Definitionsmacht und Selektivität in Zeiten neuer Kontrolltechnologien

Jan Wehrheim

Die *Critical Surveillance Studies* boomen seit einigen Jahren, obwohl ihnen – oder gerade deswegen – nach wie vor eine klare Definition ihres Gegenstandes ‚Überwachung' fehlt. Ihr Kernproblem begründet dabei gleichzeitig ihre Popularität und politische Reichweite: Überwachungsdiskussionen sind normative Diskussionen. Nicht das wie auch immer definierte Überwachen gilt per se als Problem (bei Patienten auf Intensivstationen etwa wird es nicht problematisiert, ebenso wenig, wenn es um die Herkunft von Rindfleisch geht), sondern wer wen zu welchem Zweck überwacht oder wozu Alltagstechnologien wie Handy und Internet oder alltäglich entstehende Datensammlungen potenziell verwendet werden könnten. Die normative Kritik aufgrund dieser Potenzialität bringt den Überwachungskritikern teilweise den Vorwurf ein, paranoid zu sein (Hess und Scheerer 2004). Dem entgegen schwingt des Öfteren die These der Demokratisierung von Kontrolle durch Überwachungstechnologien mit: Weil Datensammlung und Computer ubiquitär seien und somit Überwachung alle beträfe und weil Techniken nur binär – ja/nein, richtig/falsch – entscheiden könnten, spielten typischerweise zuschreibungsrelevante Variablen wie Schichtzugehörigkeit, Geschlecht, Lebensführung oder Ethnizität bei Prozessen der Kriminalisierung und Ausgrenzung keine Rolle mehr. Soziale Kontrolle wäre neutral und somit demokratisiert (vgl. dazu auch Marx 1995).

Im Folgenden soll danach gefragt werden, in wie weit Kontrolltechnologien in diesem Sinne neutrale und demokratische sind und somit klassische kriminologische Befunde etwa zur Relevanz von Definitions- und Beschwerdemacht (Feest und Blankenburg 1972) überholt wären oder aber, ob die Befunde zu Selektivität und Diskriminierung auch unter veränderten technischen Vorzeichen Bestand haben. Die damit verbundene Frage, was denn überhaupt passiert, wenn neue Kontrolltechnologien im Einsatz sind, soll schrittweise anhand der Darstellung einfacher Videoüberwachung über biometrische Erkennung bis hin zu „al-

H. Schmidt-Semisch, H. Hess (Hrsg.), *Die Sinnprovinz der Kriminalität,*
DOI 10.1007/978-3-658-03479-5_8, © Springer Fachmedien Wiesbaden 2014

gorithmischer Überwachung“ (Norris et al. 1999) komplexer Handlungsabläufe beantwortet werden.[1]

Videoüberwachung

Zunächst kann festgestellt werden: eine einfache Videokamera in einer Straße hat abgesehen von der symbolischen Dimension und dem, was die Betrachter der Kamera mit ihr verbinden, von sich aus keinen Effekt. Ohne das Bewusstsein beobachtet zu werden und ohne den Glauben daran, dass jemand beobachtet, und dies ggf. mit einer Kontrollintention, sind die Kameras für das Handeln der Beobachteten bedeutungslos. (Was nicht heißt, sie wären ohne Bedeutung für diejenigen, die mit bestimmten Intention die Bilder betrachten, später auswerten oder publizieren.) Betrachtet man die zahlreichen Studien zu präventiven, panoptischen Effekten von Videoüberwachung (Closed Circuit Television – CCTV), dann zeigt sich, dass – je nach setting – die Wirkung etwa auf Gewalt genannte Handlungen tendenziell gegen Null geht; präventive Effekte scheinen mehr oder weniger auf Kfz-Aufbrüche und -diebstähle auf überwachten Parkplätzen beschränkt zu sein (u. a. Welsh und Farrington 2002; Gill and Spriggs 2005).[2] Generelle disziplinierende Effekte sind zu bezweifeln. Ebenso wenig gibt es empirische Hinweise auf Ausgrenzung und etwa darauf, dass in innerstädtischen Fußgängerzonen als unerwünscht betrachtete Punks oder Obdachlose allein deshalb einen Ort verlassen, weil sie von einer Kamera beobachtet werden. Auch partizipieren die Menschen immer noch an Demonstrationen, obwohl diese regelmäßig (extralegal) von der Polizei gefilmt werden (zu politischem Protest und CCTV vgl. Ullrich 2011). Videokameras *alleine* führen nicht zu Ausgrenzung. Sie werden gleichwohl regelmäßig dazu genutzt, weitere Kontrollhandlungen einzuleiten, die auf Ausgrenzung und/oder Kriminalisierung zielen und diskriminierende Effekte haben können. Kameras müssen in Verbindung mit anderen Kontrollmaßnahmen analysiert werden (Wehrheim 2012).

---

1 Dieser Aufsatz geht u. a. auf zwei von der DFG geförderte Forschungsprojekte zurück. In dem noch laufenden Projekt „Biometrie als ‚soft surveillance‘“ sind u. a. auch Sylvia Kühne und Susanne Krasmann beteiligt. Der Beitrag geht zudem teilweise auf einen Vortrag auf der internationalen Tagung „Security, Ethics, and Justice: Towards a More Inclusive Security Design“ an der Universität Tübingen im Juni 2012 zurück. Ich danke v. a. Sylvia Kühne für kritische Anmerkungen zum Manuskript.

2 Dass Henner Hess und Sebastian Scheerer (2011, S.32f.) diese Empirie außen vor lassen und auf Basis eines ausführlicheren Tagungszeitungsartikels das Gegenteil behaupten und auch keine methodischen Fragen stellen, verwundert.

Hinsichtlich der Wirkungen bleibt zu fragen, was denn genau passiert, wenn Bilder von Videokameras von Beobachtern betrachtet werden. Stellen wir uns eine typische Situation vor, wie sie nach wie vor viele Kontrollräume kennzeichnet: Sicherheitspersonal sitzt vor einer Wand mit beispielsweise acht Bildschirmen, auf denen permanent Bilder von jeweils vier Kameras übertragen werden. Es liegen also 32 Settings vor. Je nachdem, welche Ausschnitte die einzelnen Kameras zeigen, sehen die Beobachter Menschen nur als kleine Punkte oder aber sie können ggf. sehen, welches Buch jemand liest – in dem Moment sehen sie aber den ganzen Rest nicht und auch nicht, was auf den 31 anderen Aufnahmen zu sehen ist. Polizisten oder andere so genannte Sicherheitsdienstleister müssen also entscheiden, worauf sie näher achten. Verdacht ist kein Ereignis, sondern ein Prozess (Norris 2005).

Um diesen Prozess zu verstehen, helfen die Befunde von Herbert Blumer (1981) und die Perspektive des interpretativen Paradigmas weiter: Menschen handeln aufgrund der Bedeutung, die Dinge für sie haben. Die Bedeutung der Dinge hängt wiederum von den Kontexten ab, innerhalb derer die Dinge wahrgenommen werden. Um komplexe soziale Situationen zu verstehen und handhabbar zu machen, greifen wir auf Typisierungen von Situationen und Personen zurück. Typisierungen sind Alltagsvorstellungen, die es uns ermöglichen, nicht jedes Mal von Neuen die kleinsten Details auszuhandeln und wechselseitig zu lernen. Sie belegen die Umwelt mit Sinn, machen uns entfernte Welten erst zugänglich und haben einen Entlastungscharakter. Um mit der Umwelt und alter ego eine Beziehung herzustellen, aufrechtzuerhalten oder auch abzubrechen, wird auf Typisierungen zurückgegriffen. Dabei nutzen wir, so Alfred Schütz (1972), ein „Rezeptwissen", das Orientierung schafft. Ein „Wissen", das sich bewährt hat. Je anonymer nun eine Situation ist, d. h. je weiter sie von unmittelbarer face-to-face-Interaktion und von bereits bekannten Personen entfernt ist, desto eher wird auf standardisierte Typisierungen zurückgegriffen (Berger und Luckmann 2000, S.1ff.). Dies kann, muss aber keineswegs darauf verweisen, was gemeinhin als Vorurteil bezeichnet wird. Typisierungen bewirken gleichwohl die Reproduktion bestehender Strukturen.

Bei CCTV ist nun die Relevanz von Typisierungen insofern gesteigert, als dass erst einmal keine sprachliche Vermittlung und Aushandlung stattfindet. Die Beobachteten sind im Schützschen Sinn eher „Mit-" als „Umwelt". Im Moment der Beobachtung findet bei Videoüberwachung keine *interaktive* Be-Deutung der Situation statt, sondern eine einseitige Zuschreibung. Videoüberwachung stellt keine Subjekt-Subjekt-Beziehung dar, sondern eine Subjekt-Objekt-Beziehung

(Rammert 2005). Die Definitionsmacht liegt bei den Beobachtern: allein ihre Interpretation ist für weitere Handlungen relevant.

Auch Polizisten verfügen über so ein „Rezeptwissen", das Situationsdeutungen überhaupt erst ermöglicht. Verdacht, so Wolfgang Keckeisen, „gründet sich auf beides: Normalitätserwartungen, in denen festgehalten ist, welche Erscheinungen an einem bestimmten Ort zu einer bestimmten Zeit als ‚normal' gelten […]; und typisierten Vorstellungen ‚unangemessener' Erscheinungen, die eine Überprüfung nahe legen" (1974, S.66). Dieser Befund bringt weit reichende Folgen mit sich, denn Polizisten reproduzieren ihre Typisierung durch ihr berufliches Alltagshandeln immer wieder von neuem: wenn ich erwarte, dass auf der Straße eher die ärmer als die wohlhabender aussehenden Personen mit Drogen handeln, dann beobachte ich diese öfter und länger und dann werden diese auch öfter mit illegalisierten Substanzen ‚erwischt'. Die Typisierung gewinnt durch die Erfahrung an Glaubwürdigkeit und führt bei weiterer Orientierung daran zu einer self-fulfilling prophecy. Solche Erfahrungen und Bilder reproduzieren Alltagstheorien und Kriminalitätsvorstellungen: niemand wird vermutlich an Drogendeals oder Raub denken, wenn er auf der Straße zwei weiße Männer in teuren Anzügen mit einer Financial Times unter dem Arm sieht. Unsere Typisierung wird vielleicht eher sein: Oberschichtsangehörige und Investment Banker. Aber dies sagt rein gar nichts über die Wahrscheinlichkeit von Drogenkonsum aus oder darüber, ob man finanzielle Nachteile zu befürchten hat – aber genau das ist es, was in beiden Fällen regelmäßig passiert: Beim Raub und bei der Nachfrage nach einem Kredit. Im ersten Fall nennen wir es Kriminalität im zweiten nicht. Neben Entdeckungswahrscheinlichkeiten variieren die Zuschreibungen.

Dass junge, männliche, subkulturell gekleidete Personen sowie Personen mit dunkler Hautfarbe bei Videoüberwachung überproportional oft und lange beobachtet werden, wie es Norris und Armstrong bereits 1999 in ihrer paradigmatischen Studie „The maximum surveillance society" nachwiesen, verweist auf Diskriminierung aufgrund von solchen Typisierungen, die ggf. auch durch Rassismus überhöht werden, aber nicht einfach darauf reduziert werden können. Der Befund verweist aber erneut noch nicht auf Exklusion oder Kriminalisierung. Knüpfen daran jedoch weitere kontrollierende Handlungen an, dann greifen wieder Typisierungen und dann auch sprachliche Aushandlungen – genau wie bei traditionellem Policing, wobei jedoch eine Vorselektion ausschließlich aufgrund visueller Merkmale stattfand. Die Interaktion bietet dann zwar die Chance, sich von der konkreten Person ein anderes Bild zu machen und dabei deren sprachlichen Äußerungen zu berücksichtigen, aber weil die Situation von den typisierenden

Schemata geleitet wird (Verdacht), sind die Situationsaushandlungen und -definitionen nun zusätzlich vorbelastet.

Bleiben verdachtsgenerierende Zuschreibungen jedoch aus, folgen auch keine intervenierenden Handlungen, wie sich anhand von zwei prominenten Videobildern, die beide in den letzten 20 Jahren als Argumente für mehr Überwachung herhielten, verdeutlichen lässt (siehe auch: Wehrheim 2008):

Bild 1: Szene in einer Shopping Mall

Quelle: http://www.guardian.co.uk/uk/2010/mar/02/james-bulger-jon-venables-prison; 12.02.1993 [2012-08-28]

Das Bild 1 stammt aus dem Jahr 1993 und zeigt den knapp dreijährigen James Bulger an der Hand des zehnjährigen Robert Thompson, der ihn kurze Zeit später zusammen mit einem Freund totschlagen wird. Obwohl die Aufnahme die Tat weder verhinderte noch zur Überführung der Jungen beitrug, galt ab dem Zeitpunkt, „an argument against CCTV was interpreted as an argument in favor of baby killers“ (Davies 1999, S.244). Dass die Tat nicht aufgrund der Kameraüberwachung verhindert wurde, ist mit Blick auf das Ausgeführte nicht überraschend: Falls überhaupt jemand die Situation am Bildschirm beobachtet haben

sollte, dann dürfte sie als Normalität im doppelten Sinn, als konform und üblich, und eben nicht als Abweichung und verdachtsleitend interpretiert worden sein. Die Alltagstypisierung wäre vielleicht ‚Kleinkind an der Hand des älteren Bruders' gewesen.

Bild 2: Szene in einem Flughafen

Quelle: http://en.wikipedia.org/wiki/File:Atta_in_airport.jpg [2012-08-28]

Ähnlich verhält es sich mit dem zweiten prominenten Bild. Es zeigt Mohammed Atta am 11.9.2001 am Flughafen von Boston kurz bevor er vermutlich das von ihm betretene Flugzeug in das World Trade Center in New York steuerte. Die Interpretation des Bildes im Moment seines Entstehens dürfte ggf. gewesen sein ‚ein namentlich bekannter und überprüfter Flugpassagier checkt ein' – und besondere Reaktionen folgen nicht. Auf als ‚normal' definierte Situationen erfolgen eben keine außergewöhnlichen Reaktionen – egal was ‚wirklich' passiert. Das ist es, was das berühmte Thomas-Theorem ausdrückt: "If men define situations as real, they are real in their consequences" (Thomas/Thomas 1928, S.572).

Biometrische Technologien

Diese Nicht-Reaktion wäre auch nicht anders gewesen, wenn das CCTV-System mit einem biometrischen Gesichtserkennungssystem kombiniert gewesen wäre. Mohammed Atta reiste unter seinem eigenen Namen, er war in keiner entsprechenden Datenbank enthalten.

Was wäre aber wenn er oder alle Menschen in so einer Datenbank erfasst wären und in Fußgängerzonen oder an Flughäfen automatische Gesichtserkennung eingesetzt würde? Wenn nicht ausnahmslose jede Person in so einer Datenbank erfasst ist – die biometrischen Fotos in unterschiedlichen Ausweisen böten dafür eine Basis – griffen zunächst übliche, i. d. R. polizeiliche oder geheimdienstliche Interpretationen von Situationen und Etikettierungen von Personen, die dann in eine Datenbank eingepflegt würden. Informationen entstehen bei und durch die Arbeit von Menschen, mit all den Typisierungen und (normativ ausgedrückt) mit all den Stärken und Schwächen, die uns tagtäglich bekannt sind.[3] Darüber hinaus ist es eine Frage der konkreten technischen Systeme, was weiter passiert. Der Biometrieboom in den letzten zehn Jahren hat zu rapiden Veränderungen und zu immer neuen Möglichkeiten und effektiveren Systemen geführt. Schaut man jedoch in die Literatur zum Thema, befragt man Nutzer von Biometrie und beobachtet sie oder nimmt man an einschlägigen Treffen von Biometrieproduzenten, -anwendern, Lobbyverbänden und Datenschützern teil, so drängt sich vor allem ein Eindruck auf: Aufgrund der technischen Komplexität, aufgrund der Vielzahl der Systeme – Gesichtserkennung etwa ist nicht gleich Gesichtserkennung – sowie aufgrund der Vielzahl konkurrierender Akteure und Interessen sind seriöse Aussagen darüber, was die biometrischen Technologien wirklich können und welche Risiken mit den komplexen Systemen verbunden sind, aktuell nicht möglich. Auch die Protagonisten selbst können es nicht wirklich beurteilen, auch wenn sie dies behaupten – egal ob sie Kritiker oder Befürworter sind. Was als verlässliches Wissen gilt, sind (interessensgeleitete) Interpretationen.

3 Wie dieser Prozess der Datenproduktion und -verarbeitung aussieht, lässt sich anhand einer aktuellen Stellenanzeige für eineR „Stadtangestellte/r in der Datenstation des Führungsstabes der Ortspolizeibehörde Bremerhaven" erahnen: „Das Aufgabengebiet umfasst im Wesentlichen: Überprüfung eingehender Vorgänge auf fachgerechte Ausfüllung und Vollständigkeit sowie Durchführung erforderlicher Ergänzungen und Berichtigungen, Eingabe von Daten in eine Datenverarbeitungsanlage (Speichern, Ergänzen, Berichtigen); Löschen von Daten sowie die Abfrage von Daten im Rahmen der Personen- und Sachfahndung und die selbstständige Recherche bei der Eingabe der Daten; Bearbeitung und Umschreibung eingehender Texte, Auswertung der Dateien/Auskunftssysteme mittels Recherche und Weitergabe der Ergebnisse/Auskunftserteilung; Überwachung des elektronischen Fernschreibsystems EPOST 810; Telefonvermittlung." (https://stellen.bremen.de/sixcms/detail.php?id=63325 v. 29.09.2012)

Daneben ist zu konstatieren, dass nicht die absolute technische Qualität des Systems ausschlaggebend ist, sondern bei Entscheidungen für den Einsatz von Biometrie Fragen nach der Alltagspraktikabilität sowie ökonomische Erwägungen vorherrschen. Bei den Entscheidungen für oder gegen Biometrie oder für oder gegen ein bestimmtes biometrisches System spielen ggf. bei Pilotprojekten identifizierte „Ausreißer" bei der Erkennungsleistung kaum eine Rolle. Dies ist bereits statistisch problematisch, denn über die Masse der Personen werden „Ausreißer" schnell zu einer relevanten Größe. Denkt man etwa an einen Einsatz bei Hunderttausenden von Flugpassagieren, so werden als niedrig geltende false acceptance (FAR) oder false rejection rates (FRR) von 0,01 % schon zu einem ernsthaften Problem.[4] Hinsichtlich eines Biometriepilotprojekts am neuen Flughafen Berlin-Brandburg, bei dem zukünftig lediglich die Tausenden von Angestellten biometrisch erfasst werden sollen, seien die „Ausreißer" Personen mit Tattoos sowie mit Hautkrankheiten gewesen, so berichtete es ein Vertreter. Diskutiert wurden sie jedoch nicht unter Fragen von Diskriminierung, sondern unter Fragen der möglichen Verzögerung von Arbeitsabläufen. Das Beispiel der Tattoos und die über die Benachteiligung entstehende neue Kategorie von Personen (Tätowierte) deutet es an: Biometrische Technologien haben bereits einen *technischen Bias* und sie selegieren dadurch. Für Fingerabdrucksysteme etwa gilt: ältere Menschen sind aufgrund von im Lebensverlauf stärker abgenutzten Fingern schwerer einzulesen und zu identifizieren als jüngere; Arbeiter mit trockenen oder stark beanspruchten Fingern wie etwa Gärtner oder Maurer sind problematischer zu erkennen als Büroangestellte. Weil Fingerprint-Systeme gewöhnlich Systeme zur Zugangsregulierung sind – Zugang zu Computern, Gebäuden, Bankkonten, Nationalstaaten – ist die Konsequenz, dass die Gruppen mit einer schlechteren Erkennungsleistung eine höhere Wahrscheinlichkeit haben, dass der Zugang verweigert wird. Dies bedeutet nicht automatisch Exklusion, wohl aber eine häufigere Erfahrung von Diskriminierung: die soziale Sichtbarkeit der Betroffenen wird erhöht, sie sind ggf. genötigt, einen alternativen Marker oder Code zu benutzen, und an Grenzen müssen sie sich eher mit ggf. rüden Beamten auseinandersetzen oder aber gar Verhöre über sich ergehen lassen.

Für Gesichtserkennungssysteme weisen Lucas Introna and David Wood (2004, S.190) in Bezug auf drei Pilotstudien aus den Jahren 2002/2003 darauf hin, dass "Asians are easier to recognize than whites, African-Americans are easier than whites, other race members are easier than whites, older people are easier than

4 FAR und FRR bedingen sich gegenseitig. Sie stellen sozial definierte Toleranzgrenzen dar: je empfindlicher das System, desto höher die FRR (v. a. in Hochsicherheitsbereichen mit wenigen Nutzern) und je toleranter die Systeme eingestellt sind, desto höher die FAR (in Bereichen, in denen viele Menschen schnell erfasst werden müssen).

young people, …". Die Erkennungsleistung soll je nach weiteren Bedingungen (Lichtverhältnisse, Alter der Ausgangsfotografien etc.) zwischen 5-10 % variiert haben. Unabhängig davon, wie groß die Unterschiede im Jahr 2012 bei technisch fortgeschritteneren Systemen und unterschiedlichen Ausgangsbedingungen sind, es bedeutet bei automatischer Erkennung, dass diese Personen eine größere Wahrscheinlichkeit haben einen automatisierten Alarm auszulösen: und zwar nicht, weil sie nicht erkannt werden, sondern weil sie (richtig oder ggf. auch falsch) erkannt wurden. D. h. bei einer Richtigerkennung ist die Wahrscheinlichkeit, erkannt zu werden, für in einer Datenbank erfasste, asiatisch aussehende Personen größer, als für „whites". Nach einem automatisierten Alarm greifen dann wieder die Typisierungen und Situationsinterpretationen von denen bereits die Rede war. Wie Introna und Wood weiter schreiben: Im Fall eines falsch positiven Alarms an einem us-amerikanischen Flughafen war es für die weitere Behandlung definitiv relevant, dass die Person, die den Alarm auslöste, aussah, „as if he might be from the Middle East" (Introna und Wood 2004, S.193). Die Definitions- und Beschwerdemacht variiert nach wie vor mit dem sozialen Status einer Person und diese Macht dürfte noch geringer sein, wenn ein automatisiertes Überwachungssystem sie zuvor als verdächtig identifiziert hat. Der technische Bias wird überhöht durch die alltagsweltliche Typisierung und durch politisch begründete Verdachtskonstruktionen. Nach dem technischen greift ein *sozialer Bias*.

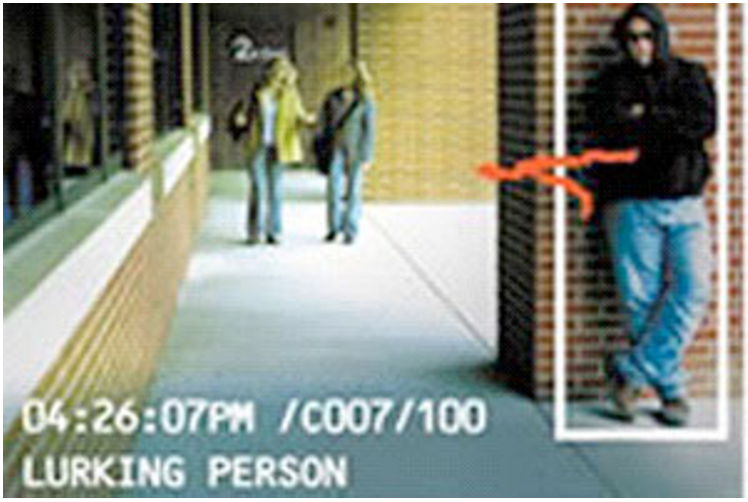
Intelligent Monitoring for Threat Detection

Bereits seit den 1990er Jahren in der Entwicklung, sind Projekte und Feldversuche, Personen automatisiert zu identifizieren oder ‚Abweichung' zu erkennen, bis heute immer wieder mehr oder weniger kläglich gescheitert. Gleichwohl ist es nach wie vor das Ziel, verdächtiges Verhalten und/oder Personen in Echtzeit sowie Kriminalität zu erkennen, *bevor* sie geschieht (für einen Überblick über aktuelle Projekte siehe Adams und Ferryman 2012). Der wesentliche Unterschied zur anfangs erwähnten konventionellen Videoüberwachung ist, dass der Verdacht wie beim Biometriebeispiel automatisiert wird, also die Definitionsmacht auf das technische System übertragen wird, das eine soziale Situation objektiv „erkennen" soll.

Software kann dabei erstens auf Basis der durch die Überwachung selbst gewonnenen statistischen Auffälligkeiten programmiert werden. Für die Londoner U-Bahn fiel etwa auf, dass Personen mit suizidalen Absichten häufig eine längere Weile unbeweglich am Gleis stehen und Züge passieren lassen, bevor sie vor einen Zug springen. Die Konsequenzen, falls jemand etwa nur betrunken ist und keineswegs springen will, scheinen bei einer Intervention verhältnismäßig unpro-

blematisch. Anders sieht es aus, wenn unmittelbare Kriminalitätszuschreibungen erfolgen: sowohl wenn sie auf statistischen Wahrscheinlichkeiten beruhen als auch wenn sie zweitens als Typisierung, als sozialer Bias, vorab in die Programmierung eingeflossen sind. Letzteres ist beim Bild 3 der Fall, das zum Projekt „Perceptrak" der Firma Smart CCTV gehört. Überwachungssoftware soll eine Gefahr erkennen, bevor etwas passiert und so möglichst ex ante Interventionen ermöglichen. Mit dieser Vorstellung werden entsprechende Systeme zumindest beworben.

Bild 3: Lurking Person



Quelle: http://derstandard.at/1227287603564 [2012-08-24]

Bei der „lurking person", die automatisch als Indikator für eine bedrohliche Situation gedeutet wird, wurden die Typisierungen und Alltagsvorstellungen bereits bei der Programmierung in das System eingeschrieben. Programme wissen nicht von alleine, wann sie Alarm schlagen sollen. Diese Vorgehensweise ist auch bei dem seit 2009 von der EU mit ca. 15 Millionen Euro finanzierte Projekt INDECT „Intelligent information system supporting observation, searching and detection for security of citizens in urban environment" der Fall.[5] INDECT soll

5 Für Hinweise auf verschiedene andere, nationale Projekte mit teilweise breiter Beteiligung von Rüstungskonzernen siehe: Monroy (2012).

nicht nur automatisiert „abnormales“ Verhalten erkennen,[6] sondern Drohnen, die Überwachung von Internetverkehr, der Abgleich biometrischer Datenbanken mit sozialen Netzwerken und vieles mehr sollen zusammen zum Einsatz kommen (das Programm PRISM der NSA berührt nur einen vergleichsweise kleinen Teil dieser ambitionierten Überwachung). Gerüchten zufolge erfolgte ein erster Test 2012 bei der Fußball EM in Polen und der Ukraine. Eine Basis der automatisierten Überwachung soll es sein, dass die Typisierungen von Polizisten in die Programmierung eingehen (INDECT Consortium 2009/2012, S.18ff.). Dafür wurden (polnische) Polizisten befragt, die – so muss man sagen – glücklicherweise nicht auf alle Fragen geantwortet haben (Tab.1):

Tabelle 1: Zuschreibungsrelevante Merkmale

A 3. How would you recognize a particular person that is of following type? Is it a dress, behaviour, what type?	
Burglar	Observes entrances and monitoring, loiters, nervous, untypical tools, luggage, frequent presence in the location, peeking through the window
Pickpocketeer	Observes people, holds cloth in ones hand, frequent presence in public transport nodes, doesn't avoid crowd, a group of perpetrators is spreading, then gathers around the victim creating artificial crowd
Thief	
Drug dealer	
Drug addict	
Lost kid	Cries, bothers other people, loiters, runs without purpose, in circles
Pedophile	
Terrorist	
Hooligan	

Quelle: INDECT 2009/2012, S.20

6 “Q1.3: What behaviour is an “abnormal” behaviour? A: As regards the definition of “abnormal behaviour”, the term is not introduced by the INDECT Project, and it was created by EC and explained in the FP7 Work Programme. This term will be always controversial. In our case we clearly understand abnormal behaviour as “criminal behaviour”, and especially as “behaviour related to terrorist acts, serious criminal activities (e. g.: murders, bank robberies, someone leaving the luggage in the airport with the bomb) or criminal activities in the Internet (e. g.: child pornography)”. We will produce the tools to avoid such situations.“ (http://www.indect-project.eu/faq#Q1.3 [2012-09-01], Herv. i. O.). Dass die INDECT-Projektmitarbeiter glauben oder behaupten, die Trennung zwischen kriminellem und nicht-kriminellem „abnormen“ Verhalten sei eindeutig, ist schon beachtlich. Beachtlich ist darüber hinaus, dass bei der weiteren Forschung viele Aktivitäten auf die Erkennung von Graffiti und Taschendiebstahl bezogen werden und nicht auf das, was als meist Terrorismus oder Bankraub bezeichnet wird.

Bereits das Antwortverhalten zeigt die Problematik der Fragestellung: Definitionstheoretisch informierte Soziologen und Kriminologen könnten nicht einmal sagen, was Terroristen, Drogenabhängige oder Pädophile genau sein sollen, geschweige denn, welche visuellen Merkmale diese und ihr Verhalten grundsätzlich charakterisieren sollen.

Weiterhin wurde beispielsweise gefragt, was die wichtigsten Symptome von „gefährlichen Bestrebungen" seien (Tab.2).

Tabelle 2: Symptome „gefährlicher Bestrebungen"

A 5. What are the most important symptoms for dangerous attempts?		
Situation	Percentage of answered YES	What type of danger can it suggest?
Looking around	76 %	
Running with looking around repeatedly	33 %	
Loitering	71 %	
Standing near the door/ car for too long	57 %	
What else in your opinion?	Staying for too long in a single place, repeatedly coming back to a place	

Quelle: INDECT 2009/2012, S.20

71 % der befragten Polizisten antworteten „Rumhängen" und 33 % nannten etwa „Laufen mit wiederholtem Umdrehen". Aber was meint „Rumhängen"? Sitzen, liegen, stehen? Für eine Stunde, für 10 Minuten oder für fünf? Auf der Straße, auf einer Bank, in einem Park? Bei Sonnenschein und lauen Lüftchen bei 25°C oder bei Hagel und 0°C? Was heißt „Rennen mit wiederholtem Umdrehen"? In welchem Kontext? Wenn ich einen Fahrradweg überqueren will, um einen Bus zu erreichen und ich gucken muss, dass ich nicht überfahren werde, weil mehrere Fahrradfahrer passieren?

Solche nahe liegenden theoretischen Einwände werden in der Praxis auch empirisch bestätigt. So kommen Robert Rothmann und Stefan Vogtenhuber (2012, S.109f.) bei ihrer ethnographischen Videoanalyse einer Anlage, die das technisch vergleichsweise relativ simple Ziel hatte, u. a. Diebstähle aus Kraftfahrzeugen in einer Parkgarage automatisch zu erkennen (Gehen von einem Auto zum nächsten und nicht direkt vom Auto zum Ausgang diente als ein Indikator), zu dem Ergebnis, dass in etwa mit 50 Fehlalarmen pro Stunde zu rechnen sei, weil ca. 20 %

der „normalen" Garagenbenutzer sich genauso verhielten, wie das zuvor als „verdächtig" programmierte Verhalten.

Eine kontextunabhängige allgemeingültige, im unmittelbaren Wortsinn berechenbare Objektivität kann es nicht geben, egal wie präzise Soft- und Hardware arbeiten, und Vorhersagen basieren eben nur auf statistischen Wahrscheinlichkeiten. Auch der Wetterbericht stimmt manchmal, und ob das Wetter dann als ‚gut' oder ‚schlecht' interpretiert wird, variiert zusätzlich mit persönlichen Vorlieben, körperlicher Konstitution, geplanten Aktivitäten, Diskursen über Hautkrebs oder „vornehme Blässe", dem Vergleich zum Wetter die Tage vorher etc. pp.

Automatisierung und die Verdoppelung der Typisierung

Michalis Lianos and Mary Douglas schrieben, eine der bedeutensten Änderungen der neuen Kontrollkultur sei, dass Automated Socio-Technical Environments[7] "*radically transform the cultural register of the societies in which they operate by introducing non-negotiative contexts of interaction"* (Lianos and Douglas 2000, S.265; Herv. i. O.). Die Definitionsmacht würde ausschließlich auf die technischen Systeme übertragen. Ein Alarm erfolgt dann unabhängig davon, ob jemand läuft und sein Laufen sowohl als Versuch, einen Bus zu erreichen, als auch als Flucht nach Aneignung einer fremden beweglichen Sache bedeutet werden könnte. Die Software kennt ggf. nur „Laufen" als zuschreibungsrelevantes Merkmal.

Zurück zur „lurking person": natürlich trägt die „lurking person" eine Sonnenbrille[8]; natürlich trägt sie einen Kapuzenpullover und selbstverständlich ist dieser schwarz; natürlich handelt es sich um eine vermutlich männliche Person und natürlich sind die potenziellen Opfer weiblich. Der soziale Bias ist in das technische System eingeschrieben und damit wird von den Produzenten geworben. Wenn das System funktioniert wie programmiert und wenn die „algorithmic surveillance" einen Alarm auslöst, dann wird jeder mit einem „hoody" und einer Sonnenbrille, der auf den Bus oder einen Freund wartet von einem Polizisten kontrolliert werden – wenn denn ein Polizist in der Nähe ist – und zwar *nur* weil er wartet und die „falsche" Kleidung trägt. Ab diesem Moment endet jedoch der automatisierte und nicht-verhandelbare Verdacht. Die neuen „Automated Socio-Technical Environments" sind den bekannten Interaktionsformen und -situatio-

7 Automated Socio-Technical Environments "… are technology-based contexts of interaction that regulate, organize or monitor human behaviour by integrating it into a pre-arranged environment, built upon a conception of 'normality' or 'regularity' that all subjects are expected to reproduce." (Lianos und Douglas 2000, S.264)

8 Auch bei nicht-automatisierten CCTV-Systemen war das Tragen einer Sonnenbrille verdachtsgenerierend. (Norris und Armstrong 1999)

nen (nur) vorgeschaltet und sie sind über ihre Programmierung und über die Ausrichtung des technischen System eben gerade auch sozial. In der ggf. folgenden face-to-face-Interaktion griffen dann wieder die oben erwähnten Typisierungen, und die Chance für die „lurking person", die Situation nicht als mutmaßlichen Überfall, sondern etwa als schlichtes Warten zu definieren, hängt u. a. davon ab, ob und wie sie sich dazu äußern kann und als wie plausibel das dann von Polizisten vor dem Hintergrund ihrer Alltagstheorien gedeutet wird. Allerdings wird die Ausgangsposition im Falle einer so gekleideten Person vermutlich noch ungünstiger, als ohne vorherigen technischen Alarm. Die *Typisierungen verdoppeln sich*: bei der Programmierung griff bereits das stereotype Bild des Muggers und bei der Interpretation des automatisierten Alarms/Verdachts wird es erneut relevant, wobei die Situation dann „technisch vorbelastet" ist. Soll der Kontrolleur nun *nicht* diskriminierenden Typisierungen folgen, dann wird von ihm nicht nur verlangt, von den beruflich mitproduzierten *Sinnprovinzen*, dem eigenen „Denken wie üblich" (Schütz 1972) und von den im Alltag ja bewährten Typisierungen abzuweichen, sondern zusätzlich müsste er entgegen seiner eigenen Vorstellung auch noch der technischen Diagnose widersprechen. Das System hatte ja Alarm geschlagen und die Annahme, dass da was dran sein muss, ist näher liegender als die gegenteilige Annahme (zumindest wenn sich noch nicht die Erfahrung eingestellt hat, dass das System 50 Fehlalarme pro Stunde produziert). Aufgrund des Status, den Technik in heutigen Gesellschaften besitzt, hat diese schon selbst eine große „Definitionsm[a]cht".[9] Datenbanken und Gesichtserkennung können generell und besonders in Situationen von Unsicherheit bzw. reduzierter Erwartbarkeit als „autoritärer" gelten, als die involvierten Personen.

Was wären also die statistischen und damit auch sozialen Konsequenzen im Fall der „lurking person"? Wenn das System tatsächlich dann Alarm schlagen würde, wenn ein Raub bevorstünde (und Sicherheitspersonal in der Nähe wäre), dann hätten Personen mit Kapuzenpullover eine höhere Wahrscheinlichkeit wegen Raubes kriminalisiert zu werden, als Personen in Anzügen. Würde das System reagieren, ohne dass später die „lurking person" die beiden Frauen überfällt, so würden Personen mit Kapuzenpullover regelmäßig Diskriminierungserfahrungen machen, denn (so Polizei vor Ort ist) die Erfahrung der polizeilichen Kontrolle ist immer unangenehm und in der Öffentlichkeit aufgrund ihrer star-

9 Das wird einem spätestens dann bewusst, wenn man bei der it-Abteilung der eigenen Universität ein technisches Problem meldet, was nach deren Einschätzung gar nicht sein kann, oder wenn man bei einer Behörde den Sachbearbeiter überzeugen will, dass der Eintrag in seiner Datenbank nicht stimmt. (Technik und Bürokratie gehen eine unheimliche Allianz ein.)

ken Symbolik beschämend und stigmatisierend.[10] Man stelle sich nun vor, die beim INDECT-Projekt befragten Polizisten hätten auf die Frage, wie Pädophile und Terroristen aussehen und handeln, geantwortet und Informatiker hätten dies dann programmiert…

„Social control technologies" und auch die ‚algorithmischen' sind nicht neutral und objektiv: erstens, weil Technologien schon in ihrem Design grundsätzlich politisch sind, denn sie basieren auf bestimmten Interessen und klammern andere aus (Introna und Wood 2004, S.179). Normalität und Abweichung sind immer durch Macht- und Herrschaftsbeziehungen mitbestimmt und Normen Produkte solcher Beziehungen. Gleiches gilt für die Definition von Risikokategorien (vgl. Wehrheim 2011). Zweitens greifen die Technologien in ihrer Anwendung Zuschreibungen und Typisierungen auf und reproduzieren diese. Für eine Theorie der Kriminalität, wie sie Henner Hess und Sebastian Scheerer (2004) vorschwebt, müssten diese Aspekte der Zugang sein, denn die fürchterlichsten Dinge werden meist gerade nicht Kriminalität genannt.[11]

Um jedoch wirklich zu wissen, was passiert, wenn Personen, die zu einer bestimmten sozialen Gruppen oder zu einer „Risikokategorie" gezählt werden, durch ein Fingerabdrucksystem nicht erkannt, durch Gesichtserkennung als „besonders" erfasst oder des Raubes aufgrund eines System automatisierter Gefahrenerkennung verdächtigt werden, bedarf es mehr Forschung und insbesondere ethnographischer Forschung. Klar scheint bisher jedoch, dass die technischen Systeme (noch) einen in ihrer Entwicklung begründeten technischen Bias haben und dieser durch einen sozialen beeinflusst und ggf. verstärkt wird. Wenig strittig dürfte auch sein, dass diese technische Vorselektion nachgehende Situationsbedeutungen beeinflusst. Eine entscheidende Frage für weitere Forschung scheint jedoch die nach der Relevanz einzelner neuer und alter zuschreibungsrelevanter Variablen wie Alter, Geschlecht, Klasse und Beruf sowie Ethnizität und Aussehen zu sein. Die Selektivität von Exklusion und Kriminalisierung beginnt beim Verdacht.

10 Einmal von dem prinzipiellen rechtlichen und ethischen Dilemma abgesehen, das entstünde, wenn die Polizei tatsächlich vor Ort wäre, *bevor* überhaupt etwas passieren konnte, und abgesehen davon, wie die letztendliche Zuschreibung ‚Raub' zustande kommt.

11 Schade ist es, dass die Autoren die Herrschaftsfrage zwar erwähnen, die Konsequenzen daraus aber weitestgehend ignorieren. Schade ist auch, dass sie davon ausgehen, die Täter würden ihr Handeln immer selbst als Kriminalität deuten. Harald Welzer (2009) etwa zeigt nachdrücklich, wie Massenmord von den Handelnden selbst zwar als unerfreuliche und ggf. auch moralisch problematische Handlung interaktiv definiert wird, aber eben gerade nicht als Kriminalität.

Literatur

Adams, A. A. und J. M. Ferryman. 2012. *The Future of Automated CCTV Analysis and Its Ethical Implication.* Vortragsmanuskript, The Fifth Biannual Surveillance and Society Conference „Watch This Space: Surveillance Futures". Sheffield.

Berger, P. L. und T. Luckmann. 2000. *Die gesellschaftliche Konstruktion der Wirklichkeit*. 5. Aufl. Frankfurt/Main: Fischer Verlag.

Blumer, H. 1981. Der methodologische Standort des Symbolischen Interaktionismus. In *Alltagswissen, Interaktion und gesellschaftliche Wirklichkeit,* hrsg Arbeitsgruppe Bielefelder Soziologen, 5. Aufl., 80-146. Reinbek bei Hamburg: Rowohlt.

Davies, S. G. 1999. CCTV: a new battleground for privacy. In *Surveillance, Closed Circuit Television and Social Control.* hrsg. Norris et al., 243-254. Aldershot/Brookfield USA/Singapore/Sydney.

Feest, J., und E. Blankenburg. 1972. *Die Definitionsmacht der Polizei.* Düsseldorf: Bertelsmann.

Gill, M., und A. Spriggs. 2005. *Home Office Research Study 292: Assessing the impact of CCTV.* http://www.homeoffice.gov.uk/rds/pdfs05/hors29.pdf. Zugegriffen: 27.02.2007.

Keckeisen, W. 1974. *Die gesellschaftliche Definition abweichenden Verhaltens*. München: Juventa Verlag.

Lianos, M., und M. Douglas. 2000. Dangerization and the End of Deviance. *British Journal of Criminology* 40, 261-278.

Hess, H., und S. Scheerer. 2004. Theorie der Kriminalität. *Kölner Zeitschrift für Soziologie und Sozialpsychologie. Sonderheft 43: Kriminalsoziologie*, hrsg. von S. Karstedt, und D. Oberwittler: 69-92.

Hess, H., und S. Scheerer. 2011. Radikale Langeweile. In *Langweiliges Verbrechen,* hrsg. H. Peters und M. Dellwing, Michael, 27-52. Wiesbaden: VS Verlag.

INDECT Consortium 2009/2012: *D1.1 Report on the collection and analysis of user requirements.* http://www.indect-project.eu/files/deliverables/public/INDECT_Deliverable _D1.1_v20091029a.pv.pdf/view Zugegriffen: 01.09.2012.

Introna, L. D., und D. Wood. 2004. Picturing Algorithmic Surveillance. The Politics of Facial Recognition Systems. *Surveillance & Society* 2 (2/3), 177-198.

Marx, G. T. 1995. The Engineering of Social Control: The Search for the Silver Bullet. In *Crime and Inequality*, hrsg. J. Hagan and R. Peterson, 225-246. Stanford: Stanford University Press.

Monroy, M. 2012. Nasenhaare in Großformat. *Telepolis.* http://www.heise.de/tp/artikel/37/37653/1.html. Zugriffen: 20.09.2012.

Norris, C. 2005: Vom persönlichen zum Digitalen. Videoüberwachung, das Panopticon und die technologische Verbindung von Verdacht und gesellschaftlicher Kontrolle. In *Bild – Raum – Kontrolle. Videoüberwachung als Zeichen gesellschaftlichen Wandels,* hrsg. L. Hempel und J. Metelmann, 360-401. Frankfurt a. M.: Suhrkamp.

Norris, C., und G. Armstrong. 1999. *The maximum surveillance society. The rise of CCTV.* Oxford/New York: Berg.

Norris, C., J. Moran, und G. Armstrong. 1999. Algorithmic surveillance: the future of automated visual surveillance. In *Surveillance, Closed Circuit Television and Social Control*, hrsg. Dies., 255-276. Aldershot/Brookfield USA/Singapore/Sydney.

Rammert, W. 2005. Gestörter Blickwechsel durch Videoüberwachung? Ambivalenzen und Asymmetrien soziotechnischer Beobachtungsordnungen. In *Bild – Raum – Kontrolle. Videoüberwachung als Zeichen gesellschaftlichen Wandels,* hrsg. L. Hempel und J. Metelmann, 342- 359. Frankfurt/Main: Suhrkamp.

Rothmann, R., und S. Vogtenhuber. 2012. Videoüberwachung*, Ereigniserkennung und Automation*. Wien: IHS. http://www.equi.at/dateien/Rothmann-Vogtenhuber_2012.pdf. Zugegriffen: 01.09.2012.

Schütz, A. 1972. Der Fremde. Ein sozialpsychologischer Versuch. In *Gesammelte Aufsätze II. Studien zur soziologischen Theorie,* hrsg. A. Schütz, 53-69. Den Haag: Nijhoff.

Thomas, W. I., und D. S. Thomas. 1928. *The Child in America. Behavior Problems and Programs.* New York: Knopf.

Ullrich, P. 2011. *Gesundheitsdiskurse und Sozialkritik – Videoüberwachung von Demonstrationen. Zwei Studien zur gegenwärtigen Regierung von sozialen Bewegungen und Protest.* München: Deutsches Jugendinstitut.

Wehrheim, J. 2008. Videoüberwachung. Das Interesse am Ungewöhnlichen im gewöhnlichen Alltag, In *Bilderatlas des 20. und beginnenden 21. Jahrhunderts, 1949-2006*, Bd. 2, hrsg. Paul, Gerhard, 622-629. Göttingen. Vandenhoeck und Ruprecht.

Wehrheim, J. 2011. Kriminologie, Sicherheit und die herrschende Normalität des ungleichen Sterbens. Zur gesellschaftlichen Funktionalität von Kriminalisierung und Securitization. In *Langweiliges Verbrechen*, hrsg. H. Peters und M. Dellwing, Michael, 53-70. Wiesbaden: VS Verlag.

Wehrheim, J. 2012. *Die überwachte Stadt. Sicherheit, Segregation und Ausgrenzung.* 3. Aufl. Opladen/Toronto: Leske und Budrich.

Welsh, B. C., und D. P: Farrington. 2002. *Crime prevention effects of closed circiut televison: a systemetic review*. London: Home Office.

Welzer, H. 2009. *Täter. Wie aus ganz normalen Menschen Massenmörder werden.* Frankfurt a. M.: Fischer Verlag.