

# WHO COINED THE PHRASE “DATA SHADOW”?

STEVEN M. BELLOVIN\*

*The phrase “data shadow” is commonly used in books and articles on privacy. The origin of the phrase, though, is mysterious. It is often attributed to Alan Westin, but it does not seem to appear in any of his writings. I show that it was used in 1972 by Dr. Kerstin Anér, a member of the Swedish parliament, as “dataskugga”. She later used the phrase in English, later in the 1970s. It was briefly popular then, but disappeared until the early 1990s. It has since become a popular and evocative phrase to describe how our activities, online and offline, follow us around.*

---

\* Steven M. Bellovin is the Percy K. and Vida L.W. Hudson Professor of Computer Science at Columbia University and affiliate faculty at Columbia Law School. Several people aided me in this research. Patrik Fältström and Peter Krantz tracked down the Swedish references. Simson Garfinkel aided in my web searches and performed the earlier Ngram search. Tarah Wheeler and Matthew Nuding helped with the *Local Regional Government* reference. “Garson O’Toole”—a pseudonym—who runs the Quote Investigator website found the *Pittsburgh Press* article. Ari Ezra Waldman and Marc Rotenberg made many useful suggestions.

Anyone who works in privacy law and policy is familiar with the phrase “data shadow”: the sum of the digital footprints we leave behind us in our daily lives as we buy things, search online, even travel. But who first used the phrase has been something of a mystery. It’s often used without citation; if there is a citation, it is frequently to Alan Westin. But was he the first? My investigation suggests not.

It is tempting to attribute it to Professor Westin; his classic book *Privacy and Freedom*<sup>1</sup> was the foundation of much of today’s discourse on privacy.<sup>2</sup> Indeed, a number of authors have used this attribution.<sup>3</sup> I’ve done it myself,<sup>4</sup> albeit with some trepidation since neither I nor several respected scholars whom I consulted could find a page number in his book. Dr. Simson Garfinkel was more circumspect, attributing it only to Westin’s writings in the 1960s;<sup>5</sup> again, though, there was no precise citation. Paul McFedries also attributes it to Prof. Westin, though he cites Dr. Garfinkel’s book as his source.<sup>6</sup> The matter was settled shortly after my 2018 submission, when I obtained the Kindle edition of *Privacy and Freedom*: the phrase “data shadow” does not appear in it. Indeed, when the word “shadow” appears—and it’s in the index—it’s only in the context of physical surveillance. A similar search of a somewhat later book co-authored by Westin<sup>7</sup> also came up empty;

---

<sup>1</sup> ALAN F. WESTIN, *PRIVACY AND FREEDOM* (1967).

<sup>2</sup> Westin’s book is primarily a report of the work of the Committee on Science and Law of the Association of the Bar of the City of New York; *see id.* at ix.

<sup>3</sup> *See, e.g.*, Carol Diamond, Melissa Goldstein, David Lansky & Stefaan Verhulst, *An Architecture for Privacy in a Networked Health Information Environment*, 17 *CAMBRIDGE Q. HEALTHCARE ETHICS* 429, 431 (2008); Dan L. Burk, *Information Ethics and the Law of Data Representations*, 10 *ETHICS & INFO. TECH.* 135, 139 (2008); Brandt Dainow, *Digital Alienation as the Foundation of Online Privacy Concerns*, 45 *SIGCAS COMPUTS. & SOC’Y* 109, 112 (Sept. 2015).

<sup>4</sup> Steven M. Bellovin, *Comments on Privacy*, L. ARCHIVE 3 (Nov. 7, 2018), <https://osf.io/preprints/lawarxiv/5s2vt> [<https://perma.cc/3UK6-6TVS>].

<sup>5</sup> SIMSON GARFINKEL, *DATABASE NATION: THE DEATH OF PRIVACY IN THE 21ST CENTURY* 70 (2000).

Alan Westin coined the term data shadow in the 1960s. Westin, a professor at Columbia University in New York, warned that credit records, bank records, insurance records, and other information that made up America’s emerging digital infrastructure could be combined to create a detailed digital dossier. The metaphor, with its slightly sinister feeling, was uncannily accurate: just as few people are aware of where their shadows fall, few data subjects in the future, Westin conjectured, would be able to keep track of their digital dossiers.

<sup>6</sup> PAUL MCFEDRIES, *WORD SPY: THE WORD LOVER’S GUIDE TO MODERN CULTURE* 166 (2004).

<sup>7</sup> *See* ALAN F. WESTIN & MICHAEL A. BAKER, *DATABANKS IN A FREE SOCIETY* (1972).

the same is true of a LexisNexis search of law review articles that contained the words “Westin” and “shadow”.

I decided to do more digital searches, with Lexis, Google Scholar, and Google Books. Dr. Garfinkel and I both performed Google Ngram searches. These all confirmed my earlier impression: the phrase was not used in the 1960s nor in most of the 1970s, and most of the hits were false positives.<sup>8</sup>

My searches did turn up one early hit, in a quotation attributed to “Kirstein Aner” [sic]:<sup>9</sup>

Every single person will be closely followed all his life by his own data shadow where everything he has ever done, learnt, bought, achieved, or failed to achieve is expressed in binary numbers for eternity. His shadow will often be taken for himself and we will have to pay for all the inaccuracies as if they were his own; he will never know who is looking at his shadow or why.

Dr. Kerstin Anér was a Swedish legislator and an early activist on privacy. In discussing the debate that led to the (Swedish) Data Act of 1973, Per Lundin describes her this way:<sup>10</sup> “The single most influential voice in the debate was the Liberal Member of Parliament, Kerstin Anér. In her book *Datamakt* [Computer Power] (1975) she exposed the ambitions of the state as well as the corporations to control, and she empathically defended the right and the need for individuality.”

At this point, I asked a Swedish colleague, Patrik Fältström, for help; he also asked Peter Krantz. What they found appears to confirm the hypothesis that Dr. Anér originated the phrase in its Swedish form, “dataskugga”, and was the first to use it in English.

The first use they found was in a 1972 magazine article by Dr. Anér, one that does not seem to be online.<sup>11</sup> She wrote, “That’s enough by looking at this vast computer picture of man, laid out before the eyes of the one who has all the legal right to it. The Social Welfare Board,

---

<sup>8</sup> See, e.g., K.A. Graf, H. Guthart & D.G. Douglas, *Refraction Effects in Turbulent Media*, 9 RADIO SCI. 777, 781 (1974) (“As the electron density in the flame increases, all these *data* show a *shadow* immediately behind the flame.”) (emphasis added).

<sup>9</sup> Judith Bloor, *Data Bank Control Begins in Sweden*, 63 NEW SCIENTIST 718, 720 (1974).

<sup>10</sup> Per Lundin, *Computers and Welfare: The Swedish Debate on the Politics of Computerization in the 1970s and the 1980s*, in HISTORY OF NORDIC COMPUTING 4, 3, 6 (Christian Gram, Per Rasmussen & Søren Duus Østergaard eds., 2014) (internal citation omitted).

<sup>11</sup> Kerstin Anér, *Dataskuggan*, VÅR LÖSEN, 1972, at 310.

e.g. What do they see? Do they see a human being, or her data shadow? Of course, the latter.”<sup>12</sup>

A more accessible rendering is given by Johan Fredrikzon.<sup>13</sup> In Fältström’s translation:

By the term “data shadow” Kerstin Anér meant “the two-dimensional image of the individual that is evoked through one or more data registers” or “the series of crosses in different boxes that a living person is transformed into in the data registers and which tends by those in power to be treated as more real and more interesting than the living man himself.”<sup>14</sup>

Fredrikzon attributes that quote to Dr. Anér’s own 1975 book *Datamakt* (Computer Power).<sup>15</sup> His footnote goes on to say “Anér, *Datamakt*, 245; Anér referenced in the article [unsigned], ‘Data legislation is to protect privacy. But we are sold as numbers’, *Dagens Nyheter*, 13 juli 1973; Ten years later the term was still in use: Calle Hård, ‘The tale of the data shadow 600412–5890’, *Aftonbladet*, 11 mars 1982.”<sup>16</sup>

Dr. Anér herself used the phrase “data shadow” in English on other occasions. For example, in 1978 she wrote “Any citizen worried about privacy and integrity in a computerized world should worry not only about who has access to his data shadow—the composite picture that all records about him make up.”<sup>17</sup>

She used the phrase in a slightly different sense in a very prescient essay in 1979.<sup>18</sup> There, she defined “data shadow” as “what appears when no one asks the person about his data any more, but pulls them out of databanks where they are already filled.”<sup>19</sup> This definition is strongly reminiscent of Plato’s Allegory of the Cave,<sup>20</sup> where prisoners chained in a cave who can see only shadows of outside objects

<sup>12</sup> *Id.* (translation by Bing Translate).

<sup>13</sup> JOHAN FREDRIKZON, KRETSLOPP AV DATA 231 (2021).

<sup>14</sup> *Id.*

<sup>15</sup> KERSTIN ANÉR, DATAMAKT 245 (1975).

<sup>16</sup> FREDRIKZON, *supra* note 13, at 497 n.29.

<sup>17</sup> Kerstin Anér, *Citizen Involvement in Decision-Making of Local and Regional Government*, in 12 LOCAL GOVERNMENT AND INFORMATION TECHNOLOGY 260, 264 (OECD Informatics Stud. 1978).

<sup>18</sup> Kerstin Anér, *Privacy and Data Protection Legislation*, in HUMAN CHOICE AND COMPUTERS 2 151, 154 (Abbe Mowshowitz ed., 1980).

<sup>19</sup> *Id.*

<sup>20</sup> PLATO, REPUBLIC: BOOK VII § 514a–520a.

interpret the shadows as the actual reality. She went on to note the danger of this paradigm: “He should also worry about who has the right to define the contours of his data shadow, because that shadow will often be treated instead of himself, and a Procrustes fit may prove very uncomfortable indeed.”<sup>21</sup> This reading is confirmed by a motion she introduced in the Swedish Riksdag (Parliament) in 1982: “the computer’s image of a person or other phenomenon in the perceptual world.”<sup>22</sup>

We have not located an earlier use of the phrase by Dr. Anér nor an explicit link to Plato. She was a prolific writer, though, and while her papers are in the Swedish national archives, most have not been digitized.<sup>23</sup>

There are earlier near-miss references. In 1964, Vance Packard wrote “[c]redit bureaus like to boast, when soliciting local business, that their records stick to a man like a shadow.”<sup>24</sup> Similarly, Myron Brenton wrote in that same year that “[r]egardless of where—or how often—you move, your credit report—good or bad—can follow you with the tenacity of your own shadow.”<sup>25</sup> A 1973 newspaper heading used the exact phrase, but with “shadows” as a verb, not a noun: “Computer Data Shadows School Children.”<sup>26</sup>

Alexander Solzhenitsyn understood the concept, though he did not use this terminology.<sup>27</sup>

As every man goes through life he fills in a number of forms for the record, each containing a number of questions. A man’s answer to one question on one form becomes a little thread, permanently connecting him to the local center of personnel records administration. There are thus hundreds of little threads radiating from

---

<sup>21</sup> Anér, *supra* note 18, at 160.

<sup>22</sup> Motion in the Swedish Riksdag (Parliament), Kerstin Anér, *De Statliga Datasystemen* 1981/82:179 (Jan. 12, 1982), [https://www.riksdagen.se/sv/dokument-och-lagar/dokument/motion/de-statliga-datasystemen\\_g502179/](https://www.riksdagen.se/sv/dokument-och-lagar/dokument/motion/de-statliga-datasystemen_g502179/) [<https://perma.cc/DSX4-DWTN>] (translation by Peter Krantz).

<sup>23</sup> Conversation with Peter Krantz (Sep. 27, 2023).

<sup>24</sup> VANCE PACKARD, *THE NAKED SOCIETY* 183 (1964).

<sup>25</sup> MYRON BRENTON, *THE PRIVACY INVADERS* 28 (1964).

<sup>26</sup> Benjamin Fine, *Computer Data Shadows School Children*, PITTSBURGH PRESS, Feb. 18, 1973, at F-16, <https://www.newspapers.com/clip/80364948/computerdatashadows/> [<https://perma.cc/J8TW-GT3Z>].

<sup>27</sup> ALEXANDER SOLZHENITSYN, *CANCER WARD* 192 (Nicholas Bethell & David Burg trans., Bantam Books 1969) (1966).

every man, millions of threads in all. If these threads were suddenly to become visible, the whole sky would look like a spider's web, and if they materialized as rubber bands, buses, trams and even people would all lose the ability to move, and the wind would be unable to carry torn-up newspapers or autumn leaves along the streets of the city. They are not visible, they are not material, but every man is constantly aware of their existence. The point is that a so-called completely clean record was almost unattainable, an ideal, like absolute truth. Something negative or suspicious can always be noted down against any man alive. Everyone is guilty of something or has something to conceal. All one has to do is to look hard enough to find out what it is.

There are some related terms, notably “digital shadow,” “data image,” “digital footprint,” “data portrait,” and “data profile.” None of these appear to have been used earlier than data shadow. Roger Clarke was, as best I can tell, the first to use “digital shadow”;<sup>28</sup> it was also cited in one court opinion describing an expert witness' qualifications.<sup>29</sup> (By contrast, “data shadow” does not appear to ever have been used by a court.) The first law review use of “digital shadow,” per Lexis, was in 2010;<sup>30</sup> it cites to a 2009 blog post.<sup>31</sup> Professor Kenneth Laudon apparently coined the phrase “data image” in a 1986 book.<sup>32</sup> It achieved some currency, but the only early usage appears to have been in 1988.<sup>33</sup> “Digital footprint” is more widely used, but again, not in the early years. Marc Rotenberg noted that “data shadow,” “data portrait,” and “data profile” are equivalent terms;<sup>34</sup> he appears to have been the first to use

---

<sup>28</sup> Roger Clarke, *The Digital Persona and Its Application to Data Surveillance*, 10 INFO. SOC'Y 77 (1994).

<sup>29</sup> *United States v. Bradbury*, No. 2:14 CR 71 PPS, 2015 WL 4627018, at \*2 (N.D. Ind. July 31, 2015).

<sup>30</sup> Danielle Keats Citron, *Fulfilling Government 2.0's Promise with Robust Privacy Protections*, 78 GEO. WASH. L. REV. 822, 835 (2010).

<sup>31</sup> Brian Solis, *Casting a Digital Shadow; Your Reputation Precedes You*, BUS. GROWTH BLOG (July 17, 2009), <https://www.briansolis.com/2009/07/casting-a-digital-shadow-your-reputation-precedes-you/> [<https://perma.cc/69NR-Y3TS>].

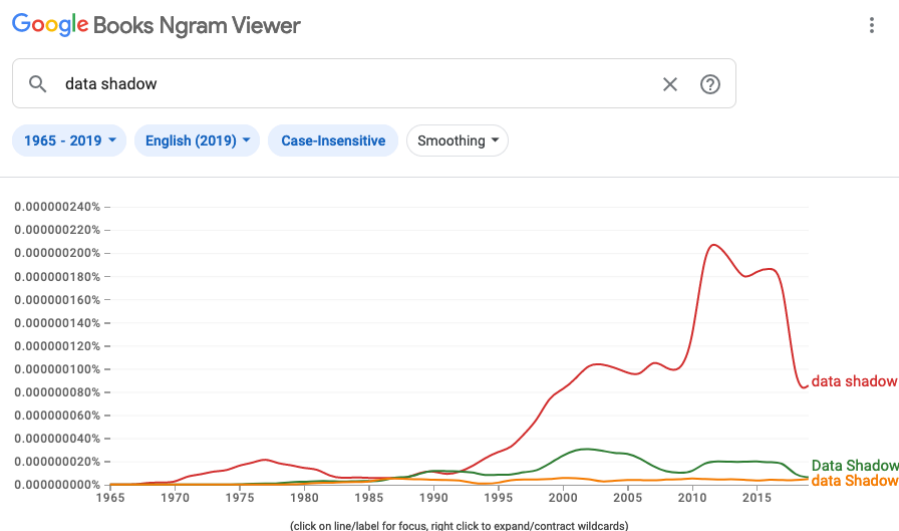
<sup>32</sup> KENNETH LAUDON, DOSSIER SOCIETY: VALUE CHOICES IN THE DESIGN OF NATIONAL INFORMATION SYSTEMS 4 (1986).

<sup>33</sup> Gary T. Marx, *The Maximum Security Society*, in *NEW TECHNOLOGIES AND PENAL JUSTICE* 468, 475 (Marc LeBlanc et al. eds., 1988).

<sup>34</sup> Marc Rotenberg, *Electronic Monitoring of Employees: Should Government Limit the Practice?: Computer Professional's View*, 1 EMP. TESTING: L. & POL'Y REP. 25 (1992).

any of those latter two phrases.<sup>35</sup>

The Google Ngram analysis of “data shadow” shows an interesting story of the usage history of “data shadow.”



There was a brief peak of usage in the 1970s, presumably following Dr. Anér’s English writings. The usage started rising again in the early 1990s. (Use of “Data Shadow” and especially “data Shadow” are likely false positives.) The question is what triggered the increased usage. One candidate is Marc Rotenberg’s 1991 testimony before a Senate subcommittee:

Third-party vendors of information, such as credit reporting agencies, augment the information available to an employer and help to create the fine-grained “data portrait” of an individual. Decisions taken solely on the basis of one’s “data profile” or “data shadow” have raised great concern in Europe where a comprehensive privacy policy is now under consideration.<sup>36</sup>

There were a few more uses in the early 1990s, in academic

<sup>35</sup> Ironically, one of my own articles uses the word “profile” in this context. See Steven M. Bellovin et al., *When Enough is Enough: Location Tracking, Mosaic Theory, and Machine Learning*, 8 N.Y.U. J.L. & LIBERTY 555, 592 (2014).

<sup>36</sup> *Privacy for Consumers and Workers Act: Hearing on S.516 Before the Subcomm. on Emp’t & Productivity of the S. Comm. on Lab. & Productivity*, 102nd Cong. 73 (1991) [hereinafter *Privacy for Consumers and Workers Act*] (statement of Marc Rotenberg, Washington Director, Comput. Pros. for Soc. Resp.).

settings.<sup>37</sup>

The first popular use of “data shadow” in the 1990s was apparently in 1993, in a series of six articles in the Toronto *Globe and Mail*. The journalist, Mary Gooderham, used the phrase in the title to her articles but does not cite a source. She defines it by usage and links it to surveillance:<sup>38</sup>

Every credit-card purchase casts a shadow. So does each entry into a security-minded workplace or store, application for health insurance, call to a phone-sex service, selection of a pay-per-view movie or movement of a cellular telephone.

It’s called a *data shadow*, and it grows longer as computer data bases record more and more of our daily activities. The image reveals who we are, where we go, whom we know, what we do and when—sort of an electronic alter ego that is required for us to obtain credit, receive welfare benefits, vote, get a job or cross a border without a hassle.

The Global Village is fast growing into Surveillance City.

By the late 1990s, the phrase was becoming common. There were continuing academic uses,<sup>39</sup> uses in legal guidance,<sup>40</sup> and even religious mentions, not only by the Church of England<sup>41</sup> but also in apocalyptic literature.<sup>42</sup>

At this remove, there does not seem to be any definitive answer to why the phrase went into eclipse and then came back. Certainly, Dr.

<sup>37</sup> Colin Beardon, *Social Citizenship in the Information Age*, in COMPUTERS AND SOCIETY 9 (Colin Beardon & Diane Whitehouse eds., 1993); Roger Clarke, *Computer Matching and Digital Identity* (3d Conf. on Comput., Freedom, & Priv., Mar. 1993), <http://cpsr.org/prevsite/conferences/cfp93/clarke.html/> [https://perma.cc/RFJ4-KBNY].

<sup>38</sup> Mary Gooderham, *Living in the Data Shadow: Farewell to the Private Life*, GLOBE & MAIL, Aug. 14, 1993, at A1, <https://perma.cc/AZU8-4L6E> (the series continued for about the next week).

<sup>39</sup> See, e.g., Crauford D. Goodwin & Alex Rogers, *Cyberpunk and Chicago*, in THE STATE OF THE HISTORY OF ECONOMICS 39, 45 (James P. Henderson ed., 1997).

<sup>40</sup> M. DREW JACKSON & TIMOTHY L. TAYLOR, THE INTERNET HANDBOOK FOR CANADIAN LAWYERS 203 (3d ed. 2000).

<sup>41</sup> CHURCH OF ENGLAND BOARD FOR SOCIAL RESPONSIBILITY, CYBERNAUTS AWAKE!: ETHICAL AND SPIRITUAL IMPLICATIONS OF COMPUTERS, INFORMATION TECHNOLOGY AND THE INTERNET 32, 73, 78 (1999).

<sup>42</sup> WILLIAM T. JAMES, FORESHADOWS OF WRATH AND REDEMPTION 270 (1999).



Anér was not a major figure in the Anglophone privacy space. A lot of work was done on privacy in the 1960s,<sup>43</sup> but a Lexis search did not turn up any uses of the phrase “data shadow” in the 1980s. Arguably, the deregulatory agenda of the Reagan era contributed. President Ronald Reagan said, in his inaugural address, “government is not the solution to our problem; government is the problem.”<sup>44</sup> In such an environment, proposals for privacy regulation might<sup>45</sup> have gotten short shrift in Washington. Legal academics, though, did not ignore privacy. A Lexis search shows 154 articles mentioning “information” and “privacy” during the 1970s, but 3,228 during the 1980s. Clearly, there can be many confounding factors in such a crude metric. Some of the retrieved articles were about Fourth Amendment or Fifth Amendment issues, rather than the sort of privacy issues encompassed by the phrase “data shadow,” but the large difference is still suggestive.<sup>46</sup> It might have been sexism, but there is no evidence for that, one way or another. Most likely, the phrase persisted in the international privacy community, especially in Europe, until Marc Rotenberg reintroduced it in the United States.<sup>47</sup> (Rotenberg recalls hearing the phrase used in the late 1980s.<sup>48</sup>) Supporting this, in 1988, Prof. Harvey Choldin was told of its use in Germany.<sup>49</sup>

The concept, by whatever name, is at the heart of the mosaic

---

<sup>43</sup> See Bellovin, *supra* note 4 (lists many early references).

<sup>44</sup> Ronald Reagan, Inaugural Address (Jan. 21, 1981) (transcript available at <https://www.reaganfoundation.org/media/128614/inaguration.pdf>).

<sup>45</sup> That said, there were some noticeable privacy laws passed, *e.g.*, Electronic Communications Privacy Act, 18 U.S.C. §§ 2510–2521, 2701–2710, 3121–2126 (1986), which extended the Wiretap Act to protect data transmissions and not just voice, protected subscriber records, and created a procedure for authorizing pen register and trap-and-trace orders; Employee Polygraph Protection Act, 29 U.S.C. §§ 2001–2009 (1988), which effectively barred private sector employers from using lie detectors on their employees; Video Privacy Protection Act, 18 U.S.C. § 2710 (1988), which protected video purchase or rental records. This latter act was passed in the wake of Judge Robert Bork’s confirmation hearings for a Supreme Court nomination, in the course of which his video records were leaked. See also *The Video Privacy Protection Act as a Model Intellectual Privacy Statute*, 131 HARV. L. REV. 1766 (2018).

<sup>46</sup> See, *e.g.*, David Dolinko, *Is There a Rationale for the Privilege Against Self-Incrimination?*, 33 UCLA L. REV. 1063 (1986); C. Ryan Reetz, *Warrant Requirement for Searches of Computerized Information*, 67 B.U. L. REV. 179 (1987).

<sup>47</sup> See *Privacy for Consumers and Workers Act*, *supra* note 36 (specifically noting European interest in privacy).

<sup>48</sup> Conversation with Marc Rotenberg (July 19, 2023).

<sup>49</sup> Harvey M. Choldin, *Government Statistics: The Conflict Between Research and Privacy*, 25 DEMOGRAPHY 145, 147 (1988).

theory of the Fourth Amendment.<sup>50</sup> Some courts have explicitly linked the mosaic theory to “profiles,” though apparently not to “data profile.”<sup>51</sup> But “data shadow” is not just the oldest term, it’s the most accurate. Phrases such as “data image,” “digital footprint,” “data portrait,” and “data profile” connote something static or localized: someone has a profile about a person or owns their portrait. But a shadow follows you around,<sup>52</sup> just as your aggregated data does today.

Still, the original source seems to have been forgotten. Academic articles continue to cite Westin.<sup>53</sup> No one else seems to have used “data shadow” in its modern sense before Dr. Kerstin Anér. She clearly used it no later than 1972 and often thereafter. She should get full credit.

---

<sup>50</sup> See, e.g., Orin S. Kerr, *The Mosaic Theory of the Fourth Amendment*, 111 MICH. L. REV. 311 (2012); Danielle K. Citron & David Gray, *A Shattered Looking Glass: The Pitfalls and Potential of the Mosaic Theory of Fourth Amendment Privacy*, 14 N.C. J.L. & TECH. 381 (2013).

<sup>51</sup> *United States v. Maynard*, 615 F.3d 544, 562 (D.C. Cir. 2010); *Commonwealth v. Perry*, 184 N.E.3d 745, 757 (Mass. 2022).

<sup>52</sup> Think of Peter Pan’s shadow, which needed to be sewn back on by Wendy Darling. J.M. BARRIE, *PETER PAN* 39 (Charles Scribner’s Sons 1911).

<sup>53</sup> See, e.g., *Diamond et al.*, *supra* note 3, at 431.