Douglas Cunningham, Petra Hofstedt, Klaus Meer, Ingo Schmitt (Hrsg.): INFORMATIK 2015
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2015

# Ein Rahmenwerk für Datenschutz-Metriken in der Cloud

Sebastian Luhn[1], Maxmilian Hils[2]

**Abstract:** Eine automatisierte Überprüfung der Erfüllung von Datenschutzanforderungen in Cloudsystemen kann Kunden von Cloudanbietern helfen, notwendigen Kontrollpflichten einfacher nachzukommen und bietet zudem die Möglichkeit, diese Überprüfungen kontinuierlich zu vollführen. Zu diesem Zweck können Datenschutzmetriken verwendet werden, die anhand einer Vielzahl von Messungen die Eigenschaften des Cloudsystems bezüglich der Anforderungen bewerten. Diese Ausarbeitung stellt ein Rahmenwerk vor, mit dem solche Metriken konstruiert werden können. Es ist theoretisch in der Messtheorie fundiert und nutzt zudem das Konzept des Konstrukts aus den Sozialwissenschaften, um die oft komplexen und nicht direkt messbaren Anforderungen des Datenschutzes messbar zu machen. Außerdem berücksichtigt es domänenspezifische Anforderungen, die sich aus der Nutzung von Metriken in IT-Unternehmen ergeben. Am Beispiel einer Standortmetrik wird gezeigt, dass das Rahmenwerk zur Konstruktion komplexer Metriken genutzt werden kann, die Aussagen über die Erfüllung von Datenschutzanforderungen ermöglichen.

**Keywords:** Automatisierte Kontrolle, Datenschutz, Messquellen, Metrik, Verarbeitungsstandort.

## 1 Einleitung

Die Möglichkeit für Unternehmen, eigene IT-Infrastrukturen in die Cloud auszulagern, erfreut sich steigender Beliebtheit. Dennoch gibt es nach wie vor Bedenken und Hinderungsgründe, die gegen den Wechsel von internen zu externen IT-Infrastrukturen sprechen. Ein wesentlicher Aspekt sind Datenschutzbedenken und Anforderungen, die sich aus dem Datenschutzrecht ergeben. So müssen sich potenzielle Cloud-Kunden auf die Angaben des Cloud-Anbieters verlassen, was etwa den Speicherort der Daten angeht. Außerdem kann dieser häufig auch nicht während des Betriebs der Cloud überwacht werden, da viele Cloud-Anbieter keine (überprüfbaren) Angaben zu Speicherorten machen[3]. Eine Auditierung des Anbieters zur Überprüfung dieser Angaben ist zwar möglich und vorgeschrieben, findet aber üblicherweise nur punktuell statt [Bo12]. Eine kontinuierliche Überprüfung ist damit nicht möglich.

In diesem Artikel wird ein Rahmenwerk präsentiert, das die automatisierte Überprüfung der Einhaltung von Datenschutzanforderungen bei Cloud-Anbietern ermöglicht. Dies geschieht durch die Konstruktion von Datenschutzmetriken, die eine Vielzahl von Datenquellen aggregieren, um Aussagen über die Einhaltung von Datenschutzanforderungen

[1] Forschungsgruppe IT-Sicherheit, Universität Münster, Leonardo-Campus 3, 48149 Münster, sebastian.luhn@wi.uni-muenster.de
[2] Forschungsgruppe IT-Sicherheit, Universität Münster, Leonardo-Campus 3, 48149 Münster, maximilian.hils@uni-muenster.de

[3] Google bietet z. B. die Möglichkeit, den Speicherort bestimmter Daten zu spezifizieren. Ob die Daten tatsächlich dort liegen, kann nicht überprüft werden. Siehe `https://cloud.google.com/storage/docs/bucket-locations`

geben zu können. Metriken werden bereits, unabhängig von Cloud-Anwendungen, häufig in IT-Unternehmen eingesetzt [So11]. In der Literatur werden diese Art von Metriken aber sehr praxisnah behandelt (vgl. [So11, AS12, CA05]) Dieser Artikel befasst sich daher in Abschnitt 2 zunächst mit den theoretischen Grundlagen der Messtheorie und Konzepten zur Messung komplexer Phänomene, wie sie im Bereich der Datenschutzanforderungen üblich sind. Anschließend werden in Abschnitt 3 domänenspezifische Anforderungen und das übliche Vorgehen bei der Verwendung von Metriken in IT-Unternehmen vorgestellt. Es folgt die Beschreibung des Rahmenwerks und der Konstruktion von Metriken in Abschnitt 4. Zudem wird es für die Bestimmung des Standorts einer virtuellen Maschine beispielhaft in Abschnitt 4.2 angewendet, bevor Abschnitt 5 mit einer Zusammenfassung und einem Ausblick den Artikel beschließt.

## 2 Theoretische Grundlagen

Die theoretische Basis des hier vorgestellten Rahmenwerks stammt im Wesentlichen aus zwei wissenschaftlichen Disziplinen. Die Messtheorie beschäftigt sich mit der Messbarkeit verschiedener Aspekte der Realität und ist damit Grundlage aller Messungen. Vor allem in den Sozialwissenschaften ist jedoch häufig das Ziel, auch nicht direkt messbare Phänomene einer Größe zuzuordnen. Dies ist bei Datenschutzmetriken ebenfalls der Fall. Eine Datenschutzmetrik misst, inwieweit ein System, bspw. eine virtuelle Maschine in der Cloud, die Datenschutzanforderungen, die aus Gesetzen und Regularien stammen, erfüllt. Beispielsweise dürfen bestimmte, personenbezogene Daten nur innerhalb des Europäischen Wirtschaftsraums (EWR) gespeichert werden. Eine entsprechende Datenschutzmetrik dazu gäbe an, inwieweit diese Anforderung bezüglicher aller betroffenen Daten erfüllt wird. Die Konzepte des Konstrukts und des Indikators, die in Abschnitt 2.2 erläutert werden, bilden das Fundament für die Approximation dieser Phänomene.

### 2.1 Messtheorie

Grundlage jeden Messens ist die Messtheorie [SM11]. Sie beschäftigt sich mit der Frage, ob Aspekte bzw. Phänomene der Realität überhaupt gemessen werden können und ist somit wesentlicher Bestandteil der Wissenschaftstheorie. Während statistische Methoden dazu dienen, von Daten auf darin enthaltene Informationen zu schließen, beschäftigt sich die Messtheorie damit, die Realität in Form von Daten abbilden zu können. Eingeführt wurde die Messtheorie von Stevens [St46], wesentliche Beiträge stammen zudem von Suppes [Su14].

Es wird zwischen zwei verschiedenen Varianten der Messtheorie unterschieden, der repräsentativen und der operationalen Messtheorie. Die repräsentative Messtheorie wird dabei auch als die klassische Messtheorie bezeichnet [SM11]. Messen im Sinne dieser Theorie bedeutet, die *Struktur der Realität* durch Zahlen abzubilden, d. h., eine *Repräsentation* der (empirischen) Realität durch Zahlen zu erstellen. Insbesondere bedeutet die Abbildung der Struktur der Realität, dass Relationen realer Objekte auf die numerischen Re-

präsentationen abgebildet werden. Das empirische Objekt und die Messmethode sind dabei unabhängig voneinander und wichtige Aufgabe der Messtheorie ist es, zu zeigen, dass eine gewählte Methode dem empirischen Objekt gerecht wird. Die der repräsentativen Messtheorie entgegengesetzte *operationale* Messtheorie hingegen besagt, dass ein Objekt allein durch die verwendete Messmethode definiert wird und damit mit dieser eine Einheit bildet. Ein Bezug zur Realität ist nicht unbedingt vonnöten, da er nicht Teil der Betrachtung dieser Theorie ist [Ha96]. Ein operationales Vorgehen bedingt, dass sich Skalen entweder von selbst ergeben oder eine Auswahl mehrerer, eventuell äquivalenter Skalen nur schwer möglich ist [Ba00]. Die Logik hinter der operationalen Messtheorie ist, dass die Messmethode Teil des gemessenen Objekts ist – und damit auch die Skala. Bemisst man die Intelligenz eines Menschen etwa mittels eines IQ-Tests und basierend auf einer anderen Skala, so wären dies nach der operationalen Messtheorie zwei verschiedene Objekte.

Um Relationen zu überprüfen, werden Vergleiche vorgenommen. Die Anzahl möglicher Vergleiche ist jedoch je nach empirischem Objekt unterschiedlich. So kann bspw. die Größe zweier Objekte verglichen werden, die dann auch Aussagen über das Größenverhältnis zulassen. Bei anderen Objekten kann hingegen u. U. nur die Aussage getroffen werden, ob sie unterschiedlich oder gleich sind.

### 2.1.1 Skalenniveaus und Gütekriterien

Die Anzahl möglicher Vergleiche drückt das *Skalenniveau* einer Messung aus. Es gibt an, wie *eindeutig* ein Objekt mittels numerischer Werte beschrieben werden kann bzw. wie viel Information eine Skala enthält. Dabei wird zwischen Nominal-, Ordinal-, Intervall- und Verhältnisskalen unterschieden.

Die am wenigsten eindeutige Skala ist die *Nominalskala*. Auf dieser Skala gemessene Objekte werden in Kategorien unterteilt, wobei die einzige mögliche Aussage über zwei Objekte unterschiedlicher Kategorien ist, dass sich die Objekte unterscheiden. Ein Beispiel hierfür sind IP-Adressen, bei denen die Zahlen lediglich der Unterscheidung verschiedener Adressen dienen. Es folgt die *Ordinalskala*, die eine Ranganordnung der Objekte zulässt, d. h., die mathematischen Operatoren $>$ und $<$ sind als Vergleich zweier Objekte zulässig. Es kann jedoch keine Aussage darüber erfolgen, wie weit zwei Objekte auseinanderliegen. Das von einem Kunden eines Cloudanbieters gewünschte Datenschutzniveau kann etwa auf einer solchen Skala gemessen werden. Zwischen den Abstufungen „normal “, „„hoch“ „sehr hoch“ ist eine eindeutige Rangfolge erkennbar. Es kann jedoch keine Aussage darüber getroffen werden, wie groß der Unterschied zwischen den einzelnen Niveaus ist. Die Möglichkeit der Addition und Subtraktion von Werten erlaubt erst die *Intervallskala*. Zwischen Objekten, die auf dieser Skala gemessen werden, kann der absolute Abstand bestimmt werden. Nur bei Messungen auf *Verhältnisskalen* können Aussagen wie „doppelt so groß wie…“ getroffen werden. Der Unterschied zwischen beiden Skalen wird am Beispiel der Temperatur deutlich: Eine Temperatur in Grad Celsius kann nicht doppelt so hoch wie eine andere sein, da sie auf einer Intervallskala, d. h. ohne absoluten Nullpunkt, gemessen wird. Das Verhältnis zweier Längen kann hingegen angegeben werden.

Ein weiterer wichtiger Aspekt einer Messung und insbesondere darauf anwendbarer statistischer Methoden ist deren *Bedeutsamkeit* [SM11]. Sie ergibt sich aus den erlaubten Transformationen, der Umkehrung der Eindeutigkeit: Je eindeutiger eine Skala ist, d. h., je mehr Aussagen über das Verhältnis zweier Objekte möglich sind, desto weniger mathematische Transformationen sind erlaubt, ohne die Struktur der Realität zu verändern. Bei Ordinalskalen ist bspw. die Bezeichnung der einzelnen Ränge mit Zahlen wahllos, d. h., es ist unerheblich, ob die Ränge nun mit 0 bis 10 oder etwa mit -5 bis 5 bezeichnet werden. Bedeutsam ist eine Messung nur dann, wenn alle erlaubten Transformationen, bspw. die Verdoppelung jedes Werts, nichts an den Relationen der einzelnen Messobjekte ändern. Für anwendbare statistische Methoden, etwa die Berechnung des arithmetischen Mittels, gilt: Das Ergebnis der Methode muss identisch sein, wenn entweder zuerst die Messdaten mittels einer erlaubten Transformation verändert werden und darauf die Methode angewendet wird oder das gleiche umgekehrt geschieht. So ist z. B. das arithmetische Mittel für Ordinalskalen nicht erlaubt, weil dieses abhängig von der Wahl der Werte unterschiedliche Ergebnisse liefert.

Ein Hauptproblem bei der Anwendung der repräsentativen Messtheorie ist, die erlaubten Transformationen der gemessenen empirischen Objekte zu erschließen. Dies ist insbesondere dann der Fall, wenn komplexe Phänomene gemessen werden sollen, aber nicht klar ist, was *genau* gemessen wird. Ein Beispiel aus den Sozialwissenschaften hierfür ist die Bemessung der Armut. Zudem ist auch die strikte Auslegung des Begriffs Bedeutsamkeit unter Umständen problematisch: So kann etwa ein arithmetisches Mittel über ordinal skalierte Objekte empirisch durchaus eine sinnvolle Bedeutung haben, selbst, wenn dies mathematisch bzw. messtheoretisch nicht erlaubt ist. Zur Lösung des letzteren Problems wird der Begriff der Bedeutsamkeit oftmals nur als *numerisch bedeutsam* verstanden [Ni94]. Dadurch kann es allerdings den empirisch sinnvollen Methoden an theoretischer Fundierung mangeln [SM11].

Wichtig bei der Erstellung von Datenschutzmetriken ist jedoch auch die Frage, ob eine (falsche) Wahl der Skala das Messergebnis beeinflussen kann oder statistische Methoden als möglich suggeriert, die tatsächlich nicht angewendet werden sollten. In der Literatur wird hierfür ein pragmatischer Ansatz vorgeschlagen, der direkt von den empirischen Daten ausgeht [Ga75]: Mittels Messmethoden wird auf Skalen gemessen, die meist durch die Methoden selbst bedingt sind. Leistungen von Schülern, die mittels Schulnoten gemessen werden, bedingen z. B. eine Ordinalskala. Wenn nun die Anwendung verfügbarer statistischer Methoden auch nach erlaubten Transformationen, etwa der Verdopplung der gemessenen Werte, zum gleichen Ergebnis kommt, kann die Skala verwendet werden, da durch sie keine falschen Schlüsse gezogen werden können.

Für Messungen existieren eine Reihe von Gütekriterien, um zu bestimmen, ob sie sinnvoll sind und zu überprüfbaren Ergebnissen führen können [Hi06]. Messungen, die in Datenschutzmetriken verwendet werden, müssen vor allem die Kriterien der Reliablität und der Validität erfüllen. Messungen, die das erste Kriterium erfüllen, sind reproduzierbar, d. h. bei wiederholter Messung desselben Sachverhalts ergibt sich das gleiche Ergebnis. Das Kriterium der Validität besagt, dass die Messung auch das misst, was tatsächlich gemessen werden soll. Die Bemessung dieses Kriteriums kann u. U. schwierig sein, wenn, wie

im letzten Abschnitt erwähnt, gar nicht klar ist, was genau gemessen wird. Hierzu muss dann auf die im nächsten Abschnitt erwähnten Konzepte zurückgegriffen werden.

### 2.2 Konstrukte und Indikatoren

Wie bereits im letzten Abschnitt erwähnt, kann die Anwendung der repräsentativen Messtheorie insbesondere bei komplexen Phänomenen schwierig sein. Außerdem kann es von Interesse sein, Aussagen über Phänomene treffen zu wollen, bei denen eindeutig feststeht, dass sie nicht direkt messbar sind. Diese Phänomene finden sich vor allem in den Sozialwissenschaften, sie können aber auch, wie später gezeigt wird, für Datenschutzmetriken in der Cloud auftreten. Es bedarf daher einer Möglichkeit, auf messtheoretischer Basis Aussagen über komplexe, nicht direkt messbare Phänomene treffen zu können.

In den Sozialwissenschaften haben sich dafür die Konzepte des *Konstrukts* und des *Indikators* etabliert. Ein Konstrukt ist das Phänomen, über das eine Aussage getroffen werden soll, das aber nicht direkt messbar ist. Zu diesem Zweck werden Indikatoren gesucht, die in einem semantischen Zusammenhang mit dem Konstrukt stehen und messbar sind. Mehrere Indikatoren werden mittels statistischer Methoden kombiniert und geben so eine indirekte Aussage über das Konstrukt.

Dieses Konzept des indirekten Messens hat zur Folge, dass immer auch von der operationalen, nicht nur der repräsentativen Messtheorie ausgegangen werden muss [Bo10]. Das ist dadurch bedingt, dass das Konstrukt per Definition nicht direkt messbar ist und die Auswahl, Gewichtung und Kombination der Indikatoren – die Operationalisierung der Indikatoren – semantischen Überlegungen folgt. Operationalisierungen können auch statistisch überprüft werden. Dabei wird der Einfluss von Indikatoren auf das zu bestimmende Konstrukt mittels Testdaten überprüft. Bei diesen Testdaten ist das Ergebnis bekannt, d. h., man weiß beispielsweise, an welchem Standort sich Daten befinden, für die eine Reihe von Messungen ausgewählter Indikatoren erfolgen. Daraus kann der Einfluss der einzelnen Indikatoren auf das Ergebnis erschlossen werden.

## 3 Domänenspezifische Anforderungen

Nachdem bisher die theoretischen Ansätze aus der Messtheorie und den Sozialwissenschaften erläutert wurden, befasst sich dieser Abschnitt mit den domänenspezifischen Anforderungen, die sich aus dem Kontext der Datenschutzmetriken in der Cloud ergeben, für die das Rahmenwerk entwickelt wurde. Dazu wird zunächst Literatur aus dem IT-Bereich vorgestellt, der sich mit Metriken beschäftigt. Aus diesen ergibt sich in Kombination mit den theoretischen Ansätzen der Forschungsansatz, der mit dem Rahmenwerk verfolgt werden soll.

### 3.1 Verwandte Arbeiten

Sowa [So11] bietet einen umfassenden Überblick über die Definition, Entwicklung und Verwendung von Metriken im IT-Umfeld. Die Definition einer Metrik orientiert sich dabei stark am National Institute for Standards and Technology (NIST) [NI08] bzw. dem entsprechenden ISO-Standard 27004 [IS09]. Danach sind Metriken Werkzeuge zur Performanzmessung und Entscheidungsunterstützung, die diese Aufgabe durch Sammeln, Analyse von Daten und Ausgabe von Informationen, d. h. kontextbezogenen Daten, erfüllen. Diese eher allgemein gehaltene Definition kann als mit einer Zielgröße versehene Approximation des in Abschnitt 2.2 erwähnten Konzepts des Konstrukts angesehen werden. Sowa beschreibt daraufhin die regulatorischen Anforderungen, aus denen heraus sich Anwendungsszenarien für Metriken ergeben können. Es folgt die Erläuterung des Messplans, der für die automatisiert erfolgenden Messungen wichtig ist. Die Skalen, auf denen diese Messungen stattfinden, entsprechen denen aus Abschnitt 2.1. Allerdings bedingt der starke Anwendungsbezug, dass Skalen hier bereits durch die Messung oder die danach stattfindenden Verarbeitungsschritte schon vorgegeben sind. Implizit wird hier von einer operationalen Messtheorie ausgegangen, auch, wenn sie nie benannt wird. Eine weitere Anforderung, die von Sowa erwähnt wird, ist die Notwendigkeit, verschieden stark aggregierte Metriken für unterschiedliche Zielgruppen innerhalb eines Unternehmens zu erstellen. Diese können sich etwa im Hinblick auf Frequenz und Detailgrad unterscheiden. Abschließend wird von Sowa beispielhaft eine Liste möglicher Metriken angegeben. In Abschnitt 3.2 werden Unterschiede dieser Art von Metriken zu den bereits aufgeführten theoretischen Ansätzen erläutert.

Ammann und Sowa [AS12] definieren wichtige Begriffe, wie sie für Metriken in IT-Unternehmen verwendet werden. Alle Messungen werden auf dem *Untersuchungsgegenstand* ausgeführt. Dieser ist beispielsweise eine virtuelle Maschine oder ein Server in der Cloud. Messungen erfolgen nach einem *Messplan*, der angibt, wie oft welche Messung durchgeführt wird und wie das Ergebnis dieser Messung aussieht. Das Ergebnis der Messungen sind *Rohdaten*. Zur Verarbeitung dieser Rohdaten, bspw. zur Erstellung von Durchschnittswerten, werden *Instrumente* verwendet. Diese sind einfache mathematische Verfahren, die auf die Rohdaten angewendet werden. Teil von Metriken im IT-Umfeld sind *analytische Modelle*, die angeben, wie Metriken berechnet werden. Außerdem geben *Entscheidungskriterien* mit *Schwellwerten* an, ob die gemessenen Daten das Kriterium erfüllen oder nicht. Jede Metrik hat einen *Zielwert*, den die Metrik erfüllen soll und der sich aus externen Anforderungen ergibt. Besonders wichtige Metriken werden *Kennzahlen* genannt. Alle diese Begriffe finden in angepasster Form Anwendung im Rahmenwerk. Die Autoren stellen außerdem ein Bottom-Up- sowie ein Top-Down-Verfahren zur Entwicklung von Metriken vor. Ersteres geht dabei von verfügbaren Messwerten aus und entwickelt aus diesen ein Messmodell, das Zusammenhänge zwischen Prozessen innerhalb von Unternehmen darstellt, aus dem heraus sich die zu verwendenden Metriken ergeben. Für das Top-Down-Verfahren werden Ziele und dazugehörige Fragen bzw. mögliche Problemstellungen formuliert, aus denen sich Metriken zur Beantwortung herleiten lassen. Dieses Verfahren heißt *Goal-Question-Paradigma*. Beide Verfahren werden in abgewandelter Form auch für das Rahmenwerk verwendet.

Einen Überblick über qualitative Sicherheitsmetriken, die in IT-Unternehmen Verwendung finden, geben Chapin und Akridge [CA05]. Dabei werden auch mögliche Probleme bei der Aussagekraft einzelner Metriken erwähnt, etwa, ob viele oder wenige entdeckte Viren innerhalb eines Systems als gut klassifiziert werden können. Im Fokus stehen auch hier vor allem Standards und regulatorische Anforderungen, denen Metriken gerecht werden sollen.

### 3.2 Forschungsansatz

Die in Abschnitt 2 erwähnten Methoden befassen sich auf theoretischer Basis mit der Art und Weise, wie Messungen vorgenommen und nicht direkt messbare Phänomene approximiert werden können. Anwendungsbeispiele finden sich häufig in den Sozialwissenschaften. Diese Grundlagen sind auch für die Erstellung von Datenschutzmetriken, die ebenfalls auf Messungen basieren und auch nicht messbare Zielgrößen umfassen, notwendig. Ziel dieses Rahmenwerks soll es sein, den speziellen Anforderungen von Datenschutzmetriken in der Cloud gerecht zu werden. Die Art und Weise, wie Metriken im IT-Umfeld verwendet werden, wurde im letzten Abschnitt beschrieben. Dabei zeigt sich, dass Metriken in IT-Unternehmen häufig direkt messbare Werte sind, selbst, wenn die eigentlich erwünschte Information nicht direkt messbar ist. Es gibt kein explizites Konzept des Konstrukts (vgl. Abschnitt 2.2). Dass Metriken Aussagen über die Phänomene von Interesse treffen können, wird entweder implizit angenommen oder rein verbal erläutert. Eine Operationalisierung (vgl. Abschnitt 2.1), also die Auswahl und Gewichtung von Indikatoren, findet nicht statt.

Das im folgenden Kapitel vorgestellte Rahmenwerk definiert Metriken hingegen so, dass sie Konstrukte approximieren, d. h., dass die Operationalisierung Teil der Metriken selbst ist. Es bringt dadurch die theoretischen Ansätze aus der Messtheorie und den Sozialwissenschaften in die Domäne IT-bezogener Metriken. Diese sind notwendig, um Aussagen über die nicht direkt messbare Erfüllung von Datenschutzanforderungen zu ermöglichen, ohne diese Information manuell und nur implizit aus einer Vielzahl von Indikatoren zu erschließen.

## 4 Ein Rahmenwerk für Metriken

In diesem Abschnitt wird das Rahmenwerk für Datenschutzmetriken in der Cloud vorgestellt. Dafür werden zunächst die Begriffe und Verknüpfungen der Elemente des Rahmenwerks erläutert und dargestellt. Es folgt die Beschreibung der Umsetzung der theoretischen und domänenspezifischen Prozesse zur Konstruktion von Metriken und der Erzeugung von Indikatoren anhand eines Anwendungsbeispiels. Abschließend werden analytische Modelle vorgestellt, die bei der Approximation der Konstrukte durch Metriken verwendet werden.

### 4.1 Begriffe und Verknüpfungen

Vorrangiges Ziel des Rahmenwerks ist es, die in den vorherigen Kapiteln vorgestellten theoretischen Konzepte sowie die domänenspezifischen Anforderungen von Datenschutzmetriken in der Cloud zu verbinden. Der erste Schritt ist dabei die Definition der Begriffe der Metrikenkonstruktion sowie deren Verknüpfungen untereinander. Eine Übersicht ist in Abbildung 1 zu sehen.

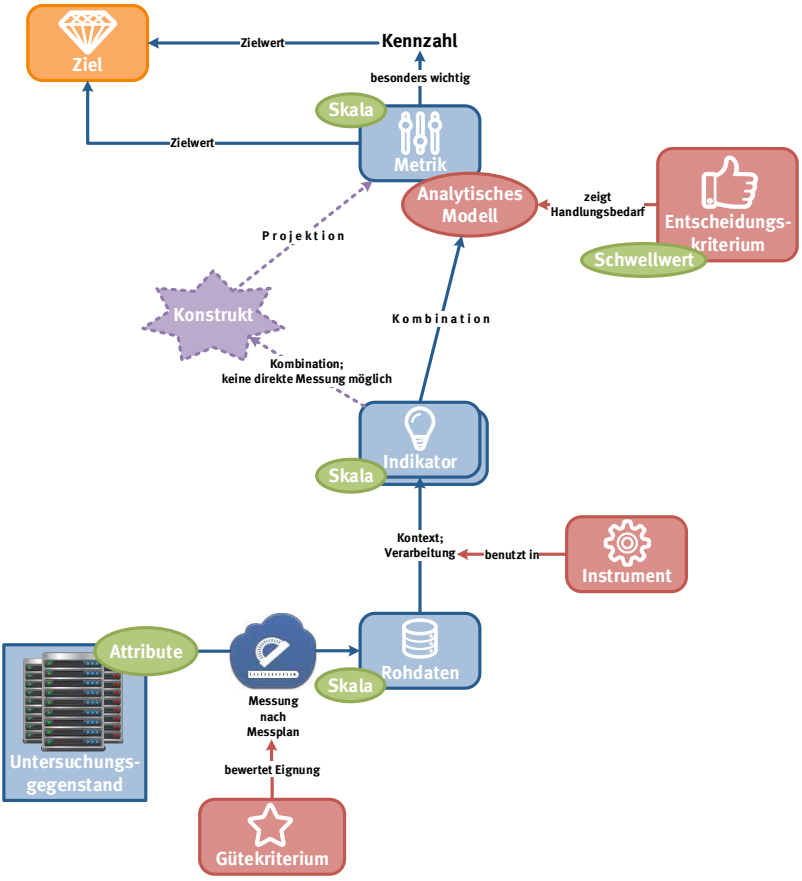


Abb. 1: Rahmenwerk für die Konstruktion von Metriken

Basis einer jeden Messung ist der *Untersuchungsgegenstand*. Dieser kann etwa eine virtuelle Maschine auf einem Server in der Cloud sein. Der Untersuchungsgegenstand besitzt Attribute, die mittels eines Messplans automatisiert gemessen werden können. *Gütekriterien* bewerten dabei die Eignung der Messung. Diese sind die bereits in Abschnitt 2.1 erwähnten messtheoretischen Kriterien wie Reliabilität, Reproduzierbarkeit, Validität und Generalisierbarkeit. Ergebnis der Messungen sind *Rohdaten*, die für die weitere Verarbeitung ge-

speichert werden. Diese Zusammenhänge ergeben sich sämtlich aus der in Abschnitt 3.1 erwähnten Literatur. Teile dieser Rohdaten können abhängig vom Kontext Indikatoren für ein zu approximierendes *Konstrukt* sein.

Das bereits in Abschnitt 2.2 eingeführte Konstrukt steht im Zentrum des Rahmenwerks und beschreibt die tatsächlich erwünschte Information. Gerade im Bereich der Datenschutzmetriken ist diese oft nicht direkt messbar, da sie sich üblicherweise aus datenschutzrechtlichen oder regulatorischen Anforderungen ergibt (vgl. Abschnitt 3.1). Ein Beispiel hierfür, die Bestimmung des Standorts einer virtuellen Maschine, wird in Abschnitt 4.2 behandelt. Aufgrund der nicht möglichen direkten Messungen muss das Konstrukt durch die Kombination messbarer Größen approximiert werden. Da die Konstrukte sich üblicherweise aus Anforderungen ergeben, geben die dazugehörigen Metriken den Erfüllungsgrad dieser Anforderung sowie einen Konfidenzwert an, der beschreibt, wie sicher die Aussage über den Erfüllungsgrad ist. *Ziel* einer Metrik ist, die Anforderung möglichst vollständig zu erfüllen. Die Notwendigkeit eines Zielwerts ergibt sich aus der vorgesehenen Anwendung des Rahmenwerks, der Messung der Erfüllung von Datenschutzanforderungen an Cloud-Dienste.

Metriken entstehen durch Kombination mehrerer *Indikatoren*. Die Art und Weise der Kombination bestimmt das *analytische Modell*, auf das in Abschnitt 4.3 näher eingegangen wird. Teil des analytischen Modells sind ein oder mehrere *Entscheidungskriterien*, mittels derer der Erfüllungsgrad der Metrik ermittelt wird.

Rohdaten, Indikatoren und Metriken werden jeweils auf – möglicherweise unterschiedlichen – Skalen gemessen (vgl. Abschnitt 2.1). Diese sind im Falle der Rohdaten und Indikatoren oft durch die Messungen selbst vorgegeben. Für die Metriken muss mittels des analytischen Modells (vgl. Abschnitt 4.3) eine passende Skala gefunden werden. Diese ergibt sich, wie in Abschnitt 2.2 erläutert, aus der operationalen Messtheorie bzw. pragmatischen Erwägungen.

Im nächsten Abschnitt wird das Vorgehen bei der Konstruktion von Metriken beispielhaft für die Bestimmung des Standortes einer virtuellen Maschine eines Cloud-Kunden dargestellt. Der Standort von Kundendaten bietet sich an, da rechtliche Anforderungen in diesem Bereich bestehen, etwa, dass personenbezogene Daten nicht außerhalb des Europäischen Wirtschaftsraums (EWR) gelagert werden dürfen. Zudem ist dieses Beispiel sehr anschaulich und es existieren eine Reihe von Indikatoren, sodass auch das analytische Modell einigermaßen umfangreich gestaltet werden kann.

### 4.2 Anwendungsbeispiel: Standortbestimmung

Metriken werden innerhalb des hier vorgestellten Rahmenwerks in einem kombinierten Top-Down-/Bottom-Up-Verfahren konstruiert. In diesem Abschnitt wird der Top-Down-Ansatz zur Konstruktion von Metriken aus Konstrukten heraus erläutert, der nächste Abschnitt befasst sich mit dem Bottom-Up-Verfahren zur Erzeugung von Indikatoren. Der kombinierte Ansatz ist notwendig, weil sich die Metriken zum einen aus den Anforderungen ergeben, die angeben, welche Eigenschaften des Untersuchungsgegenstands approxi-

miert werden sollen. Zum anderen ist eine Betrachtung der möglichen direkten Messungen notwendig.

Im Folgenden soll nun das Vorgehen der Metrikkonstruktion und die Anwendung des Rahmenwerks anhand eines Beispiels veranschaulicht werden. Als Ausgangspunkt liegt hier das Konstrukt des Speicher- und Verarbeitungsortes von Kundendaten, welches sich aus der datenschutzrechtlichen Anforderung, dass Daten nur in ausgewählten Ländern mit gesichertem Datenschutz gespeichert und bearbeitet werden dürfen, zugrunde. Es ergibt sich direkt aus den gesetzlichen Vorgaben („Wo werden die Daten gespeichert?“) und umfasst nicht etwa Mittel, mit denen diese Frage beantwortet werden könnte.

Auf Basis dieses Konstrukts lässt sich nun eine Metrik formulieren, die den Erfüllungsgrad der Anforderungen misst: „Alle Daten werden immer nur an erlaubten Standorten gespeichert.“. Zu dieser Metrik gilt es nun Indikatoren zu finden, die in einem semantischen Zusammenhang stehen.

Im Gegensatz zum bisherigen Top-Down-Verfahren vom Konstrukt zur Metrik werden die Indikatoren in einem Bottom-Up-Verfahren gebildet: Zentraler Untersuchungsgegenstand ist hier im Beispiel die virtuelle Maschine, von der aus mögliche Datenquellen ermittelt werden, die zu Indikatoren verarbeitet werden können. Beim Beispiel der Standortbestimmung lassen sich unter anderem die folgenden Daten nutzen:

- Bestimmung der minimalen Latenzzeiten zu nahegelegenen Servern zur Eingrenzung des physischen Standorts.
- Erfassung der bei Kommunikation mit Referenzservern verwendeten Kommunikationswege/Hops
- Generierung eines Fingerabdrucks der Maschinenumgebung (Andere Hosts im Netzwerk, MAC-Adressen, ...)
- Durchführung von DNS-Abfragen, welche je nach Standort unterschiedliche Antworten liefern

Diese Datenquellen gilt es in Anlehnung an das Rahmenwerk auf Informationsgehalt, Komplexität und Akzeptanz seitens des Cloudanbieters zu überprüfen. Beispielsweise lässt sich festhalten, dass die Generierung eines Fingerabdrucks zur Standortbestimmung nur begrenzt geeignet ist: Zum einen lassen sich mit einem Fingerabdruck nur Änderungen in der Netzwerkumgebung erkennen, diese können jedoch unabhängig von einer für die Metrik relevanten geographischen Änderung des Serverstandorts stattfinden. Zum anderen dürfte die Akzeptanz des Cloudanbieters für Scans der Maschinenumgebung vergleichsweise gering sein. Ziel der Indikatorenbildung ist jedoch gleichzeitig die Schaffung einer breiten Datenbasis, weshalb weniger aussagekräftige Indikatoren auch bei Vorhandensein von für die Metrik besseren/direkteren Indikatoren beibehalten werden sollen, auch um eine Kreuzvalidierung zu ermöglichen.

Im letzten Schritt sollen nun die verschiedenen Indikatoren miteinander kombiniert werden, sodass das daraus entstehende analytische Modell Aussagen darüber treffen kann, ob das Gesamtsystem die gestellten Anforderungen erfüllt. Wie anhand der beispielhaft

aufgeführten Datenquellen deutlich wird, ist eine Zusammenführung der Daten aufgrund unterschiedlicher Skalen nicht einfach.

### 4.3 Analytische Modelle

Wesentlicher Bestandteil einer Metrik ist das analytische Modell, mit dem die Indikatoren zu einem Wert kombiniert werden können. Die Auswahl des richtigen Modells, d. h. die in Abschnitt 2.2 erwähnte Operationalisierung der Indikatoren, beruht dabei hauptsächlich auf semantischen Überlegungen sowie der Validierung des gewählten Modells durch Testdaten.

Da sich die Konstrukte im Rahmen von Datenschutzmetriken für die Cloud aus datenschutzrechtlichen und regulatorischen Anforderungen ergeben, ist es Aufgabe des analytischen Modells, die gegebenen Ausprägungen der Indikatoren danach zu klassifizieren, ob das Gesamtsystem die Anforderungen erfüllt oder nicht. Wenn beispielsweise der Standort personenbezogener Daten in der Cloud bestimmt wird, gibt es, je nach den spezifischen Anforderungen, erlaubte und nicht erlaubte Standorte, an denen sich die Daten befinden können.

Allerdings gibt es keine einfache, eindeutige Möglichkeit der Bestimmung des Standortes. Daher werden eine Reihe direkter und indirekter Indikatoren für diesen Zweck herangezogen und in einem analytischen Modell kombiniert. Direkte Indikatoren geben dabei schon einen vermuteten Standort an, beispielsweise durch Datenbanken, die IP-Adressen Standorten zuweisen. Weitere Indikatoren werden benutzt, um die Zuverlässigkeit der Standortbestimmung zu erhöhen. Dabei werden auch indirekte Indikatoren benutzt. Diese geben selbst keinen direkten Hinweis auf den Standort, sondern werden als Fingerabdruck eines Standortes verwendet. Diese Indikatoren geben bspw. Auskunft über die an einem Standort in der Nähe befindlichen Geräte im gleichen Netzwerk, um diese mit vorher gewonnenen Fingerabdrücken zu vergleichen und einer Kategorie zuzuordnen.

Diese Zuordnung – sowie die spätere Bestimmung des Standorts auf Basis aller Indikatoren – geschieht mit statistischen Klassifikationsverfahren, etwa Entscheidungsbäumen und Nearest-Neighbours-Algorithmen. Gemein ist diesen Verfahren, dass, vereinfacht gesprochen, jede Sammlung von Messwerten (d. h., Messwerte aller Indikatoren zusammengefasst) mit bereits vorher in Kategorien unterteilten früheren Sammlungen verglichen wird und danach in die ähnlichste Kategorie sortiert wird. Entscheidungsbäume nehmen dafür Schwellwerte einzelner Indikatoren, die eine besonders prägnante Unterscheidung erlauben. Dieses Verfahren wird im Beispiel für die Klassifikation der indirekten Indikatoren verwendet. Bspw. könnte das Vorhandensein eines bestimmen Geräts im gleichen Netzwerk der virtuellen Maschine ein sehr prägnantes Unterscheidungsmerkmal sein. Beim Nearest-Neighbour-Verfahren wird diejenige Kategorie, d. h. derjenige Standort als Ergebnis geliefert, das die meisten ähnlichen Sammlungen von Messwerten aufweist wie die zu kategorisierende Sammlung. Das Ergebnis ist ein vermuteter Standort, der dann in einer Metrik daraufhin überprüft wird, ob er erlaubt oder nicht erlaubt ist.

## 5 Zusammenfassung und Ausblick

In diesem Artikel wurde ein Rahmenwerk für Datenschutzmetriken in der Cloud vorgestellt. Metriken finden bereits jetzt vielfältige Anwendung in IT-Unternehmen, hauptsächlich in den Bereichen IT-Sicherheit und Compliance, werden dort meist aber sehr praxisbezogen benutzt. Eine theoretische Fundierung des Messens ergibt sich aus der Messtheorie, die dabei hilft, die Messbarkeit empirischer Objekte zu untersuchen. Phänomene, die nicht direkt messbar sind, können mithilfe der Sozialwissenschaften, die das Konzept des Konstrukts für nicht direkt messbare Phänomene eingeführt haben, approximiert werden. Dieses Konzept ist für Datenschutzmetriken sehr nützlich, weil die Ziele dieser Metriken, die sich aus datenschutzrechtlichen und regulatorischen Anforderungen ergeben, oft nicht direkt messbar sind.

Der wesentliche Beitrag dieses Artikels ist die Kombination der theoretischen Ansätze mit der gängigen Praxis von Metriken im IT-Umfeld. Das Konstrukt als Repräsentation des nicht direkt messbaren Phänomens von Interesse ist hier zentraler Bestandteil. Mithilfe dieses Rahmenwerks ist es möglich, Metriken in einem kombinierten Top-Down-/Bottom-Up-Verfahren zu erstellen. Dies wurde am Beispiel der Bestimmung des Standorts einer virtuellen Maschine gezeigt.

Als nächster Schritt wird ein detailliertes analytisches Modell für die Standortbestimmung entwickelt, getestet und validiert werden. Damit kann dann gezeigt werden, inwieweit die hier präsentierte Herangehensweise an Metriken helfen kann, Datenschutzanforderungen in der Cloud automatisiert zu überprüfen.

## Literaturverzeichnis

[AS12] Ammann, F.-E.; Sowa, A.: Systematische Entwicklung von Metriken zur Beurteilung der Datensicherheit. Datenschutz und Datensicherheit – DuD, 36(4):247–251, 2012.

[Ba00] Balzer, Wolfgang: Die Wissenschaft und ihre Methoden. Grundsätze der Wissenschaftstheorie. Journal for General Philosophy of Science / Zeitschrift für Allgemeine Wissenschaftstheorie, 31(1):179–186, 2000.

[Bo10] Bohrnstedt, George W: An Overview of Measurement in the Social Sciences. In: Presentation at the National Academy of Sciences Workshop on Advancing Social Science Theory: The Importance of Common Metrics, February. S. 24–25, 2010.

[Bo12] Borges, G. et al.: Datenschutzrechtliche Lösungen für Cloud Computing. Kompetenzzentrum Trusted Cloud, Oct 2012.

[CA05] Chapin, D. A.; Akridge, S.: How Can Security Be Measured. Information Systems Control Journal, 2:43–47, 2005.

[Ga75] Gardner, Paul Leslie: Scales and Statistics. Review of Educational Research, S. 43–57, 1975.

[Ha96] Hand, David J: Statistics and the Theory of Measurement. Journal of the Royal Statistical Society. Series A (Statistics in Society), S. 445–492, 1996.

[Hi06] Himme, Alexander: Gütekriterien der Messung: Reliabilität, Validität und Generalisierbarkeit. Methodik der empirischen Forschung, Wiesbaden, S. 383–400, 2006.

[IS09] ISO: ISO 27004:2009, Information Security Management — Measurement. International Organization of Standardization (Hrsg.), 2009.

[Ni94] Niederée, Reinhard: There Is More to Measurement than Just Measurement: Measurement Theory, Symmetry, and Substantive Theorizing. Review of Foundations of Measurement. Journal of Mathematical Psychology, 38(4):527–594, 1994.

[NI08] NIST: NIST 800-55: Performance Measurement Guide for Information Security. National Institute of Standards and Technology (Hrsg.), 2008.

[SM11] Saint-Mont, Uwe: Messtheorie. In: Statistik im Forschungsprozess, S. 23–76. Physica-Verlag HD, 2011.

[So11] Sowa, A.: Metriken, der Schlüssel zum erfolgreichen Security und Compliance Monitoring. 2011.

[St46] Stevens, S. S.: On the Theory of Scales of Measurement. Science, 103(2684), 1946.

[Su14] Suppes, Patrick: Foundations of Measurement, Jgg. 2. Elsevier, 2014.