

## Verstehen: Syndrome//Muster

---

Durch das Web-Interface können die Benutzer:innen sich durch eigene Suchanfragen bestimmte Datensätze anzeigen lassen. Das System soll aber nicht nur den epidemiologischen Zugriff auf das Gesundheitsgeschehen verbessern. Vor allem soll das System auch in der Lage sein, Gefahren selbstständig zu erkennen und Warnmeldung auszugeben. Im Grunde geht es darum, unspezifische und wenig formalisierte Daten, die aber von hoher Aktualität sind, einer maschinellen statistischen Auswertung und Mustererkennung verfügbar zu machen. Die Ermöglichung dieser speziellen Kopplung zwischen den Daten und einer automatischen Durchsicht im Hinblick auf Regelmäßigkeiten und Häufungen ist das charakteristische Merkmal dieser Infrastrukturierung. Vor allem mit dieser Kopplung präsentiert syndromisches Monitoring ein Alleinstellungsmerkmal gegenüber anderen Formen des Monitorings. Wie kann diese Kopplung gelingen?

Um die Daten anschlussfähig für automatisierte Tests zu machen, müssen diese maschinenlesbar sortiert vorliegen. Ein erster entscheidender Schritt ist es daher, die Daten zu sinnvollen Kategorien zusammenzufassen und damit prozessierbar für statistische Modellierungen und schließlich für Mustererkennung zu machen.

Eine erste Entscheidung betrifft daher das Suchraster. Welche Beschwerden sind relevant für das kollektive Gesundheitsgeschehen und sollen Gegenstand des Monitorings werden? »Categorization of clinical symptoms into disease syndromes is the cornerstone of syndromic surveillance.« (Henning und Hamburg 2003, 291)

Beispielhaft im Fall von ESSENCE wurde diese Kategorisierung in einem Abwägen zwischen medizinischen und pragmatischen Gründen entschieden. Zunächst wurde eine Liste mit epidemiologisch relevanten Beschwerden erstellt, die sich aus den Daten entnehmen lassen. Zudem wurde geprüft, welche dieser Beschwerden überhaupt statistisch beobachtungsfähig sind, d.h. welche Beschwerden mit einem für Mustererkennung ausreichenden täglichen Grundvorkommen Eingang in die Daten finden. Aus beiden Erwägungen ergibt sich schließlich eine Liste mit zwölf Syndromen, auf die sich eine automatisierte Erkennung richten kann und die zugleich medizinisch relevante Vorfälle beschreiben können: 1. *Botulism-like*; 2. *Hemorrhagic Illness*; 3. *Lymphadenitis*; 4. *Localized Cutaneous Lesion*; 5. *Gastrointestinal*; 6. *Respiratory*; 7. *Rash*; 8. *Specific Infection*; 9. *Fever*; 10. *Neurological*; 11. *Severe Illness or Death potentially due to Infectious Disease*; 12. *Other* (Foster 2004).

Zu jedem dieser Syndrome ist eine Definition erstellt, die beschreibt, welche Indikatoren jeweils einbezogen oder ausgeschlossen werden. Unter das 6. Syndrom ›Atemwegsbezogen‹ [Respiratory] fallen z.B. Indikationen einer Infektion der oberen oder unteren Atemwege, aber nicht solche, die auf chronische Bronchitis oder Asthma hinweisen.

Neben der Festlegung dieser Syndrom-Kategorien muss außerdem gewährleistet sein, dass die entsprechenden Datensätze korrekt diesen Kategorien zugeordnet werden. Gesundheitsbezogene Informationen in dem System sind – wie gesehen – vor allem die bei der Aufnahme angegebenen *chief complaints*. Außerdem werden von Kliniken jeweils noch die Diagnosen gemeldet, mit denen die Patient:innen schließlich entlassen werden.

Im Fall dieser Entlassungs-Diagnose ist die Zuordnung zu den Syndrom-Kategorien unproblematisch. Gesetzlich geregelt kommt für die Dokumentation und Klassifikation medizinischer Diagnosen und Behandlungen im Gesundheitssystem der USA – wie oben gesehen – das von der Weltgesundheitsorganisation bereitgestellte Klassifizierungsschema, die so genannte *International Classification of Diseases* (ICD) zum Einsatz. Seit 2015 ist in den USA die 10. Auflage dieses Schemas (mit geringen Modifikationen) einschlägig. Das Schema wird

von dem *US National Center for Health Statistics* bereitgestellt (2020). Bekannte Krankheiten, Verletzungen und Behandlungsmethoden sind darin in einem Ordnungsschema sortiert und mit einer Zahlenfolge codiert. Die Diagnose-Daten liegen somit bereits klar strukturiert vor. Die Zuordnung der ICD-10 Kategorien zu den zwölf Syndrom-Kategorien muss einmal vorgenommen werden und kann danach automatisch erfolgen.

Der grundlegende Anspruch des Systems ist allerdings, wie bereits gesehen, signifikante Änderungen des kollektiven Gesundheitsgeschehens möglichst frühzeitig wahrzunehmen. Die bereits codierten Entlassungs-Diagnosen kommen erst nach der Behandlung und damit zwangsläufig mit Verzögerung in die Akte. Relevanter für das Versprechen auf ein beinahe Echtzeit-Monitoring sind daher die *chief complaints*, die bei der Aufnahme, d.h. teils Tage vor der Entlassungs-Diagnose vorliegen. Zudem können diese *complaints* aussagekräftiger als der reine ICD-10 Code sein, da sie offen als Freitext vorliegen. »[C]apturing surveillance information as free text does not incur the human effort, delay, or drastic reduction in information incurred by coding.« (Shapiro 2004, 95)

Allerdings macht diese Qualität des unstrukturierten Freitexts eine Zuordnung zu den Syndrom-Kategorien schwierig. So gibt es zu einer einzigen Beschwerde teils eine Reihe von Synonymen, Kliniken benutzen hauseigene Abkürzungen und unter Stress kommt es häufiger zu Rechtschreibfehlern (Shapiro 2004). Häufig kommen daher später Listen mit Zuordnungen bestimmter Begriffe zu den Syndrom-Kategorien zum Einsatz oder als Mittel der Datenbereinigung. So werden etwa Listen typischer Tippfehler erstellt, um eine automatische Kategorisierung der Freitext-Angaben zu ermöglichen. Neben »*Vomiting*« werden dann auch Angaben wie etwa »*Andvomiting*«, »*Vomiting*«, »*Vbomiting*« etc. der Syndrom-Kategorie »5. Gastrointestinal« zuordbar.

Die so den Syndromen zugeordneten Datensätze werden im zweiten Schritt nach irrelevanten Einträgen gefiltert. Beispielsweise sind unter dem Suchbegriff »*Mumps*« folgende Einträge in der Datenbank: »*possible mumps*«, »*Mumps Vaccination*«, »*fever, rash, exposure to mumps*«, »*I think she has mumps, denies fever/sore throat*«. Auf Basis einer semantischen

Liste wird hier z.B. der Datensatz ›Mumps Vaccination‹ für die weiteren Analysen aussortiert. Der Umstand einer prophylaktischen Impfung wäre nicht relevant für die Suche nach akuten Ausbruchseignissen und würde unnötige Warnungen bezüglich dieser Syndrom-Kategorie auslösen (Chen 2017).

Zutreffende Fälle sind somit auf Basis der Entlassungs-Diagnosen (ICD-10 Codes) sowie vor allem auf Basis der Freitext-Angaben (*chief complaints*) einer von zwölf beobachtungsrelevanten Syndrom-Kategorien zugeordnet.

Die so identifizierten Datensätze können nun einer Reihe von statistischen Prüfverfahren unterzogen werden. Ziel des Systems ist es, die sortierten und kategorisierten Daten in Echtzeit zu analysieren und auf mögliche Anomalien hin zu bewerten. Allerdings haben die eintreffenden Daten von sich aus eine große Variabilität. Dazu gehören saisonale oder wöchentliche Schwankungen. Eine gewöhnliche Grippe-Saison kann teils früher oder später im Jahr einsetzen. Bei auf geringe Fallzahlen gestützten Meldungen spielen zudem rechnerische Zufallseffekte eine erhebliche Rolle. Gerade die singulären Vorfälle eines bioterroristischen Anschlags oder eines neuartigen Infekts wären von besonderer Brisanz. Für solche Szenarien sind bereits Einzelfälle bedeutsam. Gerade hier ist es für automatisierte Verfahren aber besonders schwer, solche Signale gegenüber einem unbedeutenden ›Rauschen‹ in den Daten (Unterschiede im Aufkommen aufgrund von Zufallseffekten) auszumachen. Zudem führen pragmatische Gründe zu weiteren Einschränkungen. So ist die Rechenleistung, die für diese Verfahren angesetzt werden kann, durch die bei den lokalen Gesundheitsbehörden vorhandenen Geräte stark begrenzt. Entsprechend können keine aufwändigen Verfahren maschinellen Lernens eingesetzt werden (Chen 2017).

Es gilt also, im üblichen Zielkonflikt zwischen Sensitivität, Spezifität und Unmittelbarkeit (Stoto u.a. 2006, 142) unter eingeschränkten Bedingungen einen Kompromiss zu finden. Eine sensible Messung schlägt häufiger falschen Alarm, eine robuste Messung übersieht häufiger das Gesuchte, die Meldung soll jeweils möglichst unverzüglich erfolgen. Im Fall des ESSENCE-Systems wird der Kompromiss durch eine Verschränkung mehrerer statistischer Verfahren gesucht. Zudem sind

Nutzer\*innen in der Lage, bestimmte Rechenverfahren gezielt auszuwählen.

Letztlich beruht die automatisierte Mustererkennung auf üblichen Verfahren analytischer Statistik. Standardmäßig werden die Daten einer Mischung aus einer linearen Regression, einem exponentiell gewichteten Verfahren des gleitenden Mittelwerts und einer *Poisson* Regression unterzogen (Foster 2004; Chen 2017). Durch die Kombination relativ einfacher Verfahren, sollen die Berechnungen mit den technischen Mitteln der Gesundheitsbehörden vor Ort möglich bleiben. Die Kombination zielt darauf, die Schwächen der einzelnen Verfahren in der Anfälligkeit für fälschlicherweise als relevant erkannte Fälle bzw. für nicht erkannte relevante Fälle gegenseitig auszugleichen.

Die Anwendung dieser Testverfahren erlaubt schließlich eine für das System entscheidende Funktion: die automatisierte Anzeige von relevanten Ereignissen des kollektiven Gesundheitsgeschehens. Ein Kernbestandteil der Software ist neben dem Interface für gezielte Abfragen eine Zustands-Übersicht von elf der zwölf Syndrom-Kategorien (»Other« wird von dieser laufenden Übersicht ausgenommen). Am Beispiel des ESSENCE-Systems wird diese Übersicht nach dem Login als ein eigener Reiter prominent auf der Startseite angezeigt. Hier wird für jede Kategorie farblich markiert, mit welcher Überschreitungs-wahrscheinlichkeit (p-Wert) die aktuellen Fälle als normal im Hinblick auf die Entwicklung der letzten 30 Tage betrachtet werden können. Wenn diese Nullhypothese unwahrscheinlich (p-Wert zwischen 0.01 und 0.05) oder sehr unwahrscheinlich ist (p-Wert kleiner 0.01), werden die Kategorien in der Übersicht gelb oder rot ausgeflaggt (Kite-Powell, Ofori-Addo und Hamilton 2010).

Letztlich ist diese Kopplung zwischen als relevant erachteten Merkmalen des kollektiven Gesundheitsgeschehens und einer tagesaktuellen, auf einen Blick überschaubaren Tafel unterschiedlich gefärbter Kategorien, eine zentrale Errungenschaft der Infrastrukturierung *Syndromic Surveillance*.

## Vorstellung: datenbasierte Prognosen

Erneut ist es hilfreich, zunächst die Vorstellungen in den Vordergrund zu stellen, die mit diesem Aspekt der untersuchten Infrastrukturierung verbunden sind. Was versprechen sich politische Entscheidungsträger:innen, Akteure und Beobachter:innen letztlich von dieser Koppelung kategorisierter Daten und statistischer Verfahren? Welche Vorstellungen und Erwartungen stehen im Hintergrund dieser für die Systeme charakteristischen Anordnung von Elementen?

Die Zunahme von Sensoren und die gesunkenen Kosten der Speicherung haben in den letzten Dekaden zu einer Verfügbarkeit von unzähligen, häufig in nahe-Echtzeit übermittelten Daten geführt. Diese technischen Voraussetzungen sind Anlass für unterschiedliche Hoffnungen auf eine stärker datenbasierte und teilautomatisierte Erkenntnisgewinnung. Mit aktuellen Datensammlungen, die ein Phänomen ausreichend detailliert, umfassend und unmittelbar abbilden, wird ein *»reality mining«* möglich, so Nathan Eagle und Alex Pentland in einem viel beachteten Plädoyer für eine neue datengetriebene Erforschung sozialer Phänomene.

»By continually logging and time-stamping information about a user's activity, location, and proximity to other users, the large-scale dynamics of collective human behavior can be analyzed.« (Eagle und Pentland 2005, 263)

Im Hinblick auf das Verhältnis von Daten und Wissen beinhaltet dieser Vorschlag eine entscheidende Verschiebung. Mit der erstmals ausreichend plausiblen Annahme eines getreuen Abbilds des tatsächlichen Nutzerverhaltens in den Daten, kann das *»Schürfen«* [*mining*] nun in der Datensammlung selbst erfolgen. Phänomen, Datum und Erkenntnis sind diesem Vorschlag gemäß anders gereiht. Die Daten funktionieren nicht als Belege für bestimmte Annahmen zu einem Phänomen. Stattdessen wird die Datensammlung zu einem ausreichenden Substitut für das Phänomen an sich. Das Ziel ist es, Aussagen über die Regelmäßigkeit menschlicher Interaktion zu treffen. Das Objekt, das diese

Aussagen ermöglicht ist aber – dank *ubiquitous computing* – nun die Datensammlung selbst.

Dieser Vorschlag einer neuartigen datenbasierten Epistemologie, wird auch unter dem Stichwort Datafizierung diskutiert. »Datafication [...] is the transformation of social action into online quantified data, thus allowing for real-time tracking and predictive analysis.« (Van Dijck 2014, 198; siehe auch Cukier und Mayer-Schönberger 2013)

## Epistemologie von Big Data

Unter den Schlagworten *reality mining*, *Big Data*-Analysen oder Datafizierung wird vor allem auch ein Vorschlag für neue Wege der Erkenntnisgewinnung unterbreitet. Ein Vorschlag, der vielfältig skeptisch kommentiert worden ist. Dabei dreht sich die Auseinandersetzung vor allem um das Versprechen bzw. die Unmöglichkeit ›hypothesenfreier‹ Mustererkennung aus den Daten selbst. Das Zusammenfallen von Daten und Wissenobjekt ist eine notwendige Voraussetzung dieses Versprechens. Der Vorschlag eines ›Reality Mining‹ in ausreichend vollständigen Datensammlungen bedeutet eben auch einen veränderten Status von Daten. Daten sind nicht mehr als Verweis auf einen Sachverhalt interessant, sondern bereits selbst das Objekt von Erkenntnis.

»Big Data analytics enables an entirely new epistemological approach for making sense of the world; rather than testing a theory by analyzing relevant data, new data analytics seek to gain insights ›born from the data‹.« (Kitchin 2014, 2)

Apologeten unterstreichen die Vorzüge einer Mustererkennung aus den Daten (Lazer u.a. 2009). Wenn Muster in den Daten die Stelle des Wissenobjekts einnehmen, kann eine Forschung ohne den Ballast von Hypothesen über das Phänomen rein aus den Daten gemacht werden. »Prediction trumps explanation« so das provokante Credo (Siegel 2013). Gerade auch in der medizinischen Forschung verspricht die Beschränkung auf Korrelationen, die aus Beobachtungsdaten gewonnen werden, einen kostengünstigen Weg zur Erkenntnisgewinnung. Vor allem braucht es dadurch keine aufwändigen experimentellen Studien. Ent-

sprechend dominant ist dieser Weg hier inzwischen geworden. »Wir haben die Ära der Kausalität verlassen und befinden uns bereits mitten im Zeitalter der Korrelationen«, behauptet ein an einem Universitätsklinikum forschender Mediziner bereits 2018 (Antes 2018).

Gleichzeitig ist eine Vielzahl von skeptischen Einwänden gegen dieses Versprechen vorgebracht worden. Ein derartiger datenbasierter Weg zu Erkenntnisgewinnung ist an vielen Stellen wenig plausibel: So wird bei diesem Vorschlag zur ›Forschung ohne Annahmen‹ vergessen, wie stark Annahmen und Hypothesen zwangsläufig die Sammlung von Daten prägen (Gitelman und Jackson 2013; Frické 2015). Es bleiben selbst bei weitreichenden zusammengestellten Datensammlungen immer Blindstellen (Kitchin 2014; Lazer u.a. 2014). Die gefundenen Muster sind aufgrund des fehlenden Kontexts nur bedingt nützlich (Kitchin, Lauriault und McArdle 2015). Solche datenbasierten Analysen fokussieren eher nicht auf relevante Probleme, sondern tendenziell auf das, was besonders leicht in den Datensätzen sichtbar gemacht werden kann (Rabari und Storper 2015; Frade 2016; Elragal und Klischewski 2017).

### **Datensammlung als Wissensobjekt**

Unabhängig von solchen Einschätzungen zur Sinnhaftigkeit einer dezidiert datenbasierten Epistemologie geht es mir hier um die Verschiebung des Status der Datensammlung in diesem Zusammenhang. Eine Prämisse, damit ›Big Data‹ zu einer Erkenntnismöglichkeit werden kann, ist die Annahme, dass die Daten für sich selbst sprechen können und es keines weiteren Kontextwissens bedarf. Die (vermeintlich) hypothesenfreie Erkenntnisgewinnung sei strikt aus der Korrelation von Beobachtungsdaten möglich. Für die Datensammlung bedeutet dies die Aufwertung zu einem Wissensobjekt an sich. Daten sind hier nicht mehr bloß Mittel zum Zweck der Überprüfung einer Vermutung, sondern werden für sich genommen zum Objekt quantitativer Modellierung und einer automatisierten Suche nach Korrelationen.

Die technische Möglichkeit der Datensammlung, -übertragung und -speicherung macht es plausibel, die Datensammlung selbst als das Wissensojekt zu behandeln. Diese epistemologische Verschiebung wird derzeit in allen möglichen Anwendungsfällen propagiert. Im Fall des Gesundheitsmonitoring geschah dies früh und in konkreten Systemen umgesetzt. Infektionskontrolle hat – wie gesehen – im Zuge einer Versichertheitlichung von Gesundheit (siehe Kapitel »Vorstellung: Gesundheit als Sicherheitsproblem«) und einer Problematisierung von »*emerging diseases*« einen neuen Stellenwert in der Gesundheitspolitik bekommen. Die technischen Möglichkeiten wecken gerade hier Erwartungen an eine Lösung durch *Epidemic Intelligence* (Paquet u.a. 2006). Insbesondere Anstrengungen zur globalen Infektionskontrolle waren ein wichtiger, früher und häufig übersehener Treiber bei der Etablierung von datenbasierten Techniken im Bereich von Früherkennung und automatisierter Mustererkennung (Roberts und Elbe 2017). Seit Anfang der 2000er Jahre wurden auf der Ebene globaler Gesundheitspolitik unterschiedliche Projekte etabliert, um relevante Gesundheitsereignisse mit Hilfe ereignis-fokussierter Überwachung (Zeng u.a. 2010) zu erkennen. Namentlich das Global *Public Health Intelligence Network* (GPHIN) oder das Projekt *HealthMap* der WHO. Auch in der europäischen Gesundheitspolitik stellt eine ereignisorientierte, datenbasierte und automatisierte Früherkennung inzwischen eine zentrale Säule dar (Bengtsson, Borg und Rhinard 2019). Die Infrastrukturierung von *Syndromic Surveillance* in den USA ist im Kontext dieser unterschiedlichen Bemühungen ein nationaler Vorreiter.

Der Fall ist dadurch geeignet, die Auswirkungen dieser Vorstellung genauer in den Blick zu nehmen. Was bedeutet die epistemologische Verschiebung zu der Datensammlung als Wissensojekt konkret für *Public Health*? Wie verändert diese Vorstellung die Hinwendung und die Maßnahmen der Förderung des kollektiven Gesundheitsgeschehens?

## Voreinstellung: Wissensobjekt Ist-Zustand

Eine Stärke der Perspektive auf das Phänomen als Infrastrukturierung ist der Einbezug der Nutzung in die Untersuchung. Wie bereits gesehen sind die propagierten Zwecke und die in der Nutzung relevant gemachten Zwecke nicht notwendig deckungsgleich. Für das Verständnis der tatsächlichen Implikationen der sozio-technischen Einrichtung ist daher neben den Vorstellungen, d.h. den propagierten Zielen, vor allem diese Ebene der realisierten Zwecke zentral. Vor allem solche Voreinstellungen sind relevant für die Machtwirkungen der Infrastrukturierung.

Im Folgenden möchte ich entsprechend analysieren, wie die Vorstellung der Datensammlung als Wissensobjekt in der Nutzung konkretisiert und relevant gemacht wird. Nicht zuletzt beeinflusst eine solche Verschiebung des Wissensobjekts in die Datensammlung auch die generelle Auffassung von *Public Health* und Gesundheit über die konkrete Nutzung von Systemen des syndromischen Monitorings hinaus.

Wie bereits gesehen, verspricht die Infrastrukturierung einen besonderen, unmittelbaren Zugriff auf das kollektive Gesundheitsgeschehen (siehe Kapitel »Voreinstellung: Modus unmittelbarer Zugriff«). Merkmal des Wechsels zu einem ereignisbasierten Monitoring und Vorbedingung der Früherkennungs-Funktion ist außerdem das Versprechen, die Vorfälle in ihrer Gesamtheit zu erfassen. Die Kopplung der Syndrom-Kategorien mit statistischen Verfahren funktioniert erst dann, wenn das System laufend und im Grunde lückenlos den Ist-Zustand registriert. Erst mit diesem Kontrast können relevante Häufungen eines Syndroms identifiziert werden. Aus dieser Kopplung ergibt sich die charakteristische Arbeitsweise syndromischen Monitoring. Im Unterschied zu anderen Formen werden hier nicht nur bestimmte Fälle registriert, sondern idealerweise die gesamte gesundheitliche Lage (Purtle u.a. 2018, 2). Die Voraussetzung der automatisierten Erkennung ist eine Annäherung an den Ist-Zustand des kollektiven Gesundheitsgeschehens in den registrierten Daten.

## Situational awareness

Damit ändert sich nun in einer entscheidenden Weise die Qualität des Wissens, welches die Nutzer:innen von diesen Daten erwarten. Bei einem krankheitsbezogenen Monitoring ist die Erwartung an die Daten, Wissen über einen bestimmten Aspekt des Gesundheitsgeschehens zu liefern. Beispielsweise den Umfang des Vorkommens von grippeähnlichen Infektionen in der Bevölkerung im Fall des Influenza-Registers. Die entscheidende Qualität des ereignisbezogenen, syndromischen Monitoring ist im Kontrast eine allgemeine *situational awareness*. Dabei verschiebt sich die Erwartung an die Daten. Das Wissen, das nun in den Daten enthalten ist, betrifft entgegen dem vermeintlichen Fokus auf Ereignisse paradoxe Weise gerade nicht so sehr den Ausbruch oder die Störung. Weit aus relevanter beschreiben die Gesprächspartner:innen die neue Möglichkeit einer Lageeinschätzung. »The usefulness is when an event has occurred, you can use it to monitor the cause of the event.« (Davies-Cole 2012, 00:09:33) An dieser Stelle wird eine entscheidende Präzisierung möglich. Die in der Nutzung relevant gemachte Erwartung an das Wissensobjekt Datensammlung ist nicht so sehr die Warnmeldung selbst, sondern eine Rückversicherung über den Normalzustand des kollektiven Gesundheitsgeschehens.

Wie in den Gesprächen immer wieder deutlich wird, ist das frühzeitige Erkennen von Ausbrüchen durch das System eher unwahrscheinlich. Das Wissen über einen bioterroristischen Anschlag oder eine neu entstehende Infektionskrankheit lässt sich aus den Daten des Systems gar nicht erhalten. »It is not sensitive enough.« (Davies-Cole 2012, 00:09:3) Falls eine automatisierte Warnmeldung auftritt, ist diesem vermeintlichen Wissen über den Eintritt eines Bedrohungsfalls eher nicht zu trauen. Da es systematisch um die Frühphase eines relevanten Gesundheitereignis geht, sind geringe Fallzahlen unvermeidlich. Die Datenbasis, auf die sich eine automatisierte Warnung stützen kann, ist zwangsläufig schmal. Zugleich soll die Sensitivität sehr hoch sein. Schon zufällige Häufung von Fällen müssen daher eine Warnmeldung provozieren, da nicht sicher zu berechnen ist, ob sich Ausreißer anzeigen oder der Beginn eines relevanten Ausbruchsge-

schehens. Entsprechend vorsichtig gehen die Nutzer:innen mit den ausgeflagten Warnmeldungen um:

»For instance we had at one time where six people came in two different ERs [Emergency Rooms] with pneumonia and it gave an alert. But there's nothing to be panicked over. Next day it was nothing. Zero. [...] It's an anomaly. It's an aberration. So you have to have somebody who has that understanding to not always press the panic button.« (Russell 2012, 00:42:35)

Für dieses Wissen über den Eintritt eines Bedrohungsfalls ist immer noch ein aufmerksamer Arzt oder eine aufmerksame Ärztin das weit-aus bessere Erkennungsmittel, wie auch in den in den Gesprächen immer wieder betont wird. »I can assure you that when there's something serious happening, an astute physician will be the first person to call us.« (Davies-Cole 2012, 00:51:23)

Das System liefert für die Nutzer:innen aber durchaus eine verbesserte Erkenntnismöglichkeit. Allerdings bezieht sich diese Erkenntnis weniger auf den gemeldeten Vorfall. Das Wissen über den Beginn eines Vorfalls, welches das System ermöglicht, ist wahrscheinlich gar nicht mehr relevant, da der Vorfall dann womöglich längst ›händisch‹ erkannt und übermittelt worden ist. Zudem ist der Status des Wissens einer Meldung ungesichert und es bedarf in jedem Fall einer weiteren Prüfung, ob tatsächlich ein Alarm ausgelöst werden muss.

Die für die tägliche Praxis weitaus zentralere Funktion, so betonen die Gesprächspartner:innen, ist statt Früherkennung eine neue Möglichkeit zur Lageeinschätzung. Aus Perspektive der Anwender:innen sieht der Nutzen einer *situational awareness* im Vordergrund. Die Daten liefern vor allem ein Wissen über den Kontext des Vorfalls. Wie ist ein Ausbruchsereignis im Kontrast zur normalen Situation einzuschätzen? Was sind längere Trends und Entwicklungslinien? Gibt es lokale Cluster oder regionale Muster in den Daten? Früherkennung gelingt wahrscheinlich schneller und präziser durch eine Ärztin oder einen Arzt. Das System dient vor allem für den zweiten, kontrastierenden und erklärenden Blick auf das Gesundheitsgeschehen.

»Someone had already reported it [an outbreak]. And then now you have a list. So you can look at the system and see. I mean, are we seeing the same pattern in other areas? You know, so you can it can help you. Where are they? You know. And then you can then if it's dying out, you know, you will be able to manage it.« (Davies-Cole 2012, 00:09:33)

Der Erkenntnisgewinn, den das System zu leisten vermag, ist ein »societal sense of what's actually happening«, wie es eine andere Gesprächspartnerin ausdrückt.

»You know, the astute clinician might pick it up and might send up that red flag. And maybe even that red flag is earlier than what you're picking up with surveillance. It doesn't give you a societal sense of what's actually happening.« (Dugas 2012, 00:06:28)

Was das System ermöglicht, ist eine »Komplettierung des Bildes« durch die Bereitstellung von unterschiedlichen Typen von Daten (*chief complaints*, demographische Daten, Verkaufszahlen, Fehlzeiten) mit einer hohen räumlichen Auflösung. You know, so I'd like to see, is this a sustained thing? And if especially if I can see it in other data sources, then that really completes the picture.« (Russell 2012, 00:42:35)

Vor allem aus der Perspektive der Nutzer:innen ist das relevante Wissenobjekt des Monitorings gerade nicht der Moment des Ausbruchs, sondern der Normalfall des kollektiven Gesundheitsgeschehens. Schon früh wird eine durch die Fördermittel im Bereich *Biosurveillance* vorangetriebene Orientierung auf Bioterrorismus und auf Ausbrüche kritisch gesehen (Reingold 2003). Allerdings verspricht *Syndromic Surveillance* von Beginn an ein Werkzeug mit einem doppelten Nutzen zu sein. Ein Nebeneffekt des für die Suche nach Ausbrüchen notwendigen Einbezugs von umfassenden nahe-Echtzeit Daten ist die Stärkung der *situational awareness*, der gesundheitlichen Lageeinschätzung.

Dieser Nutzen wird von den Epidemiolog:innen gerade angesichts der ansonsten spärlichen Bereitstellung von Mitteln für die öffentliche Gesundheitspflege in den USA als Vorteil gesehen. In den USA macht der Bereich *Public Health* z.B. nur einen Anteil von 2,5 % aller gesund-

heitsbezogenen Ausgaben aus (TFAH 2019). Über die letzte Dekade sind in lokalen Gesundheitsbehörden der USA über 55.000 Stellen gestrichen worden (NACCHO 2018). Unter solchen Beschränkungen erscheint jede Ausgabensteigerung als Chance für die Akteure. »By not limiting its use to bioterrorism, however, syndromic surveillance can be a boon rather than a drain on traditional public health activities.« (Mostashari und Hartman 2003, 15)

In der anfänglichen Konzeption sind die Syndrom-Kategorien vor allem im Hinblick auf die Erfassung eines möglichen Infektions- und Ausbruchsgeschehens gefasst (grippaler Infekt, Nahrungsmittelvergiftung etc.). Angesichts des für die Nutzer\*innen eigentlich relevanten Wissensobjekts ›Ist-Zustand des kollektiven Gesundheitsgeschehens‹ wird diese Fokussierung auf Infektionskrankheiten zunehmend als eine hinderliche Begrenzung gesehen. Fortgesetzt werden Stimmen laut, in den Syndrom-Kategorien auch chronische, strukturelle und nicht-ausbruchshafte Aspekte der gesundheitlichen Lage abzubilden. »We wanted to include asthma, injuries. So you can use it, you know, for some other things.« (Davies-Cole 2012, 00:15:07)

»We have another Epi [Epidemiologist] looking at some of the chronic disease issues, some of the injury prevention type issues, even motor vehicle accidents [...] alcohol intoxication. I can look at that. So things like that. You know, it's not just disease specific stuff.« (Russell 2012, 00:41:15)

Mit der allmählichen Ausweitung und Etablierung syndromischen Monitorings über den Umweg des landesweiten *Biosense 2.0*-Systems nun in das *National Syndromic Surveillance Program*, sowie mit der Formierung einer *community of practice*, wird diese Verschiebung immer sichtbarer nach außen getragen (Paterson und Durrheim 2013). In einem von den CDC veröffentlichten Newsletter zum NSSP werden regelmäßig Erfolge und *Best-Practice*-Beispiele markiert. In den aktuellen Ausgaben werden hier vor allem Aspekte gesundheitlicher Lagebeobachtung hervorgehoben, die mit Hilfe syndromischer Überwachung möglich waren. So wird etwa betont, inwiefern inzwischen chronische und umweltbezogene Gesundheitsprobleme die anfängliche Fokussierung

auf Ausbruchsergebnisse abgelöst hat. Dies kann am Beispiel New Jersey veranschaulicht werden: »New Jersey moved beyond using syndromic data for communicable disease into monitoring environmental and occupational exposures, opioid classifications, tick-related illness, and waterborne illness« (NSSP 2019, 5) Die ausgeweitete Lagebeobachtung durch syndromisches Monitoring erfasst in einzelnen Fällen auch den Konsum von Betäubungsmitteln (Shekiro, Sussman und Brown 2018), die Effekte der Legalisierung vom Marijuana oder auch soziale Missstände wie Gewalt gegen Minderjährige. Syndromisches Gesundheitsmonitoring vermag inzwischen auch die Problematik der Wohnungslosigkeit zu erfassen (NSSP 2019, 8).

### **Machtwirkung: Normalisierung**

Entscheidend sind vor allem die Auswirkungen dieser Voreinstellung. Die Wirkungsweise lässt sich unter dem Schlagwort Normalisierung beschreiben: Im Effekt eine Verschiebung von Ressourcen und Aufmerksamkeit weg von strukturellen Aspekten und sozialen Determinanten von Gesundheit.

Mit der teil-automatisierten Überwachung des kollektiven Gesundheitsgeschehens durch ein Monitoring-Systems ist auch eine implizite Behauptung über dessen Normalzustand verbunden. Betriebsgrundlage einer algorithmischen Mustererkennung ist die Annahme eines funktionierenden, als unproblematisch zu erachtenden Normalzustands als Kontrastfolie. Die grundlegende Vorstellung von *Epidemic Intelligence*, nämlich dass sich Gefahren aus ungewöhnlichen Häufungen in den Daten erkennen lassen, setzt einen unproblematischen Normalzustand als Hintergrund voraus, von dem sich eine Häufung erst abheben kann.

Zudem, und hier kommt der Moment der Voreinstellung zum Tragen, versichert das System in seiner Anwendung beständig das Vorliegen dieses Normalzustands selbst. Wenn – wie an den meisten Tagen – von ESSENCE keine ungewöhnlichen Häufungen gemeldet werden, suggeriert das auch eine beruhigende Normalität. Wie gesehen konkretisiert sich die grundlegende Annahme der Datensammlung als Wis-

sensobjekt in der Nutzung. Es geht weniger um die Suche nach Auffälligkeiten in der Datensammlung als um die Datensammlung als Ausdruck des Zustands des kollektiven Gesundheitsgeschehens selbst. Die normalisierende Machtwirkung entfaltet sich vor allem dann, wenn keine automatisierten Meldungen ausgeflaggt werden, in der Suggestion eines unproblematischen Zustands.

Wie Kezia Barker herausstellt, bekommt mit der Einführung der Störung als ›Norm of norms‹ im Gegenzug gerade der Normalzustand als Abwesenheit solcher Störungen, wachsende Bedeutung. Mit der Etablierung solcher Systeme entsteht eine Disposition in den Gesundheitsbehörden, Aufmerksamkeit und Ressourcen auf die Norm zu richten, d.h. auf die Sicherstellung von Abwesenheit von Störungen zu fokussieren. Mit solchen Systemen entsteht die Möglichkeit und schließlich aber auch die Erwartung, die Abwesenheit von Störungen des Normalzustands festzustellen. Letztlich richtet sich ein wachsender Teil der verfügbaren Mittel der kollektiven Gesundheitspflege dann darauf, das Vorhandensein des Normalzustands zu versichern. Für eine aktive Verbesserung oder Änderung ist damit noch nichts erreicht. Die bloße Feststellung der Abwesenheit von Störungen wird mit den aufwändiger zu installierenden und zeitaufwändig zu bedienenden Monitoring Systemen hingegen immer kostspieliger. »Every moment of surveillance, every report of a sighting, every ›false positive‹ becomes a possible disruptive presence that requires intensive work to stabilise as an absence.« (Barker 2014, 4) Die Sicherstellung des Nicht-Ereignisses bindet Kraft, Geld und Aufmerksamkeit. Dadurch bekommen die nicht-ereignishaften, strukturellen Aspekte von Gesundheit im Gegenzug weniger Aufmerksamkeit. Strukturelle Determinanten von Gesundheit oder Gesundheit als ein nicht pathogenetisch definiertes Ziel geraten tendenziell aus dem Blick.

Permanente Gesundheitsbelastungen in der Stadt wie Lärm, Luftverschmutzung oder fehlender Zugang zu gesunder Ernährung erscheinen als Bestandteil dieses Normalzustandes ebenso wie chronische Krankheiten und lebensstil-bezogene Gesundheitsprobleme. Anders als der ›Verzicht auf Annahmen‹ suggeriert, ist mit der Implementierung eines datengetriebenen Monitorings eine grundlegende

Annahme bezüglich der öffentlichen Gesundheit (funktionierender Ist-Zustand) und von bearbeitungswürdigen Gesundheitsproblemen (störende Ereignisse) verbunden. Die Warnmeldung ist zumeist nur ein Indiz für weitere Recherche. Allerdings orientiert das Monitoring in dieser Form grundsätzlich auf den Kontrast zwischen Normalzustand und plötzlich auftretenden Störungen (siehe dazu auch Hyde u.a. 2006).

Die Nutzung des Monitorings impliziert in diesem Sinne eine Normalisierung. Anders als etwa in der von Michel Foucault am Beispiel der so genannten Disziplinargesellschaften prominent herausgearbeiteten Normalisierung geht es in diesem Fall wohlgerne nicht um das Aufprägen bestimmter erwünschter Verhaltensweisen. Normalisierung als Effekt der Verfahren automatisierter Mustererkennung bedeutet im Kontrast die Reduktion unterschiedlicher Fälle und Ausprägungen auf Modulationen eines einzigen grundlegenden Schemas. William Bogart spricht etwa von einer ›Supernormalisierung‹ durch Code, um den Kontrast zu der im Kontext der Disziplinargesellschaft beobachteten ›Normalisierung‹ zu markieren (Bogard 1996). »No longer conformity to a historically variable and continuously contested system of norms, but rather, if you will, production from and return to a singular, universal ›norm of norms.‹« (Bogard 1996, 30)

So wird im Zuge der Nutzung von syndromischen Monitoring das Faktum eines gesundheitlichen Vorfalls von statistischer Bedeutung als eine übergreifende Norm mit eingeführt. Jenseits der qualitativ sehr unterschiedlichen Syndrome erscheint der ›Meldungsfall‹ als eine übergreifende Kategorie. Die statistische Berechnung ermittelt für jede der zwölf Syndrom-Kategorien getrennt, ob es an einem bestimmten Ort zu einer außergewöhnlichen Häufung von Meldungen kommt. Die teils sehr unterschiedliche Brisanz einer solchen Meldung ist durch das System zunächst nivelliert. Eine Warnmeldung zeigt gleichermaßen alle ungewöhnlichen Häufungen in einer Kategorie an, gleich ob es dabei um Hautausschlag (*rash*), Lebensmittelvergiftung (*botulism-like*) oder Todesfälle im erwarteten Zusammenhang mit einer ansteckenden Infektionskrankheit (*severe illness or death potentially due to Infectious disease*) geht. Erst die Nutzer:innen qualifizieren die identi-

schen roten Flaggen je nach Syndrom-Kategorie anders und reagieren unterschiedlich darauf. Gemäß der *›norm of norms‹* des Systems werden nur Normalzustand und Vorfälle unterschieden.

Die hier verwendete analytische Strategie zielt auf die gesellschaftlichen Implikationen eines sozio-technischen Arrangements, indem relevante Kopplungen zum Ausgangspunkt genommen werden. Die dauerhafte Anordnung von Elementen ist weder rein zufällig noch trivial. Es bedarf bestimmter Gründe, um die benötigten Ressourcen zu kanalieren und eine solche Kopplung auf Dauer zu stellen. Dadurch sind solche Kopplungen ein fruchtbare Anknüpfungspunkt, um sich zugrundeliegende Vorstellungen vor Augen zu führen. Zugleich verändert eine solche Kopplung das Feld möglichen Handelns, erleichtert bestimmte Praktiken und macht andere weniger wahrscheinlich. Machtwirkungen der Infrastrukturierung gehen insbesondere von den Zwecken aus, die in der tatsächlichen Nutzung relevant gemacht werden.

Wie gesehen ist syndromisches Monitoring exemplarisch für die Vorstellung einer Datensammlung als Wissenobjekt. Allgemein unter dem Schlagwort *Big Data* und für den Bereich der Infektionskontrolle konkreter unter dem Schlagwort *Epidemic Intelligence* werden zunehmend solche Vorschläge zu einer Erkenntnisgewinnung *›aus den Daten‹* diskutiert. Statt die Datensammlung als Archiv von Messungen zu bestimmten Aspekten eines Wissenobjekts zu verstehen, machen es aktuelle technische Bedingungen möglich, die Datensammlung selbst als ein Substitut für das Wissenobjekt zu untersuchen. Syndromisches Monitoring ist als Beispiel für ein ereignisbezogenes Monitoring auch ein Beispiel für die dabei charakteristische epistemologische Verschiebung der Datensammlung zum Wissenobjekt selbst. Die zugrundeliegende Vorstellung beinhaltet eine solche Verschiebung im Verhältnis von Phänomen, Daten und Erkenntnis.

Aus den Gesprächen ist nun deutlicher geworden, in welcher Weise diese Vorstellung in der Praxis relevant gemacht wird: weniger als ein Sensorium für ungewöhnliche Ereignisse, sondern eher als eine nahe-Echtzeit Repräsentation des Ist-Zustands des kollektiven Gesundheitsgeschehens. In der epidemiologischen Praxis in den Gesundheitsämtern wird das System vor allem als ein Instrument der *situational awareness*

*reness* geschätzt. Für die Nutzer:innen ist an dem ereignisbezogenen Monitoring weniger das automatisch ausgeflaggte Ereignis relevant als vielmehr die Verfügbarkeit über den Normalzustand in den Daten. Der zentrale Gewinn ist die Bereitstellung einer Datenbank, in dem der Ist-Zustand des kollektiven Gesundheitsgeschehens abgelegt ist.

In der Verschränkung von Vorstellung und Nutzungen sind somit weitergehende Voreinstellungen festgelegt. Damit sind »vermittelte Zwecke« gefasst, die sich nicht deckungsgleich aus den programmatischen Zielen ergeben. Die Nutzung als Instrument einer *situational awareness* impliziert ein Aufgreifen der gesammelten Daten als ein ausreichendes Substitut für den Ist-Zustand des Gesundheitsgeschehens. »You just click a button and see« (Davies-Cole 2012, 00:24:04) ist nicht nur auf die Unmittelbarkeit bezogen, mit der dieser Zugriff erfolgt. Der Ausspruch verweist auch auf die Vorstellung, mit dem Mausklick das Objekt des Interesses sichtbar machen zu können. Eine Möglichkeit, die dadurch gegeben ist, dass das Objekt des Interesses nun in der Datensammlung selbst enthalten ist. In diesem Sinn ändert sich die Vorstellung der Datensammlung als Erkenntnisobjekt durch die Nutzung.

## **Gerichtetheit: Rekontextualisierung**

Von den Voreinstellungen unterscheide ich die Gerichtetheit einer Infrastrukturierung, um vor allem auch den Aspekt der Verschränkung der Materialität der angeordneten Elemente mit den Nutzungsweisen in den Blick zu bekommen. Neben der Verbindung von Vorstellungen und Nutzungen zu bestimmten Voreinstellungen, sind auch Pfadabhängigkeiten und Einstreuungen der relevant gemachten Elemente und Techniken entscheidend. Eine auf Dauer gestellte Anordnung und Kopplung von Elementen hat eine Reihe von Voraussetzungen und strukturiert zugleich andere Elemente (siehe Kapitel »Infrastrukturierung«). Die relativ dauerhaft angeordneten Elemente geben einen Rahmen vor, verändern das Handlungsrepertoire von Nutzer:innen und schaffen neue Routinen und Möglichkeiten. Machtwirkungen

einer Infrastrukturierung sind gerade auch mit dieser Gerichtetheit verbunden.

Im Folgenden untersuche ich die für die Infrastrukturierung charakteristische Kopplung von Syndrom-Kategorien und Mustererkennung im Hinblick auf diesen Aspekt der Gerichtetheit. Welche Rolle spielen die angeordneten Elemente und Techniken in ihrer Nutzung über die propagierten Zwecke hinaus? Welche Rückwirkungen hat diese Gerichtetheit auf die Auffassung und die Praxis von *Public Health* in dem untersuchten Fall?

Ausgehend von der Kopplung von Syndrom-Kategorien und statistischen Verfahren ist die besondere Relevanz der Datensammlung deutlich geworden. Die Aufwertung zum Wissensobjekt und die normalisierende Machtwirkung wurden artikulierbar. Mit der Frage nach der Gerichtetheit verschiebt sich der Fokus auf den Effekt der Hervorbringung zwischen den Elementen. Die Annahme eines solchen ›Hervorgebrachtwerdens‹ der Elemente erweist sich in diesem Zusammenhang für eine kritische Analyse besonders nützlich. Zentraler Teil der Infrastrukturierung ist gerade eine als objektiv behauptete Sammlung von Daten. Es lohnt sich angesichts dies ›Objektivitätsvermutung‹, einen Standpunkt für eine kritische Metaperspektive zu haben. Die Heuristik der Infrastrukturierung erlaubt dies.

In der kritischen Auseinandersetzung ist eine allzu naive Annahme bezüglich der Objektivität von Daten bereits mit einer Reihe von Argumenten hinterfragt worden. Daten sind niemals roh, sondern von den Bedingungen der Messung, Kategorisierung und Speicherung beeinflusst (Gitelman und Jackson 2013). So listet etwa Rob Kitchin eine Reihe von möglichen Verzerrungen auf, die den Status von Daten als gespeichertes Resultat einer rein objektiven Messung in Frage stellen:

»[Data] are both a representation and a sample, shaped by the technology and platform used, the data ontology employed, the regulatory environment, and are subject to sampling bias.« (Kitchin 2014, 4)

Machtwirkungen der Infrastrukturierung ergeben sich vor allem auch aus Art und Weise, in der die Daten als Elemente hervorgebracht werden. Die prozedurale Perspektive auf die Verschränkung von Elementen

und Nutzung lenkt den Blick genau auf die vielfältigen Voraussetzungen, die ein vermeintlich objektiv gegebenes ›Datum‹ [lat. ›das Gegebene‹] erst möglich machen. Aus dieser Perspektive steht im Vordergrund, inwiefern Daten vor allem auch von den Strukturen, Mitteln und Praktiken ihrer Erhebung, Sammlung und Zusammenstellung geprägt sind. Die Anstrengungen zur Fixierung und Stabilisierung von Aspekten eines Phänomens als eine abrufbare Datensammlung ist in diesem Sinne erheblich voraussetzungsvoll (Beer 2016). Zugleich hat die Übersetzung in eine Datensammlung Rückwirkungen auf die Art und Weise des Umgangs mit einem Phänomen. Die als Daten fixierten Aussagen wirken zurück auf die Konstitution von Problemen und damit auf die Möglichkeit ihrer Bearbeitung.

»[D]ata demand and build the human, organizational, and infrastructural worlds around them – enforcing a burden of care and work that disappears beneath (but ultimately, constitutes) [their] possibilities.« (Ribes und Jackson 2013, 164)

Mit der Perspektive auf die Hervorbringung der Daten treten insbesondere Momente der De- und Rekontextualisierung in den Vordergrund. Das Abspalten und Anreichern von Kontext ist für die Überführung gesundheitsbezogener Informationen zu systematisch verarbeitbaren Daten zentral. Unter Einbezug der Nutzungen wird hier zudem die Bedeutung einer ›freihändigen‹ Rekontextualisierung durch die Nutzer:innen deutlich.

### **Chimäre immaterieller Information**

Grundsätzlich bedeutet die Nutzung des Systems eine Behandlung der gespeicherten und aggregierten Informationen als wahrheitsfähige, objektive Daten. In der Nutzung werden die Daten im Sinne ›immaterieller Information‹ verwendet, wie es Martin French ausdrückt. Ähnlich wie die elektronische Patientenakte, die French untersucht hat, basiert auch das Versprechen des syndromischen Monitoring auf der Vorstellung, Informationen als etwas Immaterielles behandeln, transportieren und aggregieren zu können. Allerdings gerät

mit dieser Vorstellung aus dem Blick, inwiefern eigentlich vielfältige Formen der Säuberung, Abstraktion und Übersetzung nötig sind, um die gesundheitsbezogenen Informationen in solche entmaterialisiert gedachten Datensätze zu übertragen. French verweist entsprechend auf die »questionable assumption that IT-mediated information is immaterial. Such assumptions have the effect of eliding the amount of work involved in translating information from one context to another« (French 2009, 31).

Für das Funktionieren ist syndromisches Monitoring grundsätzlich auf eine solche Stabilisierung der Information über unterschiedliche Kontexte hinweg angewiesen. Eingängig spricht Bruno Latour hier von »immutable mobiles« um diesen typischen Moment des Sozialen auszudrücken. Bei allen möglichen Gegenständen bedarf es einer solchen Doppelbewegung von Anstrengungen der Stabilisierung, um über Kontexte hinweg transportierbar zu sein (Latour 1987). Der Vorschlag Latours, diesen Moment ins Zentrum zu stellen, ist eine der formierenden Ideen der so genannten Actor-Network Theorie gewesen.

»This concern with transport and the work of holding configurations together and in shape – with so-called »immutable mobiles« [...] – was to lead to what became known as actor-network theory.« (Law und Mol 2001, 611)

Auch mit der Heuristik der Infrastrukturierung rücken solche Fragen des Hervorbringens von Elementen und Anstrengungen zu ihrer Stabilisierung in den Vordergrund. Die Frage nach der Kopplung von Elementen ist auch die Frage nach den Anstrengungen die nötig sind, um die Elemente über Kontexte und über die Zeit hinweg in Beziehung setzen zu können. Analog zu dem Ansatz der ANT verschiebt sich der Fokus von den Elementen selbst zu den Bedingungen der Möglichkeit ihres Stabil-bleibens. Der Gewinn ist eine stärkere Aufmerksamkeit für die Verluste und Verzerrungen, die mit einer solchen Fixierung eines Elements als Vorbedingung der Kopplung einhergehen. Im konkreten Beispiel geht es um die Hervorbringung und Stabilisierung der Informationen als eine dekontextualisierte Sammlung von Daten. Dazu werden die Informationen als relativ einfach zu transportierende Daten

stabilisiert, denen der jeweilige Kontext der Erhebung – etwa in einer bestimmten Notaufnahme oder an der Kasse eines bestimmten Drogenmarkts – abgespalten wurde. Die Behandlung der Datensammlung als eine Repräsentation des Ist-Zustands des kollektiven Gesundheitsgeschehens bedeutet vor allem auch – gemäß dem Begriff von French – die Suggestion immaterieller Information. Diese Suggestion ist für das Funktionieren der syndromischen Überwachung entscheidend.

»[T]he health threat is narrowed from a wide variety of uncontrollable events to a series of numbers that can be transmitted without change and one that refers to the same thing regardless of context.« (Cakici und Sanches 2014, 405)

Die Überführung des komplexen Phänomens kollektives Gesundheitsgeschehen in eine hinreichend repräsentative Datensammlung bedarf einer ganzen Reihe von Reduzierungen, etwa im Hinblick auf die überhaupt einbezogenen Datenquellen, bei der Festlegung des Formats der Speicherung, bei der Bereinigung von fehlerhaft registrierten Freitext-Kommentaren, oder bei der Festlegung der Syndrom-Kategorien, auf die hin aggregiert wird etc.

Abstrakter lassen sich diese vielfältigen Herausforderungen als Bemühen um die Hervorbringung von *immutable mobiles* beschreiben, also von über unterschiedliche Kontexte hinweg stabilen und transportfähigen Einheiten. Die einzelnen Schritte der Überführung, z.B. die Aussage einer Patientin in der Notaufnahme des *Kaiser Permanente Harbor Medical Center* in Maryland in einen Datensatz auf dem ESSENCE-Sever, beinhalten jeweils ein Abspalten von Kontextinformationen. Das komplexe körperliche Wohlbefinden der Patientin muss für die Möglichkeit der Erfassung und Weitergabe reduziert werden, in dem Fall auf demographische Informationen einerseits, sowie auf *chief complaints* und später auf die ärztliche Diagnose in Form von ICD-10 Codes andererseits. Unterschiedliche *chief complaints* werden sodann als Ausprägungen einer der zwölf Syndrom-Kategorien vereinheitlicht etc.

Wichtig für die Akzeptanz der so angefertigten Datensammlung als eine hinreichende Repräsentation ist es, dass diese Kette der Reduktionen auch zurückverfolgt werden kann. Von einem Syndrom, das in ei-

nem Bezirk als besonders hervorstechend angezeigt wird, können die Nutzer:innen sich die einzelnen Vorfälle anzeigen lassen. Der einzelne Vorfall wiederum kann bis auf Ort und Zeit hin zurückverfolgt werden, an dem die Aufnahme in das System erfolgt ist. Über einen Telefonanruf dort kann schließlich die einzelne Patientin identifiziert werden etc. »you know, the age, you know, the ZIP code. So we can call the hospital, say who was this patient?« (Russell 2012, 00:33:57)

Die Bedeutung einer Rückverfolgbarkeit hat Latour in seinen Ausführungen ebenfalls deutlich unterstrichen. Das Beispiel der sukzessiven Übertragung von komplexen gesundheitlichen Vorgängen in der Bevölkerung in eine Datensammlung versinnbildlicht genau jene ›zirkulierende Referenz‹ die Latour etwa anhand der Übersetzung von Bodenproben in eine wissenschaftliche Aussage verdeutlicht. Auf ein Objekt wird in einer Referenz Bezug genommen. Und diese Referenz ist im nächsten Schritt wieder ein Objekt für eine weitere Referenz (Latour 2002). Diese Figur der zirkulierenden Referenz ist nach Latour generell eine gute Beschreibung dafür, wie aus Beobachtungen wissenschaftliches Wissen formiert wird: kein einziger großer Schritt von Gegenstand zu Geist, sondern eine verkettete Serie. Die ›Wahrheitsfähigkeit‹ dieser Kette hängt nun entscheidend davon ab, dass diese Verkettung auch zurückverfolgt werden kann.

»Wichtig ist, daß diese Kette *reversibel* bleibt. Die Nachvollziehbarkeit der Schritte muß es im Prinzip erlauben, sie in beiden Richtungen auszuführen. Unterbricht man sie an irgendeinem Punkt, so ist auch der Transport, die Produktion, die Konstruktion, gewissermaßen die Leitfähigkeit des Wahren unterbrochen.« (Latour 2002, 85)

Beide Momente dieser Idee der zirkulierenden Referenz sind auch für die untersuchte Infrastrukturierung prägend. Zum einen kommt es zu einer solchen Verkettung von Referenzen, zu einer sukzessiven Abspaltung von Kontextinformationen bei den einzelnen Schritten der ›Datafizierung‹ des kollektiven Gesundheitsgeschehens. Zum anderen wird über bestimmte Vorkehrungen des Systems – bestimmte Filter oder Auswahlmöglichkeiten über die Kontaktinformationen der jeweiligen Notaufnahme, an der ein Datensatz registriert wurde etc.

- die Rückverfolgbarkeit gewährleistet. Die Wahrheitsfähigkeit der Datensammlung wird durch die ›Leitfähigkeit des Wahren‹, wie es Latour ausdrückt, bedingt: die Möglichkeit der Bewegung in beide Richtungen der Referenzkette ist sichergestellt.

Gleichwohl ist die Reduzierung von Beobachtung zu wissenschaftlichem Wissen selbst in Form einer verketteten Referenz niemals bruchlos möglich. Die Glieder der Kette sind durch einen »gemeinsamen Operator [verbunden] der sich vom folgenden Schritt durch einen Bruch unterscheidet, durch ein *gap*, das durch keinerlei Ähnlichkeit überbrückt werden kann.« (Latour 2002, 84) Die Metapher des Bruchs, die Latour hier verwendet, verweist auf einen Teil des Kontexts, der bei den einzelnen Schritten irreversibel entfällt.

Ein erster solcher *gap* entsteht bereits an der Stelle, an der das kollektive Gesundheitsgeschehen gemessen wird. Trotz der charakteristisch breiten Integration von diversen Datenquellen bleiben hier notwendig eine Reihe von Lücken. Eine Reihe von Vorfällen, die relevant sind, haben gar keine Chance, als Datensatz registriert und für das Monitoring sichtbar zu werden. Vorfälle, die weder mit einem Einkauf in einem der angebundenen Drogeriemärkte, noch mit einem abwesenden Schulkind, noch mit einem Besuch der Notaufnahme verbunden sind, bleiben systematisch unsichtbar. Ein Interviewpartner verdeutlicht das am Beispiel von hitzebedingten Todesfällen. Von Seiten der Politik gibt es ein formuliertes Interesse an dieser Gefährdung, das Monitoring ist allerdings in dieser Richtung blind.

»We had an elderly person die of heat, heat exhaustion. And they were like, well, did you see it? Did you see it coming? I didn't see it indeed. They never took them to the ER [Emergency Room]. So if they don't take him to the ER, I don't know. As far as this screen is concerned, it doesn't happen, you know, so it [ESSENCE] has big gaps.« (Russell 2012, 00:57:04)

Bei der Übersetzung des Gesundheitsgeschehens in die Form von transportierbaren und aggregierbaren Daten kommt es zu einer ersten grundlegenden Reduzierung des Informationsreichtums bereits durch die blinden Flecken der einbezogenen Meldestellen. Ein bestimmter

Teil des kollektiven Gesundheitsgeschehens hat systematisch keine Chance, als Datum erfasst und damit Teil der zirkulierenden Referenz des syndromischen Monitorings zu werden.

Selbst im Fall von registrierten Vorfällen kommt es für die Möglichkeit des Transports über Kontexte hinweg zu einer Dekontextualisierung, die eben nur begrenzt reversibel ist. Auch in der weiteren Aggregation der Daten werden bestimmte Kontextinformationen unwiederbringlich abgespalten.

Verdeutlichen lässt sich das am Beispiel der Notaufnahmen. Wie gesehen stammen die Daten für das System letztlich vor allem aus der täglichen Meldung der Notaufnahmen der angeschlossenen Kliniken. Dieser Kontext der Datenerhebung hat einen entscheidenden Einfluss auf das, was überhaupt Eingang in das System findet. Allerdings ist die Notaufnahme an sich durch die institutionelle Struktur der Gesundheitsversorgung, durch Zugangsbarrieren und unterschiedlich verteilte gesellschaftliche Möglichkeiten eine von vornherein verzerrte Linse auf das kollektive Gesundheitsgeschehen. Durch die Funktion der Notaufnahme in dem institutionellen Gefüge der gesundheitlichen Versorgung in den USA findet hier eine grundsätzliche Verzerrung statt. So stellt die Notaufnahme für Patient:innen mit geringem sozio-ökonomischen Status häufig den einzigen möglichen Zugang zu einer formalen Gesundheitsversorgung dar. Knapp 28 Millionen Menschen im arbeitsfähigen Alter waren 2018 in den USA ohne Krankenversicherung. Überwiegend sind das Personen ohne festes Einkommen, oder mit geringem Verdienst (Tolbert u.a. 2019). Ein niedriger sozio-ökonomischer Status bedingt eine messbar geringe Nutzung von Angeboten der gesundheitlichen Versorgung und Früherkennung (Fiscella u.a. 2000); teils auch dann, wenn ein fehlender Versicherungsschutz durch staatliche Programme wie Medicare oder Medicaid ausgeglichen werden könnte. Es ist anzunehmen, dass selbst ein erheblicher Teil von eigentlich antragsberechtigten Personen auf diese Möglichkeit verzichtet, aus Scheu vor dem verbundenen bürokratischen Prozedere (Adler und Newman 2002).

Faktisch stellt der Arztbesuch für Personen mit geringem sozio-ökonomischen Status eine hohe finanzielle Hürde dar und wird als

Möglichkeit der Gesundheitsversorgung selten genutzt. Stattdessen wird die zunächst kostenfreie Behandlung in der Notaufnahme der Krankenhäuser in solchen Fällen zur ersten Anlaufstelle (Kite-Powell, Ofori-Addo und Hamilton 2010). Das syndromische Monitoring registriert durch diesen Umstand überproportional viele Vorfälle milder Symptome aus einem bestimmten sozio-ökonomischen Personenkreis.

»Certain people use the E.R. as their first line of defense because of socioeconomic issues. [...] So I see often people going in with their chief complaint is that they have a fever and they have diarrhea. I don't want to go to the ER for that.« (Russell 2012, 00:45:47)

Sozio-ökonomisch benachteiligte Personen werden hier erwartungsgemäß überproportional vorstellig (»it is their first line of defense«) und entsprechend auch mit dem gesamten Arsenal möglicher Beschwerden von sehr milden gesundheitlichen Symptomen bis hin zu lebensbedrohlichen Vorfällen. Solche Effekte der strukturellen Einbettung des Messinstruments »Notaufnahme« sind in Rechnung zu stellen, um die Relevanz eines an diesem Messpunkt registrierten Vorfalls für das kollektive Gesundheitsgeschehen tatsächlich zu entscheiden. Allerdings gibt es in dem System keine Vorkehrungen, dieses relevante Kontextwissen reversibel zu halten. Mit der Transformation der Aussagen der Patient:innen in die Freitext-Felder des Datensatzes ist der sozio-ökonomische Status abgespalten. Die im System registrierten personenbezogene Attribute sind demographische Merkmale (Alter, Geschlecht, ethnischer Hintergrund) (Office of Information and Regulatory Affairs 2020). Der Versicherungs-Status oder Angaben zur sozio-ökonomischen Lage werden in der Meldung nicht erfasst. So auch die Klage eines Gesprächspartners: »I can give you demographic information about their age, sex, but I can't tell you anything about [...] insurance.« (Russell 2012, 00:12:40)

Wie gesehen ermöglicht der im System gespeicherte Aufnahmezeitpunkt der Daten den Nutzer:innen einen Umweg, um die Vorfälle durch einen Anruf in der betreffenden Klinik händisch zu repersonalisieren. »If it's a public health event of significant concern, we have the authority to step in. They could give us the name« (Russell 2012, 00:12:38). Das ist

aber nur in Ausnahmefällen gangbar. Zudem wäre damit allenfalls der Versicherungsstatus in Erfahrung zu bringen, die sozio-ökonomische Lage hingegen auch dann nicht. Ein bestimmter Aspekt des Kontexts der erhobenen Daten ist von vornherein abgespalten und bleibt irreversibel entfernt.

### Händische Rekontextualisierung

Der hier gewählte Ansatz erlaubt eine entscheidende Erweiterung. Die Hinwendung zu Hervorbringung von sozio-technischen Einrichtungen in der Verschränkung von Elementen und Nutzungen öffnet den Blick für entscheidende weitere Wirkungen. Das System wird hier nicht allein aus seiner programmatischen Funktion heraus beschrieben, sondern in der Verschränkung von relativ dauerhaften Anordnungen und Nutzungen betrachtet.

Vor dem Hintergrund dieses offeneren Rahmens wird ein bestimmter Vorgang einer ›Anreicherung mit Kontext‹ durch die Nutzer:innen sichtbar, der typisch für Verfahren der automatisierten Erkenntnisgewinnung erscheint, aber häufig nicht explizit betrachtet wird. Es geht dabei um eine quasi ›freihändige‹ Rekontextualisierung durch die Nutzer:innen, jenseits der verketteten Referenzen. Diesen Aspekt betont bereits Eugene Thacker in einer frühen Auseinandersetzung mit syndromischem Monitoring.

»Despite the technophilic quality of many biosurveillance projects, their most interesting network properties come not from the ›automated detection systems,‹ but from the ways in which a multiplicity of human agencies produces a intentional yet indeterminate aggregate effect.« (Thacker 2005, 8)

Die Aussagen über das kollektive Gesundheitsgeschehen, die mit Hilfe der Infrastrukturierung nun möglich werden, sind entscheidend durch solche Kontexte geprägt, welche den Daten bei der Benutzung wieder beigefügt werden. Nutzer:innen sind sich häufig sehr wohl über die Beschränkungen der Datenerhebung und der teils irreversiblen Abspaltung von Kontext bewusst. Wie oben gesehen, werden die von dem Sys-

tem automatisch generierten Warnungen deswegen auch mit einer gewissen Skepsis betrachtet. Allerdings bietet das System durchaus die Möglichkeit – so die Aussage in den Gesprächen, einer Sache auf den Grund zu gehen. »The usefulness is when an event has occurred, you can use it to monitor the cause of the event« (Davies-Cole 2012, 00:09:33) Der hauptsächliche Nutzen besteht in der Möglichkeit zur Nachforschung, meist im Anschluss an eine durch das System erhaltene, oder von außen eintreffende Warnung.

Für diesen Zweck werden die gemeldeten Datensätze durch unterschiedliche Strategien validiert. Diese Entscheidung, inwiefern der im System gemeldete Ausreißer ein relevantes Signal ist, oder nur ein zufälliges Rauschen bedeutet, macht einen zentralen Aspekt der praktischen Verwendung aus. Die Nutzer:innen greifen dazu auf die Freitext-Suche zurück, um einer bestimmten Vermutung nachzugehen oder lassen sich die gesamten Daten einzelner Notaufnahmen anzeigen, um die Güte der Daten einzuschätzen. Datensätze aus unterschiedlichen Quellen werden bewusst verschnitten, um auffällige Ballungen an einem bestimmten Ort in einen Kontext zu setzen. »And especially if I can see it in other data sources, then that really completes the picture.« (Russell 2012, 00:42:35) Oder Meldungen von anderen Orten werden als Kontext hinzugezogen. »If there is actually something going on here, I would also want to know if other places are seeing it in this way.« (Davies-Cole 2012, 00:20:34).

Für alle diese Strategien der Validierung sind bestimmte Hypothesen und Annahmen der Nutzer:innen entscheidend. Sowohl das fachliche Vorwissen ihre Erfahrung und ihre Reflexion über die Verzerrungen und Beschränkungen der im System abgelegten Daten spielt an dieser Stelle eine Rolle. Letztlich imaginieren die Nutzer:innen die Daten erneut in einem Kontext und überprüfen diese Annahme. In der aktiven Benutzung für den Zweck der *situational awareness* kommt es zu einer Wiederanreicherung der Daten mit Kontext durch die Nutzer:innen. Beispielsweise wird ein:e Epidemiolog:in in Utah die Spitzen von gemeldeten Abwesenheiten in den Schulen am ersten und letzten Tag der Jagdsaison mit einer aus Erfahrung bezogenen Gelassenheit betrachten (Livnat u.a. 2013).

Diese Rekontextualisierung ist als notwendiger Bestandteil einer *Datafizierung* selten explizit thematisiert. Das Wiederanreichern mit Kontext ist eine notwendige Folge der Überführung von gesundheitsbezogenen Informationen in stabilisierte aber daher auch entkontextualisierte Daten. Dieses Wiederanreichern mit Kontext soll und kann Blindstellen und Verzerrungen ausgleichen, die bei dieser Operation entstehen. Zugleich ist die Wiederanreicherung aber auch selbst eine eigene Quelle für Verzerrungen.

Ein Beispiel aus dem Material soll diese Problematik greifbar machen: Den Nutzer:innen in den Gesundheitsbehörden ist die Überbetonung von milden Symptomen aufgrund der strukturellen Situation der Notaufnahmen als erste Anlaufstelle der Gesundheitsversorgung durchaus bewusst. Um die tatsächliche Relevanz solcher typischerweise verzerrter Daten einzuschätzen, greifen sie als Korrektiv meist auf andere zentrale Datenquellen zurück, oftmals auf die Verkaufszahlen der großen Drogeriemarkt-Ketten. Hier werden bestimmte Annahmen leitend. Auch aus eigener Erfahrung erscheint es für die Nutzer:innen plausibler, das Vorfälle von Lebensmittelvergiftung oder einer infektiösen Durchfallerkrankung nicht in der Notaufnahme auftauchen, sondern eher in einem massiven Anstieg des Bedarfs an bestimmten Medikamenten sichtbar werden. Die Datensätze der Verkaufszahlen werden damit als der verlässlichere Indikator angesehen, wie ein Gesprächspartner deutlich macht. »You know, I've had food poisoning. I never went to the ER. So I am assuming that medication sales should alert first.« (Russell 2012, 00:47:47)

Durch die Verschneidung der Daten, in dem Fall die bewusste Privilegierung der Drogeriemarkt-Daten für bestimmte Aussagen, wird aktiv eine der reflektierten Verzerrungen bei der Erhebung der Daten ausgeglichen. Allerdings werden die Daten dabei auch als kontextualisiert vorgestellt und entsprechend ernster genommen (Medikamenten-Verkäufe) oder als weniger bedeutsam erachtet (Notaufnahme-Daten). Das Herantragen von Kontext an die Daten ist in diesem Fall selbst eine Quelle von Verzerrungen. Das Suchraster, das hier zur Anwendung kommt, ergibt sich im Grunde aus dem eigenen gesundheitlichen Verhalten der regelmäßig gut gebildeten und zumeist fest an-

gestellten Mitarbeiter:innen in den Gesundheitsbehörden. Lebensmittelvergiftung, so die Erwartung, wird vor allem als unerwarteter Anstieg der Verkäufe des Medikaments Pepto-Bismol sichtbar werden. Allerdings ist diese Vorstellung vor allem aus einer bestimmten sozio-ökonomischen Position her gedacht. Für Menschen in einer anderen strukturellen Lage stimmt die zugrundeliegende Erwartung häufig gar nicht. In bestimmten Wohngegenden ist der Zugang zu Drogerimärkten möglicherweise erschwert: »Not a lot of pharmacies in the area« (Russell 2012, 00:50:44). Zudem legen kulturelle Gepflogenheiten und Traditionen bei einem bestimmten Personenkreis in bestimmten Fällen ganz andere Formen der Behandlung nahe. »They may use homeopathic things like >grandma always made me this<. And that's what they do.« (Russell 2012, 00:50:48) Es kommt trotz Gesundheitsproblem dann gar nicht zu einem Kauf von Pharmazeutika, den das System registrieren könnte. Das System ist für solche Fälle dann generell blind, oder die Notaufnahme wäre der Ort einer Registrierung solcher Vorfälle. Durch den Kontext, vor den die Nutzer:innen des Systems die Daten in diesem Fall stellen, können hier Vorfälle unsichtbar werden, selbst wenn das System solche Fälle registriert und die automatisierte Durchsicht der Syndrome eine Warnung ausgegeben hat. Ein Gesprächspartner stellt diese Verzerrung durch die Situiertheit der auswertenden Person explizit als ein *>major mistake<* heraus.

»So assuming that medication sales should alert first, I think people who were like me, who are looking at this, they say, OK, well, I didn't see over-the-counter medications sales alert. So I'm going to be a little bit dismissive of this. I think that's a major mistake.« (Russell 2012, 00:47:47)

Die Überführung des kollektiven Gesundheitsgeschehens in die Form aggregierbarer, speicherbarer und berechenbarer Daten bringt als typischen Effekt einer Datafizierung auch eine Entkontextualisierung der Daten mit sich. Die Auseinandersetzung mit der Nutzung macht zusätzlich deutlich, inwiefern damit auch die Notwendigkeit einer situativen Rekontextualisierung der Daten in bestimmten Nutzungssituationen einhergeht. Die dekontextualisierten Daten machen eine Validie-

rung notwendig, bei der vor allem auch die impliziten Vorannahmen der Nutzer:innen zentrale Relevanz bekommen.

Die detaillierte Aufschlüsselung der beiden unterschiedlichen Momente des Abspaltens und Wiederanreicherns der Daten mit Kontext – zum einen auf Basis der reversiblen Referenzketten, zum anderen auf Basis der Erfahrungen und Kompetenzen der Nutzer:innen – verweist auf ein typisches Muster, das über das konkrete Fallbeispiel hinaus von Bedeutung sein kann.

In allen möglichen Bereichen gewinnen derzeit datenbasierte, automatische Verfahren zur Sortierung und Mustererkennung an Relevanz.

»Operations, decisions and choices previously left to humans are increasingly delegated to algorithms, which may advise, if not decide, about how data should be interpreted and what actions should be taken as a result.« (Mittelstadt u.a. 2016, 1)

Allerdings findet die kritische Auseinandersetzung häufig strikt entlang der Funktionsweise der Systeme selbst statt. Der Fokus liegt auf den Unzulänglichkeiten und Verzerrungen bei der Erhebung der Daten, auf den Analyseverfahren und der Reichweite der dadurch gewonnenen Aussagen und schließlich bei den Effekten solcher Systeme automatisierter Entscheidungsvorbereitung. Häufig kommt es dabei zu einer übermäßigen Betonung der Eigenständigkeit automatischer Systeme oder der Behauptung einer faktischen Übernahme von Entscheidung durch Algorithmen. Ein solch isolierter Blick auf die Ebene automatischer Verfahren selbst, verdeckt aber die Bedeutung der Verschränkung mit menschlicher Praxis.

»If human and machinic capabilities and functions are distributed in new sensor/processor/actuator affiliations, then distinctions such as ›automatic‹, ›autonomous‹ or ›unmanned‹ are misleading due to their imposition of absolutes.« (Crandall 2010, 83f)

Die hier verfolgte Perspektive der Infrastrukturierung zielt unter anderem darauf, eine solche ›absolute‹ Hinwendung zu Werkzeugen entweder als eigenständig oder als passiv zu vermeiden. Erst in der Ver-

schränkung von Systemen und Nutzungsweisen lassen sich die tatsächlichen Implikationen erfassen.

## Syndromisches Verstehen

In der Auseinandersetzung mit der zweiten symptomatischen Koppelung von Syndrom-Kategorien mit Verfahren der Mustererkennung werden weitere Auswirkungen des Systems auf die *Public Health* Praxis bestimmbar. Hier werden vor allem Momente der Normalisierung sichtbar, die mittels der Infrastrukturierung etabliert werden.

Eine Normalisierung in mindestens zwei Bezügen: Zum einen ist das epistemologische Konstrukt, auf dem die Mustererkennung ›aus den Daten‹ aufgebaut ist, grundsätzlich auf die Trennung eines Normalzustands von den zu identifizierenden Ereignissen abhängig. Auch wenn die Nutzer:innen der automatisierten Erkennung nicht blind vertrauen, so ist die Gegenüberstellung von einem unproblematischen Normalzustand und den zu identifizierenden Ereignissen mit dem System grundlegend gesetzt. Erneut bedeutet die Infrastrukturierung hier die Förderung einer *Public Health* Strategie, die in Richtung der Bewahrung eines gesundheitlichen Ist-Zustands zu intervenieren versucht. Ansätze einer proaktiven Gesundheitsförderung, etwa eine Sorge für die Verringerung gesundheitlicher Verwundbarkeit durch eine eher strukturelle Intervention, werden in dieser ereignisfixierten Perspektive auf Gesundheit zweitrangig. Das System verkörpert in diesem Sinn das grundlegende Muster, das Allen Feldbaum im Anschluss an Jacques Rancière als postpolitische Form ausmacht: die Trennung zwischen dem dystopischen Ereignis und dem funktionierenden Normalzustand. Dem ›Nicht-Ereignis‹ wird dadurch implizit Richtigkeit und Funktionsfähigkeit unterstellt und es geht letztlich vor allem um die Bewahrung dieses, nun perfekter erscheinenden, Normalzustands. ›[T]hese public safety wars are not wars of utopia, but wars of dystopia that assume that ›perfected‹ liberal democracies are threatened by an invisible, infiltrating menace.‹ (Feldman 2004, 332)

Ein zweites Moment der Normalisierung betrifft zum anderen die Art und Weise, in der die Informationen des Systems in handlungsrelevante Erkenntnisse übersetzt werden. Wie oben detailliert auseinandergesetzt, wird die für das System notwendige Komplexitätsreduktion bei der Erhebung, Aggregation und Auswertung von einer aktiven Anreicherung mit Komplexität durch die Nutzer:innen begleitet und kompensiert. Praxisrelevant werden die Daten erst nach einer zusätzlichen Anreicherung mit Kontext, den die Nutzer:innen zum Ausgleich reflektierter Lücken und Verzerrungen den prozessierten Daten wieder zuschreiben.

Bei dieser Anreicherung kommen vielfältige bewusste und unbewusste Annahmen zum Tragen. Fragen von Klasse, Positionalität und sozialer Hierarchisierung spielen hier eine entscheidende Rolle. Die Positionalität, aus der heraus diese Zuschreibungen erfolgen, ist dabei häufig eng begrenzt. Die Möglichkeit zur Nutzung und Interpretation, beispielsweise des ESSENCE-Systems, setzt eine Anstellung als Epidemiolog:in in den Gesundheitsbehörden der *Counties*, der Bundesstaaten oder an den CDC voraus. Dadurch sind einige Hürden in Bezug auf Bildung und sozialen Status gesetzt, die den Kreis der Nutzer:innen einschränken.

Hier wird das zweite Moment der Normalisierung sichtbar. Einerseits geben die getroffenen Vorkehrungen zur Stabilisierung von Referenzketten und zur Ermöglichung der ‚Leitfähigkeit des Wahren‘ in solchen strikt automatisierten und datenbasierten Systemen den Nutzer:innen hinreichend Sicherheit, in der Datensammlung eine Repräsentation des kollektiven Gesundheitsgeschehens in seinem Ist-Zustand zu sehen. Das trifft noch mehr auf Außenstehende zu, die mit den Aussagen des Systems konfrontiert werden. Andererseits wird hier auch die Bedeutung von Rekontextualisierung sichtbar, die solche Systeme in der praktischen Anwendung immer auch erfordern. Verzerrungen sind stets händisch zu korrigieren oder vermutete Artefakte bzw. falsche Positive zu ignorieren etc. Diese händische Korrektur bedeutet häufig eine Normalisierung – eine Interpretation aus einer spezifischen Positionalität heraus. Eine Häufung von verdauungsbezogenen Symptomen wird auch im Hinblick auf das eigene Verhalten mit

den Verkaufszahlen bestimmter Medikamente validiert. Meldungen über Fehlzeiten in bestimmten Schulen können hingegen, auch vor dem Hintergrund von Zuschreibung bestimmter ›Problemschulen‹, vernachlässigt werden. »You have certain schools especially certain high schools. They're always above 10 percent absenteeism.« (Russell 2012, 00:27:09)

Erst auf Basis solcher Erfahrungen werden die Daten handlungsrelevant oder eben nicht. Erfahrung und Sachkunde auf der einen Seite, aber auch »unconscious motivations, particular emotions, deliberate choices, socio-economic determinations, geographic or demographic influences« (Hildebrandt 2011, 376) auf der anderen Seite werden relevante Faktoren für die ›objektiven‹ Meldungen des Systems.

Ruha Benjamin spricht in Anlehnung an die so genannten Jim-Crow-Gesetze von einem ›New Jim Code‹, um auf die neue soziotechnische Komponente hinzuweisen, mit der zunehmend bestehende soziale Hierarchien auf Dauer gestellt oder verstärkt werden (Benjamin 2019). Ihr geht es vor allem darum, dass bestimmte Herrschaftseffekte in Gestalt vermeintlich objektiver und unbeteiligter technischer Verfahren schwieriger wahrzunehmen sind.

Ein soziales Bias automatisierter Systeme wird zunehmend problematisiert (Eubanks 2019) und das hier untersuchte technische Verfahren des Gesundheitsmonitoring macht beispielhaft einen der dabei wirksamen Mechanismen greifbar. Wie das Beispiel zeigt, bleibt bei automatisierten Systemen der Kategorisierung, Mustererkennung und Entscheidungshilfe in der Praxis regelmäßig die Notwendigkeit, Mängel der Daten aktiv auszugleichen. Im Zuge der so genannten ›Big Data‹ nimmt auch die ›Struppigkeit‹ der Daten zu, durch die wachsende Zahl diverser Datenquellen und den nicht einheitlichen Schemata, nach denen die Daten kategorisiert sind. Die Positionalität der Nutzer:innen bekommt hier häufig relevanten Einfluss in den Prozessen. Zugleich bekommen die Ergebnisse durch das ›Wahrheitsversprechen‹ von automatisierten, datenbasierten Verfahren eine besondere Autorität.

