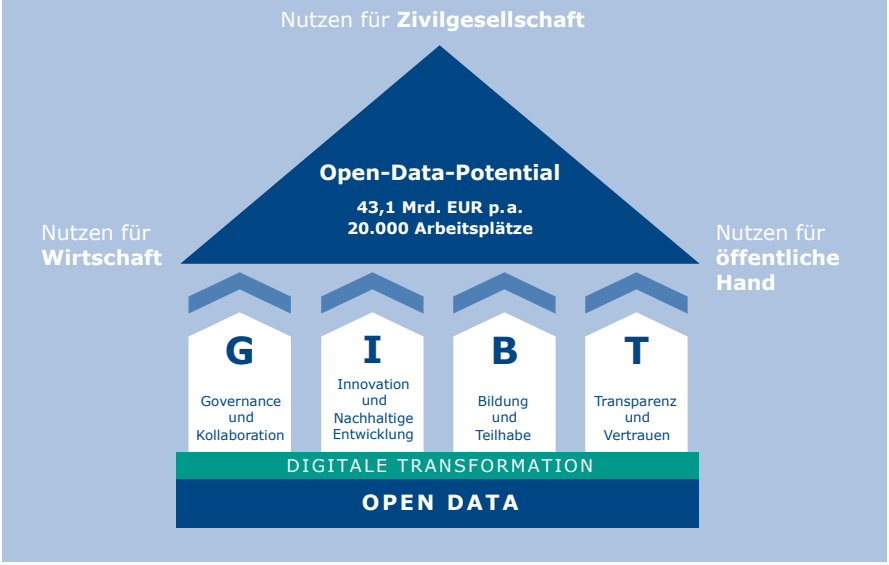
Das Ergebnis sind drei mögliche Zielbereiche für Deutschlands Open Data Ecosystem, die – in Abhängigkeit von der Gestaltung der politischen Rahmenbedingungen – in den kommenden zehn Jahren realisierbar erscheinen.

- **Im konservativen Fall** ist ein Potential von 12,1 Mrd. EUR p.a. realistisch. Dies bedeutet, dass Deutschland eine reaktive Strategie verfolgt, bei der bereits begonnene Aktivitäten im Bereich Open Government Data ausgebaut werden, um andernorts realisierte Potentiale auch in Deutschland zu heben.
- **Im ambitionierten Fall** ist ein Potential von 43,1 Mrd. EUR p.a. realistisch. Dies bedeutet, dass Deutschland eine proaktive Strategie („open by default") verfolgt. Ziel ist es, den Nutzenumfang durch die zusätzliche Öffnung von Business und Personal Data massiv zu erhöhen und so international aufzuschließen. Durch Open Data können in diesem Szenario bis zu 20.000 neue Arbeitsplätze[1] entstehen.
- **Der optimistische Bereich** kann bis in die Höhe von 131,1 Mrd. EUR p.a. reichen. Dies bedeutet, Open Data als Kernkomponente einer nationalen Strategie für das 21. Jahrhundert zu positionieren und sein Potential für große gesellschaftliche Herausforderungen zu identifizieren und zu mobilisieren. Ein derartiger Plan ginge über die digitale Transformation hinaus: Es ginge darum, sich als Nation in Bezug auf den Aufbau eines nationalen offenen Ecosystems selbst herauszufordern und dadurch zur weltweit führende Open-Data-Nation zu werden.

Praktisch bedeutet diese Bandbreite, dass die Art und Weise der Potentialausschöpfung von entscheidender Bedeutung ist. Unterschiede in Breite und Tiefe der Ausgestaltung administrativer, wirtschaftlicher, gesellschaftlicher und vor allem politischer Rahmenbedingungen können die Position Deutschlands innerhalb dieser Bandbreite erheblich beeinflussen.

Auch wenn die Frage nach „Zahlen" für ein Potential verständlich ist, so ist diese Perspektive für sich allein genommen zu verengt, denn nicht alles wird messbar sein. Um nicht nur das ökonomische, sondern das disruptive Potential von Open Data ganzheitlich bewerten zu können, ist eine breitere Sicht – im Sinne eines nicht-quantifizierbaren Nutzens für Zivilgesellschaft, Wirtschaft und öffentliche Hand – notwendig.

ABB. 1 **Open Data als Treiber des gesellschaftlichen Wandels**



Ausgangspunkt dieser Überlegungen ist die Frage, was für die Aktivierung des in Open Data schlummernden Potentials sorgen wird. Die Digitalisierung der Gesellschaft schreitet voran und ist dabei, in ihrer Wirkung einen Wendepunkt zu erreichen, an dem der Effekt von inkrementell auf transformativ springt. Dieser Punkt der „digital transformierten Gesellschaft" wird erreicht, wenn alle Prozesse in allen Organisationen aller Gesellschaftsbereiche vollständig digitalisiert und hochgradig vernetzt sind (Abb. 1).

Das Potential von Open Data wird durch die digitale Transformation aktiviert und treibt den Wandel der öffentlichen Hand, der Wirtschaft und der Zivilgesellschaft an. Diese Potentialentladung bzw. Nutzenentfaltung geschieht über vier Potentialträger:

- **Potentialträger: Governance und Kollaboration.** Erst durch das Öffnen von Daten wird es möglich, Informationsflüsse so umzulenken, dass neuartige Prozesse entstehen können, denen heute die Silo-Ressort-Haltung entgegensteht. Erst durch das Öffnen entscheidungsrelevanter Daten aus Institutionen und Organisationen auf breiter Basis wird die Partizipation und Kollaboration neuer, großer Anspruchsgruppen praktisch ermöglicht.

- **Potentialträger: Innovation und Nachhaltige Entwicklung.** Erst durch das Öffnen von Struktur- und Produktionsdaten aus Institutionen und Organisationen werden neue gesellschaftliche Gruppen – außerhalb der traditionellen Wertschöpfungsketten – nicht nur Beteiligte an der Generierung und Umsetzung neuer Leistungen und Dienste, sondern zu neuen, unabhängigen Anbietern. Erst durch das Öffnen von Daten in allen Fachbereichen ist es möglich, viele Anspruchsgruppen zu befähigen, zu Lösungen komplexer gesellschaftlicher Herausforderungen beizutragen.
- **Potentialträger: Bildung und Teilhabe.** Erst durch das Öffnen von Bildungsinformationen werden der Zugang zu Bildung in allen Bereichen massiv verbreitert und die Landschaft der Bildungsinstitutionen größer und vielfältiger. Erst durch neue digitale Schlüsselkompetenzen wie Grundfähigkeiten im Umgang mit Daten (data literacy) wird die politische und gesellschaftliche Teilhabe im 21. Jahrhundert effektiv möglich.
- **Potentialträger: Transparenz und Vertrauen.** Erst durch das Öffnen von Strukturdaten wird das Agieren politischer Institutionen und privater Organisationen nachvollziehbar und kontrollierbar. Erst diese Öffnung kann eine Transparenz erzeugen, die mittelfristig das Vertrauen in große Institutionen und Firmen stärkt.

Die konsequente Ausschöpfung des Potentials – die Entwicklung entlang der Potentialträger „aufwärts" in Abb. 1 – wird über die Zeit zur Entstehung eines umfassenden offenen Ecosystems in Deutschland führen. Darunter versteht man jenes Gefüge, das durch Erzeugung, Aufbereitung und Nutzung offener Daten in verbundenen Akteursnetzen wirtschaftlichen und gesellschaftlichen Mehrwert generiert. Um diesen Zustand zu erreichen, stellt jeder Potentialträger ein Handlungsfeld dar, in dem politische Entscheidungsträger aktiv werden müssen.

Summary

The digital transformation has come with the rapid growth of data and offers completely new possibilities in all areas of society. Data can be seen as an important resource in the 21st century and presents huge potential for the public and private sector, as well as civil society. In addition to "big data" and sensor data (the 'Internet of things'), the idea of "open data" offers another approach for tapping into this potential.

Open data are data that can be accessed and used by anyone without any constraints. According to Francoli (2011), open data supports important objectives within the public administration sector. In addition to strengthening democracy through online participation and other social benefits, open data can also bring economic benefits. Gurin (2014) has noted several approaches that could bring about economic benefits: company founders can create new business models that were not previously possible; public administrations can offer central data that can be used by companies; investors can use data for risk analysis and risk reduction; public administrations and companies can become more transparent and thus increase the trust of citizens and customers; scientific work can be accelerated; prepared data can improve the decision-making of customers, companies, and public administrations.

Regarding the establishment and development of national open data ecosystem, Germany finds itself in the midfield. According to the Open Data Census ranking, Germany was relegated from position 11 down to 27 between 2013 and 2015. The reason for this was not a decrease of open data initiatives in Germany, but rather the rapid increases of sixteen other countries over the past two years. Germany currently risks falling behind in regard to international developments. In order to make a contribution to this important discussion, this study aims to answer the following question: What is the economic potential of open data for Germany?

This study's estimation of the economic potential of open data in Germany is based on a large set of existing studies. The authors examine what Germany can realistically achieve over the next ten years and present three different scenarios for moving towards a national open data ecosystem. Which of the three scenarios will ultimately be achieved depends mainly on the political will to shape the necessary framework conditions.

- **In the conservative scenario,** a potential of 12.1 bn EUR p. a. is realistic. This would mean that Germany pursues a reactive strategy in order to catch up to other countries by nationwide extension of the already existing activities in the context of open government data.
- **In the ambitious scenario,** a potential of 43.1 bn EUR p. a. is realistic. To achieve this, Germany would need to follow a proactive strategy (open by default). The aim is to increase the benefits through an additional opening of business and personal data and hence, catch up internationally. Up to 20,000 new jobs[2] could be created under this scenario.
- **In the optimistic scenario,** the potential is as high as 131.1 bn EUR p. a. It requires open data to be a central part of an overall national strategy for the 21st century. The potential of open data would have to be identified and mobilized to address major social challenges. Such a plan would go beyond the digital transformation: Germany would challenge itself as a nation with regard to the development of an encompassing national open data ecosystem, rendering Germany the leading open data nation.

These three scenarios offer a wide spectrum of possibilities. Consequently, the way in which potential will be exploited is of crucial importance. Differences in the design of administrative, commercial and civil framework conditions will directly shape Germany's position within this spectrum.

The question of economic potential is important, yet considering this perspective alone leads to an isolated analysis because various types of benefits cannot be expressed in quantitative terms. In order to estimate not only the economic, but also the wider disruptive potential of open data for society, a more holistic perspective is necessary. Such a complete perspective should consider both quantifiable and non-quantifiable benefits for the public and private sector, as well as for civil society.

What will activate the potential of open data? The digital transformation progresses continuously and will eventually reach a turning point where systemic changes will shift from an incremental to a transformative effect. This point in a digitally transformed society is achieved when business processes of organizations are fully digitized and interconnected. Thus, the potential of open data will be activated

by the digital transformation enabling changes in the public and private sectors, and in civil society. A realization of the benefits of open data will occur within the four 'pillars' described below:

- **Governance and Collaboration.** Only the opening of data will allow the rerouting of information flows and the redesign of business processes within government that current 'silo' work practices prevent. The opening of data for decision-making from institutions and organizations will enable massive participation and collaboration by outside stakeholder groups.
- **Innovation and Sustainable Development.** The opening of infrastructure and production data from institutions and organizations will empower new stakeholder groups outside the traditional value chains to become independent providers of new services. The massive publication of relevant data will enable larger numbers of stakeholders to contribute ideas and solutions to pressing societal challenges, e.g. the Sustainable Development Goals as defined by the United Nations.
- **Education and Participation.** The opening of education data will massively broaden access to education and make the institutional landscape larger and more diverse. Only by promoting data literacy and other crucial digital competencies will political and societal inclusion in this century be effective.
- **Transparency and Trust.** The opening of infrastructure data will make decisions and actions by political institutions and private organizations open to scrutiny and democratic control. Only such openness can create the level of transparency that will once again strengthen trust in large institutions and companies.

The consequent exploitation of the potential of open data will eventually lead to the development of an encompassing open data ecosystem in Germany. We understand such an ecosystem as a system that generates economic and social value through the creation, preparation and usage of open data in a large network of diverse stakeholders. In order to achieve this, each of the four 'pillars' for potential represents a field of action in which political decision-makers at all federal levels need to become active.

1 | Einleitung

Innerhalb weniger Jahre hat die Open-Data-Bewegung breite Unterstützung erfahren. Die Open Government Partnership,[3] 2011 von acht Staaten gegründet, ist inzwischen auf 66 angewachsen – allerdings ohne Deutschland. Im Juni 2013 haben die G8-Regierungschefs eine eigene Open Data Charta[4] unterzeichnet; auch Deutschland hat im Nachgang einen Nationalen Open-Data-Aktionsplan vorgelegt.[5] Gleichzeitig beeindrucken Studien durch große Zahlen: Auf 3 Billionen USD schätzt die Unternehmensberatung McKinsey das weltweite volkswirtschaftliche Potential (McKinsey Global Institute, 2013), während die Europäische Kommission einen Wert von 140 Milliarden EUR pro Jahr für die EU veranschlagt (European Commission, 2011).

Die digitalen Eigenschaften von Daten machen sie zu einer ungewöhnlichen Ressource. Einmal öffentlich zur Verfügung gestellt, (a) werden sie nicht knapp, (b) können sie zu Grenzkosten nahe null kopiert und beliebig verteilt werden und (c) eignen sie sich für eine potentiell globale Wiederverwendung, Neu- und Umnutzung. Die Entwicklung vergleichbarer offener digitaler Ressourcen in den letzten Jahren, wie Betriebssysteme (Linux), Enzyklopädien (Wikipedia) oder Karten (OpenStreetMap), lässt die Vermutung zu, dass Ausmaß von Nutzen und Wirkung von Open Data zukünftig noch signifikant steigen wird.

Noch handelt es sich bei Open Data um ein junges, wenig erforschtes Gebiet. Sowohl die theoretische Grundlagenarbeit als auch die empirische Datenlage stecken noch in den Kinderschuhen. Einige Potential-Modelle wurden in Studien vorgeschlagen, und es wurden bereits Werte für die wirtschaftliche Bedeutung hochgerechnet. Bedingt durch die noch immer kleine Datenbasis handelt es sich jedoch häufig um wenig belastbare Ergebnisse.

Diese Studie präsentiert eine quantitative Schätzung des volkswirtschaftlichen Potentials von Open Data für Deutschland. Den Schwerpunkt der Studie bilden offene bzw. potentiell offenzulegende Daten in der öffentlichen Verwaltung, sogenannte Open Government Data. Das berechnete Potential ist aus Sicht der Wirtschaft und der Zivilgesellschaft quantifiziert. Darüber hinaus werden qualitative Aspekte des Nutzenpotentials für die öffentliche Verwaltung beschrieben.

Diese Studie beantwortet dazu drei Fragenkomplexe:

A. Wie kann Open Data im öffentlichen und privaten Sektor Mehrwert stiften? Wie erfolgt die Nutzung konkret, wie wird durch Open Data Mehrwert geschaffen, und wie lässt sich dieses volkswirtschaftliche Potential messen?
B. Aus welchen Komponenten setzt sich das volkswirtschaftliche Potential von Open Data zusammen? Welcher methodische Ansatz ist für dessen Ermittlung zu wählen?
C. Welches nachvollziehbare Projektionsmodell kann zur Quantifizierung eingesetzt werden?

Die Beantwortung der Fragen findet sich in der Struktur der Studie wieder: Kapitel 2 skizziert die Grundlagen und die Begriffsdefinition von Open Data und legt somit die Basis zur Beantwortung der Fragen A und B. Darauf aufbauend vertieft Kapitel 3 die Fragen A und B und stellt den Mehrwert dar, der durch Open Data geschaffen werden kann. Kapitel 4 zeigt auf, welcher methodische Ansatz sich unter den Gegebenheiten in Deutschland in Bezug auf die verfügbare Datenlage am besten zur Berechnung des Open Data Potentials (Frage C) eignet.

Zudem werden in diesem Kapitel einschlägige Studien aufgeführt, die sich mit der Open-Data-Potentialberechnung in anderen Ländern beschäftigen, um dadurch ein Bild verschiedener Ansätze zu erhalten. Basierend auf den Annahmen und dem im Kapitel 4 vorgestellten Vorgehen wird im Kapitel 5 das Potential von Open Data berechnet (Frage C). Kapitel 6 schließt mit einer kritischen Reflektion und einem Fazit.

ÖFFNEN

Insgesamt ist „Open Data" ein sehr junges Politikfeld, das bereits auf eine dynamische Geschichte zurückblicken kann. Daten werden entweder zu Transparenzzwecken „geöffnet" oder zum Erschließen neuer Nutzenpotentiale. Verschiedene Datenkategorien (Finanzen, Umwelt usw.) besitzen unterschiedlich ausgeprägte Potentiale.

Daten „öffnen" bedeutet, sie unter minimalen rechtlichen und technischen Restriktionen der Allgemeinheit gratis zur Verfügung zu stellen. In der Praxis sind noch einige weitere Prinzipien einzuhalten (zum Beispiel Maschinenlesbarkeit), damit Daten als „Open Data" gelten.

2 | Kurzeinführung

2.1 | DIE KURZE UND DYNAMISCHE GESCHICHTE VON OPEN DATA

Kaum ein Regulierungsthema hat es in so kurzer Zeit von der ersten Idee bis zur Umsetzung in Gesetze geschafft wie Open (Government) Data. 2007 trafen sich in Sebastopol (USA) Internetaktivisten, Experten und Entrepreneure, um Prinzipien eines neuen Terminus „Open Government Data" zu entwickeln.[6] Bereits ein Jahr später fand der Begriff Einzug in das Wahlprogramm eines US-Präsidentschaftskandidaten, und im Januar 2009 verabschiedete US-Präsident Obama das Memorandum über Transparenz und Open Government.[7] Im selben Jahr wurde in den USA, 2010 in Großbritannien, das erste nationale „Open-Data-Portal" ins Leben gerufen. 2010 wurden die Prinzipien von „Open Government Data" durch die Sunlight Foundation, eine Transparenz-NGO, konsolidiert.[8] Und ein Jahr später wurde die internationale Open Government Partnership gegründet, ein Zusammenschluss von anfangs acht Staaten mit dem Ziel, Open Data und Open-Government-Prinzipien auf nationaler und internationaler Ebene voranzubringen.[9] Im Mai 2013 erließ Präsident Obama eine executive order, die Offenheit und Maschinenlesbarkeit von Regierungsdaten als neuen Standard für amerikanische Behörden vorschreibt.[10]

Etwa zwei Jahre versetzt zu den USA begann die Bewegung auch in der EU Fuß zu fassen. Hier wurde der Begriff „Public Sector Information (PSI)" schon deutlich früher diskutiert, aber der Aspekt „open" wurde erst im Zuge der US-Entwicklungen aufgegriffen.

Zum Vergleich, folgende Meilensteine der letzten Jahre sind auf EU-Ebene bedeutsam:

2003	DIRECTIVE of the European Parliament and of the Council on the re-use of public sector information 2003/98/EC
2011	COMMUNICATION FROM THE COMMISSION Open data – An engine for innovation, growth and transparent governance COM(2011) 882 final
2012	Launch des Portals der EU-Institutionen http://www.open-data.europa.eu/
2013	DIRECTIVE of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information 2013/37/EU erstmalige Erwähnung des Begriffs und des Konzepts „Open Government Data"

2014	COMMISSION NOTICE Guidelines on recommended standard licences, datasets and charging for the reuse of documents (2014/C 240/01) COMMUNICATION FROM THE COMMISSION Towards a thriving data-driven economy COM(2014) 442 final (SWD(2014) 214 final)
2015	COMMUNICATION FROM THE COMMISSION A Digital Single Market Strategy for Europe COM/2015/0192 final Launch des paneuropäischen Open-Data-Portals https://open-data.europa.eu

Eingebettet in diese Entwicklungen in den USA und der EU haben sich die G8-Staaten im Juni 2013 der G8-Open-Data-Charta.[11] verpflichtet. Die Charta enthält als Absichtserklärung fünf Prinzipien für den Umgang mit Open Government Data sowie Richtlinien für die Umsetzung von Open-Data-Vorhaben.

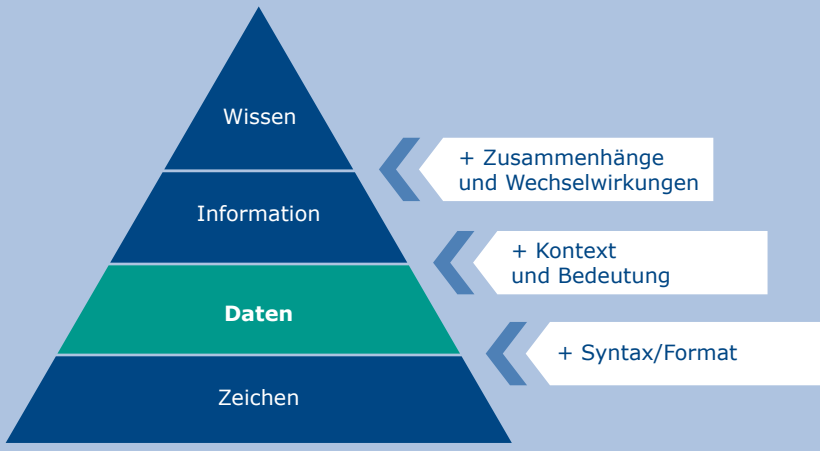
Das erste Prinzip Open Data by Default soll in der Praxis der öffentlichen Verwaltung gelebt werden, d. h. Daten sollen grundsätzlich öffentlich gestellt werden. Das zweite Prinzip Quality and Quantity soll dafür Sorge tragen, dass Daten rechtzeitig, umfassend und fehlerfrei öffentlich gestellt werden. Das dritte Prinzip Usable by All stellt die Einigung der Staaten darüber dar, dass Open Data ohne Nutzungsbarrieren in Form von Gebühren oder einschränkenden Lizenzen sowie administrativen und bürokratischen Hürden, wie z. B. Registrierung und Anmeldung zur Nutzung, zur Verfügung gestellt werden soll, damit die Daten für jedermann jederzeit zugänglich sind. Das vierte Prinzip Releasing Data for Improved Governance stellt die Verpflichtung der Staaten zum Austausch von Erfahrung und technischer Expertise im Bereich Open Data sowie zur Offenlegung und Transparenz sämtlicher damit verbundener Prozesse in den Vordergrund. Dazu zählen etwa die Datensammlung, der Einsatz einschlägiger Standards oder der Veröffentlichungsprozess. Die Verwaltung insgesamt soll dadurch verbessert, d. h. die demokratischen Institutionen gestärkt und die Ausrichtung der Politik auf die Bedürfnisse der Bürger gefördert werden. Das Prinzip Releasing Data for Innovation erkennt die Bedeutung von Vielfalt als Stimulus von Kreativität und Innovation in Wirtschaft und Gesellschaft und betont die Förderung von Open Data Literacy (Bildung mit Fokus auf Open Data) der Bevölkerung, damit die Basis für ein aktives zivilgesellschaftliches Engagement bei der Entfaltung der Nutzenpotentiale von Open Data geschaffen werden kann.

Diese fünf Prinzipien sind deshalb von hoher Bedeutung, weil sie als Grundlage für die Realisierung von Open-Data-Vorhaben der G8-Staaten dienen und entsprechend als Commitment auf nationaler Ebene zu betrachten sind. Die nur wenigen Jahre zwischen der Entstehung der gesellschaftlich-politischen Idee bis zur regulatorischen Verankerung und Umsetzung durch staatliche Strukturen sind ein weiteres Indiz für die hohe Bedeutung in der politischen Wahrnehmung einiger Länder.

2.2 | BEGRIFFSDEFINITION UND ABGRENZUNG

Durch die allgegenwärtige Verwendung des Begriffs „Daten" ist eine Abgrenzung zu weiteren Begriffen wie „Information" und „Wissen" vonnöten. In einer verallgemeinerten Form lassen sich die Begriffe wie folgt abgrenzen (vgl. Abbildung 2). Zeichen (Buchstaben, Ziffern usw.) bilden die erste Ebene der Wissenspyramide. Diese sind ohne Syntax bzw. Format jedoch nur eingeschränkt weiter nutzbar.

ABB. 2 **Daten als Grundlage von Wissen**



Rehäuser/Krcmar (1996,6) und Krcmar (2015,11)

Um die Nutzbarkeit von Zeichen zu erhöhen, wird eine Syntax bzw. ein Format definiert, und so entstehen aus den Zeichen Daten. Wenn nun diese Daten in Kontext gesetzt bzw. mit Bedeutung (Semantik) versehen werden, können mit Hilfe von Daten Informationen entstehen. Auf der höchsten Ebene der Wissenspyramide können Informationen angereichert werden, indem sie um Zusammenhänge und Wechselwirkungen ergänzt werden, woraus Wissen entsteht.

Laut Definition des IT-Planungsrats soll der Begriff Daten in dieser Studie „reine ‚Fakten' bezeichnen – unabhängig von Bedeutung, Interpretation und Kontext. Erst indem solche ‚Daten' (oder ‚Fakten') in einem konkreten Bedeutungskontext interpretiert werden, werden aus ihnen ‚Informationen'. Dies bedeutet auch, dass durch eine Weitergabe von Daten und die daraus resultierende Einbindung in andere Kontexte neue Informationen entstehen können."[12]

Im Rahmen dieser Studie verwenden wir für Open Data die gängige Definition des Open Data Handbook:[13] „Offene Daten sind Daten, die von jedermann frei benutzt, weiterverwendet und geteilt werden können – die einzige Einschränkung betrifft die Verpflichtung zur Nennung des Urhebers."

Für Open Government Data verwenden wir die deutsche Definition von von Lucke/Geiger (2010): „Offene Verwaltungsdaten sind jene Datenbestände des öffentlichen Sektors, die von Staat und Verwaltung im Interesse der Allgemeinheit ohne jedwede Einschränkung zur freien Nutzung, zur Weiterverbreitung und zur freien Weiterverwendung frei zugänglich gemacht werden."

In Analogie und Ergänzung werden in dieser Studie unter „Open Business Data (OBD)" geöffnete Daten der Wirtschaft und unter „Open Personal Data (OPD)" von Einzelnen veröffentlichte, persönliche Daten verstanden. Unter dem Kürzel „OxD" werden alle drei genannten Arten von Open Data subsumiert. Diese Unterscheidung ist bei der Projektion von Relevanz (vgl. Kap. 5).

Des Weiteren entstanden um das Thema „Daten" verschiedene Begriffe. Da sie im Open-Data-Diskurs häufig genannt werden, erfolgt hier eine Abgrenzung zu einigen dieser Konzepte.

- **Public Sector Information (PSI).** Begriff, den die EU bis 2013 nutzte, bevor der aus den USA stammende Begriff Open (Government) Data populär wurde.
- **Open-Shared-Closed.** Das Open Data Institute entwickelt das Konzept eines Datenspektrums,[14] in dem versucht wird, verschiedene Grade an Offenheit in einem Spektrum darzustellen. Open Data sind Daten, auf die jede/r zugreifen kann und die jede/r nutzen und teilen kann. Closed Data sind Daten, die nur dem Eigentümer/Halter zugänglich sind. Shared Data lassen sich in drei Varianten unterteilen: (a) Named-access-Daten, die nur mit benannten Personen/Organisationen geteilt werden; (b) Attribute-based-access-Daten, die nur Gruppen gemäß spezifischer Kriterien zugänglich sind; (c) Public-access-Daten, die allen zugänglich sind, jedoch nicht zu offenen Bedingungen.
- **Big Data.** Bedeutet in der Regel, dass eine Organisation (meist ein Unternehmen) persönliche Daten zu Zwecken der Marktforschung und Produktentwicklung sammelt – nicht immer mit dem expliziten Einverständnis der Kunden. „Big" bezieht sich auf die in Aggregation riesigen Datenmengen, die technisch erhöhte Anforderungen an die Analyse stellen (Rechenpower, statistische Methoden u. ä.). Es handelt sich nicht um offene Daten.
- **MyData/MiData.** Ist als Gegenbewegung zu Big Data zu verstehen und hat zum Ziel, persönliche Daten vollständig – rechtlich und technisch – unter die Kontrolle ihrer Eigentümer zu stellen, indem organisatorische Maßnahmen in Kombination mit technischen getroffen werden, um informationelle Selbstbestimmung auch de facto ausüben zu können.[15]

Die Menge neuer Begriffe verdeutlicht, dass sich ein neues Bewusstsein für und ein neuer Diskurs über Daten bildet. Da sich dieses Feld noch in einer dynamischen Phase befindet, ist davon auszugehen, dass unterschiedliche Begriffe auch in Zukunft Verwendung finden werden. Die vorangestellten Definitionen dienen der Abgrenzung zum Begriff Open Data sowie dem Verständnis, wie er in dieser Studie verwendet wird.

NUTZEN

Der Wert von Daten liegt in deren Nutzung. Um den Nutzen zu maximieren, muss der Zugang maximiert werden: Offene Daten bieten maximalen Zugang und damit ein maximal mögliches Nutzenpotential.

Die Wirkung von Open Data besteht in der Nachnutzung durch andere Gruppen außerhalb der Verwaltung, die sie freigibt. Die Heterogenität und Größe dieser Gruppen prägen die Vielfalt in der Nachnutzung und der Wertschöpfung.

3 | Wertschöpfung durch Open Data

Die Verwaltung erfasst, bearbeitet und nutzt Daten, um ihre hoheitlichen Aufgaben wahrzunehmen. Diese „Datenverarbeitung" geschieht, seit es Verwaltungen gibt, und geschah bislang mit verwaltungsinternem Fokus, während der externe Datenaustausch mit Formularen erfolgte. Erst seit dem Aufkommen digitaler Möglichkeiten der Datenverarbeitung in der Bevölkerung – indem universelle Rechenmaschinen („Computer"), durchgängige Netze („Internet") und (Open Source) Software für einzelne BürgerInnen zugänglich und damit nutzbar wurden – kann die Frage gestellt werden, ob die vormals unzugänglichen Daten der Verwaltung in diesem neuen Umfeld Nutzenpotentiale entfalten können, die in der Jahrhunderte andauernden analogen Vorzeit schlicht nicht denkbar waren.

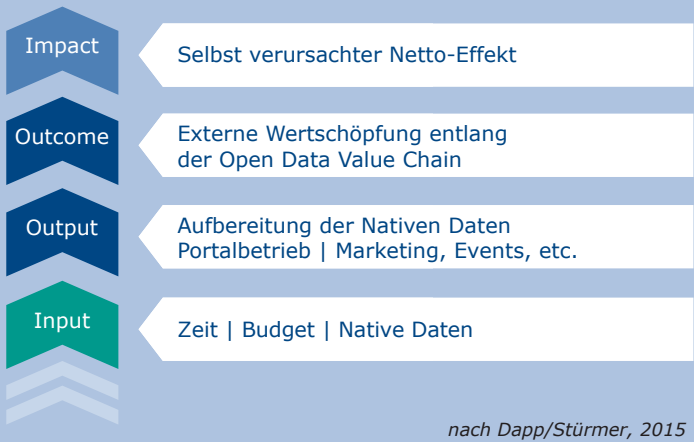
Die Ausgangsfrage dieses Kapitels ist daher: Wie entfaltet sich die Wirkung von Open Data? Die einfachste Antwort darauf ist: Open Data wirkt nicht durch die Veröffentlichung der Daten, sondern (nur) durch die Vielfältigkeit der Nachnutzung. Mit anderen Worten: Die Verwaltung erbringt nicht die Wertschöpfung, sondern sie ermöglicht sie – für alle anderen.

3.1 | WIE OPEN DATA WIRKUNG ENTFALTET

Eine Möglichkeit, die Wirkung (impact) von Open Data darzustellen, bietet die Social Return on Investment (SROI)-Methodik, die bereits für Open Government Data angewendet wurde (vgl. Dapp/Stürmer, 2015). In einer für die SROI-Methodik üblicherweise formulierten „Theorie des Wandels" wird in vier Stufen beschrieben, wie man entlang einer Wirkungskette von einem Input zur gewünschten Wirkung gelangt (Abbildung 3).

Input sind die Ressourcen (Zeit, Budget, native Daten), mit deren Einsatz ein Output (Aufbereitung der nativen Daten, Portalbetrieb, Marketing etc.) ermöglicht werden kann. Outcome ist die vielfältige Gesamtheit der Aktivitäten aller anderen Stakeholder. Diese Gesamtheit übersteigt in der Regel den eigenen Outcome um ein Vielfaches. Der Impact ist schließlich die Differenz dessen, was an Outcome auch ohne eigenes Zutun passiert wäre – also die selbst verantwortete Nettowirkung, die durch die Veröffentlichung von Open Data erzielt wurde. [16]

ABB. 3 **Theorie des Wandels von Open Data gemäß SROI**



nach Dapp/Stürmer, 2015

Zwei Erkenntnisse können aus dieser Betrachtung abgeleitet werden:

- Ein Datenbereitsteller kann nur die ersten beiden Stufen – Input und Output – direkt steuern. Das breite Spektrum des Outcome und damit die langfristige Wirkung lassen sich zwar beeinflussen, aber nicht direkt steuern. Aus diesem Grund ist die isolierte Publikation von Daten, ohne weitere flankierende Maßnahmen und Strategien dahinter, ein Ansatz, der zu wenig sichtbarem Nutzen von Datenportalen führt.
- Die einfache Messbarkeit der Wirkung nimmt entlang der vier Stufen rapide ab. Während der Datenbereitsteller den Input und Output in der Regel durch interne Messungen quantifizieren kann, sind auf der Stufe Outcome nur noch Abschätzungen möglich, die zudem auf Angaben vieler beteiligter Akteure beruhen. Der Aufwand für eine Wirkungsmessung nach dieser Methode ist groß.

Kapitel 3.2 erläutert, wie Datenbereitsteller die Stufen Input und Output – also die Entstehung und Publikation von Open Data – effektiv gestalten können. Kapitel 3.3 zeigt auf, wie die komplexe Stufe Outcome – im Fall von Open Data durch die Heterogenität der Nachnutzung durch viele Akteure und Nachnutzungsmöglichkeiten gekennzeichnet – strukturiert werden kann. Damit sind die Grundlagen gelegt, um in Kapitel 3.4 eine ökonomische Betrachtung von Open Data vornehmen zu können.

3.2 | ENTSTEHUNG UND PUBLIKATION VON OPEN DATA

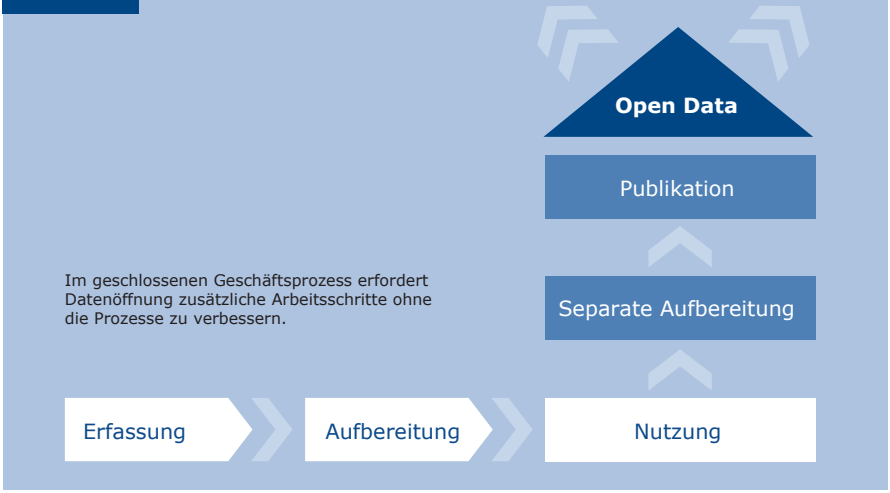
In den Input fließen alle vom Datenbereitsteller eingebrachten Ressourcen, um Open Data umzusetzen und die erwünschten Wirkungen zu erzielen. Dazu gehören beispielsweise finanzielle Mittel, Personal und Zeit. Basierend auf den Ressourcen kann der Output – meist in Form eines Datenportals oder veröffentlichter Daten – erreicht werden. Noch wichtiger sind aber die Datenbestände, die als Open Data veröffentlicht werden sollen. Hierfür schlagen Dapp/Stürmer den Begriff „native Daten" vor, weil sie „die ursprünglichen Daten der Verwaltung, sozusagen im Naturzustand, darstellen". Sie argumentieren, dass dieser „Naturzustand (...) deshalb ein wichtiger Faktor ist, weil er direkt den notwendigen Bedarf an Zeit und Geld bestimmt, um die Daten in das Portal zu überführen. Typischerweise liegen native Daten in vielen unterschiedlichen Datenformaten vor, was einen beträchtlichen Aufwand für die Konversion und Aufbereitung bedeuten kann. Die Formate sind in vielen Fällen an die Hersteller gebunden, welche die Informationssysteme entwickelt haben. Solche geschlossenen Datenformate stellen besondere technische und rechtliche Hürden und somit Kosten für eine Aufbereitung als Offene Daten dar (...) Das Problem besteht in der Regel deshalb, weil historisch keine organisationsweit einheitlichen Datenstandards bei Beschaffungen verlangt wurden, die auch die Offenheit der Datenformate als Anforderung enthielten" Dapp/Stürmer (2015, 14).

Hinzu kommt, dass die Verwaltung traditionell funktional nach internen Zuständigkeiten organisiert ist und nicht aus der durchgängigen Prozesssicht eines Bürgers oder eines Unternehmens. Zu technischen Aufwänden kommen noch organisatorische Aufwände hinzu, wenn Open Data voll integriert werden soll.

Wenn sich die Verwaltung der Publikationsfrage widmet, wie sie aus eigenen, nativen Daten Open Data machen kann, eröffnet sich ein ganzes Spektrum an Umsetzungsmöglichkeiten.

Das eine Ende des Spektrums beschreibt die heute am weitesten verbreitete Praxis (Abb. 4): Daten werden einzeln und fallweise aus bestehenden Geschäftsprozessen extrahiert, konvertiert, bereinigt und schließlich veröffentlicht. Dafür

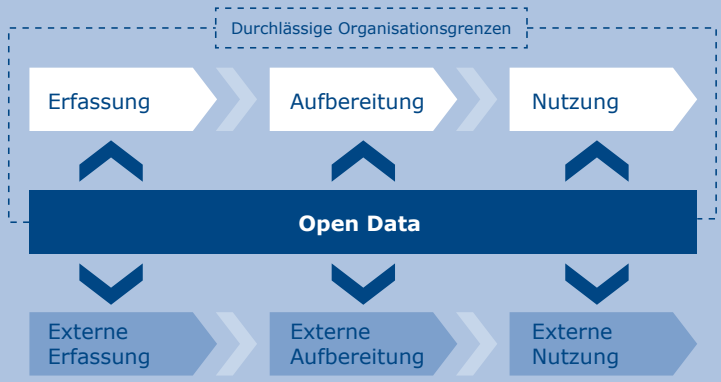
ABB. 4 **Publikation von Open Data in nicht-integrierter Form**



werden weder Geschäftsprozesse noch Methoden der Datenerfassung innerhalb der Verwaltung angepasst. Ohne parallele Neugestaltung von Prozessen wird das Potential von Open Data intern nicht ausgeschöpft, sondern nur als zusätzliche Belastung ohne direkt sichtbaren Nutzen empfunden. Diese Belastung entsteht deshalb, weil für jeden einzelnen Datensatz verschiedene Schritte zusätzlich ausgeführt werden müssen: Daten konvertieren und Maschinenlesbarkeit herstellen, Inventarisieren, Freigaben ggf. im Einzelfall abstimmen (u. a. Lizenzierung, Aktualisierung, Pflege etc.), Daten veröffentlichen und klären, ob und wie mit Rückmeldungen umzugehen ist etc. Diesen Zusatzaufwand für jeden Datensatz ohne einen unmittelbar sichtbaren Nutzen leisten zu müssen, ist im Sinne einer Kosten-Nutzen-Analyse schwer vermittelbar.

Mit einer Prozessintegration werden Daten hingegen ab der Erfassung als Open Data eingeordnet und durchweg so behandelt (Abb. 5). Die zuvor beschriebenen Schritte der Transformation sind weitgehend automatisiert. Durch Vorgaben sind Einzelfallentscheidungen nur selten erforderlich. Dies macht den Prozess insgesamt weniger aufwändig und schneller.

ABB. 5 **Publikation von Open Data in einem offenen, integrierten Prozess-Design**



Die durch Open Data ermöglichte Öffnung von Geschäftsprozessen erlaubt Verbesserungen durch kollaborative Arbeitsschritte.

Vor allem erlaubt dieses Design, dass einzelne Prozessteile teilautomatisiert geöffnet und einzelne Schritte ausgelagert werden können, weil die Verarbeitung der Daten nicht rechtlich oder technisch eingeschränkt ist. So können die Schritte der Qualitätsprüfung, Bereinigung, aber auch der Veredelung und Verknüpfung mit anderen Daten aus der Verwaltung nach außen gegeben werden. Über traditionelle Kooperationsmethoden der IT (zum Beispiel Public-Private-Partnerships) hinaus sind somit neue kollaborative Ansätze mit der Bevölkerung möglich: Digitales Bürgerengagement kann zu digitaler Ko-Kreation und Ko-Produktion werden – zu gemeinschaftlichem Gestalten und Erbringen von Leistungen (vgl. Benkler, 2006).

3.3 | AKTEURE UND AKTIVITÄTEN IN DER NACHNUTZUNG

Einmal freigegeben, stehen offene Daten für die Nachnutzung durch die Öffentlichkeit zur Verfügung. Dies geschieht durch das Hochladen auf eine Website, auf ein eigenes oder geteiltes Open-Data-Portal, das ggf. so mit weiteren Portalen verbunden ist, dass eine Datensuche über alle Portale möglich ist. Was passiert dann mit den Daten, wie werden sie technisch genutzt? Wer lädt sie für solch eine Nachnutzung herunter?

Was wird mit diesen Daten konkret gemacht? Davies (2010) beschreibt fünf Arten, wie eine Programmiererin Daten technisch weiterverwendet:

- sie extrahiert Fakten aus den Daten, um diese bestätigen zu können („fact-checking");
- sie analysiert Daten, um daraus Informationen zu erhalten, die dann visualisiert oder anders aufbereitet werden können;
- sie schreibt eine Programmierschnittstelle, um die Daten interaktiv verwenden zu können;
- sie bietet die Daten aufbereitet über eine eigene Schnittstelle für andere Nutzer an;
- sie bindet die Daten via Programmcode im Hintergrund ein, um einen eigenen Online-Dienst anzubieten.

Wer sind die Nachnutzer? Die Akteure lassen sich nicht vollständig aufzählen, sie stellen eine sehr heterogene Gruppe dar. Nutzer von Open Data können Hobby-Entwickler, politisch interessierte BürgerInnen, Startups mit datenbasierten Geschäftsmodellen, NGOs mit Datenbedürfnissen, Unternehmen verschiedener Sektoren, die Wissenschaft und die Verwaltung selbst sein. Die Heterogenität der Akteure ist auch einer der Gründe, weshalb es schwierig ist, die Wirkung von Open Data zu messen.

Wie entsteht aus Nachnutzung Wertschöpfung? Bei der Nachnutzung können verschiedene Akteure in Gruppen zusammenarbeiten. Die Kooperationsformen, um gemeinsam Open Data zu nutzen, sind so vielfältig wie die Akteurskreise. Typische Formen sind Hackathons, Datenjournalismus-Projekte oder Startups, die auf Open Data aufbauen. Das OpenData500 Projekt des GovLab (New York University) bietet zum Beispiel Analysen für verschiedene Länder, die zeigen, welche Open Data von welchen Firmen/Startups genutzt werden.[17]

3.4 | ÖKONOMISCHE BETRACHTUNG

Open Data unterscheiden sich, wie alle digitalen Artefakte, durch eine Reihe von Eigenschaften fundamental von physischen Gütern.

- Open Data können als öffentliche Güter betrachtet werden.
- Die Grenzkosten, d. h. die Kosten für die Bereitstellung eines zusätzlichen Datensatzes, sind fast null („zero marginal costs").
- Beim Konsum der Daten besteht keine Rivalität, d. h. Verwendung der offenen Daten durch eine Person hindert andere nicht, diese ebenfalls zu verwenden, da einfach Kopien erstellt werden können. Bei nicht-offenen Modellen, wie zum Beispiel proprietärer Software, gibt es deswegen Kopiersperren und eng gefasste Urheberrechtslizenzen.
- Es ist unmöglich, andere Individuen von der Nutzung des Angebots auszuschließen. Sobald Daten einmal veröffentlicht wurden, können sie nicht zurückgenommen werden.
- Hinzu kommen sogenannte positive Netzwerkexternalitäten: Der Nutzen der Daten steigt, je mehr Nutzer der Daten es gibt.

In der Vergangenheit wurden Daten, sogenannte Public Sector Information (PSI), nicht gemäß dieser Logik behandelt. Es wurden Preise für den Verkauf von Daten erhoben, die sich nicht an den Grenzkosten orientierten. Zudem wurde der Zugang und somit die Nutzung ausgeschlossen, indem Daten nicht veröffentlicht wurden, oder begrenzt, indem sie nicht umfänglich zur Verfügung gestellt wurden. Als Folge wurden die positiven Netzwerkexternalitäten von Daten eingeschränkt, und der Effekt, dass der Nutzen der Daten steigt, je mehr Nutzer es gibt, wurde limitiert.

Erst der Wechsel der Perspektive von geschlossenen Daten im Sinne von PSI zu „Open Data" trägt diesen Eigenschaften Rechnung. Daten zu öffnen, heißt die Eigenschaften, dass Daten Grenzkosten von null haben, nicht konkurrierend und schwer exkludierbar sind, nicht als Marktversagen zu verstehen, das durch rechtliche Instrumente wie geistiges Eigentum gelöst werden muss, sondern das Potential der positiven Netzwerkexternalitäten zu nutzen. Das ist die ökonomische

Erklärung dafür, dass der Nutzen von Open Data erst durch die große Gruppe der Nachnutzer realisiert wird. Darauf baut die Erwartung jener Verwaltungen, die Daten offen zur Verfügung stellen.

Nutzen von Open Data. Der gesamtwirtschaftliche Nutzen von Open Data speist sich aus unterschiedlichen ökonomischen Effekten:

- Einsparungen (z. B. von Energie, CO_2, effizientere Verwaltung);
- Wachstumspotenziale, die sich aus neuen, datenbasierten Geschäftsmodellen sowie verbesserten Services ergeben;
- gesamtgesellschaftliche Effekte (erhöhte Lebensqualität, höhere Ausbildung, Flexibilität, Transparenz, Zufriedenheit).

Der Nutzen tritt dabei in allen Sektoren auf, jedoch in unterschiedlicher Stärke. Anhand der Auswirkung lässt sich der Nutzen in drei Arten einteilen (vgl. z. B. (Pollock 2010), (Deloitte 2013), (McKinsey 2013), (Lateral Economics 2014), (Preische 2014)):

- **Der direkte Nutzen** ist unmittelbar zurechenbar (messbar) und es besteht eine Korrelation zwischen Einflussfaktoren und Auswirkungen, die hinreichend nachvollziehbar sind. Das heißt, bei einer Nutzenrechnung ist auch eine Rückwärtsrechnung möglich. Beispiele:
  - direkte neue Arbeitsplätze durch Open Data, direkte Erlöse;
  - bessere Ausbildung durch Open Data;
  - detaillierte Arbeitsmarktdaten.

- **Der indirekte Nutzen** ist mittelbar zurechenbar. Eine Korrelation zwischen Einflussfaktoren und Auswirkungen ist argumentativ darstellbar, aber nicht mehr direkt nachvollziehbar, d. h. Umfang und Anzahl der Annahmen nehmen zu. Eine Rückwärtsrechnung ist noch möglich, allerdings mit Abweichungen in den Ergebnissen. Beispiele:
  - Erlöse durch Weiterverarbeitung von Daten;
  - ein höheres Einkommen durch bessere Ausbildung;
  - bessere Jobvermittlungen durch detaillierte Arbeitsmarktdaten.

- **Der weiterreichende bzw. nachgeordnete Nutzen** ist nur mittelbar zurechenbar. Eine Korrelation zwischen Einflussfaktoren und Auswirkungen ist argumentativ nur unter weiteren Annahmen und einer positivistischen Wahrnehmung der Welt darstellbar (sogenannte „Wohlfahrtseffekte"). Eine Rückwärtsrechnung ist nicht mehr möglich, und Zurechenbarkeiten aus Ursachen beruhen auf Annahmen. Beispiele:
  - bessere Ausbildung und höheres Einkommen führen zu besserer Gesundheit;
  - bessere Jobvermittlung erhöht die Arbeitszufriedenheit.

Es ist anzunehmen, dass alle drei Nutzenarten noch durch Netzwerkeffekte verstärkt werden können, u. a. aufgrund der Kombination mehrerer Datenquellen und somit Schaffung neuer, angereicherter Daten und darauf aufbauenden Geschäftsfeldern.

Zur Veranschaulichung der verschiedenen Nutzenkategorien und deren unterschiedlichen Spannweiten dienen die US-Wetterdaten und der GPS-Markt als Beispiele:

- Offene Wetterdaten in den USA unterstützen Anwendungen im Wert von 1.5 Mrd. USD[18] im Sekundärversicherungsmarkt, verhelfen dem Unternehmen Weather Channel zu einem Verkaufswert von 3.5 Mrd. USD[19] bzw. The ClimateCorp von 1 Mrd. USD und sparen durch die darauf basierenden Wettervorhersagen 30 Mrd. USD[20] jährlich.
- Der Wert der GPS-Daten wird auf ca. 90 Mrd. USD[21] geschätzt. Garmin, eine Firma die GPS-Produkte herstellt, hat einen Marktwert von 6 Mrd. USD.[22] Bei der Freischaltung von GPS 1983[23] erwarteten nur sehr Wenige solch ein Potential.

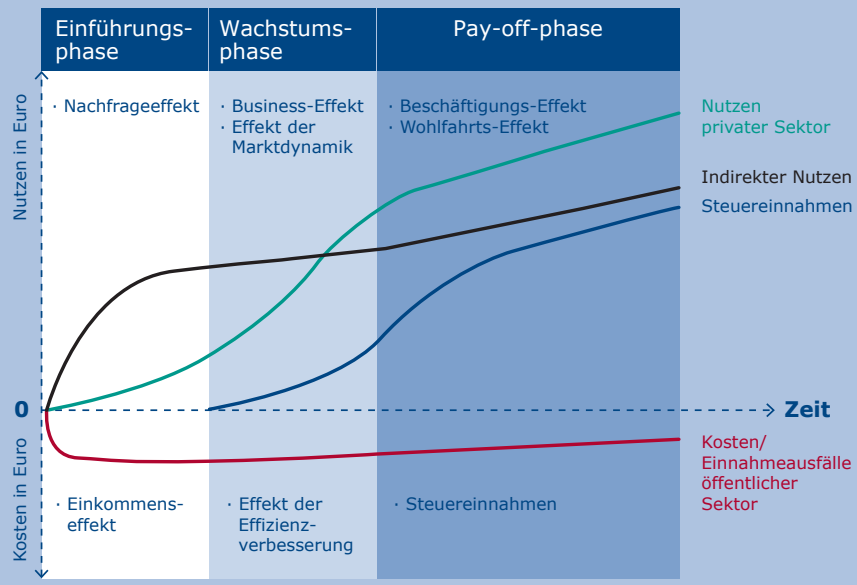
Die genaue Beurteilung des Nutzenbeitrags bedarf eines besseren Verständnisses der zusätzlich verursachten Kosten bei weiteren Datenfreigaben sowie der Unterscheidung zwischen den dadurch entstehenden Produzenten- und Konsumentenrenten (Wert oder Nutzen, den der Verbraucher genießt, der oberhalb des Kaufpreises liegt). Es zeigt sich, dass die Konsumentenrente ein Vielfaches der Produzentenrente ist (Deloitte, 2013, 110).

Maximale Ausschöpfung der Nutzenpotentiale. Ökonomisch betrachtet besitzt Open Data ein großes Potential, das aber zunächst einmal nur theoretisch vorliegt: Wie stark es ausgeschöpft wird, hängt von der politischen und administrativen Führung im Management dieser Öffnungsprozesse ab. Die ökonomische Argumentation legt nahe, ein Maximum anzustreben.

Dateninfrastruktur. Da durch Netzwerkeffekte der Nutzen von Daten davon abhängt, wie viele andere Daten schon verfügbar sind, spielt auch der Reifegrad einer zukünftigen Dateninfrastruktur eine Rolle. So wie die physische Infrastruktur (Strom, Wasser, Telekommunikation etc.) für die analoge Ökonomie von großer Bedeutung ist, wird es die Dateninfrastruktur (Daten zu Energie, Transport oder Gesundheit etc.) für die digitale Ökonomie bzw. die Informationsgesellschaft sein.[24] Trotzdem werden stark unterschiedlich hohe Ausgaben für diese beiden Infrastrukturen getätigt: Den Ausgaben für Straßenbau und Eisenbahn von rund 11,5 Mrd. EUR in Deutschland 2013 stehen Ausgaben für die Bereitstellung von Open Data von rund 0,002 Mrd. EUR gegenüber (von Lucke, 2013).

Nutzenverlauf über die Zeit. Einschätzungen zu treffen, wann sich konkrete Nutzenpotentiale von bestimmten offenen Daten realisieren, sind nur mit Annahmen möglich. Abbildung 6 gibt eine schematische, nach Phasen gegliederte, Darstellung über das Eintreten verschiedener ökonomischer Effekte bei Open Data von Preische (2014) wieder. Wie daraus ersichtlich, ist zu erwarten, dass der Nutzenverlauf für Open Data dem typischen S-Kurven-Konzept folgt. Während der Einführungsphase, in der sich Deutschland heute befindet, ist vor allem ein Anstieg des indirekten Nutzens bei höheren Investitionskosten des öffentlichen Sektors zu erwarten. Die Wachstumsphase wird anschließend durch Innovationen und andere Marktdynamiken sowie Effizienzverbesserungen innerhalb der Verwaltung geprägt sein.

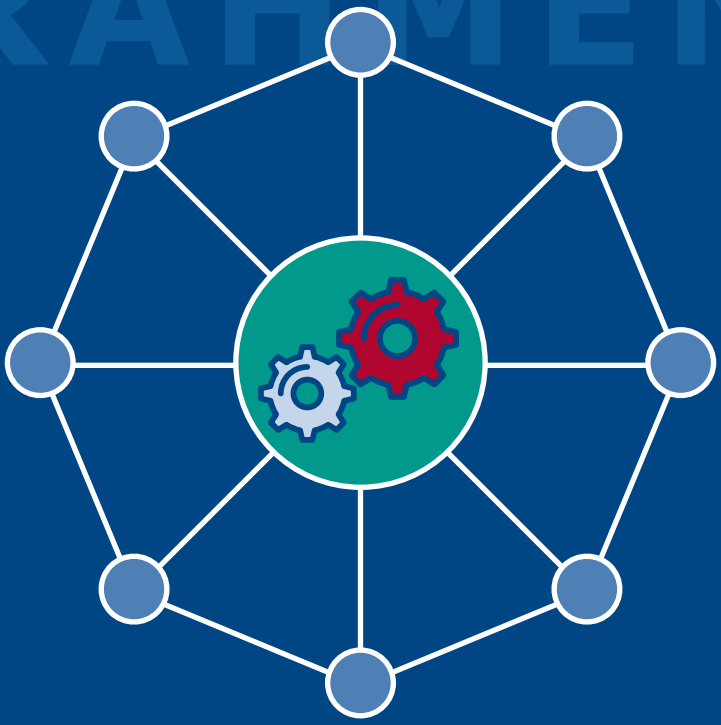
ABB. 6 **Ökonomische Effekte durch Datenöffnung im Zeitverlauf**



Quelle: Preische (2014)

Erst in der pay-off-Phase wird der Nutzen der geöffneten Daten breiter verankert, und die Anfangsinvestitionen der Verwaltung werden durch höhere Steuereinnahmen und Beschäftigungs- bzw. Wohlfahrtseffekte wieder ausgeglichen. Der Prozess zur Erreichung des vollen Potentials von Open Data in Deutschland wird abhängig von der Frequenz und dem Umfang von Datenfreigaben und der entstehenden Netzwerkeffekte mehrere Jahren benötigen. Für die Berechnungen in Kapitel 5 wird von einem Zeitraum von zehn Jahren bis zum Einsetzen der pay-off-Phase ausgegangen.

RAHMEN



Open Data befindet sich noch in einem relativ frühen Entwicklungsstadium. Geringe Datenmengen, wenige Praxisbeispiele und viele Annahmen stellen erhebliche methodische Herausforderungen an eine quantitative Schätzung des wirtschaftlichen Potentials. Das Untersuchungsdesign schlägt daher vor, Daten aus bestehenden Studien zu verwenden und das für andere Länder berechnete Potential auf Deutschland zu adaptieren. Die Werte erscheinen also entlang eines Spektrums, das in Kapitel 5 mit konkreten Zahlen belegt wird. Grundsätzlich wird davon ausgegangen, dass die Schaffung optimaler politischer und rechtlicher Rahmenbedingungen für Open Data in Deutschland rund zehn Jahre dauern wird.

Demgegenüber steht der Status quo in Deutschland: Der sehr tiefe heutige Stand wird veranschaulicht, indem eigene Daten aus „Open-Data-Kommunen" erhoben und exemplarisch auf Deutschland hochprojiziert werden. Dieser Status quo zeigt auf, wie wenig die Infrastruktur entwickelt ist und, dass Deutschland bei Open Data noch nicht an die Spitze aufgeschlossen hat.

4 | Methodik der Potentialschätzung

4.1 | ZWEI ANSÄTZE: NATIONALE SEKUNDÄR- UND KOMMUNALE PRIMÄRDATEN

Der Fokus dieser Studie liegt auf einer quantitativen Datenauswertung sowie einer monetären Bewertung des Potentials von Open Data für Deutschland. Das Untersuchungsdesign stellt eine Kombination methodischer Ansätze dar. Existierende quantitative Studien aus unterschiedlichen geographischen Regionen werden gesammelt, analysiert und daraus Sekundärdaten erhoben (vgl. Kap. 4.2). Auf Basis der Sekundärdaten wird dann eine Adaption für Deutschland (Projektion) errechnet (vgl. Kap. 5.2).

Ergänzend werden deutsche Kommunen nach ihren bisherigen Open-Data-Kosten befragt (vgl. Kap. 4.3). Ausgehend von der Annahme, dass getätigte Investitionskosten den Durchschnittskosten gleichgesetzt werden können, wird aus diesen Primärdaten eine Hochrechnung vorgenommen (vgl. Kap. 5.3): Welches Potential für Deutschland ergäbe sich, wenn alle Kommunen ähnlich in Open Data investieren würden wie die derzeit bei Open Data aktiven Städte?

Warum dieses Vorgehen? Wie bereits anhand von mehreren Beispielen verdeutlicht, ist die Wertermittlung und Potentialabschätzung ein sehr aufwändiges Unterfangen, da zum einen die Natur von Open Data als immaterielles Gut keine eindeutige Zuordnung von Kosten und Nutzen zulässt und zum anderen eine umfassende Primärdatenerhebung mit einem beträchtlichen Aufwand verbunden ist.[25] Hinzu kommt, dass die Daten, die als Grundlage für Berechnungen in existierenden Studien gedient haben, für Deutschland teilweise nicht verfügbar sind, weil sich Open Data hierzulande in einem frühen Entwicklungsstadium befindet. Es bleibt lediglich die Möglichkeit, die in existierenden Studien berechneten Werte über gemeinsame Umrechnungsbasen – etwa das Bruttoinlandsprodukt und die strukturelle Ähnlichkeit zwischen den geographischen Regionen wie etwa EU oder OECD – auf Deutschland zu adaptieren.

Ziel ist es, ein plausibles Wertspektrum zu schaffen, um daraus eine realistische Abschätzung für das Open-Data-Potential für Deutschland in den kommenden zehn Jahren zu geben. Während der erste Ansatz eine langfristige Potentialabschätzung erlaubt, soll der zweite einen Anhaltspunkt liefern, wo Deutschland

stünde, wenn die Durchdringung in den Kommunen jener der heutigen „Open-Data-Kommunen" entspräche.

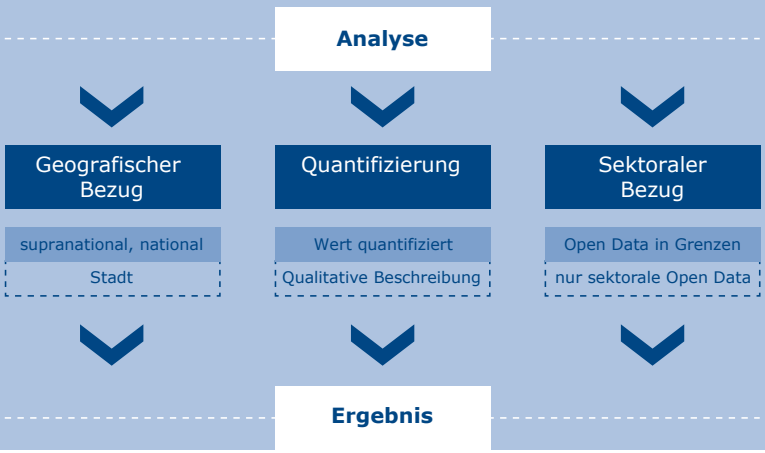
4.2 | POTENTIALSCHÄTZUNG DURCH INTERNATIONALE SEKUNDÄRDATEN

Es werden alle Studien seit dem Jahr 2000 ausgewertet, die einen Bezug zur Abschätzung des Open Data-Potentials haben. Die untersuchten Studien wurden zuerst auf Zusammenhänge untereinander analysiert: In einer „Studienlandkarte" (s. Selektion Abb. 8) wurde für jede einzelne Studie festgehalten, ob sie eine andere Studie referenziert, die Methodik einer anderen Studie verfolgt, ihre Daten aus einer anderen Studie bezieht oder einen komplett eigenen Ansatz mit eigenen Daten und eigener Methodik verfolgt.

Allgemein lässt sich feststellen, dass Studien vor dem Jahr 2013 mit dem Begriff „Informationen des öffentlichen Dienstes" (engl. Public Service Information, kurz PSI) arbeiten und dessen Potential abschätzen. Dies ist darauf zurückzuführen, dass der Begriff „Offene [Regierungs-] Daten" (engl. Open [Government] Data, O[G]D) noch nicht etabliert war. Studien ab 2013 führen den Begriff OGD ein und setzen ihn in Bezug auf Berechnungen mit PSI gleich. Es ist sicherlich nicht unkritisch, beide Begriffe gleichzusetzen, da es sich bei OGD und PSI nicht um identische Konzepte handelt. Wie allerdings bereits im Kap. 2 aufgeführt, ist eine scharfe Trennung der Begriffe im Umfeld von Open Data nicht immer möglich. Unserer Ansicht nach besteht kein nennenswertes Risiko bei der Übertragung der Potentialschätzung von PSI auf OGD. Der Hauptunterschied ist, dass in früheren Studien von einem Kommerzialisierungspotential von PSI die Rede ist, während bei OGD die Befolgung der Open-Data-Prinzipien als Ziel gesetzt wird, d. h. eine Kommerzialisierung seitens der öffentlichen Hand keine Option und kein vorrangiges Ziel ist. Diese Studie berücksichtigt Sekundärdaten auch von PSI-Studien für die Abschätzung des Open-Data-Potentials.

Die weitere Analyse der Studien wurde anhand der Kriterien „geographischer Bezug", „Quantifizierung des Open-Data-Potentials" sowie „sektoraler Bezug der Studie" durchgeführt, und darauf basierend wurden die Studien in Kategorien eingeteilt (vgl. Abbildung 7). Die Kriterien werden im Folgenden kurz beschrieben.

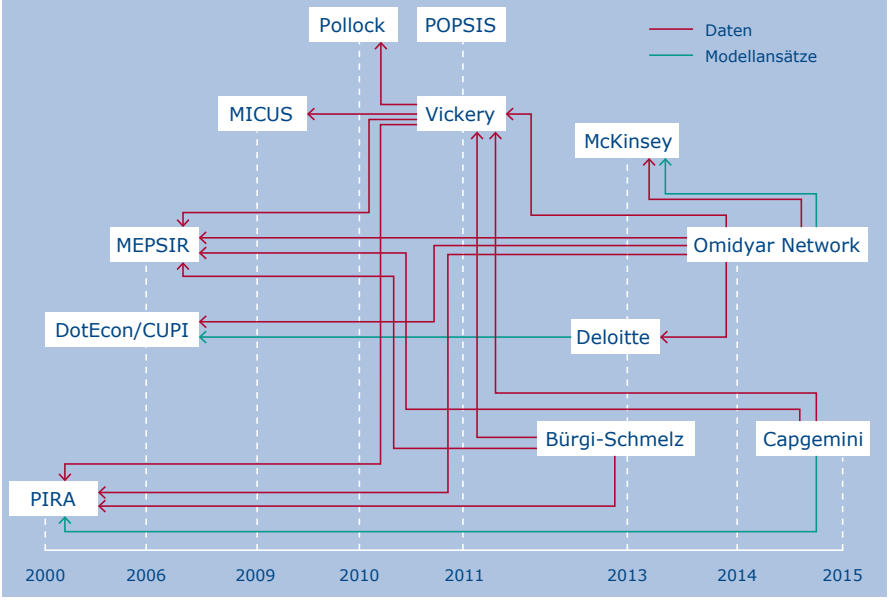
ABB. 7 **Auswahlschema der Studienselektion**



Unter dem Kriterium „geographischer Bezug" wurden Studien gefunden, die entweder eine Gruppierung von Ländern (supranational, z. B. EU-15) oder einzelne Länder (national, z. B. UK) oder einzelne Städte (z.B. Wien) abdecken. Durch den Fokus unserer Studie auf eine Adaption vorhandener Studien für das volkswirtschaftliche Potential von Open Data für Deutschland wurden zwei Kategorien gebildet. Die erste Kategorie enthält 25 Studien, die einen nationalen oder supranationalen Fokus haben. Die zweite Kategorie besteht aus zwei Studien, die eine kleinere Einheit fokussieren, z. B. eine Stadt. Für die Adaption werden Studien nur aus der ersten Kategorie gewählt, damit die Größenverhältnisse zwischen den errechneten Potentialen vergleichbar sind.

Eine Unterscheidung anhand des Kriteriums „Quantifizierung des Open-Data-Potentials" hat zu zwei Kategorien geführt. In der ersten Kategorie wurden 15 Studien erfasst, die das Potential von Open Data quantifizieren, d. h. die Studien adaptieren oder durch eigene Datenerhebung einen konkreten Wert oder einen Wertspektrum herleiten. Die zweite Kategorie umfasst zwölf Studien, die das Potential von Open Data anhand von Fallstudien oder argumentativ herleiten und dabei entweder dieses lediglich qualitativ beschreiben oder auf quantitative Studien referenzieren und das dort berechnete Potential zitieren. Obwohl der so beschriebene Wert von Open Data u. U. greifbarer wirkt, lassen sich kaum neue

ABB. 8 **Ausgewählte Studien für die Potentialschätzung**

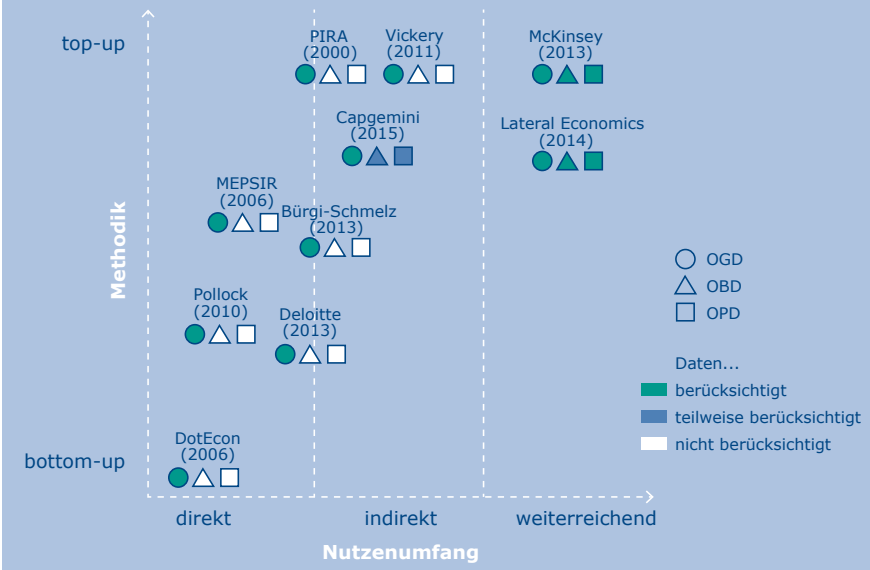


Erkenntnisse für dessen Quantifizierung herleiten. Daher wurden nur Studien in der weiteren Analyse berücksichtigt, die der ersten Kategorie angehören.

Beim Kriterium „sektoraler Bezug" wird unterschieden, ob Open Data (ggf. unterteilt in OGD, OBD und OPD) als Ganzes oder ob einzelne Sektoren (Geo-Daten, Wetterdaten usw.) fokussiert worden sind. Die erste Kategorie umfasst nur 25 Studien, die Open Data als Ganzes betrachten. Die zweite Kategorie umfasst zwei Studien, die einzelne Sektoren betrachten. Die Studien in dieser Kategorie wurden nicht weiter analysiert und für die Adaption nicht berücksichtigt, da wir Open Data als Ganzes betrachten und keine Annahmen über die sektorale Zusammensetzung des Potentials treffen.

Das Ergebnis der Selektion ist eine Zahl von zehn Studien, die in Abbildung 8 dargestellt sind. Sie sind als Netzwerk abgebildet, um die Zusammenhänge zu veranschaulichen: Verbindungen zeigen, welche Modellansätze (M) adaptiert wurden und welche Daten (D) berücksichtigt worden sind.

ABB. 9 **Übersicht der Studien nach Methodik und Nutzenumfang**



Die zehn Studien unterscheiden sich anhand mehrerer Kriterien: unterschiedliche methodische Ansätze, eigens erhobene Daten (Primärdaten) oder Daten anderer Studien (Sekundärdaten), im Nutzenumfang (direkt, indirekt, weiterreichend), im betrachteten Umfang von Open Data: Open Government Data/OGD, Open Business Data/OBD oder Open Personal Data/OPD oder einem Mix (OxD). Eine Übersicht über diese Unterscheidung stellt Abbildung 9 dar.

Eine Studie kann zudem auf einer top-down- oder einer bottom-up-Methodik basieren,[26] d. h. sie kann auf unterschiedlichen Aggregationsstufen die Berechnung des Potentials durchführen. Bei top-down findet eine Übertragung bzw. Adaption von Werten, die auf einem hohen Aggregationslevel (volkswirtschaftliche Angaben wie das Bruttoinlandsprodukt) basieren, auf Open Data statt. Das Gesamtvolumen ist nachvollziehbar, anteilige Zurechenbarkeit von Open Data ist nur durch Annahmen möglich, die durch Analogien argumentiert werden können. Folglich liegt eine Verallgemeinerung in der Natur dieser Methodik, und somit sind Studien, die dieser Methodik folgen, als eher optimistische bezüglich des geschätzten Potentials einzustufen. Ein Beispiel ist die PIRA-Studie (PIRA 2000).

Bei bottom-up wird der Wert basierend auf einem niedrigen Aggregationslevel (eigene Datenerhebung von Einzelwerten und -fällen, wie z. B. Fahrzeit einer Person und die Gruppierung von Personen nach der Dauer der Fahrzeit) erhoben, um daraus eine Hochrechnung des Potentials von Open Data zu errechnen. So kommen hier Daten einer direkten Messung von Einzelfällen zum Einsatz, und die Zurechenbarkeit ist nachvollziehbar, d. h. welcher Mehrwert in der Praxis konkret entsteht. Allerdings besteht die Gefahr, dass nicht sämtliche Komponenten des Nutzens von Open Data berücksichtigt werden können und dass das berechnete Potential damit nur einen Teil des tatsächlichen Potentials darstellt. Daher sind Studien dieser Methodik als eher konservativ bezüglich des errechneten Potentials einzustufen. Ein Beispiel ist die DotEcon (2006) Studie.

Daten in Studien werden nach Primär- oder Sekundärdaten unterschieden.[27] Primäre Daten werden direkt im Rahmen einer Studie erhoben – beispielsweise mittels Umfragen, Recherche von Dokumenten (Jahresabschlüsse, offizielle Statistiken usw.) oder Expertenschätzungen aus der Open-Data-Praxis. Sekundäre Daten werden mit Hilfe von Ergebnissen anderer Studien erhoben; es handelt sich um bereits vorhandene Ergebnisse, die auf einen neuen Kontext übertragen werden. In manchen Studien dient eine Mischung aus primären und sekundären Daten als Basis für die Berechnung des Potentials. Meistens werden in diesem Fall die Daten nicht „vermischt", sondern dienen zur Berechnung von einzelnen Werten, die mit Hilfe von Annahmen zu jeder verfügbaren Datenbasis argumentativ dargestellt und im Zusammenhang erklärt werden. Diese Unterscheidung ist wichtig, denn daraus lässt sich ablesen, ob ein Potentialwert eher optimistisch oder konservativ geschätzt ist. Die Studie von Vickery (2011) ist ein Beispiel, wie die Potentialschätzung anhand von Sekundärdaten erfolgen kann, während die Studie von Deloitte (2013) einen Mix verwendet, wobei jedoch eher Primärdaten als Basis für die Potentialschätzung verwendet werden.

Hinsichtlich des Umfangs von Open Data – OGD, OBD, OPD und OxD – gibt es bisher wenige Studien, die den Nutzen und das Potential von Open Data in der gesamten Breite (OxD) schätzen; Ausnahmen sind McKinsey (2013) sowie Lateral Economics (2014). Die meisten Studien decken den Nutzen von OGD ab, der aktuell eine große Überlagerung mit jenen von Open Data hat und deshalb als Annäherungswert herangezogen werden kann. Trotzdem ist zu beachten, dass

TAB. 1	Kennwerte der Potentialschätzungen der ausgewählten Studien			
Autor (Jahr)	**Geo-grafischer Bezug**	**Potential-spektrum absolut**	**Potential-spektrum in BIP-Anteil**	**Für die Adaption verwendete Werte: Poten-tial, BIP-Anteil**
PIRA (2000)	*EU15*	*28 – 134 Mrd. EUR*	*0,30 – 1,46 %*	*68 Mrd. EUR 0,74 %*
MEPSIR (2006)	*EU25+ Norwegen*	*10 – 48 Mrd. EUR*	*0,09 – 0,44 %*	*27 Mrd. EUR 0,25 %*
DotEcon (2006)	*Groß-britannien*	*0,59 – 1,1 Mrd. GBP*	*0,04 – 0,08 %*	*1,1 Mrd. GBP 0,08 %*
Pollock (2010)	*Groß-britannien*	*1,6 – 6 Mrd. GBP*	*0,11 – 0,40 %*	*4,5 Mrd. GBP 0,30 %*
Vickery (2011)	*EU27*	*28 – 200 Mrd. EUR*	*0,22 – 1,56 %*	*140 Mrd. EUR 1,1 %*
McKinsey (2013)	*Welt*	*3.200 – 5.400 Mrd. USD*	*4,26 – 7,19 %*	*3.380,8 Mrd. USD 4,26 %*
Bürgi-Schmelz (2013)	*Schweiz*	*0,86 – 6,6 Mrd. CHF*	*0,14 – 1,01 %*	*2,2 Mrd. CHF 0,35 %*
Deloitte (2013)	*Groß-britannien*	*1,2 – 7,2 Mrd. GBP*	*0,08 – 0,46 %*	*6 Mrd. GBP 0,39 %*
Lateral Economics (2014)	*Australien*	*21 – 64 Mrd. AUD*	*1.30 – 3.90 %*	*22,6 Mrd. AUD 1,40 %*
Capgemini (2015)	*EU28+*	*51 – 229 Mrd. EUR*	*0,37 – 1,58 %*	*188 Mrd. EUR 1,30 %*

Studien mit OxD-Fokus tendenziell ein optimistischeres Potential im Vergleich zu Studien mit OGD-Fokus berechnen.

Zusammenfassend gibt Tabelle 1 einen Überblick über die Studien und die darin berechneten Potentiale von Open Data, die als Basis für die Adaption auf Deutschland dienen werden.

Die unterschiedlichen Währungen hängen vom geographischen Bezug der Studie ab.[28] Die Werte für das Bruttoinlandprodukt der jeweiligen geographischen Region wird – falls angegeben – direkt aus dem Studientext übernommen. Falls diese Werte nicht angegeben waren, wurden sie über Statistiken der Weltbank, Eurostat oder des jeweiligen Landes ermittelt.

Bei der Adaption der Werte als Berechnungsbasis wurde der von der jeweiligen Studie vorgeschlagene Wert (z. B. der Mittelwert des Spektrums) übernommen. So wird beispielsweise in der Studie von Lateral Economics (2014) für Australien argumentiert, dass – anhand der von McKinsey (2013) errechneten Werte sowie unter Berücksichtigung weiterer Studien über das volkswirtschaftliche Potential einzelner Datensektoren – in den nächsten fünf Jahren ein kumulierter Wert von 1,40 Prozent des BIP erreicht werden kann. Unter Berücksichtigung der Tatsache, dass Deutschland und Australien OECD-Länder sind, eine vergleichbare wirtschaftliche Leistung aufweisen und demokratisch regiert werden, konnte der vorgeschlagene Wert für Deutschland – allerdings konservativer geschätzt auf ein Zeithorizont von zehn Jahren – übernommen werden.

Falls eine Studie keinen Wert vorgeschlagen hat oder unrealistische Annahmen getroffen worden sind,[29] wurde ein Wert aus dem angegebenen Spektrum unter Berücksichtigung und Analyse von in der Studie getroffenen Annahmen und Rahmenbedingungen ermittelt. So wurde in der DotEcon-Studie (2006) der höchste Wert im ermittelten Spektrum gewählt, da dieser als erreichbar bezeichnet wird, wenn entsprechende gesetzliche Rahmenbedingungen existieren. Die veröffentlichte Evaluation dieser Studie und dieser Bedingungen zeigt (DotEcon, 2015), dass die Rahmenbedingungen durchaus geschaffen werden konnten. Daher ist der gewählte Wert als realistisch anzusehen.

Ein weiteres Beispiel ist die als optimistisch eingestufte Studie von McKinsey (2013), die zwar einen Wert angibt, dieser aber für Deutschland in den nächsten zehn Jahren als nicht erreichbar erscheint, da die entsprechenden Einflussfaktoren – etwa der gesetzliche Rahmen, die Aufmerksamkeit von Entscheidern sowie die Agilität der Zivilgesellschaft und Wirtschaft in Bezug auf Open Data – noch nicht das Niveau erreicht haben, das in der Studie als Basis der Berechnung angenommen wird.[30]

Ein drittes Beispiel ist die Studie von Bürgi-Schmelz (2013) über das Potential von Open Data für die Schweiz. Dort wird zwar ein Wert angegeben, aber auch Argumente aufgeführt, dass die untere Grenze des Spektrums nur einen Teil des Nutzens durch Open Data erfasst und so z. B. Zeitersparnis nicht berücksichtigt. Daher wurde ein höherer Wert aus dem berechneten Spektrum übernommen.

Zusammengefasst zeigen die Studien die folgenden Aspekte auf: Der direkte Nutzen von Open Data ist signifikant (Deloitte (2013), Vickery (2011)), jedoch ist der Großteil des Nutzens indirekt und, wenn auch schwer zu quantifizieren, weiterreichend (McKinsey (2013), Lateral Economics (2014)). Die untersuchten Studien, insbesondere die qualitativen mit vielen Fallbeispielen, zeigen auf, dass Open Data schon heute neue Nutzer und Märkte erzeugt (z. B. (Bonina 2013), (Ubaldi 2013), (Hogge 2015)). Das geschätzte volkswirtschaftliche Potential ist, auch wenn die Schätzungen zum Teil stark differieren, klar auf reale Beispiele zurückzuführen und nicht imaginär. Der Wert der Daten liegt dabei nicht in den Daten selbst, sondern in der Verknüpfung mit anderen Daten, Dienstleistungen oder Ideen. Die Schwierigkeit bei der Verkürzung der Spannbreite des Potentialspektrums ist auf unscharfe Trennlinien bei der Berechnung des gewählten methodischen Ansatzes, die Datenbasis oder den Datenumfang zurückzuführen. Diese Schwierigkeit resultiert allerdings nicht aus mangelnder wissenschaftlicher Rigorosität der Studien, sondern aus der Natur von Open Data als immateriellem Gut, das sich noch in einem frühen Entwicklungsstadium befindet. Daraus folgt zum Beispiel, dass die Berechnung des Nutzens für die öffentliche Hand schwierig ist und in noch weniger erschlossenen Bereichen wie OBD und OPD nur durch Expertenschätzungen abzubilden wäre. Eine Adaption und Berücksichtigung der Wiederverwertung von OxD-Daten und eine breitere Bemessung des Nutzens von Open Data sollten im Laufe der Zeit allerdings möglich sein.[31]

4.3 | HOCHRECHNUNG KOMMUNALER OPEN-DATA-KOSTEN

Während das Hauptaugenmerk dieser Studie auf der Potentialabschätzung entsprechend der im vorangegangenen Kapitel beschriebenen Methodik liegt, wird im Folgenden auch ein weiterer methodischer Ansatz angewendet. Dabei wird eine Primärdatenerhebung durchgeführt, die ausgehend von den Open-Data-Kosten

TAB. 2 **Kennwerte der Potentialschätzungen der Studien auf Stadtebene**

Autor (Jahr)	Geografischer Bezug	Potentialspektrum absolut	Potentialspektrum in BIP-Anteil	Für die Adaption verwendete Werte: Potential, BIP-Anteil
Fuchs et al. (2013)	*Wien (AT)*	*2,7 Mio. EUR*	*0,003375 %*	*2,7 Mio. EUR 0,003 %*
Preische (2014)	*Berlin (DE)*	*21 – 54 Mio. EUR*	*0,01 % – 0,05 %*	*43,2 Mio. EUR 0,04 %*
Pollock (2010)	*(UK)*	*1.6 – 6 Mrd. GBP*	*0,11 % – 0,40 %*	*4,5 Mrd. GBP 0,30 %*

einiger Kommunen ein Potential errechnet. Dieses Kapitel erläutert das methodische Vorgehen.

Deutschland ist bei Open Data auf allen föderalen Ebenen aktiv – jedoch mit unterschiedlicher Intensität. Kommunen nehmen heute bei der Weiterentwicklung sowie der Nutzenschaffung von Open Data für Deutschland eine entscheidende Rolle ein. Die Open-Data-Pioniere befinden sich vor allem auf kommunaler Ebene: Einige Kommunen sind Vorreiter beim Betrieb von Datenportalen, der nachfrageorientierten Publikation von Daten, der Durchführung von Hackathons zur Entwicklung von Anwendungen und bei der Umsetzung eigener Open-Data-Strategien.

Es ist davon auszugehen, dass Städte weiterhin die Orte sein werden, an denen die größte Nutzung von Open Data stattfindet. Dafür gibt es verschiedene Gründe: Auf kleinem Raum kommen viele Menschen zusammen, die in direktem Kontakt zur Verbesserung der Lebensbedingungen in der Stadtgesellschaft zunehmend auch datenbasierte Ideen und Lösungen ersinnen; die lokale Verankerung durch den Lebensmittelpunkt erzeugt eine direktere Betroffenheit und Motivation; der Kontakt zur kommunalen Ebene ist einfacher; nicht zuletzt können größere Städte eine höhere Datendichte zur Verfügung stellen. Diese Aspekte unterstützen eine allmähliche Vergrößerung des Nutzenspektrums.

ABB. 10 **Modell für die Berechnung des Open-Data-Potentials bezogen auf eine Kommune**

(Preische, 2014, 34)

Zwei Studien aus der „Studienlandschaft" stechen hervor, weil sie das Potential von Open Data auf der Ebene Stadt berechnen: Die Studie von Fuchs et al. (2013) berechnet das Potential für Wien, die von Preische (2014) das Potential für Berlin. Als methodischer Ansatz kommen bei beiden eine bottom-up-Projektion zum Einsatz sowie eine Adaption existierender Studien inklusive ihrer Annahmen. Beide Studien basieren auf der Modellierung des volkswirtschaftlichen Potentials von PSI, wie in der Studie von Pollock (2010) beschrieben. Die errechneten Werte der drei Studien sind in Tabelle 2 zusammengefasst.

Auffallend ist, dass die prozentualen BIP-Anteile sehr unterschiedlich ausfallen. Zwischen Preische und Pollock ist der Faktor ca. 10, zwischen Fuchs und Pollock sogar 100. Eine nähere Betrachtung des verwendeten Modells und der Wahl der Modellparameter gibt Hinweise darauf, wie diese Unterschiede entstanden sind.

Das Modell von Preische (2014) verwendet die in Abbildung 10 aufgeführten Parameterdefinitionen. Die für diese Studie adaptierten Werte sind am Ende der jeweiligen Parameterbeschreibung erklärt:

■ **Nutzen U** stellt den errechneten Nutzen von Open Data dar. Dabei ist zu beachten, dass der Nutzen hier je nach Definition und Annahmen bzgl. der Kosten- und Nutzenmultiplikatoren einen unterschiedlichen Umfang aus semantischer Sicht hat, d. h. die enthaltenen Nutzenarten können zwischen direktem, indirektem und weiterreichendem Nutzen variieren.

■ Die **Investitionskosten F** für Open Data der Stadt umfassen die Kosten der Bereitstellung von Open Data, die für eine Kommune (hier Berlin und Wien) anfallen. Die Überlegung dabei ist, dass diese Kosten nicht direkt, z. B. durch Verkauf der Daten, als Einkommen zurück an die Kommune fließen. Für diesen Basisnutzen (der Nutzer zahlt nichts dafür, schafft aber durch die Verwendung einen Mehrwert) wird nur die Bereitstellung als Open Data berechnet, da die Kosten der Datenermittlung allgemeine Verwaltungsaufgaben sind, die ohnehin anfallen. Die Autoren dieser zwei Studien nehmen an, dass wenn die veröffentlichten Daten auf einem freien Markt bepreist werden könnten, dieser Preis den ermittelten Kosten[32] entspräche. Diese Deutung von F weicht allerdings von dem ursprünglichen Modell von Pollock (2010) ab, da dort F eine Adaption auf Open Data basierend auf tatsächlichen Angaben über das erzielte Einkommen durch den Verkauf von PSI darstellt. Kosten kommen dabei nicht in Betracht, da diese u. U. zu gering ausfallen können und somit nicht den Durchschnittskosten von Open Data entsprechen, sondern lediglich die „Anlaufkosten" für die Bereitstellung einer initialen Datenmenge darstellen. Rückblickend auf den Prozess der Entstehung und Nutzung von Open Data (vgl. Kap. 3.2) ist festzustellen, dass der größere Anteil der Kosten in Zusammenhang mit der Vergrößerung der zu veröffentlichenden Datenmenge steht. Dieser Prozess ist deutlich kostenintensiver (z. B. Personalkosten) im Vergleich zu der bloßen technischen Bereitstellung der Daten über ein (oder mehrere) Datenportale.

■ Da Preische (2014) sowie Fuchs et al. (2013) aus Datenmangel nicht auf erzieltes Einkommen,[33] sondern auf ermittelte Kosten für die Open-Data-Bereitstellung zurückgreifen, entsteht eine Verzerrung, zu deren Ausgleich die Autoren die Einführung eines **Kostenmultiplikators λ1** vorschlagen. Der Nutzen von Open Data wird verzerrt/nur teilweise erfasst, weil bei manchen bereitgestellten Daten die Annahme, dass sie zu einem den Kosten entsprechenden Preis auf einem freien Markt abgesetzt werden können, nicht immer zutrifft. So gibt es Daten (Arbeits-

marktdaten, Daten zur Bevölkerungsentwicklung usw.), die von einem privatwirtschaftlichen Unternehmen nicht mit verhältnismäßigen Kosten erhoben werden können und somit kein freier Markt bzw. kein realistischer Preis auf einem freien Markt existiert. Das bedeutet, dass die imaginäre Bereitschaft von Nachfragern, einen dem Nutzen entsprechenden Preis für bestimmte Daten auf dem Markt zu zahlen – statt die Daten selbst zu erheben – nicht im Faktor F enthalten ist. In dieser Studie wird der originale Faktorwert von λ1=1,5 aus Fuchs et al. (2013) und Preische (2014) übernommen, der sich dort allerdings lediglich auf Expertenschätzungen stützt (vgl. Fuchs et al., 2013).

• Der **direkte Nutzenmultiplikator λ2** kann auf die theoretischen Grundlagen der Wertschöpfung von Open Data und die damit verbundenen Nutzenarten zurückgeführt werden.[34] Die Erwartung ist, dass Unternehmen einen Mehrwert basierend auf Open Data schaffen. Dieser Mehrwert wird allerdings durch die bisher betrachteten Faktoren nicht abgebildet, da „zum einen der Verkauf der Daten häufig an weiterverarbeitende Unternehmen stattfindet, wodurch der wertsteigernde Schritt der Verarbeitung durch Unternehmen und der anschließende Verkauf an Endkunden nicht berücksichtigt wird" und „[z]um anderen weisen konventionelle Nachfragekurven die Eigenschaft auf, statisch zu sein und dadurch mögliche zukünftige Veränderungen zu ignorieren" (Preische, 2014, 36). Eine Nichtbetrachtung der beschriebenen Effekte würde die Modellrechnung verzerren. Daher werden Schätzungen für diesen Nutzenmultiplikator eingeführt, die mit Analogien aus anderen Studien begründet werden.[35] Eine Schätzung auf diesem Weg ist deswegen wichtig und möglich, weil „die unterschiedenen Wirkungen der Umsatzsteigerungen, Effizienzverbesserungen, Steuereinnahmen und Einnahmeausfälle als Folge einer Umsetzung von Open Data gleichzeitig auftreten und in Summe den direkten Nutzeneffekt abbilden. Eine Wechselwirkung der Effekte untereinander ist vermutlich vorhanden, deren Relation aber unbekannt […]" (Preische, 2014, 37). Im Rahmen dieser Studie wird derselbe Schätzwert λ2=10 übernommen.

• Ähnlich λ2 kann der **indirekte Nutzenmultiplikator λ3** auf die theoretischen Grundlagen der Wertschöpfung von Open Data und die damit verbundenen Nutzenarten zurückgeführt werden.[36] Die Überlegung ist, dass mittelbarer Nutzen entsteht, der nicht direkt auf Open Data zurückzuführen ist (z. B. verbesserte

Bildung durch Datenöffnung im Bildungssystem und dadurch höhere Gehälter für die ausgebildeten Personen), aber Open Data trotzdem als ein Treiber solcher Entwicklungen verstanden werden kann.[37] Es ist allerdings unmöglich, eine direkte Beobachtung solcher Prozesse durchzuführen. Noch weniger ist eine direkte Quantifizierung der Auswirkungen möglich. Daher wird bei der Festlegung dieses Multiplikators auf Analogien zurückgegriffen – auf „Your Better Life Index" der OECD[38] sowie auf Web-Index der World Wide Web Foundation.[39] Diese Studie folgt diesen Annahmen und übernimmt $\lambda 3=1{,}2$ für den indirekten Nutzenmultiplikator.

- Die **Nachfrageelastizität ε** in der Nutzenfunktion gibt an, wie die Nachfragemenge vom Preis eines Guts abhängt. Da bei Open Data der Preis null ist, kann eine Elastizität nur anhand von Praxisbeispielen durch Analogien oder Schätzungen ermittelt werden (vgl. Pollock, 2010). Preische (2014) verwendet beispielsweise Angaben über die Entwicklung der Nutzerzahlen bei Wetterdienstdaten in den Niederlanden und Norwegen, die nach Öffnung der Daten ermittelt worden sind. Die Idee ist, dass manche Daten einer breiten Masse von Nutzern erst gar nicht zugänglich sind, solange diese Daten nicht offengelegt werden. Pollock (2010) gründet seine Argumentation auf PSI, die bereits verkauft werden, und die möglichen Auswirkungen, wenn diese als Open Data angeboten würden. Da keine sektorale Betrachtung von Open Data in deutschen Kommunen im Rahmen dieser Studie vorgenommen worden ist, orientiert sich die Berechnung an den Schätzungen und Annahmen von Fuchs et al. (2013) sowie Preische (2014). Analog zu Preische (2014) wird die gleiche Bandbreite verwendet, jedoch mit Gewichtung basierend auf der Berücksichtigung der Nachfragewirkung anhand der Einwohnerzahl einer Kommune. Die Annahme lautet hierfür, dass in Großstädten offene Daten Skaleneffekten haben, d. h. die Elastizität ist in Kommunen mit größerer Einwohnerzahl höher im Vergleich zu Kommunen mit geringerer Einwohnerzahl, bei denen die Elastizität entsprechend niedrig ausfällt. Als Bandbreite für die Elastizität werden in dieser Studie die Werte von Preische (2014) angesetzt: Elastizität von 10 im unteren Bereich des Spektrums, 15 im mittleren Bereich und 25 im oberen Bereich.

Die Unterschiede zwischen den Studien können aus den Definitionen der einzelnen Variablen der Nutzenfunktion wie folgt erklärt werden. Pollock verwendet Daten über reale Verkäufe und Einkommen aus PSI, die er dann auf die Gesamtbedeutung von Open Data überträgt. Aufgrund fehlender Daten können die Studien über Wien und Berlin nicht auf eine einkommensbasierte Berechnung zurückgreifen. Daher versuchen sie, die Schätzung über Kosten und weitere, teilweise neue, Multiplikatoren herzuleiten. Dies verzerrt den berechneten Wert, da die Kosten für Open Data ggf. zu gering angegeben sein könnten, was im Modell bedeuten würde, dass der imaginäre Marktpreis von Open Data zu niedrig ist. Daher versuchen beide Studien durch Multiplikatoren, die bei Pollock mit anderen Werten oder gar nicht eingesetzt werden, eine realistischere Schätzung abzugeben. Allerdings – auch im Vergleich mit den in Kapitel 4.1 genannten Studien – erscheinen die berechneten Potentiale für Wien und Berlin als deutlich zu gering. Dennoch: Methodisch sind die Modelle möglich – unter der Prämisse, dass sehr viel konservativere Schätzungen für das Potential von Open Data herauskommen.

Aufgrund der beschriebenen methodischen Überlegungen und weil Wirtschaft und Zivilgesellschaft in Deutschland nur zu einem kleinen Teil von Open Data durchdrungen sind, wurden von einigen Kommunen Open-Data-Kosten erhoben, um sie nach dieser Methodik in einem bottom-up-Ansatz für eine Potentialabschätzung zu nutzen.

Die Erhebung der Investitionskosten erfolgte in zwei Schritten. Im ersten Schritt wurden Kommunen per E-Mail kontaktiert und um Preisgabe ihrer Daten zu den getätigten Open-Data-Investitionskosten in den Jahren 2013, 2014 sowie 2015 gebeten. In diesem Schritt wurden 83 Kommunen kontaktiert mit dem Ergebnis, dass die Rückmeldungen für eine Potentialhochrechnung nicht ausreichten. Die Gründe sind – neben fehlenden offiziellen Ansprechpartnern – vor allem auf die fehlende Möglichkeit der eindeutigen Zuordnung von Investitionskosten zu Open Data zurückzuführen. Mehrere Kommunen haben mitgeteilt, dass sie keine offizielle Auskunft diesbezüglich geben können.

Um diese Probleme zu umgehen, wurde im zweiten Schritt eine anonyme Online-Umfrage gestartet. Kontaktiert wurden 24 Kommunen, die über ein eigenes funktionierendes Open-Data-Portal verfügen und die Kontaktdaten eines eindeutigen Ansprechpartners im Portal angeben. Es konnten 12 Rückmeldungen gesammelt werden. Diese Rückmeldungen wurden mit den verwertbaren Daten aus dem ersten Schritt – weitere fünf Rückmeldungen – ergänzt, wobei die Anonymität der Antworten gewahrt blieb. Die sehr geringe Zahl von Rückmeldungen – auch nach dem zweiten Schritt – zeigt, in welch frühem Stadium Open-Data-Initiativen auf kommunaler Ebene sind.

Um die Anonymität der Kommunen zu gewährleisten und Flexibilität bei den Einschätzungen zu ermöglichen[40] wurden nur Kostenbandbreiten abgefragt. Für die später in Kapitel 5.3 durchgeführten Berechnungen mussten einige Annahmen getroffen werden:

- Als Datenart wird primär Open Government Data betrachtet. Wie bereits im Kapitel 4.3 diskutiert, stellen die im Modell verwendeten Multiplikatoren Schätzungen dar, die nicht auf das breitere OxD übertragbar sind.
- Es wird angenommen, dass sehr kleine Kommunen kein eigenes Datenportal betreiben werden, sondern sich zusammenschließen.
- Es gibt anfangs eine Adaptionshürde für Pioniere, um zum Beispiel gemeinsam nutzbare Datenportale zu erstellen. D. h. initial muss mit vergleichsweise höheren Kosten gerechnet werden.
- Gleichzeitig können sich Skaleneffekte auswirken, indem
  - größere Verwaltungseinheiten bzw. Kommunen mit vielen Einwohnern leichteren Zugang zu IT-Spezialisten, Open-Data-affinen Mitarbeitern usw. haben; und
  - bei potentieller millionenfacher Nachfrage nach einem Datensatz (z. B. in der Großstadt Berlin) dieser einen höheren Nutzen generiert als derselbe Datensatz in einer kleineren Kommune. Daraus resultiert eine zusätzliche Gewichtung der Einwohnerzahl: für mittlere Städte eine Reduktion auf 1/4 im Vergleich zu großen Städten und für kleinere Städte eine weitere Reduktion auf 1/3 im Vergleich zu mittleren Städten.

Neben der Erhebung von quantitativen Daten mittels der Umfrage wurden Fragen nach der aktuellen Nutzung von Open Data, nach bestehenden Herausforderungen und erkennbaren Handlungsbedarfen gestellt. Die Zusammenfassung der Antworten lautet: Die dominierenden Hemmnisse bei der Bereitstellung und Realisierung von OGD sind vor allem Personalkapazitäten, Kostenfaktoren sowie die Gefahr möglicher Urheberrechtsverletzungen. Im Vergleich zu früheren Umfragen[41] gibt es bei den Faktoren Sicherheitsbedenken, Missbrauch der Daten sowie Fehlinterpretationen einen leichten Rückgang. Dies könnte als positiver Lerneffekt interpretiert werden.

Die Erhebung weiterer Daten wurde aus mehreren Gründen verworfen. Sektorale Daten bzw. Schätzungen zu dem damit verbunden Einkommen würden dem Berechnungsansatz nicht gerecht werden, da nur ein Teil der möglichen Sektoren betrachten werden kann[42] und keine begründeten Annahmen zu der prozentualen Aufteilung der Sektoren im gesamten Potential von Open Data getroffen werden können. Zudem ist es zum Teil strittig, welche Daten tatsächlich den Grundsätzen von Open Data entsprechen könnten bzw. welches Einkommen nur mit Daten in nicht aggregierter Form erzielt worden ist. Die zweite Überlegung ist, die Investitionskosten auf den föderalen Ebenen Bund und Länder zu ermitteln. Allerdings sind die Investitionskosten ganzheitlich betrachtet vernachlässigbar,[43] da sie – vereinfacht ausgedrückt – nur an 17 Stellen anfallen: Ein Mal für den Bund und 16 Mal für die Länder, und zwar nur für die Bereitstellung der Daten. Die Erhebung der Daten, die zu veröffentlichen sind, findet weiterhin zum größten Teil in den Kommunen statt.[44] Die dritte verworfene Überlegung ist, Daten bei Unternehmen zu erheben, die ihr Geschäftsmodell auf Open Data stützen. Allerdings steckt Open Data in Deutschland hier noch viel mehr als in der Verwaltung in den Kinderschuhen, d. h. es gibt vergleichsweise wenige Unternehmen mit einem funktionierenden Geschäftsmodell auf dem Markt[45] und somit sind darauf basierende Schätzungen kaum zu adaptieren.

POTENTIAL

Open Data stellt für Deutschland ein volkswirtschaftliches Potential von 43,1 Mrd Euro p.a. dar – ein ambitionierter, aber gleichzeitig realistischer Wert für einen Zeithorizont von zehn Jahren. Wie nahe Deutschland an das Potential herankommt, hängt entscheidend von der Art und Weise ab, wie die Öffnung von Daten der öffentlichen Verwaltung vorangetrieben wird. Die Umsetzung reaktiver Strategien führt zu einer konservativen Schätzung von 12,1 Mrd. Euro, wohingegen eine weitreichende Mobilisierung und Einbindung der Zivilgesellschaft und der Wirtschaft zu einem optimistischen Szenario und einem Potential von 131,1 Mrd. Euro führen kann.

5 | Projektion des Potentials für Deutschland

5.1 | DIE QUANTIFIZIERUNG DES POTENTIALS

Die Quantifizierung des volkswirtschaftlichen Potentials von Open Data für Deutschland erfolgt mittels Adaption von Berechnungen und Werten existierender Studien (vgl. Kap. 5.2). Die berechneten Werte liegen zum Teil weit auseinander, was einerseits auf die im methodischen Ansatz verwendeten Faktoren zurückzuführen ist. Andererseits ergeben sich Unterschiede aufgrund der Tatsache, dass Deutschland im Vergleich zu anderen Open-Data-Nationen noch deutlich zurückliegt bzw. sich langsamer entwickelt. Die Projektionen liefern insgesamt ein Spektrum: Unterschiedliche Arten von Nutzen (direkt, indirekt, weiterreichend) können nur dann abgerufen werden, wenn Rahmenbedingungen entsprechend gestaltet sind.

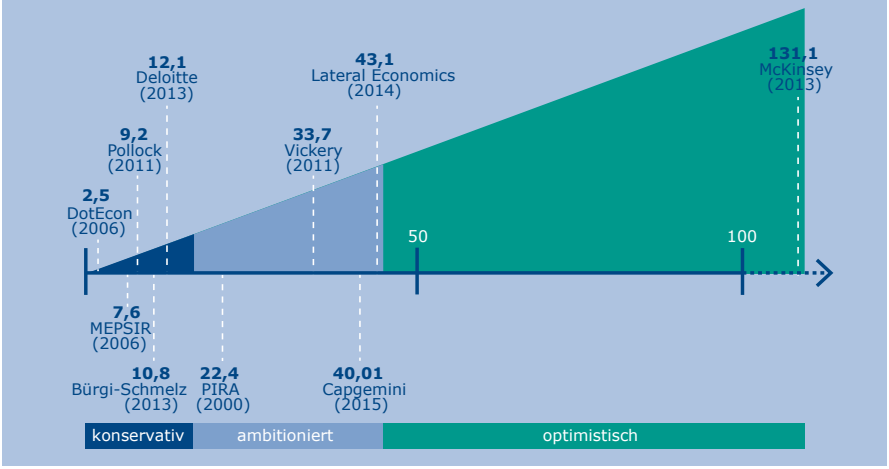
5.2 | KONSERVATIV, AMBITIONIERT ODER OPTIMISTISCH?

Für die Schätzung des volkswirtschaftlichen Potentials von Open Data für Deutschland wurden die in Kapitel 4.2 vorgestellten Studien verwendet. Für die Umsetzung der notwendigen Rahmenbedingungen und die Ausschöpfung des Potentials wurde ein Zeithorizont von zehn Jahren zugrunde gelegt. Die für Deutschland umgerechneten Werte sind als Spektrum dargestellt (vgl. Abbildung 11).

Das aufgespannte Spektrum lässt sich in drei Bereiche gliedern – den konservativen, den ambitionierten und den optimistischen. Jeder Bereich wird durch Werte aus den adaptierten Studien festgelegt, jedoch ist das in den Studien verschieden abgebildete Potential nicht immer direkt vergleichbar. Unterschiede gibt es beim Umfang der Nutzenarten (direkt, indirekt, weiterreichend) oder den berücksichtigten Datenarten (OGD, OBD, OPD oder alle drei: „OxD"). Daher werden die Bereichsgrenzen durch die berechneten Werte, die jeweils getroffenen Annahmen für die Realisierung des Potentials sowie die Methodik der jeweiligen Studie festgelegt.

Der konservative Bereich fasst Studien mit einem für Deutschland adaptierten Wert zwischen 2,5 Mrd. EUR und 12,1 Mrd. EUR p.a. zusammen. Diese Studien sind durch folgende Aspekte charakterisiert: (a) Sie treffen Annahmen für die

ABB. 11 **Berechnungsspektrum des Potentials von Open Data für Deutschland in (Mrd. EUR p.a.)**



Weiterentwicklung des Open Data-Potentials, fokussieren aber auf die kommerzielle „Verkaufbarkeit von PSI", die im Fall von Open Data so nicht gegeben ist; (b) sie fokussieren eng auf OGD als Datenart und betrachten nicht OxD; (c) es werden eher direkte und indirekte Nutzen quantitativ erfasst. Insgesamt bedeutet das: Die Annahmen basieren auf der Umsetzung einer reaktiven Strategie, d.h. bisherige Entwicklungen in der öffentlichen Hand werden betrachtet (einschlägige Rechtsnormen, politischer Wille usw.) und inkrementelle Verbesserungen vorgeschlagen. (d) Die angewendete Methodik ist überwiegend auf einen bottom-up-Ansatz zurückzuführen.

Als konservative Schätzung für das volkswirtschaftliche Potential von Open Data für Deutschland sind 12,1 Mrd. EUR p.a. realistisch. Die adaptierte Studie von Deloitte (2013) verwendet eine bottom-up-Methodik und eine Kombination aus Primär- und Sekundärdaten zur Berechnung des Potentials für Großbritannien. Die Wahl dieser Studie beruht auf folgenden Argumenten. Die Studie von Bürgi-Schmelz (2013) für die Schweiz orientiert sich stärker an Sekundärdaten als Vickery (2011) und trifft Annahmen bei der Wertermittlung über Zeitersparnisse, die sich nicht direkt auf Deutschland übertragen lassen[46]. Die Studie von Pollock (2010) basiert auf Daten über das Einkommen durch den Verkauf von PSI in Großbritannien und erfasst deshalb nicht vollumfänglich den direkten und indirek-

ten Nutzen durch kostenfreie Open Data. Daher wurden Studien, deren Schätzungen unterhalb der Werte von Pollock liegen (vgl. Abbildung 11) nicht miteinbezogen.

Was bedeutet ein Wert von 12,1 Mrd. EUR p. a. aus dem konservativen Bereich für Deutschland? Dieser Wert kann in einem Zeitraum von zehn Jahren realisiert werden, wenn die Rahmenbedingungen dafür geschaffen werden. Deutschland hat bereits erste Schritte gemacht – etwa den Nationalen Aktionsplan der Bundesregierung[47] – und müsste dieser Entwicklung weiter folgen, z. B. durch Umsetzung der in der Open-Government-Partnerschaft[48] gesetzten Ziele. Dieser Zielkorridor würde bedeuten, dass Deutschland eine reaktive Strategie des Aufholens zu Open-Data-Vorreitern wie Großbritannien, USA oder auch Taiwan verfolgt. Eine Vorreiterrolle ist mit dieser Strategie nicht erreichbar.

Der ambitionierte Bereich fasst Studien mit einem adaptierten Wert zwischen 22,4 Mrd. EUR und 43,1 Mrd. EUR zusammen. Die Studien in diesem Bereich treffen unterschiedliche Annahmen für die Weiterentwicklung des Open-Data-Potentials. Die PIRA-Studie (2000) kann als Pionier bezeichnet werden, da sie sich sehr früh mit der Potentialschätzung von PSI für die EU beschäftigt hat. Es handelt sich um eine Studie, die nur OGD fokussiert, eine eher top-down Methodik und Sekundärdaten verwendet und somit an der Grenze zwischen konservativ und ambitioniert liegt.[49] Die Studien von Vickery (2011) und Capgemini (2015) über die EU fokussieren stärker OGD, während die Studie von Lateral Economics (2014) über Australien ihren Fokus breit auf OxD setzt. In diesen Studien werden alle möglichen Nutzenarten adressiert. Die Festlegung der Spektren in den einzelnen Studien deutet darauf hin, dass an den unteren Grenzen jeweils eher der direkte Nutzen, bei den oberen Grenzen aber auch weiterreichender Nutzen berücksichtigt wurde. Für das adaptierte Berechnungsspektrum wurden Werte berücksichtigt, die sämtliche Nutzenarten – direkt, indirekt und weiterreichend – umfassen. Die Studie von Vickery (2011) fokussiert PSI, während die Studien von Lateral Economics (2014) und Capgemini (2015) Open Data adressieren. Die verwendete Methodik aller Studien in diesem Bereich ist überwiegend top-down.

Als ambitionierte Schätzung für das volkswirtschaftliche Potential von Open Data für Deutschland sind 43,1 Mrd. EUR p. a. als realistisch einzuschätzen. Dieser

Wert wird anhand der Studie von Lateral Economics (2014) adaptiert. Die Studie berechnet mit Hilfe einer top-down-Methodik und einer Kombination aus Primär- und Sekundärdaten das Potential von Open Data für Australien. Dabei verwendet sie ausgewählte Fallstudien – z. B. über sektorale Daten wie Geo-Daten und das damit verbundene volkswirtschaftliche Potential – sowie Annahmen über die Bedeutung und Entwicklung von Open Data in sämtlichen gesellschaftlichen Bereichen wie in der Studie von McKinsey (2013) beschrieben. Die Wahl dieser Studie für die Adaption eines realistischen Wertes beruht auf folgenden Argumenten:

Im Vergleich zu den anderen Studien in diesem Bereich bezieht sich die Lateral-Studie auf OxD und berücksichtigt damit als einzige das zusätzliche Potential von OBD und OPD. Die Studie relativiert die in McKinsey (2013) berechneten Werte durch Analogien mit sektoralen Daten und bildet diese Werte auf die Praxis in Australien ab. Dagegen fokussieren PIRA (2000), Vickery (2011) und Capgemini (2015) vornehmlich OGD. Zudem ist eine eindeutige Verkettung der berechneten Potentiale dieser Studien zu erkennen, d. h. aus der PIRA-Studie (2000) werden die Ergebnisse in Vickery (2011) adaptiert, die wiederum in Capgemini (2015)[50] adaptiert und dort auf einen Zeitraum von fünf Jahren berechnet werden, um auf einen kumulierten Anteil am BIP zu gelangen. Daraus folgt, dass die Lateral-Studie näher an das realistische Potential von OxD heranreicht.

Was bedeutet ein Wert von 43,1 Mrd. EUR p. a. aus dem ambitionierten Bereich für Deutschland? Der ambitionierte Wert[51] ist in zehn Jahren realisierbar, wenn in diesem Zeitraum die dafür notwendigen Rahmenbedingungen geschaffen werden. Von den Entscheidungsträgern verlangt dies, eine proaktive Strategie des Gleichziehens zu verfolgen. Dafür hat Deutschland den Anschluss an die erwähnten Spitzengruppen in den nächsten fünf Jahren zu erreichen, um spätestens dann die Generierung von Nutzen nicht nur durch OGD, sondern auch durch OxD zu ermöglichen. Dazu gehören die vollständige Umsetzung der bisher gesetzten Ziele [52] und eines Zielbildes, das durch „open by default" charakterisiert werden kann: Anders als heute wird dann das Nichtveröffentlichen von Daten im Einzelfall öffentlich zu begründen sein. Dafür ist ein ambitionierter Plan aufzustellen, der politisch getragen und operativ konsequent umgesetzt wird. Das Einbeziehen von OBD und OPD in die Ausschöpfung des Open-Data-Potentials schließt weiter ein,

dass Rahmenbedingungen nicht nur für die öffentliche Hand, sondern auch für die Wirtschaft und die Zivilgesellschaft zu schaffen sind. So sind einerseits Anreize für die freiwillige Veröffentlichung von Daten zu schaffen, andererseits ist die Öffnung bestimmter Daten in relevanten Sektoren wie Verkehr, Umwelt u. a. an gesetzliche Vorgaben zu binden.

Der optimistische Bereich ist durch eine einzige Studie mit einem für Deutschland adaptierten Wert von 131,1 Mrd. EUR repräsentiert. Mit einem geschätzten BIP-Anteil von 4,26 bis 7,19 Prozent kann die McKinsey Studie (2013) als Ausreißer bezeichnet werden. Allerdings heißt das nicht zwingend, dass die berechneten Werte unrealistisch sind, wie die Lateral-Studie argumentiert: Zum einen umfasst das berechnete Potential nicht nur OGD, sondern den gesamten OxD-Raum – und geht auch von der Annahme aus, dass die Rahmenbedingungen sich weiterentwickeln und diese Potentialausschöpfung ermöglichen. Außerdem umfasst die Studie neben direktem und indirektem auch den weiterreichenden Nutzen von Open Data.

Als optimistische Schätzung für das volkswirtschaftliche Potential von Open Data für Deutschland sind 131,1 Mrd. EUR p. a. ein zu realisierender Wert. Er wird adaptiert, in dem wir die untere Grenze von 4,26 Prozent BIP-Anteil der McKinsey-Studie ansetzen. Da die Studie aus methodischer Sicht als top-down bezeichnet werden kann und auf Expertenschätzungen basiert,[53] erscheint nur diese untere Grenze als eine realistische Schätzung. Zudem gibt die Studie Hinweise darauf, dass das Potential zwischen einzelnen Weltteilen unterschiedlich aufzuteilen ist, was unterschiedliche Annahmen für die Ausgestaltung der Rahmenbedingungen vermuten lässt. Das größte Potential wird den USA zugeteilt (McKinsey, 2013, 6).

Was bedeutet ein Wert von 131,1 Mrd. EUR p. a. aus dem optimistischen Bereich für Deutschland? Der für Deutschland adaptierte Wert bedeutet,[54] dass Deutschland sich selbst in Bezug auf ein nationales digitales offenes Ecosystem herausfordert und durchaus auch den Anspruch erhebt, die führende Open-Data-Nation zu werden. Solch ein Plan ginge weit über die derzeit diskutierte Digitale Transformation hinaus. Es ginge darum, Open Data als Kernelement einer umfassenden nationalen Strategie für das 21. Jahrhundert zu positionieren und sein Potential

für nicht-digitale gesellschaftliche Herausforderungen in den nächsten zehn Jahren zu mobilisieren. Welche Rolle Open Data in Politikfeldern spielen kann, die für die nachhaltige Entwicklung des Landes in den kommenden Dekaden von Bedeutung sind (Klima, Migration, Energieversorgung, Bildung, Gesundheits- oder Transportwesen) ist eine wichtige strategische Frage in diesem Szenario.

5.3 | KOMMUNALE OPEN-DATA-KOSTEN ALS BERECHNUNGSBASIS

In Ergänzung wird eine Hochrechnung präsentiert, die auf der in Kapitel 4.3 verwendeten Methode von „Open Data-Investitionskosten von Kommunen" beruht, wie sie in zwei untersuchten Studien (Wien und Berlin) durchgeführt wurde. Mit Hilfe der gesammelten Daten aus der Online-Umfrage (vgl. Kap. 4.3) und unter Berücksichtigung der dort getroffenen Annahmen liegen der Berechnung die in Tabelle 3 notierten Daten zugrunde. Das mit diese Methode berechnete Potential bewegt sich im Bereich von 0,68 Mrd. bis 1,69 Mrd. EUR, was ca. 0,02 bis 0,05 Prozent des BIP entspricht.

Vergleicht man diese Berechnung unter Berücksichtigung der Anpassungsfaktoren für Großstädte und kleinere Kommunen mit den anderen drei Studien (Fuchs et al., Preische, Pollock), die diesen Modellierungsansatz verfolgen (vgl. Kap. 5.3), ergibt die Hochrechnung für Deutschland Werte, die in etwa einer Hochrechnung der Stadt Berlin (Preische 2014) entsprechen.

Im Vergleich zu den Potentialwerten aus Kapitel 5.2 ist dies ein sehr großer Ausreißer nach unten. Er überrascht jedoch nicht, weil die Methodik – wie in Kapitel 4.3 bereits ausgeführt – zu konservativen Schätzungen tendiert, wenn nicht Vollkosten für die Berechnung angesetzt werden. Die gemeldeten Kostendaten aus der Umfrage sind – trotz Anonymität – jedoch keine Vollkosten, sondern vornehmlich die leicht(er) messbaren Teilkosten, die aus dem Bau und Betrieb des Datenportals und ggf. flankierenden Maßnahmen (z. B. Hackathons) resultieren. POPSIS (2011) untersucht Berechnungsmodelle für PSI und rät davon ab, eine kostenneutrale Bepreisung vorzunehmen, weil eine zuverlässige Zuordnung von Kosten nicht hinreichend möglich ist. Dieses Zuordnungsproblem wird bei dieser Methodik auf verschiedene Weise deutlich. Erstens fehlen Personalkosten, die nicht einfach zugeordnet werden können: Es gibt derzeit nur wenige Kommunen,

TAB. 3 **Mengengerüst für die Hochrechnung auf Basis Kommunen**

Kennzahlen Anzahl Einwohner	Anzahl Kommunen[55]	Durchschnittliche Investitionskosten p. a.	Einwohnerzahl total	Einwohnerzahl prozentual
mehr als 500 Tsd.	*12*	*68,0 Tsd. EUR*	*5,74 Mio.*	*7,09 %*
100 – 500 Tsd.	*64*	*36,0 Tsd. EUR*	*11,91 Mio.*	*14,71 %*
weniger als 100 Tsd.	*11.237*	*2,5 Tsd. EUR*	*63,40 Mio.*	*78,27 %*

die bereits dediziertes Personal für Open Data eingestellt haben, das Thema ist zum Beispiel bei eGovernment angesiedelt und wird dort mitbearbeitet. Die anteiligen Personalkosten gehen also nicht in die Investitionskosten ein. Zweitens ist der Publikationsprozess selbst ein schwer zuordenbarer Kostenblock: Die Arbeitsaufwände für Identifikation, Auswahl, Extraktion, Konversion und Publikation eines Datensatzes verteilen sich meist auf unterschiedliche Personen in verschiedenen Abteilungen. Drittens sind die Kosten für die ursprüngliche Erhebung der Daten, die auch eingerechnet werden müssten, sehr schwer zu eruieren und zuzuordnen. Was hat es beispielsweise gekostet, die Daten für den Haushalt, die Feinstaubwerte oder den Zustand der Radwege zu erfassen? Diese Kosten können beträchtlich sein und würden – nach dieser Methodik – das berechnete Potential deutlich nach oben verschieben. Für Open Data spräche natürlich das Gegenteil: Sawyer/de Vries (2012) stellen zum Beispiel dar, dass der Nutzen der Freigabe von Geo-Daten von Satellitenfotos die (sehr hohen) Kosten klar übersteigt. Nutzen durch Open Data entsteht nicht durch eine Erhöhung der Investitionskosten, sondern dadurch, dass die Investitionen effektiv und effizient eingesetzt werden. Deswegen wird dieser Ansatz nicht weiter verfolgt.

Auf alle Argumente wirkt sich zusätzlich der tiefe Reifegrad der Open-Data-Aktivitäten aus. Solange Daten nicht automatisch aus „open-by-default"-Geschäftsprozessen extrahiert werden können (vgl. Kapitel 3.2), fällt der Aufwand zusätzlich zum Tagesgeschäft an und kann unter Umständen erheblich sein: Kosten für Abstimmung, Kommunikation und technische Zusatzarbeiten fallen für jeden Datensatz neu an, werden aber kostenmäßig nicht erfasst.

6 | Kritische Reflektion und Fazit

Auch wenn die Frage nach „Zahlen" für ein Potential verständlich ist und einen wichtigen Schritt darstellt, so ist diese Perspektive für sich allein genommen zu verengt, denn nicht alles wird messbar sein. Um nicht nur das ökonomische, sondern das disruptive Potential von Open Data ganzheitlich bewerten zu können, ist eine breitere Sicht – im Sinne eines nicht-quantifizierbaren Nutzens für Zivilgesellschaft, Wirtschaft und öffentliche Hand – notwendig.

Selbst wenn alle Vollkosten in geeigneter Weise verfügbar wären – was unrealistisch erscheint – bliebe die methodische Konsequenz, dass die Substitution von Einnahmen (Pollock 2010) durch „Kosten plus Multiplikatoren" (Preische (2014), Fuchs et al. (2013)) zu konservativ-verzerrten Hochrechnungen führt.

Deshalb erfolgte die Quantifizierung des volkswirtschaftlichen Potentials von Open Data für Deutschland mittels Adaption von Berechnungen und Werten existierender Studien. Open Data stellt für Deutschland ein volkswirtschaftliches Potential von 43,1 Mrd Euro p.a. dar – ein ambitionierter, aber gleichzeitig realistischer Wert für einen Zeithorizont von zehn Jahren. Wie nahe Deutschland an das Potential herankommt, hängt entschieden von der Art und Weise ab, wie die Öffnung von Daten der öffentlichen Verwaltung vorangetrieben wird.

Die Annahmen und Projektionen in dieser Studie über die mögliche Entwicklung von Open Data in Deutschland hängen von folgenden Faktoren ab:

- Es ist nicht sicher, dass die massive Ausweitung des Bestands an offenen Daten tatsächlich die Nutzerzahlen und den Aufbau neuer Dienste in gleichem Maße steigern wird. Die Verarbeitung von fachlichen Infrastrukturdaten erfordert neben technisch-mathematischem Know-how auch Fachwissen, das unter Umständen auch außerhalb der Verwaltung nicht weit verbreitet ist.

- Das Öffnen von Daten bedeutet für die Verwaltung einen Zusatzaufwand an Personal und finanziellen Mitteln. Die Knappheit der Finanzressourcen (vor allem auf kommunaler Ebene) ist bereits heute ein Thema, und die Bewältigung des Migrationsstroms stellt aktuell erhöhte Anforderungen an das Personal auf allen Ebenen. Außerdem ist die Menge an einschlägigem Know-how zu Daten-

Management meist nur in den IT-Abteilungen vorhanden, die sich nicht mehr viele Behörden inhouse leisten können und deren Fachkräfte auch am freien Markt stark nachgefragt sind.

- Große Mengen von veröffentlichten Daten könnten durch unvorhergesehene Möglichkeiten der Verknüpfung zu ungewollten Analysen führen.

Die Entwicklung der Open-Data-Landschaft zu einem offenen Ecosystem benötigt daher das Setzen klarer politischer Weichenstellungen, was in Ländern wie den USA oder UK bereits vor einigen Jahren geschah. Insofern könnte sich eine nationale Open-Data-Strategie an vielen bewährten Ideen und Ansätzen für die Umsetzung an diesen Ländern orientieren.

Anmerkungen

1 | Basierend auf der Adaption für Deutschland der Capgemini Studie (2015).

2 | Based on Capgemini (2015).

3 | Vgl. http://www.opengovpartnership.org/countries, Gründungsmitglieder u. a. Großbritannien, Vereinigte Staaten von Amerika, Norwegen.

4 | Vgl. https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex

5 | Vgl. http://www.bmi.bund.de/SharedDocs/Downloads/DE/Broschueren/2014/aktionsplan-open-data.pdf

6 | Vgl. https://public.resource.org/8_principles.html, vom 18.12.2015.

7 | Vgl. http://www.whitehouse.gov/the_press_office/Transparency_and_Open_Government/, vom 18.12.2015.

8 | Vgl. http://sunlightfoundation.com/policy/documents/ten-open-data-principles/, vom 18.12.2015.

9 | Vgl. http://www.opengovpartnership.org/, vom 18.12.2015. Obwohl gewünscht und geplant, ist Deutschland bis dato kein Mitglied dieser Partnerschaft.

10 | Vgl. https://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-

11 | Vgl. https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex, vom 18.12.2015.

12 | Vgl. Steuerungsprojekt „Open Government" des IT-Planungsrates, https://www.govdata.de/documents/10156/18448/Eckpunkte_Open_Government.pdf

13 | Vgl. http://opendatahandbook.org/guide/de/what-is-open-data/, vom 18.12.2015. Diese Definition findet eine breite Verwendung in Wissenschaft und Praxis. In der Literatur sind allerdings mehrere Definitionen von Open Data zu finden (u. a. (Huber/Kurnikowski 2013), (BMI 2012), (Davies 2010), (von Lucke/Geiger 2010)).

14 | Mehr zum Datenspektrum-Konzept: http://theodi.org/blog/closed-shared-open-data-whats-in-a-name, vom 18.12.2015.

15 | Es gibt verschiedene Initiativen in Europa: (a) UK Department for Business, Innovation & Skills, report „The midata innovation opportunity"; (b) Finland Ministry of transport and communications, MyData – A Nordic Model for human-centered personal data management and processing, (c) das bereits gestartete midata.coop Projekt in der Schweiz.

16 | Für eine ausführliche Darstellung zur Anwendung der Methodik auf Open Data vgl. Dapp/Stürmer (2015).

17 | Vgl. http://opendata500.com/, letzter Zugriff vom 18.12.2015.

18 | http://www.opendatanow.com/2013/10/new-whats-the-value-of-open-data vom 18.12.2015.

19 | http://www.nytimes.com/2008/07/07/business/media/07weather.html vom 18.12.2015.

20 | http://www.ppi.noaa.gov/wp-content/uploads/PPI-Weather-Econ-Stats-10-13-11.pdf vom 18.12.2015.

21 | http://www.scientificamerican.com/article/how-to-kick-start-innovaton/ vom 18.12.2015.

22 | Marktwert im Januar 2016 siehe http://finance.yahoo.com/q?s=GRMN vom 18.12.2015.

23 | http://www.economist.com/news/business/21578084-making-official-data-public-could-spur-lots-innovation-new-goldmine vom 18.12.2015.

24 | Open Data Institute, Who owns our data infrastructure? May 2015, http://theodi.org/who-owns-our-data-infrastructure, vom 10.02.2016.

25 | Zu ähnlichen Erkenntnissen kommen auch u. a. Deloitte (2013) und Preische (2014).

26 | Dies ist eine idealtypische Unterscheidung, die der einfachen Veranschaulichung dienen soll und keinen Anspruch auf Vollständigkeit erhebt. Studien werden demnach einem der beiden Typen zugeordnet, wobei sie praktisch eine Mischung aus beiden Idealtypen methodisch verfolgen.

27 | Auch hier bildet die Unterscheidung eine Idealtypisierung, d. h. die Studien verwenden praktisch immer Primär- und Sekundärdaten als Datenbasis. Daher gilt auch hier – ähnlich wie bei der Unterscheidung nach eingesetzter Methodik – die Typisierung soll einer Veranschaulichung dienen, hat keinen Anspruch auf Vollständigkeit und gibt lediglich eine Schätzung darüber ab, ob eine Studie eher auf selbst erhobenen Daten oder auf Daten anderer Studien basiert.

28 | Eine Umrechnung in EUR zum aktuellen Tauschkurs, findet im Rahmen der Adaption der Werte bei der Potentialrechnung für Deutschland statt.

29 | Ein Beispiel hierfür sind Annahmen, die auf der Entwicklungsgeschwindigkeit von Open Data-Pionieren wie USA und Großbritannien basieren. Wir nehmen an, dass eine solche Entwicklung und eine solche Durchdringung von Open Data in die öffentliche Hand, Zivilgesell-

schaft und die Wirtschaft in einem Zeitspektrum unter 10 Jahren – auch mit optimistischen Schätzungen – nicht zu erreichen ist.

30 | Für eine detaillierte Argumentation bzgl. des Umgangs mit den in der McKinsey Studie (2013) berechneten Werten vgl. die Studie von Lateral Economics (2014) sowie die kritischen Kommentare bei der Vorstellung der berechneten Ergebnisse der McKinsey Studie bei der Regierung in Kanada http://www.parl.gc.ca/content/hoc/Committee/412/OGGO/Reports/RP6670517/oggorp05/oggorp05-e.pdf, vom 18.12.2015.

31 | Als Analogie hierzu sollen z. B. die Berechnung des Potentials von IT und die damit verbundenen methodischen Schwierigkeiten herangezogen werden – ein Phänomen mit einer Historie von mehreren Jahrzehnten. Obwohl IT heutzutage nicht mehr aus unserem Alltag wegzudenken scheint, ist eine Potentialschätzung nur bedingt möglich.

32 | Die Kosten werden mit den Durchschnittskosten für die Bereitstellung von Open Data gleichgesetzt und der zu erzielende Preis soll lediglich diese abdecken, da die öffentliche Hand nicht gewinnorientiert, sondern kostendeckend diesbezüglich handeln sollte. Eine detaillierte Diskussion dieser Überlegungen ist in Newbery et al. (2008) zu finden.

33 | Pollock (2010) berechnet das Potential basierend auf dem ermittelten Einkommen durch PSI.

34 | Vgl. Kapitel 3; An dieser Stelle ist allerdings anzumerken, dass die im Modell von Preische (2014) verwendeten Begriffe von unserer Unterscheidung nach Nutzenarten abweichen. Unter dem direkten Nutzenmultiplikator werden im Sinne dieser Studie der direkte sowie der indirekte Nutzen zusammengefasst.

35 | Einfluss auf den Multiplikator haben z. B. die Einnahmeausfälle der Verwaltung, entstehende Steuereinnahmen, Umsatzsteigerungen bei Nutzern und Anwendern sowie Effizienzverbesserungen; vgl. hierzu Preische (2014, 36).

36 | Vgl. Kapitel 3; unter diesem Multiplikator erfasst Preische (2014) den weiterreichenden Nutzen.

37 | Z. B. Förderung der Demokratie, Transparenz, Wohlfahrtseffekte; vgl. für weitere Beispiele (McKinsey 2013), (Deloitte 2013) sowie (Preische 2014).

38 | „analysiert das alltägliche Leben von Menschen in verschiedenen Staaten" (Preische 2014, S. 38).

39 | „ein multi-dimensionales Maß für die Nützlichkeit und Auswirkungen des Webs auf Menschen und Nationen sowie für die Nutzung des Webs" (Preische 2014, 39).

40 | Aus dem ersten Datenerhebungsschritt wurde die Erkenntnis gewonnen, dass genaue Werte – aufgrund der schwer exakt berechenbaren bzw. zurechenbaren Kosten – nicht ermittelt werden können bzw. nicht sinnvoll miteinander verglichen werden können.

41 | http://www.open-data-studie.de/OGD-Studie2012_web.pdf, vom 18.12.2015.

42 | So greift z. B. Pollock (2010) auf eine Vielzahl von sektoralen Daten und dem daraus resultierenden Einkommen zurück, um seine Potentialhochrechnung durchzuführen.

43 | Die Kosten für den Aufbau und Betrieb von GovData.de haben im Vergleich zu den kumulierten Kosten aller Kommunen in Deutschland nur eine marginale Bedeutung, obwohl eine zentrale Stelle sämtlicher offener Daten die Entwicklung von Anwendungen deutlich vereinfachen sollte.

44 | Es stimmt zwar, dass es sicherlich wichtige und nützliche Daten gibt, die nur auf Bundes- oder Landesebene erhoben und veröffentlicht werden. Allerdings ist es nach heutigem Stand so, dass – ähnlich wie bei Kommunen – die damit verbundenen Kosten nicht zurechenbar sind und keinesfalls mit den Kosten für den Aufbau eines der Länderportale oder des Bundesportals verglichen werden können.

45 | Nach eigener Recherche und durch Experteninterviews mit der Technologiestiftung Berlin.

46 | In der Studie werden – ohne weitere Details – Ersparnisse von wenigen Minuten pro Tag für jeden einzelnen Bürger durch die Öffnung von Daten angenommen. Da die in der Studie verfolgte Argumentationskette nicht direkt auf Deutschland übertragbar ist, sind auch die angenommenen Zeitersparnisse kaum übertragbar.

47 | Vgl. Umsetzung der Open-Data-Charta der G8, https://www.bmi.bund.de/SharedDocs/Downloads/DE/Broschueren/2014/aktionsplan-open-data.pdf?__blob=publicationFile, vom 18.12.2015.

48 | Vgl. http://www.opengovpartnership.org/, vom 18.12.2015.

49 | Vgl. hierzu (te Velde 2009). Die PIRA Studie(2000) wird in einigen Quellen als eher ambitioniert oder gar optimistisch eingeschätzt, allerdings ist der Bezug PSI bzw. OGD. Zudem liegt die Vickery-Studie (2011) eine Abschätzung vor, die u. a. auf der PIRA-Studie(2000) basiert.

50 | Die Adaption der Capgemini-Studie auf Deutschland führt zu 20.000 ausgewiesenen Open Data-Jobs.

51 | Eine detaillierte Veranschaulichung der Bedeutung dieses Werts ist Anhang B.3 zu entnehmen.

52 | Vgl. hierzu den Nationalen Aktionsplan der Bundesregierung zu Umsetzung der Open Data Charta.

53 | Zu den verwendeten Datenquellen und dem Rechenweg sind keine detaillierten Angaben in der Studie zu finden.

54 | Zur Bedeutung dieses Werts siehe Anhang B.

55 | Es existieren insgesamt 11.313 Kommunen in Deutschland, Stand 31.08.2015, Quelle: http://www.gemeindeverzeichnis.de/dtland/dtland.htm, vom 18.12.2015.

Verzeichnis

Abbildung 1 | Open Data als Treiber des gesellschaftlichen Wandels 11

Abbildung 2 | Daten als Grundlage von Wissen
(Rehäuser/Krcmar (1996,6) und Krcmar (2015, 11)) 21

Abbildung 3 | Theorie des Wandels von Open Data
gemäß SROI (nach Dapp/Stürmer, 2015) 26

Abbildung 4 | Publikation von Open Data in nicht-integrierter Form 28

Abbildung 5 | Publikation von Open Data in einem offenen,
integrierten Prozess-Design 29

Abbildung 6 | Ökonomische Effekte durch Datenöffnung
im Zeitverlauf (Quelle: Preische (2014)) 35

Abbildung 7 | Auswahlschema der Studienselektion 39

Abbildung 8 | Ausgewählte Studien für die Potentialschätzung 40

Abbildung 9 | Übersicht der Studien nach Methodik und Nutzenumfang 41

Abbildung 10 | Modell für die Berechnung des Open-Data-Potentials
bezogen auf eine Kommune (Preische, 2014, 34) 47

Abbildung 11 | Berechnungsspektrum des Potentials von Open Data
für Deutschland 56

Tabelle 1 | Kennwerte der Potentialschätzungen
der ausgewählten Studien 43

Tabelle 2 | Kennwerte der Potentialschätzungen
der Studien auf Stadtebene 46

Tabelle 3 | Mengengerüst für die Hochrechnung
auf Basis Kommunen 61

Wir danken den folgenden Personen für Interviewzeit, inhaltlichen Input oder anderweitig hilfreiche Kommentare: Jan-Ole Beyer, Marco Brunzel, André Golliez, Christian Heise, Christian Horn, Martin Kaltenböck, Max Kasy, Dietmar Offenhuber, Rufus Pollock, Benjamin Seibel, Nicolas Zimmer.

Desweiteren bedanken wir uns bei den Vertretern jener deutschen Open-Data-Kommunen, die in der anonymen Umfrage Kostendaten geteilt haben. Nicht zuletzt danken wir Janine Rosenbaum für die wertvolle Unterstützung bei den Recherchearbeiten.

Ansprechpartner in der Konrad-Adenauer-Stiftung e.V.

Tobias Wangermann
Leiter Projektteam Digitalisierung
Hauptabteilung Politik und Beratung
Postanschrift: Konrad-Adenauer-Stiftung e.V., 10907 Berlin
Telefon: +49 30 26996-3380
E-Mail: tobias.wangermann@kas.de

Dr. Pencho Kuzev
Projektmitarbeiter, Projektteam Digitalisierung
Hauptabteilung Politik und Beratung
Postanschrift: Konrad-Adenauer-Stiftung e.V., 10907 Berlin
Telefon: +49 30 26996-3247
E-Mail: pencho.kuzev@kas.de

Konrad
Adenauer
Stiftung

www.kas.de