Algorithmic surveillance and the political life of error

Aradau, C.; Blanke, T.

DOI
10.5334/jhk.42

Publication date
2021

Document Version
Final published version

Published in
Journal for the History of Knowledge

License
CC BY

Citation for published version (APA):
Aradau, C., & Blanke, T. (2021). Algorithmic surveillance and the political life of error. Journal for the History of Knowledge, 2(1), Article 10. https://doi.org/10.5334/jhk.42

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Concerns with errors, mistakes, and inaccuracies have shaped political debates about what technologies do, where and how certain technologies can be used, and for which purposes. However, error has received scant attention in the emerging field of ignorance studies. In this article, we analyze how errors have been mobilized in scientific and public controversies over surveillance technologies. In juxtaposing nineteenth-century debates about the errors of biometric technologies for policing and surveillance to current criticisms of facial recognition systems, we trace a transformation of error and its political life. We argue that the modern preoccupation with error and the intellectual habits inculcated to eliminate or tame it have been transformed with machine learning. Machine learning algorithms do not eliminate or tame error, but they optimize it. Therefore, despite reports by digital rights activists, civil liberties organizations, and academics highlighting algorithmic bias and error, facial recognition systems have continued to be rolled out. Drawing on a landmark legal case around facial recognition in the UK, we show how optimizing error also remakes the conditions for a critique of surveillance.

This article is part of a special issue entitled “Histories of Ignorance,” edited by Lukas M. Verburgt and Peter Burke.

**Keywords:** error; ignorance; surveillance; biometrics; algorithms

"All epistemology is born in fear: fear of the several sorts of errors that can corrupt, undermine, or impede knowledge," historian of science Lorraine Daston reminds us. ¹ Diagnosing and reducing—if not eliminating error—have shaped the quest for valid knowledge and truth. The horizon of error, however, has not been limited to knowledge production in science and philosophy. How errors emerge, how they are discovered, to whom they are attributed, and how they are to be tackled have been deeply political questions. Concerns with errors, mistakes, and inaccuracies have shaped political debates about what technologies do and where and how certain technologies can be used and for which purposes. Yet, unlike ignorance, uncertainty, or secrecy, error has received scant attention in the emerging field of ignorance studies.² Studies in the history of knowledge have proposed to explore the “dark side” of knowledge through failure and ignorance rather than error.³ The inattention to error might be partly due to epistemologies that associate error with an obstacle that must be surmounted in the quest for knowledge and truth.

Despite attempts to eliminate, reduce, or neutralize error, errors emerge again and again. They have increasingly become everyday occurrences, displayed on our computer screens and other devices. Errors are now inherent to vernacular modes of knowledge and mundane practices of human-machine interaction. Yet, errors can still give rise to public mobilization against the development and deployment of new technologies. Science and Technology Studies (STS) scholar Rebecca Slayton has shown how arguments about computational error and failure entered public debates about missile defense in the US in the 1970s at the

---

¹ Daston, “Scientific Error,” 4.
² Gross and McGoey, Handbook of Ignorance Studies.
³ Dupré and Somsen, “History of Knowledge.”
same time as computer experts became increasingly able to “speak authoritatively.” While in the twentieth century, error was mobilized in public debates in relation to weapon technologies in particular, error has recently emerged as a key argument against the use of another computational technology: automated facial recognition.

The inaccuracies of facial recognition and the errors of algorithmic processing of facial images that have resulted in racial and gender bias have created a highly visible political problematization of error analysis. Revelations of high error rates have led to numerous inquiries and even public outcries. A report by the civil liberties organization Big Brother Watch in the UK has shown that facial recognition systems used by the police are erroneous nine times out of ten. An academic evaluation of facial recognition for the London Metropolitan Police Service also found high error rates, showing the inaccuracy of “watchlists” and the ambiguities of defining whom they should include.

In the US, an earlier report by the American Civil Liberties Union (ACLU) concluded that facial recognition software similar to Amazon Rekognition incorrectly identified twenty-eight Congresspersons as having been arrested for a crime. The ACLU Attorney Jacob Snow pointed out that the incorrect matches were “disproportionately of people of color, including six members of the Congressional Black Caucus.” In the wake of the 2020 global Black Lives Matter protests, IBM ended the development of facial recognition technology, while Microsoft committed to not selling its technology to law enforcement, and Amazon set a one-year moratorium on the police’s use of their technologies. Despite calls for more bans, however, the “big players” of facial recognition have not been involved in these moratoriums. Even as US cities like San Francisco and Oakland have banned the use of facial surveillance by the police and other agencies, the Metropolitan Police and the South Wales Police have continued to trial it in the UK, while the Indian government has recently approved the deployment of facial recognition across the country.

The binaries of accuracy and error, precision and bias, fairness and discrimination underpin these public controversies over facial recognition for surveillance and policing. Like biometric technologies, facial recognition raises questions about the treatment of error in statistics and machine learning. How has machine learning transformed the understanding of error and how have problematizations of error translated into public controversies about facial recognition? We argue that engineers and scientists work with a machine learning epistemology of error that is often difficult to reconcile with public approaches to error.

To unpack these distinctions, we situate machine learning epistemologies of error in relation to earlier discussions of error. We start by locating error in the history of knowledge and ignorance and then trace its role in the controversy over anthropometry and fingerprinting for policing and surveillance in the nineteenth century. Secondly, we analyze how facial recognition has been transformed through machine learning algorithms that optimize error. These historical moments allow us to understand the specificity of machine learning epistemologies of error and trace the limitations of mobilizing a critique of error in algorithmic surveillance today by focusing on a landmark legal case in the UK and related public contestations of facial recognition. Facial recognition is especially suited to capture this transformation, as it one of the few machine learning algorithms that is already widely applied in real-life scenarios. As it is so widely deployed, its errors have also attracted public attention and extensive debate. This article explains how, rather than eliminating error or neutralizing it, machine learning algorithms multiply the measurement of errors to optimize them.

Our contribution to the history of ignorance is theoretical and political. Theoretically, we propose a “history of the present” approach to error in order to understand how it “becomes a problem, raises discussion and debate, incites new reactions, and induces a crisis in the previously silent behaviour, habits, practices, and institutions.” Error is not limited to laboratories or scientists’ debates, but it is invoked in arguments that shape public debates and the deployment of biometric and algorithmic technologies for surveillance. Politically, analyzing facial recognition through the prism of error can shed light on the limits of arguments about error in contesting surveillance technologies today. Different arguments about error

---

4 Slayton, Arguments that Count, 225.
5 Big Brother Watch, “Face Off.”
6 Fussey and Murray, “Independent Report.”
7 Snow, “Amazon’s Face Recognition.”
8 Ibid.
9 Ovide, “Banning Facial Recognition.”
10 Fowler, “Black Lives Matter.”
11 Cowan, “San Francisco.”
12 Foucault, “Fearless Speech,” 74.
can be mobilized in various social fields, as epistemologies of error move from laboratories and textbooks to public controversies and court rooms.

Knowledge, Error, Ignorance

From scientific to judicial error and from technical to human error, errors have been traced, tracked, neutralized, or litigated. In his reflections on the differential treatment of error in science and technology, Peter Galison highlighted the centrality of error to technological systems: “When engineers make a major error it is common for a failure inquiry to be established, staffed by heavyweights, properly funded and well publicized.”13 However, similar inquiries into scientific errors would be “almost unthinkable—barring the suspicion of gross malfeasance.”14 According to Galison, these asymmetries cannot be simply explained by the practical consequences technical errors can entail. Rather, they are due to the differences in the social worlds that scientists and engineers inhabit as well as the different questions they ask. For engineers, questions of malfunction and liability are entangled with how error is understood and managed. For physicists, errors are tied to questions about truth. Error is differently mobilized across time, disciplines, and social fields.

While Galison rightly underscores distinctions between the treatment of error by scientists and engineers, historians and philosophers of science have increasingly drawn attention to the ubiquity of error in scientific practice. As Daston’s quote with which we started this article suggests, error and scientific epistemologies have been historically entwined. Indeed, Daston argues that, since the seventeenth century, epistemology has consisted in an “elaborate nosology of errors: what their species and varieties are, and how they may best be avoided or cured.”15 Diagnosing, eliminating, or, at a minimum, “taming” error have shaped scientific practice and the formation of vigilant scientific subjects.16 Error remains an obstacle to be prospectively avoided or a mistake to be retrospectively corrected. The path to truthful knowledge is shaped by carefully controlling, if not eliminating, error. Scientific fields develop varied techniques for diagnosing and controlling error. Philosophers of science analyze different types of errors and their taxonomies across disciplines, with experimental and statistical errors garnering most attention.17 They concur that errors have, however, remained largely “residual” in the philosophy of scientific knowledge. These ambiguities of error as avoidable/unavoidable, random/systematic, knowable/unknowable, prospective/retrospective have shaped debates about science, technology, and politics.

Although errors have often featured in public debates about weapons, medical, and other technologies affecting the lives of individuals and populations, analyses of error remain sparse in comparison to studies of knowledge production and circulation. Even as the interdisciplinary field of “ignorance studies” has recently proposed to supplement work on the production of knowledge by the “other side of knowledge,” error has remained largely absent.18 This might be partly due to the use of “ignorance” as the over-arching term for a research agenda. Ignorance can be understood as both different from and encompassing heterogenous forms and sources of non-knowledge. Thus, analyses of ignorance can focus on the absence or otherwise hindering of knowledge. In this approach, error and ignorance are often used interchangeably.19 Charles Mills’ coinage of white ignorance comprises “both false belief and absence of true belief,” both error and ignorance.20 Other scholars differentiate error and ignorance, with “error” located as a subcategory of ignorance understood as “distortion.”21 While ignorance points towards absence, error is connected to the presence of a falsehood, distortion, or aberration.

Historians and STS scholars have paid more attention to the transformations of what counts as error or is identified as such. According to historian David Bates, enlightenment knowledge was “structured by error, of error as the site of both risk and promise.”22 As any knowledge could be potentially erroneous, everything harbored the risk of error. At the same time, the discovery of error promised horizons of new knowledge. Bates’ analysis is framed in the context of the “rise of statistical thinking.”23 According to probabilistic reasoning, errors could no longer be eliminated in advance but needed to be tamed. As Bates

---

13 Galison, “Author of Error,” 66–67.
14 Ibid., 68.
15 Daston, “Scientific Error,” 7.
16 “Taming” error paraphrases Ian Hacking’s famous formulation expressed in the title of his book The Taming of Chance.
17 Boumans et al., Error and Uncertainty.
18 Proctor and Schiebinger, Agnotology; Gross and McGoey, Handbook of Ignorance Studies.
19 Proctor, “A Missing Term.”
20 Mills, “White Ignorance,” 233.
21 Smithson, Ignorance and Uncertainty.
22 Bates, Enlightenment Aberrations, ix.
23 Porter, Statistical Thinking.
put it, “Enlightenment masters the unknown not so much by eliminating it but by controlling it.” Bates highlighted how the problem of error was not limited to statistical reasoning, but it also underpinned the thinking of political revolutionaries for whom “political forms had to be constructed not according to some model of truth but with the problem of error at the center.”

Different approaches to error became the object of controversy, as scientific practices were tied to different political goals and circulated across national boundaries. Donald MacKenzie has shown the uneasy relationship of Francis Galton’s eugenicist variability analysis with statistical error theory. “For the error theorists,” MacKenzie argues, “variability (‘error’) was something to be eliminated, or at least in practice to be controlled and measured.” For Galton, variability could be potentially desirable rather than eliminable. Approaches to error vary across different scientific fields and are shaped by political and national values and practices. Graeme Gooday distinguished between a German and a British approach to error in nineteenth-century electrical measurement. Building on Gaussian theory, German scientists and engineers “considered it essential to undertake a large number of measurements to decrease the error to a certain minimum threshold.” The British “devoted more effort to maximizing accuracy by enhancing the number of antecedent precautionary measures against experimental error.”

In Physics as a Calling, historian of science Kathryn Olesko has shed light on how error analysis reshaped the discipline of physics when it shifted from a qualitative approach that “was still only a matter of talking about the crude instruments, inaccurate scales, and the effect of disturbing environmental conditions, such as air pressure, temperature, and humidity” to a quantitative one. The quantification of error, developed in nineteenth-century Germany by physicist Franz Neumann through the method of least squares, transformed the relations between physics and mathematics, theory and experiment, exactitude and uncertainty. Olesko’s distinction between qualitative and quantitative understandings of error and Gooday’s differentiation of precautionary and statistical approaches to error are indicative of the role that probabilistic reasoning introduced for error analysis.

If the causes, analysis, and manifestation of error were at the heart of nineteenth-century scientific controversies, error was also mobilized in twentieth-century public debates about the development and implementation of technologies, particularly those that could have lethal consequences for individuals and populations. Slayton has shown how complex missile weapons systems came to be seen as liable to catastrophic computer error. While in the 1950s debates about complex weapons systems were shaped by the “disciplinary repertoire” of physics and electrical engineering, in the 1960s another argument was introduced to the public debate—that missile weapons systems would “lead to an unprecedented reliance on complex, failure-prone computers.” Slayton’s analysis of “arguments that count” has highlighted the increasing disciplinary authority of computer science and the ability of computer scientists and software engineers to speak about the risks of complex systems. This ability was developed through different classifications of error: errors that can be calculated and errors that are due to the unpredictability of human practices and social institutions. The software engineers’ distinction between reliability and safety resonates with that between error and failure. As philosopher John Roberts explained, errors refer to “missteps, omissions, oversights and mistakes involved in the execution of a particular activity,” while failure renders “the dissolution, collapse, breakdown of a given programme, project or systematic endeavour.” Whether errors are connected to catastrophic failures, preventable mistakes, unexpected bugs, the craft of the engineer, or the precision of instruments is an epistemological as much as a political question.

Drawing on Slayton’s analysis of “arguments that count,” we suggest that these distinctions gain political significance because they resonate or are ignored or silenced in public controversies. As the question of error circulates between social and scientific worlds, its ambiguities come to be exposed through contestation by different actors who inhabit these worlds. This was the case of DNA forensic technologies when they came to be used in US courts, and it is true for current population surveillance technologies. Facial recognition has not only become one of the most successful biometric technologies but also one of the most hotly

---

24 Bates, Enlightenment Aberrations, 10.
25 Ibid, 31.
26 MacKenzie, Statistics in Britain, 58.
27 Gooday, Morals of Measurement, 74.
28 Olesko, Physics as a Calling, 233.
29 Ibid.
30 Slayton, Arguments That Count, 106.
31 Roberts, Necessity of Errors, 190.
32 Derksen, “Sociology of Measurement.”
debated for its errors. More recently, facial recognition has become possible at scale through advancements in machine learning. Its wide deployment has led to high-profile public debates and controversies about the errors it produces that have resulted in bias and discrimination.

The political life of error entails attention to how questions of knowledge and non-knowledge are mobilized in scientific and public controversies rather than to definitions or taxonomies. We approach error as one of the forms of non-knowledge whose cognitive ecology infuses our social and political lives. In the following two sections, we retrace two different moments in the development of surveillance technologies: the biometric technologies developed by Bertillon and Galton in the nineteenth century and algorithmic processing of facial images in the twenty-first century.

Taming Error and Biometric Surveillance

How has the political life of error shaped arguments about technologies of biometric surveillance such as fingerprinting? Histories of surveillance technologies trace the development of devices for and methods of metricizing bodies to Alphonse Bertillon’s anthropometry in nineteenth-century France and Francis Galton’s fingerprints in Britain. Error has remained residual in historical analyses of biometric technology, probably as statistical reasoning shifted from an emphasis on the language of the “law of error” and error curve to that of the normal distribution. Measuring the body entailed all the problems of accuracy and error—training, instrumentation, body pose, or recording. As criminologist Simon A. Cole explained, “[t]he recording of anthropometric measurements was an elaborate dance, in which the movements of both operator and prisoner had been strictly choreographed by Bertillon himself.” Controlling error required the human craft of using instruments in a standardized fashion.

Based on Adolphe Quetelet’s law of error, Bertillon held that people were distributed along the normal curve and therefore individuals would be differentiated by specific measurements from each other. For Quetelet, “any deviations from the golden mean of l’homme moyen were mere imperfections, even errors.” For Bertillon, the variation could help distinguish between various suspects and criminals who were concealing their identities. He deployed statistical reasoning to decide on what constituted an accurate recording of a bodily measurement and claimed that measurement “within the limit of possible error” means that “the probability of identity becomes very high, and it is equivalent to almost certainty.” Repeated measurements converged towards the average, which became the measure of truthful knowledge. The combination of bodily measurements would then render individuals uniquely identifiable.

In this quest for accuracy, Bertillon located two forms of error: an error of measurement and an error of interpretation. The differences between measurements, he cautioned the operators, should not exceed “the approximation” indicated for each measure. Francis Galton, the proponent of fingerprinting and adversary to Bertillon, also acknowledged the importance of trained judgments to the success of Bertillon’s system in Paris, given a designated space, “where numerous clerks, under careful inspection, working day by day, have acquired a remarkable degree of sureness, of deftness in their work.” A second source of error was deceit (tricherie) or ill will (mauvaise volonté). These errors emerged in relation to the operator’s subjectivity—both that of skill in maneuvering the instruments of measurement and in reading the “suspect.” They were errors of craft and judgment.

For biometric surveillance, error was not to be eliminated but tamed. Error within the bounds of approximation was acceptable. If it surpassed these bounds, error could discredit a technology, a craft, or even a theory. While there are multiple reasons for the replacement of Bertillon’s anthropometry with Galton’s fingerprinting, error was problematized in the competition between their approaches. Galton thought anthropometry “would result in an unacceptable high rate of false identification because no account was taken between different bodily characteristics.” Moreover, facial images did not lend themselves to

---

33 Beck and Wehling prefer the use of not knowing or non-knowledge to that of ignorance “mainly in order to avoid the moral devaluation that might be linked to ignorance.” Aradau argues that non-knowledge draws attention to a range of vocabularies from uncertainty and ambiguity to secrecy, error, contingency, or ignorance. Beck and Wehling, “The Politics of Non-Knowing,” 62; Aradau, “Assembling (Non)Knowledge.”
34 Cole, Suspect Identities, 36.
35 Porter, Statistical Thinking, 60.
36 Bertillon, Identification Anthropométrique, ix, translation ours.
37 Galton, “Bertillon System,” 569.
38 Bertillon, Identification Anthropométrique, ix, translation ours.
39 Higgs, Information State, 114.
metrification, and Bertillon still had to resort to verbal description (portrait parlé) and photographs (both frontal and side). This introduced further possibilities for error.

To buttress his method of fingerprinting, Galton relied on statistical error calculations. His *Finger Prints* dedicated a chapter to their evidential value. He asked: “[G]iven two finger prints, which are alike in their minutiae, what is the chance that they were made by different persons?”40 While Galton offered a detailed statistical reasoning for the infinitesimal rate of error, the book is traversed by concerns about how to eliminate errors from the wider socio-technical apparatus required for fingerprinting. From the quality of the ink, the type of card required to take a fingerprint to the optician’s lens to study them, Galton traced a standardizing and disciplining apparatus harnessed towards the taming of error. Thus, fingerprinting is not readily subsumed to the strictures of “mechanical objectivity,” even as fingerprints appeared to eschew the subjectivity of the operator and the human craft required by Bertillon’s anthropometry.41 Yet, it was the absence of error that circulated in public imaginaries. As one commentator put it at the time, “[t]here is no possible margin of error, as fingerprints are absolute impressions taken from the body itself.”42 Galton himself extolled the accuracy of fingerprints in an interview and declared “the danger of making a mistake is so slight that it is not worth considering.”43 This was the fragile effect of an apparatus carefully calibrated to control error.

The controversy between anthropometry and fingerprinting also placed error within colonial imaginaries of ignorance and subjectivity. As Galton expounded in a letter to *The Times*, the introduction of fingerprinting in India was required by “the large proportion of their illiterate populations, who make marks but cannot write, partly on account of the inability felt by most Europeans in accurately distinguishing the features of men of the darker races, and partly on account of the prevalence of false witnesses among them.”44 Colonial understandings of race permeated his claims for the efficacy of fingerprints, even as his quest to find race-based statistical patterns failed. While anthropometry was deployed in the metropole, it was not until much later that fingerprinting became a general technique of identifying populations as it had been met with resistance given its association with abjection and criminality.45

More recently, fingerprinting has been supplemented by other biometric technologies, from facial to iris recognition and even behavioral biometrics. Biometric surveillance has also been transformed by developments in computer science and machine learning. Galton himself experimented with basic techniques that are still relevant to facial recognition. He tried to create a composite image of the “average man” by superimposing facial images of members of a group.46 It is to these transformations and their effects on the political life of error that the next section turns.

**Optimizing Error: Facial Recognition as Algorithmic Surveillance**

Galton had to rely on training the human eye to read differences between the minutiae of fingerprints. He used analogue techniques and detailed manual labor to superimpose faces. With computers, images and fingerprints could become data to be read by machines and subsequently processed by algorithms. In the 1960s, computer scientist Woody Bledsoe began to translate patterns of faces into data.47 Bledsoe, who later became the president of the Association for the Advancement of Artificial Intelligence, used two thousand images in a book of police mugshots as his “database” for making comparisons with new photographs to detect similarity. To this end, he marked facial “features” and their locations such as the mouth, nose, or eyes using a so-called RAND tablet, which could record coordinates on a grid. A list of twenty distances were calculated and stored in a computer together with a person’s identification. Facial recognition became an algebraic comparison of distances between facial features. Largely funded by the CIA, Bledsoe’s approach tried to create “a fully automated Bertillon system for the face.”48

Bledsoe’s successors argued that computers could identify faces better than humans and with fewer errors.49 In the 1970s, facial recognition added more features to Bledsoe’s approach. Statistical hypothesis

---

40 Galton, *Finger Prints*, 100.
41 Daston and Galison, *Objectivity*.
42 Fosdick, “Passing of the Bertillon System,” 367.
43 Galton, “Bertillon System,” 569.
44 Galton, “Personal Identification.”
45 Breckenridge, *Biometric State*; Higgs, *Information State*.
46 Stigler, *Seven Pillars*, 35.
47 Bledsoe, “Facial Recognition System.”
48 Raviv, “History of Facial Recognition.”
49 Goldstein et al., “Identification of Human Faces.”
testing was employed to minimize the risk of a false generalization from a small sample. Error analysis in statistical testing targets a null hypothesis against an alternative hypothesis, assuming an ideal world where the null hypothesis is the rule. The error measures how much surprise and evidence to the contrary are allowed in such an ideal world. Statistical hypothesis testing was used to allow for faces to diverge from how facial features identify them in order to keep trust in the overall methodology. A computer could not "see" facial images.

Once images became digital, computers could "see" faces by counting pixels. With images as data, computers can identify, for example, objects in an image, by measuring shades of pixel color. As collections of pixels, images are not especially complex digital objects, but are highly dimensional, memory-intensive, and difficult to process. With images as data, biometrics could be automated and no longer had to rely on human measurements. Machine learning provided new tools to deal with very large image data. At the end of the 1980s and the beginning of the 1990s, so-called "eigenfaces" provided a new way to automatically recognize faces by reducing the statistical extensions of a facial image. Eigenfaces were a technique to deal with the high dimensionality of facial images and reduce errors by focusing on the important pixel values, but not all at the same time. To this end, they represented faces as vectors in a feature space, which meant that computers could start to identify relevant features and process images much faster. A variation of the eigenface approach attracted public attention in January 2001 during a trial implementation at the Super Bowl in the US to identify faces from surveillance images and compare them to digital mugshots.

Since then, machine learning and datafication have further transformed epistemologies of error. According to Matthew Jones, the practices of machine learning stem "more from an engineering culture of predictive utility than from a scientific culture of truth." This has implications for how errors are problematized, as machine learning has begun to tame error by incorporating automated means of error optimization. With machine learning, error has become part of the way a machine learns by itself. Error is still tamed, but the taming is at the same time an optimization. Any remaining error means that the machine could learn more and better. Each new piece of facial data is not seen as either fitting the existing model or not, but as enabling the model to learn permanently. New machine learning models have been built iteratively around what data is available at any moment in time. As one of the foremost promoters of machine learning, Andrew Ng, has put it, "[m]achine learning is a highly iterative process: You may try many dozens of ideas before finding one that you're satisfied with."  

Error could play a more optimistic role and stand for ever better optimization, as the data has become big enough to make it possible to reveal new answers to existing errors. For a long time, collections of images had been rare beyond the police databases of mugshots. These police photographs required a fixed perspective and pose from arrested individuals that were only possible in a standardized controlled environment. Wider collections of facial images had to be created at great expense by inviting volunteers to have their photographs taken at dedicated research institutions like universities. Image data collections changed drastically once digital cameras made possible the large-scale production of images, often from private collections that were shared online. Social media collections like Flickr meant that digital facial images could become big data. The now almost-forgotten Flickr and social media kickstarted a new era of surveillance through machine-learning algorithms that could process biometric data.

Large-scale real-world applications of facial recognition could only have been developed in the past fifteen years with the new approaches to image collection and neural network processing. In 2007, computer scientist Fei Fei Li and her colleagues began to assemble ImageNet, a very large database of labeled images including faces. ImageNet has become an important resource, as it enables facial recognition experts and other developers of image technologies to test and compare their latest algorithms at the yearly ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Another breakthrough came with the resurgence of neural networks. In the 2010s, a good error rate on ILSVRC and similar competitions was around 25 percent. In 2012, a new "deep convolutional neural net" (CNN) called AlexNet entailed a jump in the reduction of the error rate to 15.3 percent. AlexNet won ILSVRC with an error rate of less than half that of the second-best performing algorithm. Deep neural networks really got the attention of researchers and only two years later...
later another neural network called “Residual Network” managed to surpass human recognition of images, realizing the dream of Bledsoe and his successors.

For neural networks, errors mean that more data could be needed. Criticisms of high-error rate in the recognition of darker-skinned men and women led to large-scale efforts to collect digital facial images in Africa and Asia.\footnote{Yang and Murgia, “Facial Recognition.”} The recent success of Chinese companies in the facial recognition market has also been due to the fact that they could get access to existing databases in collections such as ImageNet. To this they added a home market of images of underrepresented Asian faces, including ones acquired through surveillance of the Uyghur population.\footnote{Yang and Murgia, “Data Leak.”} Earlier statistics calculations of facial distances did not include facial diversity deliberately, so that facial recognition could concentrate on pre-defined features and not be “distracted” by other elements such as skin color. As Goldstein et al. explained, “[t]he population was deliberately made homogeneous to the following extent: all persons were white males, aged between 20 and 50, beardless, without eyeglasses, and had no obvious abnormal features (e.g., scars).”\footnote{Goldstein et al., “Identification of Human Faces,” 749.} Compared to this statistical challenge to generalize from a small homogeneous dataset, big data introduces a shift towards summarizing and adapting to facial diversity.

Goldstein’s mugshots only represent a sample of a population’s faces, which encompasses structural inequalities and discriminations. Snapshots of everyday faces on social media promised to overcome (some of) these limitations. Yet, they also increased error rates, as computers had been trained on particular types of data, and social media and other internet images were deemed to be “wild” images that disrupted the machine learning of police mugshots.\footnote{Grother et al., “Face Recognition Vendor Test.”} Because all machine learning is trained on one specific task, more data is always needed to accommodate change. Databases like Flickr and ImageNet are only first steps in the computational dream of turning all populations into data to find answers to all possible errors. As Stephanie Dick stated, machine learning ”aims to develop algorithms that take a huge amount of data as input to a neural network and output a prediction rule....”\footnote{Dick, “Artificial Intelligence,” 5.} Earlier statistical attempts at taming error relied on features that a human like Bledsoe could make out in face images. Neural networks do not need this input anymore and can learn new features from data on their own. It is only now that we can say that computers have come to “see” faces in their own unique way. Error analysis is key to this, as it measures the distance between the model’s output and the data. Neural networks are optimized by closing this distance.

In the data-driven world of machine learning, errors are not just relevant to control and testing but are enabling through optimization. The optimization of error becomes automated as the process of adjusting machine learning models to achieve the best possible performance within a particular use case such as facial recognition proceeds. Automated optimization is at the heart of a tension between what we can know from past data and what we don’t know about algorithm performance against yet unknown data. Whereas computer scientists do not know how a model will perform in new situations, they “control” how the model optimizes with regard to the past data used to train it. This trade-off involves everyday political decisions. A surveillance system can be built, for instance, to capture all suspect faces, but it will also include many innocent people. One can also be designed to minimize the impact on the innocent, but this risks missing out on some suspects. Both are error functions that machine learning could consider. As computer scientist Anil Jain explained in the expert report submitted to the Cardiff High Court, “[t]o the best of my knowledge, no [automated facial recognition] system is without some error rates.”\footnote{Jain, “First Expert Report,” 11.}

What are the implications of these transformations of error for how surveillance technologies can become publicly contested? The final section shows how different approaches to error vie for credibility in contestations over facial recognition used for surveillance by the Metropolitan Police and the South Wales Police in the UK.

**Political Life of Error: Contesting Surveillance**

In a landmark case brought by the civil liberties campaigner Edward Bridges against the use of Automated Facial Recognition (AFR) by the South Wales Police in the UK, the Cardiff High Court of Justice mobilized arguments about accuracy in its decision that the use of the technology was both legal and proportionate. The judges highlighted the novelty of facial recognition, which is characterized by the algorithmic processing of digital data, making it possible to “indicate matches between faces captured through the CCTV recording...”\footnote{Grother et al., “Face Recognition Vendor Test.”}
and those held on the watchlist.” Facial recognition is assumed to be highly accurate, as it “enables the extraction of unique information and identifiers about an individual allowing his or her identification with precision in a wide range of circumstances.” While acknowledging the presence of errors, the judges differentiated error by algorithms and error by humans and thus closed off debates about error in the deployment of facial recognition for surveillance. As a police officer checks the algorithmic match, they conclude that there is a human “safeguard” against algorithmic bias and error.

Yet, the question of error is not put to rest so easily. Unlike the court decision, which subsumed algorithmic error to the corrective capacities of police officers, digital rights activists and NGOs have questioned the accuracy of these surveillance technologies and their effects for rights and civil liberties. Facial recognition, they point out, risks eliding the distinction between guilt and innocence and extending suspicion to more and more categories of the population. What captured the most public and academic attention was the high error rate, which translated into racial and gender bias. As AFR became used by police and other private actors in the UK, several NGOs published reports criticizing its use, which were followed by reports by the Information Commissioner, the Biometrics Commission, the House of Commons Science and Technology Committee, and a parliamentary debate. Following an announcement by the Metropolitan Police Service that AFR would be deployed in London, twenty-five rights, race, and equality organizations signed a petition asking for the use of the technology to be stopped.

The UK House of Commons held debates about AFR while Edward Bridges began an appeal against the High Court decision. The question of error and (in)accuracy traverses much of these debates and intersects with arguments about legality, individual rights, and state-citizen relations.

In their written evidence to the Science and Technology Committee, Big Brother Watch emphasized that facial recognition algorithms have “demographic accuracy biases—that is that they misidentify some demographic groups, particularly women and people of colour, at higher rates than others, such as white men.” They requested the Metropolitan Police to “carry out or commission demographic accuracy bias testing, and they told us that they would not because they did not view it as an issue.” The Information Commissioner Office also reinforced that the technology should not be deployed “until the current concerns over the technology’s effectiveness and potential bias have been fully resolved.” The criticism of high error rates and misidentification of innocent citizens was iterated across the political spectrum in the House of Commons debates on facial recognition.

Such unacceptability of error has emerged through systematic rather than random occurrence: not only is such high error unacceptable, but it has been attached to particular groups as racial bias. The police have rejected the criticisms of error and bias. Like the South Wales Police, the Metropolitan Police have argued that they mitigate “the potential impact of this [bias] within the operational context, where it should be noted, additional checks and balances are in place and the final decision is by a human operator.” Thus, it is the human gaze that justifies the diminution or neutralization of error, even as facial recognition algorithms have been deemed “superior” to humans since Bledsoe.

Police forces have also justified the deployment of facial recognition as “trials” or “pilots.” The Financial Times noted that Londoners were becoming “guinea pigs in a police experiment.” Facial recognition relies on the “live” deployment of trials and the continuous rendition of surveillance as “trial.” As the chair of the Science and Technology Committee, MP Darren Jones observed, the “trials” have been going on for many years. The Information Commissioner also pointed out the indefinite temporality of trials: “Despite over 50 deployments, in the case of SWP, there is no clear articulation of what the police consider to be ‘effective’ or at what point the piloting phase may end.” While trials are supposed to allow for error and its subsequent correction, machine learning presupposes a different problematization of error. With machine learning,
surveillance becomes permanent experimentation through error so that error is calibrated and "optimized" rather than corrected or eliminated.

Contestations of surveillance often rely on assumptions that error can be reduced to an absolute minimum, if not eliminated. While NGOs like Liberty and Big Brother Watch base their call for banning facial recognition on the high error rate, they are also wary about the implications of minimized or eliminated error. Big Brother Watch reinforced the view that "even if live facial recognition technology improves in demographic and general accuracy it remains too great a risk to civil liberties, dangerously imbalances power between citizen and the state, and constitutes a fundamental threat to the right to privacy." In the continuous experimentation of machine learning algorithms, it is unclear when errors can become high enough to reverse or undo the development or deployment of surveillance technologies. Error is continuously calibrated and optimized, feeding back into new models.

Moreover, as Jain highlighted in the expert report in the R. (Bridges) v CCSP legal challenge, the parameters of error in facial recognition depend on "the intended use of the AFR system." While this is usually suggested by technology manufacturers, he pointed out that most AFR systems allow changes by the end user. Optimal error can be very different for border control than for iPhone applications. In the case of R. (Bridges) v SWP, the Court of Appeal reversed the initial judgment on error and discrimination, not by adopting a particular epistemology but because of the evidence that "SWP have never sought to satisfy themselves, either directly or by way of independent verification, that the software program in this case does not have an unacceptable bias on grounds of race or sex."

The form and acceptability of error in surveillance technologies remain indebted to the controversies around avoidable/unavoidable, systematic/random errors that have shaped statistical thinking on biometric technologies and error. Problematizing error in public debates would entail raising questions about how errors are optimized in algorithmic surveillance and the social imaginaries of error acceptability in machine learning algorithms.

**Conclusion**

Facial recognition for surveillance and policing has problematized algorithmic error and bias, particularly given the criticisms by NGOs, academics, and digital rights activists about the intensification of racialized and gendered discrimination. Facial recognition appeared especially egregious given its high rate of error and disparities in accuracy between lighter-skinned men and darker-skinned women. Yet, despite wide-ranging critiques and mobilization against the use of facial recognition for policing, so far there has been limited rollback of facial recognition technologies. While moves to ban facial recognition have been most successful in the wake of the Black Life Matters protests in the US, facial recognition has continued to be rolled out around the world despite criticism and litigation.

To understand the mobilization of error in algorithmic surveillance and its limitations, we have proposed to analyze the political life of error by comparing nineteenth-century developments in biometric technologies with machine learning for facial recognition. Machine learning algorithms integrate error within their performance, while computer scientists define what is acceptable error according to specific "domains." Taking error as an object of analysis alerts us to the multiple facets of non-knowledge, to the historical problematizations of what counts as knowledge, and what counts as ignorance. Error—and related concepts—have been deeply entangled in social and political controversies over technologies of surveillance. Histories of ignorance need to attend both to the multiplicity of non-knowledge and controversies over what counts as an acceptable practice of not-knowing and what does not. The political life of error helps shed light on the different understandings of error mobilized in public arguments against surveillance technologies. The optimization of error has entailed an insidious challenge for public controversies over surveillance. Even as error is mobilized to criticize algorithmic surveillance, a machine learning epistemology of error risks undoing its critical power. Highlighting error might become a continually receding intervention, as errors can be recalibrated, but not eliminated, and bias can be reduced, but not avoided. Problematizing the errors of machine learning algorithms cannot be limited to highlighting high-error rates and discrepancies across racial and gender boundaries. It needs to be supplemented by questions about decisions on what counts as acceptable error, for which areas of social life, and based on which decisions and social imaginaries of optimization.

---

75 Big Brother Watch, "Submission," 13.
76 Jain, "First Expert Report," 7.
77 R (Bridges) v CCSWP and SSHD, "Court of Appeal Judgement."
Acknowledgements
We would like to thank the special issue editors for the invitation to contribute to a conversation about histories of ignorance. We are grateful to the anonymous reviewers for their comments and generous engagement with this text. Claudia Aradau’s work was supported by funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (SECURITY FLOWS, grant agreement No 819213).

Competing Interests
The authors have no competing interests to declare.

Bibliography
Aradau, Claudia. “Assembling (Non)Knowledge: Security, Law, and Surveillance in a Digital World.” International Political Sociology 11, no. 4 (2017): 327–42.
Bates, David William. Enlightenment Aberrations: Error and Revolution in France. Ithaca: Cornell University Press, 2002. DOI: https://doi.org/10.7591/9781501726811
Beck, Ulrich, and Peter Wehling. “The Politics of Non-Knowing: An Emerging Area of Social and Political Conflict in Reflexive Modernity.” In The Politics of Knowledge, edited by Fernando Domínguez Rubio and Patrick Baert, 33–57. London: Routledge, 2012.
Bertillon, Alphonse. Identification Anthropométrique: Instructions Signalétiques. Ministère de l’Intérieur, Paris: Melun, 1885.
Big Brother Watch. “Face Off: The Lawless Growth of Facial Recognition in UK Policing.” May 2018. https://bigbrotherwatch.org.uk/wp-content/uploads/2018/05/Face-Off-final-digital-1.pdf (accessed August 4, 2021).
Big Brother Watch. “Submission to the Science and Technology Committee on the Inquiry into the Work of the Biometrics Commissioner and the Forensic Science Regulator.” UK Parliament, May 2019. http://data.parliament.uk/writenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/the-work-of-the-biometrics-commissioner-and-the-forensic-science-regulator/written/102823.pdf (accessed April 16, 2020).
Big Brother Watch. “Joint Statement on Police and Private Company Use of Facial Recognition Surveillance in the UK.” September 2019. https://bigbrotherwatch.org.uk/wp-content/uploads/2019/09/Statement-to-stop-live-facial-recognition-surveillance-BBW-September-2019-1.pdf (accessed April 16, 2020).
Bledsoe, Woodrow Wilson. “A Man-Machine Facial Recognition System—Some Preliminary Results.” Panoramic Research, Inc, Palo Alto, California. Technical Report 19 (1965).
Boumans, Marcel, Giora Hon, and Arthur C. Petersen, eds. Error and Uncertainty in Scientific Practice. London: Routledge, 2015. DOI: https://doi.org/10.4324/9781315654577
Brekenridge, Keith. Biometric State: The Global Politics of Identification and Surveillance in South Africa, 1850 to the Present. Cambridge, UK: Cambridge University Press, 2014. DOI: https://doi.org/10.1017/CBO97811393993946
Cole, Simon A. Suspect Identities: A History of Fingerprinting and Criminal Identification. Cambridge, MA: Harvard University Press, 2002.
Cowan, Jill. “San Francisco Banned Facial Recognition: Will California Follow?” New York Times, July 1, 2019.
Daston, Lorraine. “Scientific Error and the Ethos of Belief.” Social Research 72, no. 1 (2005): 1–28.
Daston, Lorraine, and Peter Galison. Objectivity. Cambridge, MA: MIT Press, 2007.
Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. “ImageNet: A Large-Scale Hierarchical Image Database.” Paper presented at the 2009 IEEE Conference on Computer Vision and Pattern Recognition. 20–25 June 2009. DOI: https://doi.org/10.1109/CVPR.2009.5206848
Derksen, Linda. “Towards a Sociology of Measurement: The Meaning of Measurement Error in the Case of DNA Profiling.” Social Studies of Science 30, no. 6 (2000): 803–45. DOI: https://doi.org/10.1177/030631200030006001
Dick, Stephanie. “Artificial Intelligence.” Harvard Data Science Review 1, no. 1. (2019). DOI: https://doi.org/10.1162/99608f92.92fe150c
Dupré, Sven, and Geert Somsen. “The History of Knowledge and the Future of Knowledge Societies.” In “History of Science or History of Knowledge?,” edited by Christian Joas, Fabian Krämer, and Käriq Nickelsen. Special issue of Berichte zur Wissenschaftsgeschichte 42, nos. 2–3 (2019): 186–99. DOI: https://doi.org/10.1002/bewi.201900006
