
— PROTECTED & SENSITIVE WHISTLEBLOWER DISCLOSURE —

July 6, 2022

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

**Re: Protected Disclosures of Federal Trade Commission Act
Violations, Material Misrepresentations and Omissions, and Fraud by
Twitter, Inc. (NASDAQ: TWTR) and CEO Parag Agrawal, SEC TCR#**

[REDACTED]

Whistleblower Aid is a U.S. tax-exempt, 501(c)(3) organization, EIN 26-4716045.

<https://WhistleblowerAid.org> — Anonymously via Tor Browser:

<http://p6ufg73qskew53cglxt6hktyt35rbl46yultzyuytq3tvicywa3pclid.onion>

Contact via **SecureDrop** over Tor: <http://whistlebloweraid.securedrop.tor.onion> — via **Signal App**: +1 201-773-1371

To the Securities and Exchange Commission (“SEC”), Federal Trade Commission (“FTC”), and Department of Justice (“DOJ”):

I. Legal Violations and Deceit

1. We are lawyers representing **Peiter “Mudge”¹ Zatko**, who was employed as “Security Lead”, a member of the senior executive team responsible for Information Security, Privacy, Physical Security, Information Technology, and “Twitter Service” (the corporate division responsible for global content moderation enforcement) at **Twitter, Inc.** from November 16, 2020,² until the morning of January 19, 2022, when CEO Parag Agrawal terminated Mudge. During Mudge’s employment, he uncovered extreme, egregious deficiencies by Twitter in every area of his mandate including (as described in detail below) user privacy, digital and physical security, and platform integrity / content moderation. As described below, these deficiencies are the basis for our client’s reasonable belief in extensive legal violations by Twitter, Inc.
2. In this submission, Mudge makes protected, lawful disclosures of original evidence³ showing that the corporation, CEO Parag Agrawal, particular senior executives, and particular members of its Board of Directors, **since 2011 and on an ongoing basis**, have engaged in:
 - a. Extensive, repeated, uninterrupted violations of the Federal Trade Commission Act by making **false and misleading statements** to users and

¹ Decades ago, anonymous articles signed with the pseudonym “Mudge” began to appear, exposing security vulnerabilities. These articles helped people to understand and correct the problems, but they were contentious. Companies called out by Mudge did everything except fix the problems: They denied everything, threatened litigation, even sought to get Mudge fired from jobs elsewhere. Over the course of years of pioneering research on buffer overflow, code injection and other fundamental vulnerabilities, Mudge became recognized as a luminary in the information security field. Mudge’s real identity remained obscured until the late 1990s when Mudge was invited to meet the President, and White House staff accidentally disclosed that Mudge was in fact Peiter Zatko.

² Before joining Twitter, Mudge held senior positions at Google and Stripe, and within the Department of Defense, where he was authorized to access Top Secret / Special Compartmented Information for work on programs at the bleeding edge of both offensive and defensive cyber operations. Mudge has been formally recognized by the CIA, White House, U.S. Army, and The Office of the Secretary of Defense bestowed upon Mudge the Exceptional Public Service Award, the highest medal honor available to civilian, non-career officials. See https://en.wikipedia.org/wiki/Peiter_Zatko. Former Twitter CEO Jack Dorsey cited this track record of speaking truth to power as a primary reason for recruiting Mudge.

³ As described in more detail below, we have worked carefully to ensure that no Attorney-Client Privileged materials or communications are included in this disclosure or exhibits.

-
- the FTC about, *inter alia*, the Twitter platform's **security, privacy, and integrity**;
- b. Violations of **SEC rules governing public companies** including, *inter alia*, auditing requirements;
 - c. **Fraudulent and material misrepresentations** in communications with the Board of Directors and investors, constituting securities law violations;
 - d. Negligence and even complicity with respect to efforts by **foreign governments to infiltrate, control, exploit, surveil and/or censor** the company's platform, staff, and operations.
3. **Particular episodes of fraud** and deliberate efforts to mislead include, among other examples:
- a. In or around February 2021, after Mudge had prepared comprehensive written materials to educate the Board on his findings about the company's extensive security, privacy and integrity problems, Mudge was instructed **not to send them** to the Board of Directors.
 - b. On multiple occasions during 2021, described in greater detail below, Mudge witnessed senior executives engaging in deceitful and/or misleading communications affecting Board members, users and shareholders. In contrast, Mudge spent 2021 designing and implementing a long-term strategy to reform and address Twitter's privacy, security and integrity vulnerabilities. On December 14, 2021, against Mudge's recommendation, CEO Agrawal explicitly **instructed Mudge to provide documents which both of them knew to be false and misleading**, regarding vital information security matters, to the Risk Committee of Twitter's Board of Directors.⁴
 - c. In January 2022, Mudge began working to document evidence of fraud. Twitter's Chief Compliance Officer opened a fraud investigation based on

⁴ Before Agrawal was appointed CEO on November 29, 2021, he had served over four years as Twitter's Chief Technology Officer. Agrawal's hiring as CEO had been contentious, with some Board Directors opposed. Our client reasonably believes that Agrawal became defensive about many of the problems that our client identified, because Agrawal had caused them, or allowed them to fester, in his role as CTO.

Mudge's allegations. On January 18, **CEO Agrawal** lied about Mudge's efforts to rectify the previous month's fraud.⁵

d. Agrawal terminated Mudge the next day, January 19.

4. **Astonishingly**, hours after Twitter terminated Mudge's employment, including immediately denying him access to corporate systems, Twitter's Chief Compliance Officer began emailing Mudge at his personal gmail account, seeking to obtain his latest disclosures of fraud. The Compliance Officer's reference to "your conversation this morning" was the video call in which Mudge had been terminated, and the "matters already under investigation" was Agrawal's instructions to knowingly present inaccurate materials to the Board:⁶

On Wed, Jan 19, 2022 at 11:59 PM [REDACTED] <[REDACTED]@twitter.com> wrote:
Mudge,

[REDACTED] advised me that during your conversation this morning you mentioned the concerns you raised about information shared with the Risk Committee in December. If you were referring to matters that are not already under investigation, please let me know so we can schedule time to talk right away.

We appreciate that you raised your concerns and want to be sure they are fully and appropriately addressed. While my investigation is not complete, we intend to bring all of the concerns you raised with me and Omid, as well as the document you have been preparing in response to my email of January 11, to the Audit Committee and the full Risk Committee in the coming days. If there is anything else you want to include or recommend we do with respect to this issue so it concludes to your satisfaction, please let me know.

I'm available to hear any further concerns you may have as well as any additional thoughts you may have to resolve this matter.

Regards,
[REDACTED]

5. **Apparently, Twitter's own compliance officers understood the gravity** of a situation in which the CEO had deliberately misled the Board. (Twitter's compliance team could face personal liability for letting fraud allegations go unaddressed.) Between Mudge's termination on January 19 and January 27, Twitter's Chief Compliance Officer emailed Mudge at least five times to obtain his corrected disclosures concerning information security.
6. Mudge ultimately **worked at least 150 hours—after he was terminated, without pay, and without access to his Twitter accounts or laptop**—to do his best to document the underlying facts about information security, and the fraud he had identified. Details of these events, including Mudge's emails with Twitter's Chief

⁵ When Mudge tried to correct the record, Board Member [REDACTED] interrupted, and refused to let Mudge speak or provide facts.

⁶ See Exhibit 16; These post-termination communications, made while the parties were adverse and without any expectation of confidentiality, are not subject to Attorney-Client Privilege, as explained below.

Compliance Officer and his final report to the Board to articulate specific fraud he was identifying, are all included with this disclosure.⁷

7. **No privileged materials included:** Mudge has carefully limited disclosure of internal corporate documents to those relevant and “reasonably necessary” to demonstrate Twitter’s legal violations.⁸ In order to identify any materials subject to a claim of attorney-client privilege,⁹ with assistance of independent filter counsel, we conducted a review of every Exhibit to this disclosure. We determined that none of the exhibits included in this disclosure are protected by attorney-client privilege:

- a. Some Exhibits include the words **“Privileged and Confidential”** or a similar designation. These labels, which Twitter staff often applied indiscriminately and without legal guidance, do not determine whether a document is in fact subject to a valid claim of privilege. Therefore, after a careful review of the documents and the applicable law, we included some documents that contain such a label but nevertheless do not contain privileged communications.
- b. Similarly, the **mere presence of a lawyer** in a communication does not mean the communication is covered by the attorney-client privilege. A lawyer may be operating in a non-legal capacity, or may have a dual role that encompasses legal as well as business or operational functions. Even when operating in a legal capacity, a lawyer’s communication may not be related to the request for, or provision of, legal advice. And emails between lawyers and their clients are not necessarily privileged if they are not made “in confidence.” After a review of the applicable law and the documents

⁷ See Exhibits 1 and 16

⁸ Cf. *Cafasso v. General Dynamics C4 Systems, Inc.*, 637 F.3d 1047, 1062 (9th Cir. 2011) (dicta suggesting that relators under the False Claims Act should limit disclosure of internal corporate documents to those documents “reasonably necessary” to pursue their whistleblower claim).

⁹ The attorney-client privilege “protects a confidential communication between attorney and client if that communication was made for the purpose of obtaining or providing legal advice to the client.” *In re Kellogg Brown & Root, Inc.*, 756 F.3d 754, 757 (D.C. Cir. 2014). See also *United States v. Mejia*, 655 F.3d 126, 132 (2d Cir.2011) (“The attorney-client privilege protects communications (1) between a client and his or her attorney (2) that are intended to be, and in fact were, kept confidential (3) for the purpose of obtaining or providing legal advice.”). See also *Restatement Third, The Law Governing Lawyers*, § 68, 2000.

-
- themselves, we included some email communications in which a lawyer is present on the chain.
- c. Further, “[t]he protection of the privilege extends only to communications, and **not to facts**. A fact is one thing and a communication concerning that fact is an entirely different thing.”¹⁰ For example, facts about a privileged communication, including its existence, are not themselves privileged.
 - d. No attorney-client privilege attaches to the set of **post-termination communications** between Mudge and Twitter counsel. The attorney-client privilege can, depending on the circumstances, cover some communications between in-house counsel and a former employee. After a review of the documents and the case law, we determined that the privilege does not apply to the post-termination communications included here. After Twitter abruptly terminated Mudge, their interests were not aligned, but rather adverse (and in fact, during the termination call, Mudge explicitly raised the possibility that the circumstances of his termination could create a legal risk for Twitter).¹¹ Under these circumstances, Twitter could not reasonably expect that its interaction with Mudge was privileged.¹²
 - e. Finally, we made redactions on a number of Exhibits to obscure some portions over which Twitter might claim privilege. (The fact that we redacted a portion of an Exhibit is not an admission that the redacted portion is in fact privileged.) Many times, we redacted material out of an abundance of caution even when we determined that the privilege would not properly apply.

8. The Work Product Doctrine does not affect this disclosure. The Work Product Doctrine applies in the context of “rule[s] dealing with discovery” requests in civil litigation and other adversarial proceedings. Hickman v. Taylor, 329 U.S. 495, 509 (1947). The rule states that “[o]rdinarily, a party may not discover documents and

¹⁰ Upjohn Co. v. United States, 449 U.S. 383 at 395-96 (1981) (internal citations omitted).

¹¹ In fact, Mudge had been so concerned with Agrawal’s conduct since December 2020 that he had already retained lawyers to advise him on how to follow whistleblower laws and, if necessary, pursue claims of unlawful retaliation. (Please note that undersigned counsel did not become involved until later.)

¹² The majority of these post-termination communications were with Twitter’s Chief Compliance Officer. Although this individual has a law license, for the purposes of these communications, this person was acting in an operational, non-legal capacity. The Chief Compliance Officer’s non-legal role in these communications is another reason why the communications are not protected by the attorney-client privilege.

tangible things that are prepared in anticipation of litigation or for trial by or for another party or its representative (including the other party's attorney, consultant, ... or agent)." Fed R. Civ. P. 26(b)(3)(A). So, "[t]he purpose of the work-product rule is not to protect the evidence from disclosure to the outside world but rather to protect it only from the knowledge of opposing counsel and his client, thereby preventing its use against the lawyer gathering the materials." The Work-Product Rule—Matters Protected by the Work-Product Rule, 8 Fed. Prac. & Proc. Civ. § 2024 (3d ed.).¹³

9. As an example, if our client were to sue Twitter, the Work Product Doctrine might allow Twitter to refuse to give certain documents (that it could show were actually prepared in anticipation of the hypothetical lawsuit) to our client during the discovery process. But there is no such lawsuit, and nobody, including our client, is making any production demands on Twitter. We are not aware of any authority suggesting that the Doctrine affects voluntary, protected disclosures under the Dodd-Frank Act. (Even if the Work Product Doctrine applied here, we have not identified any documents that Twitter, under the applicable case law, prepared in "anticipation of litigation.")
10. **Ethical Disclosure Dilemma:** Mudge is proceeding with these disclosures quite reluctantly. Mudge comes out of, even helped to create, the modern information security community of responsible security disclosures. While criminal hackers break and steal, independent security researchers (also known as "ethical hackers") use their skills to inform people about specific vulnerabilities, strengthen security and advance human rights and democracy. When ethical researchers find a vulnerability that bad actors can exploit, they first make a quiet "responsible disclosure" so that the affected company or government can fix it. But sometimes, the vulnerable institution doesn't want to hear the truth, or fix the problem. In those cases, ethical researchers are forced to weigh the risks of wider disclosure: Exposing vulnerabilities tips off bad actors, but it also allows users of a service to

¹³ Thus, the work-product doctrine, unlike the attorney-client privilege, "does not exist to protect a confidential relationship, but rather to promote the adversary system by safeguarding the fruits of an attorney's trial preparations from the discovery attempts of the opponent." *United States v. Am. Tel. and Tel. Co.*, 642 F.2d 1285, 1299 (D.C. Cir. 1980).

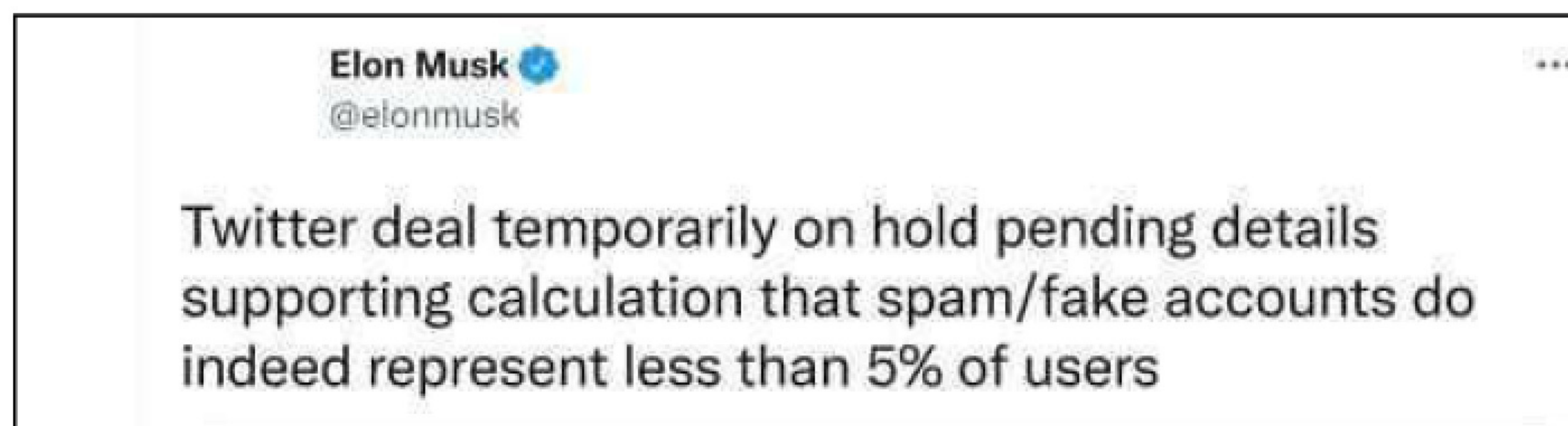
make more informed decisions, and can push the service to improve.¹⁴ Mudge made a personal commitment to Dorsey, the Twitter Board, the greater public, and to himself, that he would do his best to help fix Twitter. Mudge spent about 14 months pushing improvements from the inside, and was terminated for his efforts. With a heavy heart, Mudge has concluded that these lawful disclosures are his ethical obligation.

[Disclosure continues next page]

¹⁴ "In computer security, **coordinated vulnerability disclosure**, or "CVD" (formerly known as responsible disclosure) is a vulnerability disclosure model in which a vulnerability or an issue is disclosed to the public only after the responsible parties have been allowed sufficient time to patch or remedy the vulnerability or issue. This coordination distinguishes the CVD model from the "full disclosure" model." "Coordinated vulnerability disclosure - Wikipedia." https://en.wikipedia.org/wiki/Coordinated_vulnerability_disclosure.

II. Lying about Bots to Elon Musk

11. A recent example of misrepresentations by Twitter concerns Elon Musk's high-profile takeover attempt since April 2022.¹⁵ On May 13, Mr. Musk expressed doubts about the accuracy of Twitter's claim in legal filings that <5% of accounts are "bots," or automated spam accounts that spread propaganda and hurt the experience of real users:¹⁶



12. In response, on May 16, 2022, **CEO Agrawal tweeted false and misleading statements about Twitter's handling of bots** on the platform, starting with this:¹⁷



13. Agrawal's tweet was a lie. In fact, Agrawal knows very well that Twitter executives are **not incentivized to accurately "detect"** or report total spam bots on the platform. Here's why:

¹⁵ Please note that Mudge began preparing these disclosures in early March 2022, well before Mr. Musk expressed any interest in acquiring Twitter, and has not communicated these disclosures to anyone with a financial interest in Twitter. As a senior executive, Mudge was awarded Twitter stock, for which he previously created (and has followed) an Automatic Securities Disposition Plan pursuant to SEC rules codified at 17 C.F.R. § 240.10b5-1(c).

¹⁶ Elon Musk's Personal Twitter Page: https://twitter.com/elonmusk/with_replies?lang=en

¹⁷ Parag Agrawal's Personal Twitter Page: <https://twitter.com/paraga>

-
14. Until 2019, Twitter reported total monthly users, but stopped because the number was subject to negative swings for a variety of reasons, including situations such as the removal of large numbers of inappropriate accounts and botnets.¹⁸ Instead, Twitter announced a new, proprietary, opaque metric they called “**mDAU**” or “**Monetizable Daily Active Users**,” defined as valid user accounts that *might* click through ads and actually buy a product.¹⁹ From Twitter’s perspective, “mDAU” was an improvement because it could internally define the mDAU formula, and thereby report numbers that would reassure shareholders and advertisers. Executives’ bonuses (which can exceed \$10 million) are tied to growing mDAU.
15. Executives are incentivized to avoid counting spam bots as **mDAU**, because mDAU is reported to advertisers, and advertisers use it to calculate the effectiveness of ads. If mDAU includes spam bots that do not click through ads to buy products, then advertisers conclude the ads are less effective, and might shift their ad spending away from Twitter to other platforms with higher perceived effectiveness.
16. However there are many **millions of active accounts** that are not considered “mDAU,” either because they are spam bots, or because Twitter does not believe it can monetize them. These millions of non-mDAU accounts are part of the median user’s experience on the platform. And for this vast set of non-mDAU active accounts, Musk is correct: Twitter executives have little or no personal incentive to accurately “detect” or measure the prevalence of spam bots.
17. In fact, Mudge learned deliberate ignorance was the norm amongst the executive leadership team. In early 2021, as a new executive, Mudge asked the Head of Site Integrity (responsible for addressing platform manipulation including spam and botnets), what the underlying spam bot numbers were. Their response was “**we**

¹⁸ “Twitter...said it would stop reporting monthly active users (MAUs) after Q1 2019 as it would switch to a new metric called monetizable daily active users (mDAUs)...”

https://www.business-standard.com/article/news-ians/twitter-says-will-stop-reporting-monthly-active-users-119020701161_1.html But even after the switch, Twitter overcounted mDAU users, see <https://techcrunch.com/2022/04/28/twitter-says-it-overcounted-its-users-over-the-past-3-years-by-as-much-as-1-9m/>.

¹⁹ “We define mDAU as people, organizations, or other accounts who logged in or were otherwise authenticated and accessed Twitter on any given day through twitter.com or Twitter applications that are able to show ads.” See

<https://www.sec.gov/Archives/edgar/data/0001418091/000141809120000037/twtr-20191231.htm>. Twitter has stated that mDAU is “not comparable to current disclosures from other companies.” *Digging Into Twitter’s First Daily User Disclosure*, 7 Feb. 2019, <https://www.fool.com/investing/2019/02/07/digging-into-twitters-first-daily-user-disclosure.aspx>.

don't really know." The company could not even provide an accurate upper bound on the total number of spam bots on the platform. The site integrity team gave three reasons for this failure: (1) they did not know how to measure; (2) they were buried under constant firefighting and could not keep up with reacting to bots and other platform abuse; and, most troubling, (3) **senior management had no appetite to properly measure the prevalence of bot accounts**—because as Mudge later learned from a different sensitive source, they were concerned that if accurate measurements ever became public, it would harm the image and valuation of the company.

18. Even the Board of Directors understood the counterproductive incentives in place: In or about the Q3 2021 Board Risk Committee meeting, a Director asked why more progress has not been made around bots and related harmful content on the platform. Our client remembers an executive of the company **admitting to Board members that the company had "intentionally and knowingly deprioritized" platform health** to focus on growing mDAU. Afterwards, a different Twitter leader who had witnessed the exchange commented to Mudge, in reference to this admission, "it is very strange what this company does not share with board members, and then some of the statements that they do."


19. **Repeated Efforts to Disable ROPO:** "ROPO," which stands for "Read-Only Phone Only," is probably Twitter's most volumetrically-effective mechanism for identifying and blocking spam bots. If a script identifies an account as possibly spam and triggers ROPO, the account is placed into a "Read Only" mode and is unable to post content to the platform. Twitter sends a text message to the associated phone number, with a one-time code that the recipient needs to manually enter to regain account access. Shortly into Mudge's time at Twitter, **a senior executive (with primary responsibility for growing mDAU) proposed disabling ROPO worldwide**, based on an anecdote of a small number of unsolicited DMs (text messages) he had personally received in which users claimed they were incorrectly denied access by ROPO.²⁰ The Lead of Site Integrity told Mudge that executives responsible for growing mDAU had proposed disabling ROPO several times before. The Site Integrity Lead pleaded with Mudge, as a

²⁰ Executives at the company receive a near continuous stream of messages directed to them, complaining about the service and other requests like demanding malicious accounts be reinstated. Some percentage of the time the complaints were valid, but more often not.

senior executive, to prevent the other executives from disabling ROPO. Research later performed at Mudge's direction showed ROPO was effectively blocking more than 10-12 million bots each month with a surprisingly low rate (<1%) of false positives.²¹

20. Therefore Musk's suspicions are on target: senior executives earn bonuses not for cutting spam, but for growing mDAU. In fact, Twitter **created the mDAU metric precisely to avoid** having to honestly answer the very questions Mr. Musk raised.

21. The rest of Agrawal's May 16 tweets aren't out-and-out lies but they rely on wordplay to distract and mislead Mr. Musk, and everyone else. Musk appears to be asking a valid and intuitive question, *what percent of accounts encountered by the median user are actually bots?*

Elon Musk  @elonmusk


Twitter deal temporarily on hold pending details supporting calculation that spam/fake accounts do indeed represent less than 5% of users

Elon Musk  @elonmusk · May 15

Replying to @PPathole

Exactly. I have yet to see *any* analysis that has fake/spam/duplicates at <5%.

379 493 6,335

Elon Musk  @elonmusk · May 21

Replying to @teslaownersSV and @WholeMarsBlog

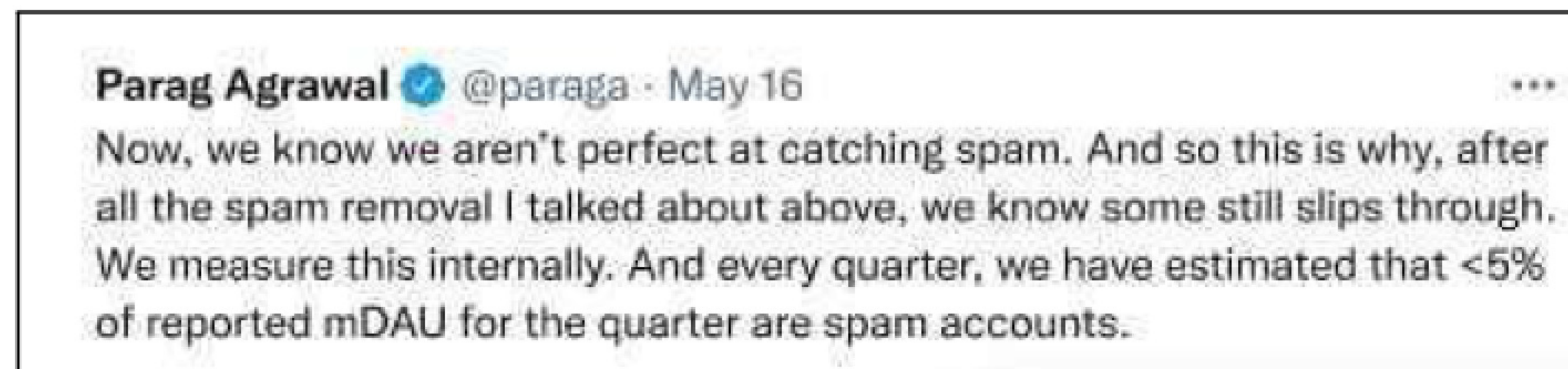
No, they still refuse to explain how they calculate that 5% of daily users are fake/spam! Very suspicious.

940 1,117 8,393

²¹ Mudge does not recall whether the 10-12 million number was per month or per week as Twitter revoked his access to the notes and data on this topic when he was terminated. Here we provide the benefit of the doubt and present the smaller of the two numbers.



22. While pretending he is answering Musk's question, in fact Agrawal is answering a very different one, namely, *Are there fewer than 5% bots in the set of mDAU accounts, as defined in secret by Twitter?* Agrawal's reasoning might appear a bit circular since, by definition, mDAU is more or less Twitter's best approximation of the set of accounts that aren't bots. And Agrawal is not exactly trying to help readers understand the bait-and-switch nature of his answer:





23. Unless you're a Twitter engineer responsible for calculating mDAU, you probably wouldn't know what Agrawal is talking about. He is *not* saying that fewer than 5% of all accounts on the platform are spam. He's saying, more or less, that Twitter starts with all the accounts on the platform, tries to automatically put all the human accounts that could be convinced by advertisers to buy products (but no spam accounts) into mDAU, and then uses humans to estimate the error rate of spam accounts that nevertheless slip through into mDAU. And naturally, Twitter "can't share" its special sauce for determining mDAU.

24. In mathematical terms, Mr. Musk is asking whether the following proposition holds:

$$\frac{\text{spam bot accounts}}{\text{Total active}^{22} \text{ accounts}} < 5\%$$

25. To which Agrawal responds by affirming a rather different proposition:

$$\frac{\begin{aligned} &\{\text{human estimate of spam bots that slip through into} \\ &(\text{mDAU, i.e. Twitter's secret automated estimate of Total active accounts} \\ &\text{minus spam bot and other worthless accounts})\} \end{aligned}}{\begin{aligned} &(\text{mDAU, i.e. Twitter's secret automated estimate of Total active accounts} \\ &\text{minus spam bot and other worthless accounts}) \end{aligned}} < 5\%$$

²² The qualifier "active" is meant to exclude accounts of users who have died or no longer use the service, etc..

26. A more meaningful and honest answer to Mr. Musk's question would be trivial for Twitter to calculate, given that Twitter is already doing a decent job excluding spam bots and other worthless accounts from its calculation of mDAU. But this number is likely to be meaningfully higher than 5%:²³

$$\frac{\text{Total active accounts minus mDAU accounts}}{\text{Total active accounts}} \approx \% \text{ Spam and other worthless accounts}$$

27. Agrawal goes on to provide raw numbers of takedowns - again without context:



28. Is half a million a day a lot or a little, for a platform as vast as Twitter? No one knows, because there is no denominator provided for context. Is Twitter *missing* a hundred thousand new spam bots every day? No one knows. Mudge attended every Board of Directors and relevant Board subcommittee meeting in 2021, where he saw this strategy regularly deployed: executives reported items, such as bot takedowns and other metrics, as **raw numbers, without context**—never in a more useful format (e.g. percentages with well-defined numerators and denominators) that would permit Board members to understand the overall prevalence of fake accounts.

29. More broadly, Agrawal's tweets and Twitter's previous blog posts misleadingly imply that Twitter employs proactive, sophisticated systems to measure and block spam bots. Mudge discovered the reality: mostly outdated, unmonitored, simple

²³ To be fair, this formula doesn't precisely measure spam bots. Rather it lumps bots in with human accounts that Twitter, for whatever reason, believes it can't monetize, perhaps because, e.g., they are not selling ads for their region, or the company has no capacity in that user's language. Relatedly, our client notes that lack of language capacity is a significant shortcoming at Twitter. Because Twitter lacks language capacity the platform is disturbingly deficient in regards to integrity efforts in dozens of countries worldwide, permitting well-established harms like disinformation and adverse electoral effects to fester largely unaddressed.

scripts plus overworked, inefficient, understaffed, and reactive human teams. The scripts were largely un-owned by any person or team, and their results were not tracked. Furthermore no effort was made to compare costs to benefits of the scripts, nor approaches, nor their veracity.²⁴

30. #Protect Initiative: Mudge was so concerned about this situation, and Twitter’s overall cybersecurity state, that during the 2021 calendar year, he developed and presented to the Board of Directors a sweeping, 3-year Board-supervised objective called “#Protect Initiative.”²⁵ Elements of the initiative would have assigned responsibility for properly measuring spam bot prevalence. The entire senior leadership team and Board of Directors received and approved Mudge’s #Protect Initiative plan. If Twitter was already accurately measuring and estimating spam bot prevalence on the platform, this issue would not have reached the Board and been a specific part of Mudge’s 2022 plans. This excerpt from the 3-year plan shows that Mudge intended to lead Twitter Services (a corporate division abbreviated as “TwS”) to inventory, obtain measurements, and improve anti-spam efforts:²⁶

4. TwS - inventory of bots, measurements, continuity of action across bot bounces, improve automation and accuracy by >25%

²⁴ In a September 2021 Twitter blog post, Twitter stated “it’s not the number of bots, (around 5%, a number Twitter reports quarterly) but the impact they have on the conversation.” Five percent of what? This was apparently an attempt to distract and mislead users about the bot problem. The post goes on to state: “First Truth: Don’t assume an account with a peculiar name must be a bot.” <https://blog.twitter.com/common-thread/en/topics/stories/2021/four-truths-about-bots> . Twitter is arguing that account names consisting of random, auto-generated sets of letters and numbers aren’t always bots. But this is a straw man. The blog post does not cite any data on what percentage of random-character handles were in fact bots, because the Site Integrity team did not have a dedicated data scientist, and this data, though available within the company, was poorly maintained and largely un-measured.

²⁵ Mudge lost access to the detailed #Protect Initiative documentation when he was terminated, but investigators should be able to acquire it easily. *But* see exhibit 8, Mudge’s #Protect board presentation.

²⁶ Confusingly, because one of Twitter’s main automation tools was called “Botmaker”, Twitter staff also used the term “bots” to describe the company’s automated scripts to identify spam bots. Whether a document’s reference to “bots” means a spam account or a Twitter script depends on the context.

31. All this is conveyed in a damning independent report²⁷ on platform integrity, produced in or about May or June of 2021, from [REDACTED]

²⁸

Here are a few of its findings (“SI”= Site Integrity, whose mandate included hunting spam bots):

Tools available to Site Integrity to work on these issues are often outdated, “hacked together,” or difficult to use, limiting Twitter’s ability to effectively enforce policies at scale. A lack of automation and sophisticated tooling means that Twitter relies on human capabilities, which are not adequately staffed or resourced, to address the misinformation and disinformation problem.

3.1.1.2 -- There are components of Twitter that are part of the disinformation and misinformation detection or response that are outside of Site Integrity / Security, and Site Integrity / Security have no access or authority to use these tools absent the good will of other teams.

3.1.1.3 -- Twitter does not have aligned incentives across the organization, and, as a result, priorities with regards to Product Safety.

3.1.1.4 -- SI relies on functions that have no accountability to SI in order to piece together solutions.

3.2.3 -- SI does not have dedicated engineering support for their tools, so even minor upgrades or changes to existing tools can take months or years to complete.

3.2.4 -- SI lacks sufficient dedicated data science support and staff with technical skills.

²⁷ Independent “filter counsel” advised us that the [REDACTED] report is not subject to Attorney-Client Privilege: (1) The report does not contain or discuss legal advice or exposure, nor does it discuss legal or regulatory options or contain legal citations. (2) It was written by non-lawyers at [REDACTED] for the non-lawyer executive tasked with security, privacy, and content moderation at scale, Mudge. Cf. Guo Wengui v. Clark Hill, PLC, 338 F.R.D. 7 (D.D.C. 2021) (a cybersecurity report created by a non-legal consulting firm is not privileged, even when consultant was hired by outside counsel, because the claimant’s goal was “gleaning [consultant’s] expertise in cybersecurity, not in obtaining legal advice from [its] lawyer,” *id.* at 13, internal citations omitted).

²⁸ See Exhibit 2

3.3.3 -- There are existing internal tools in other parts of Twitter that would be useful for the misinformation and disinformation use case, but SI analysts do not have access to them. Analysts also lack access to externally available tools or datastreams that would allow them to do more proactive cross-platform analysis.

3.4.1 -- SI does not have a knowledge management system to track and store findings and data. As a result, SI does not have the ability to monitor threat actors or identify changes in their tactics, techniques, and procedures (TTPs) over time, or to measure the impact of SI's work.

3.7.3 -- Policies to address misinformation/disinformation often do not address repeat offenders and are applied on a case-by-case basis, leading to a lack of scalability.

3.8.2 -- The process for labelling disinformation and misinformation content is largely manual, requires the use of multiple tools, and usually needs to be done on a case-by-case basis.

32. Unfortunately, as detailed in the rest of this disclosure, Agrawal's misrepresentations about spam bots are just the tip of the iceberg.

[Disclosure continues next page]

III. 2011 FTC Consent Order and 2020's "Largest Social Media Hack in History": Dorsey Recruits Mudge

33. Since Twitter's 2006 launch, the platform has earned a **reputation for problems**²⁹ with security, privacy and integrity (a broad term that includes disinformation, spam bots, election interference and other content-related abuses).

34. **2011 FTC Complaint:** In 2011, the FTC had filed a complaint against Twitter for its failure to properly protect nonpublic consumer information, which included users' email addresses, Internet Protocol ("IP") addresses, telephone numbers, and nonpublic information exchanged on the platform.³⁰ The complaint alleged that, from 2006 to 2009, far too many Twitter employees exercised administrative ("God mode") control within Twitter's internal systems and user data, thereby allowing any attacker with access to an employee account to easily compromise Twitter systems. And Twitter's systems were, and are, full of highly sensitive personal user data that enable a hostile government to find precise geo-location(s) for a specific user or group, and target them for arrest or violence.

35. **Consent Order:** As a result of the complaint, the FTC and Twitter entered into a consent decree in March 2011, which has the force of law for future violations.³¹ The FTC ordered Twitter to: "establish and implement, and thereafter maintain a comprehensive information security program that is reasonably designed to protect the security, privacy, confidentiality, and integrity of nonpublic consumer information." Components of this comprehensive information security program included identifying security risks and preventing, detecting, and effectively responding to cyberattacks.³² The order imposed various reporting requirements

²⁹ Eric Geller, *Twitter's security holes are now the nation's problem*,

<https://www.politico.com/news/2020/07/16/twitter-security-hack-congress-366771>

³⁰ U.S. Fed. Trade Comm'n, *In the Matter of Twitter, Inc.*, (No. C-4316), Compl. (Mar. 2, 2011), available at: <https://www.ftc.gov/sites/default/files/documents/cases/2011/03/110311twittercmpt.pdf>.

³¹ U.S. Fed. Trade Comm'n, *In the Matter of Twitter, Inc.*, (No. C-4316), Decision & Order (Mar. 2, 2011), available at: <https://www.ftc.gov/sites/default/files/documents/cases/2011/03/110311twitterdo.pdf>.

³² Per the FTC order, Twitter was required to include various components of this program, including, among other things:

- 1) designating employee(s) to be accountable for the information security program;
- 2) identifying reasonably foreseeable security risks that could expose or compromise nonpublic consumer information, taking into account various considerations, such as employee training and management, information systems, and prevention, detection and responses to various system failures, such as attacks and account takeovers;

upon Twitter to ensure it was keeping the FTC informed of its progress on its security system for ten years following issuance of the order, which had a final termination date of March 2, 2031.

36. **Hacked by Teenagers:** In July 2020—following nine years of supposed fixes, investments, compliance policies, and reports to the FTC by Twitter—the company was hacked by a 17-year old, then-recent high school graduate from Florida and his friends. The hackers managed to take over the accounts of former **President Barack Obama, then-Presidential candidate Joseph Biden, and high-profile business leaders including, but not limited to, Jeff Bezos, Bill Gates, and Elon Musk.** As part of the account takeovers, the hackers urged their tens of millions of followers to send Bitcoin cryptocurrency to an account they created.³³

37. The 2020 hack was then the **largest hack of a social media platform in history,**³⁴ and triggered a global security incident.³⁵ Moreover, the hack did not involve malware, zero-day exploits, supercomputers brute-forcing their way past encryption, or any other sophisticated approach. In fact, it was pretty simple.³⁶ Pretending to be Twitter IT support, the teenage hackers simply called some Twitter employees and asked them for their passwords. A few employees were duped and complied and—given systemic flaws in Twitter’s access controls—those credentials were enough to achieve “God Mode,” where the teenagers could imposter-tweet from any account they wanted. Twitter’s solution was to impose a **system-wide shutdown of system access to all of its employees, lasting days.** For about a

-
- 3) designing and implementing reasonable safeguards to control the risks identified through the risk assessment, and carrying out regular testing and monitoring of these safeguards; and,
 - 4) evaluating and adjusting its information security program and circumstances that may have a material impact on the effectiveness of its information security program.

Id.

³³ Kif Leswing, *Hackers targeted Twitter employees to hijack accounts of Elon Musk, Joe Biden and others in digital currency scam*, CNBC, July 15, 2020,

<https://www.cnbc.com/2020/07/31/twitter-bitcoin-scam-masterminded-by-17-year-old.html>.

³⁴ "Major US Twitter accounts hacked in Bitcoin scam - BBC News." July 16, 2020,

<https://www.bbc.com/news/technology-53425822>.

³⁵ Brain Fung, *Twitter's massive hack could be even worse than it seems*, CNN, July 17, 2020,

<https://www.cnn.com/2020/07/16/tech/twitter-hack-security-analysis/index.html>.

³⁶ See New York State Dep’t of Financial Servs., *Twitter Investigation Report* (Oct. 14, 2020), https://www.dfs.ny.gov/Twitter_Report (“In the hands of a dangerous adversary, the same access obtained by the Hackers—the ability to take control of any Twitter users’ account—could cause even greater harm.”).

month, hiring was paused and the company essentially shut down many basic operations to diagnose the symptoms, not the causes, of the hack.

38. Security experts agreed this extreme response demonstrated that Twitter did not have proper systems in place to understand what had happened, let alone remediate and reconstitute to a safe state.³⁷ These failures in Twitter's security raised alarms about more serious breaches that could occur in the future, especially because 2020 was a presidential election year. Bad actors with more sophisticated tools than what was used by a recent high school graduate could easily take advantage of Twitter's poor security, creating detrimental consequences for the country. As aptly framed in the *Wired* article about this incident,³⁸

But if a teenager with access to an admin panel can
bring the company to its knees, just imagine what Vladimir Putin could do.

39. Soon, Twitter's situation with the FTC got even worse. On July 28, 2020, the FTC filed a draft complaint alleging Twitter engaged in violations of the 2011 order.³⁹ Specifically the draft FTC complaint charged that from 2013 to 2019, Twitter misused users' phone number and/or email address data for targeted advertising when users had provided this information for safety and security purposes only. This implied Twitter still lacked basic understandings about how, what, and where its data lived, and how to responsibly protect and handle it. On May 25, 2022, the FTC announced a \$150 million fine against Twitter.⁴⁰

40. At the time of the hack and the new FTC draft complaint, Twitter had neither an executive versed in information security and privacy engineering (the executive-level Security Lead role Mudge would fill in November 2020), nor even a Chief Information Security Officer.⁴¹ As a result, **Parag Agrawal, then Twitter's Chief**

³⁷ See *id.* (former Facebook CISO explained his surprise that a phishing scheme led to a total shutdown at Twitter and a former Twitter employee stated that the company did not have the right systems in place to address such an attack, which led to this extreme response).

³⁸ "Inside the Twitter Hack—and What Happened Next | WIRED." 24 Sep. 2020, <https://www.wired.com/story/inside-twitter-hack-election-plan/>.

³⁹ https://www.ftc.gov/system/files/ftc_gov/pdf/2023062TwitterComplaint.pdf

⁴⁰ "Twitter Fined in Privacy Settlement, as Musk Commits More Equity" 25 May. 2022, <https://www.nytimes.com/2022/05/25/technology/twitter-fined-ftc-doj-privacy.html>.

⁴¹ Twitter's failure to employ a Chief Information Security Officer constitutes an independent violation of the 2011 Consent Order, which required "the designation of an employee or employees to coordinate and

Technology Officer (CTO), was the ultimate decision-maker for correcting the security vulnerabilities exposed by the hack.⁴² Mr. Agrawal made statements that acknowledged the problem that the FTC had precisely identified nine years earlier: too many Twitter staff and contractor accounts had too much access to too much user data. Here are quotes from Mr. Agrawal in Wired magazine⁴³ meant to assure the public—as we explain later, these were false and misleading (and therefore constitute legal violations):

But one of the first things Twitter realized in the immediate aftermath was that too many people had too much access to too many things. “It’s more about how much trust you’re putting in each individual, and in how many people do you have broad-based trust,” Agrawal says. “The amount of access, the amount of trust granted to individuals with access to these tools, is substantially lower today.”

41. Twitter’s CEO at the time, Jack Dorsey, realized the company had serious problems. To demonstrate he was serious about fixing things, Dorsey began recruiting Mudge. Mudge had other very desirable, well-paid, high-profile job opportunities. But Mudge also understood that Twitter is a critical global public resource that can build bridges between different communities and parts of the world, and serve the public conversation. Mudge recognized that Twitter’s platform could also cause real harm, and understood that it would take a lot of work to get Twitter on track. He was up for the challenge. At the request of Dorsey, and with the promise of Dorsey’s support, Mudge accepted the offer and expected to spend the rest of his career at Twitter. Mudge never expected or wanted to become a whistleblower. He was convinced that the executives and board were ready to deal with long overdue security and privacy challenges.

be accountable for the information security program.” See

<https://www.ftc.gov/sites/default/files/documents/cases/2011/03/110311twitterdo.pdf>

⁴² See Nicholas Thompson & Brian Barrett, *How Twitter Survives Its Biggest Hack—and Plans to Stop the Next One*, WIRED (Sept. 24, 2020), <https://www.wired.com/story/inside-twitter-hack-election-plan/> (detailing Mr. Agrawal’s role in responding to the hack).

⁴³ *Id.*; see also Parag Agrawal & Damien Kieran, Twitter Blog, *Our continued work to keep Twitter secure* (Sept. 24, 2020), https://blog.twitter.com/en_us/topics/company/2020/our-continued-work-to-keep-twitter-secure (“We have strict principles around who is allowed access to which tools and at what time, and require specific justifications for customer data to be accessed.”).

42. On **November 16, 2020**, four months after the summer's publicly humiliating incident, Mudge began his new job as **Security / Integrity Lead** with Twitter. **Mudge's hire was applauded** by cybersecurity industry executives and experts, and allowed Twitter to claim credit for challenging cybersecurity problems aggressively. Kevin O'Brien, co-founder and CEO of GreatHorn, a cybersecurity firm, stated, "My take is that [Mudge] remains one of the best security minds on the planet today – Dorsey bringing him [in] speaks well for their focus on security."⁴⁴ Dan Kaufman, Mudge's former supervisor at DARPA, stated Mudge would "be at the top of [his] list" of names of people who could fix Twitter's current abysmal security state.⁴⁵ Similarly, Alex Stamos, former Facebook Chief Security Officer, stated that Mudge was a great fit for the Company because of his ability to find creative solutions to security problems.⁴⁶
43. **Portfolio:** CEO Jack Dorsey assigned Mudge a vast portfolio, responsible for some of the hardest problems, with hundreds of staff and thousands of contractors in chains that reported up to him:
- a. **Information Security** - the integrity and security of all Twitter systems and data;
 - b. **Privacy** - creating the privacy policies and processes, plus engineering and executing them across all Twitter systems and data, to avoid liability with the FTC and build systems and processes that respect people's data;
 - c. **Corporate Security** - responsible for the physical security and safety of employees, offices, and data centers;
 - d. **Information Technology** - running the internal systems for finance, HR, and internal corporate technologies and communications;

⁴⁴ David Jones, *Famed hack Mudge to lead Twitter security after summer of attacks*, Cybersecurity Dive (Nov. 17, 2020), <https://www.cybersecuritydive.com/news/twitter-mudge-security/589177/>; see also *id.* (Doug Britton, CTO of RunSafe Security, commented that Mudge was "a good hire for Twitter, because the security issues are macro and micro.").

⁴⁵ Joseph Menn, *Twitter names famed hacker "Mudge" as head of security*, CNBC (Nov. 16, 2020), <https://www.cnbc.com/2020/11/16/twitter-names-famed-hacker-mudge-as-head-of-security.html>.

⁴⁶ *Id.*

- e. **“Twitter Service”** - the company’s internal name for the division tasked with operational enforcement of global content moderation at scale, including processing and the removal of various spam and spam bots.

44. After arriving, Mudge spent two months performing an in-depth evaluation to understand how things worked, or didn’t work, at Twitter. Mudge conducted in-depth interviews of about 40 employees – from members of the executive team to engineers to salespeople – to gain a better understanding of the Company’s current status with regards to security, perceived security needs, and what employees understood about Twitter’s security.⁴⁷ He attended engineering meetings, reviewed internal technical documents, and directly evaluated some of Twitter’s key computer systems and servers. Even though the company had been under the FTC consent decree since 2011, requiring by law that Twitter address fundamental security and privacy issues, Mudge remembers early in his tenure hearing Mr. Agrawal stating to the executive team that “Twitter has 10 years of unpaid security bills.”

[Disclosure continues next page]

⁴⁷ The results of Mudge’s inquiries were reflected comprehensively in a Google Sheet that he lost access to when he was terminated. Mudge can assist investigators in identifying and finding this important document.

IV. Mudge Discovers Egregious Deficiencies, Negligence, Willful Ignorance, and Threats to National Security & Democracy

45. **Mudge's findings were dire.** Nearly a decade after the FTC Consent Order, with total users growing to almost 400 million and daily users totaling 206 million,⁴⁸ Twitter had made little meaningful progress on basic security, integrity, and privacy systems. Years of regulatory filings in multiple countries were misleading, at best. In many ways, the situation was even worse than Dorsey feared, as the company haphazardly expanded into contentious international areas without even following existing (albeit deficient) corporate policies.

46. Mudge's reports, all highly-experienced experts and intimately familiar with Twitter's problems with the FTC, told Mudge unequivocally that **Twitter had never been in compliance** with the 2011 FTC Consent Order, and was **not on track to ever achieve full compliance**. Twitter's deficiencies are described in greater detail later, and in the exhibits.⁴⁹ But at a high level, Mudge found **serious deficiencies** in:

a. **Privacy**, including

- i. **Ignorance and misuse of vast internal data sets**, with only about 20% of Twitter's huge data sets registered and managed,⁵⁰
- ii. **Mishandling Personally Identifiable Information (PII)**, including repeated marketing campaigns improperly based on user email addresses and phone numbers designated for security purposes only;⁵¹

⁴⁸ Brian Dean, *How Many People Use Twitter in 2022? [New Twitter Stats]*, Backlinko, Jan. 5, 2022, <https://backlinko.com/twitter-users>.

⁴⁹ See Exhibits 1, 2, 3, 4, 20, 23, and 26

⁵⁰ See Exhibits 1 and 4

⁵¹ This was the problem identified in the FTC's July 2020 Draft Complaint. In mid-2021, in the midst of negotiations with the FTC, Twitter *did it again*: the product sales team saw a data set, and (in the absence of any data tracking) just started using it for ad targeting. When one Twitter executive learned that, even after the 2011 Consent Decree and 2020 Draft Complaint, this was happening again, he said: "So we only started to address the problem, and then got side tracked and forgot about it? We do that for everything." (This may have been Twitter SIM 144.) Around the same time, the CFO complained to Mudge that his request to send a large collection of user emails to an advertiser was being blocked by a few engineers. Mudge explained that the engineers were right to be blocking it, because Twitter did not have any understanding of data-lineage and there was no indication whether Twitter sending this data to a customer would be violating the FTC consent decree. In a further irony, better data lineage and enforced handling would not only have made the company compliant, but would have enabled the company to better monetize data, a double win.

iii. **Misusing security cookies** for functionality and marketing,⁵²

iv. **Misrepresentations to the FTC** on these matters;⁵³

b. **Information Security (InfoSec)**, including

- i. **Server vulnerabilities**, with over 50% of Twitter's 500,000 data center servers with non-compliant kernels or operating systems, and many unable to support encryption at rest,
- ii. **Employee computers exposed**, with over 30% of devices reporting they had disabled software and security updates,⁵⁴

⁵² In December 2021, the French CNIL (*Commission Nationale de l'Informatique et des Libertés*) demanded Twitter comply with their regulations. Up until Q2/Q3 2021 Twitter did not have sufficient understanding of how, and what, cookies were used for. Cookies were used for multiple functions, such as ad tracking and session security. It was apparent Twitter was in violation of international data requirements across many regions of the world. The new Twitter Privacy Engineer team had worked tirelessly with product to disentangle cookies and permit some form of user choice and control in regards to cookies and tracking performed by Twitter. On December 31, 2021, the fix was rolled out exclusively to France and then, because Twitter lacked separate testing environments, encountered a problem and it was almost immediately rolled back and disabled. The bug was fixed in a matter of hours, but product and legal blocked rolling out the fix for another month, until January 31, 2021, in order to extract maximum profit from French users before rolling out the fix. Mudge challenged executives to claim this as anything other than an effort to prioritize incremental profits over user privacy and legal data privacy requirements. The senior leaders in that meeting confessed that Mudge was correct. Twitter even launched a proactive court case attempting to claim that all cookies were by definition critical and required, because the platform is powered by advertisements. During internal conversations, Mudge heard Twitter product staff admitted that their argument was false and made in bad faith.

⁵³ In years past, the FTC had asked Twitter whether the data of users who canceled their accounts was properly "deleted." Twitter had determined that not only had the data not been properly deleted, but that data couldn't even be accounted for. Instead of answering the question that was asked, Twitter assured the FTC that the accounts were "deactivated," hoping FTC officials wouldn't notice the difference. Mudge learned about this historical practice in 2021, and was told that fines could be \$3 million each month plus 2% of revenue

⁵⁴ Twitter did not actively monitor what employees were doing on their computers. Although against policy, it was commonplace for people to install whatever software they wanted on their work systems. Twitter employees were repeatedly found to be intentionally installing spyware on their work computers at the request of external organizations. Twitter learned of this several times only by accident, or because of employee self-reporting. In other words, in addition to a large portion of the employee computers having software updates disabled, system firewalls turned off, and remote desktop enabled for non-approved purposes, it was repeatedly demonstrated that until Twitter leadership would stumble across end-point (employee computer) problems, external people or organizations had more awareness of activity on some Twitter employee computers than Twitter itself had.

- iii. **No Mobile Device Management (MDM)** for employee phones, leaving the company with no visibility or control over thousands of devices used to access core company systems;⁵⁵
- iv. **Insider Threats** were virtually unmonitored, and when found the company did not take corrective actions;⁵⁶

c. **Fundamental architecture** including

- i. **lack of development and testing environments** for all software development and testing (highly anomalous for a large tech company),⁵⁷ where engineers use live production data and test directly on the commercial service, leading to regular service disruptions,
- ii. **serious access control problems**, with far too many staff (about half of Twitter's 10,000 employees, and growing) given access to sensitive live production systems and user data in order to do their jobs, the subject of

⁵⁵ It was well known at the time that governments were targeting the cell phones of activists, journalists, and executives, yet Twitter lacked basic abilities to identify or defend against this. See "Pegasus: Spyware sold to governments 'targets activists'" 19 July, 2021.

<https://www.bbc.com/news/technology-57881364>.

⁵⁶ In 2019 two Twitter employees were accused of being Saudi government agents. Ellen Nakashima & Greg Bensinger, *Former Twitter employees charged with spying for Saudi Arabia*, Wash. Post, Nov. 6, 2019,

https://www.washingtonpost.com/national-security/former-twitter-employees-charged-with-spying-for-saudi-arabia-by-digging-into-the-accounts-of-kingdom-critics/2019/11/06/2e9593da-00a0-11ea-8bab-0fc209e065a8_story.html

⁵⁷ A fundamental engineering and security principle is that access to live production environments should be limited as much as possible. Engineers should mostly work in separate development, test, and/or staging environments, using test data (not live customer data). Over a decade prior, companies like Google moved development to segregated test systems. But at Twitter, engineers built, tested, and developed new software directly in production with access to live customer data and other sensitive information in Twitter's system. This ongoing arrangement, almost unheard of at modern tech companies, causes repeated problems for Twitter in bad software deployments and significantly reduces the work an attacker needs to do to acquire credentials with extremely sensitive access. Twitter's practice was a huge red flag for job candidates, who universally expressed disbelief. One particular candidate for Vice President of Information Technology considered withdrawing his application on the (accurate) rationale that Twitter's lack of basic engineering hygiene in their arrangement presaged major headaches.

specific misrepresentations in 2020⁵⁸ by then-Chief Technology Officer Parag Agrawal;

- iii. **Insufficient data center redundancy**,⁵⁹ without a plan to cold-boot or recover from even minor overlapping data center failure, raising the risk of a brief outage to that of a catastrophic and existential risk for Twitter's survival.

47. Unsurprisingly, given these and other deficiencies, Twitter suffered from an **anomalously high rate of security incidents**⁶⁰—approximately one security incident each week serious enough that Twitter was required to report it to government agencies like the FTC and SEC, or foreign agencies like Ireland's Data Protection Commission.⁶¹ In 2020 alone, Twitter had more than 40 security incidents, 70% of which were access control-related. These included 20 incidents defined as breaches; all but two of which were access control related.⁶² Mudge identified there were several exposures and vulnerabilities at the scale of the 2020 incident waiting to be discovered,⁶³ and reasonably feared Twitter could suffer an Equifax-level hack.⁶⁴

⁵⁸ In particular, Agrawal had misrepresented the truth when he told Wired magazine that “[t]he amount of access, the amount of trust granted to individuals with access to these tools, is substantially lower today.” Nicholas Thompson & Brian Barrett, *How Twitter Survives Its Biggest Hack—and Plans to Stop the Next One*, Wired (Sept. 24, 2020), <https://www.wired.com/story/inside-twitter-hack-election-plan/>; see also Parag Agrawal & Damien Kieran, Twitter Blog, *Our continued work to keep Twitter secure* (Sept. 24, 2020), https://blog.twitter.com/en_us/topics/company/2020/our-continued-work-to-keep-twitter-secure (“We have strict principles around who is allowed access to which tools and at what time, and require specific justifications for customer data to be accessed.”).

⁵⁹ Although Twitter “is relatively coy about its current data center footprint” (because that footprint is vulnerable), it has been publicly reported that during the 2020-21 timeframe, Twitter had “[a] facility in Sacramento, and ... [a] data center in Atlanta.” *Twitter plans to build out new data center as platform grows*, Sebastian Moss, February 7, 2020 <https://www.datacenterdynamics.com/en/news/twitter-plans-build-new-data-center-platform-grows/>.

When Mudge joined, Twitter had recently begun some amount of load shifting between data centers, but the process was both manual, and buggy.

⁶⁰ A security incident is an incident significant enough to trigger interruptions to work and redirect teams to track down the incident, determine the scope of the incident, and, if required, report it to the government.

⁶¹ A security incident that may need to be reported would include exposure of sensitive user information like emails, passwords, phone numbers or users' credit card data. An incident that might not need reporting might be a code bug.

⁶² See Exhibit 3; Mudge noted that internal reports stated more than 200 million customers and more than 20,000 employees (current and past) were impacted or involved in such breaches.

⁶³ See Exhibit 17.

⁶⁴ A 2017 hack of Equifax exposed the data of 147 million U.S. persons (fewer than those affected by Twitter's deficiencies), and led to a \$575 million fine. See U.S. Federal Trade Comm'n, *Equifax to Pay \$575 Million as Part of Settlement with FTC, CFPB, and States Related to 2017 Data Breach* (July 22,

48. **January 6 Capitol Attack:** When a violent mob attacked and invaded the U.S. Capitol Building in an attempt to prevent Congress from certifying the election, Mudge quickly went to the executive in charge of engineering and asked “how do we seal the production environment?” Not knowing if there would be acts of internal protest aligned with the rioters, Mudge did not want any employees accessing, or potentially damaging the production environment. It was at this point when he learned that it was impossible to protect the production environment. All engineers had access. There was no logging of who went into the environment or what they did. When Mudge asked what could be done to protect the integrity and stability of the service from a rogue or disgruntled engineer during this heightened period of risk he learned it was basically nothing. There were no logs, nobody knew where data lived or whether it was critical, and all engineers had some form of critical access to the production environment. (Later on January 6 after the Capitol attack, the incoming administration offered Mudge a day-one appointed position as Chief Information Security Officer for the United States; Mudge turned the position down on the grounds that he thought he could have more positive impact fixing Twitter.)
49. **Initial report:** Mudge presented his initial findings to the senior executive team in February 2021, about one week before the Q1 Board meeting.⁶⁵ Jack Dorsey had specifically recruited Mudge for his reputation of speaking truth to power, and told Mudge to not hold back. And Twitter’s other senior leaders knew they had security problems. But even so, the rest of the executives were stunned to hear Mudge tell them just how bad things were. While Mudge highlighted some positive aspects of Twitter’s security processes, such as the Company’s well exercised (but understaffed) team tasked with scrambling to react to crises, the overall picture was dire.
50. **Defensiveness and denial from Agrawal:** Even at the first executive team meeting where Mudge shared his initial findings, Mudge got stiff pushback. In particular, Twitter’s CTO Parag Agrawal vehemently challenged Mudge’s assessment that Twitter faced a non-negligible existential risk of even brief simultaneous, catastrophic data center failure, and had no workable disaster recovery plan.

2019),
<https://www.ftc.gov/news-events/news/press-releases/2019/07/equifax-pay-575-million-part-settlement-ftc-cfpb-states-related-2017-data-breach>.

⁶⁵See Exhibit 3

Twitter's most senior engineers had told Mudge they did not know whether, or on what time frame, Twitter could recover from such an outage. Perhaps Agrawal's defensiveness should not have been surprising—as a senior engineer later promoted to then Chief Technology officer for years, Twitter's problems had developed under Agrawal's watch.

51. Below are excerpts from Mudge's notes prepared for the February meeting, on the particular issue of simultaneous catastrophic data center failure. While catastrophic loss of data centers would understandably be problematic, take note of the last item. Mudge was shocked to learn that even a temporary but overlapping outage of a small number of datacenters would likely result in the service going offline for weeks, months, or permanently. This was even more disturbing as small outages were not uncommon due to bad software pushes from the engineers. On top of this all engineers had some form of access to the data centers, the majority of the systems in the data centers were running out of date software no longer supported by vendors, and there was minimal visibility due to extremely poor logging. This meant that of the four threats cited below what would normally be viewed as the least surprising, and was the statistically most likely, issue carried the greatest damage to the company; an existential company ending event:

With all of the above helping to provide context around our environment, and some of what is slowing us down or making it difficult to execute on our strategy and operations, let me share the existential threat that surprised me.

Threat matrix of effect:

- [REDACTED] data centers physically destroyed
 - Twitter unable to do business - full stop (not surprising)
- [REDACTED] goes down (hard or soft)
 - Twitter continues to run out of [REDACTED]
- [REDACTED] goes down (hard or soft)
 - Twitter operates, but impaired - and more impaired as time goes on
- [REDACTED] data centers gracefully go down and come back up
 - We don't know - best guess is weeks to months to bring the service back online
 - We can't boot(?)
 - Known unknown we really should know

52. Instructions to withhold information from Board: After the executive team meeting, Mudge was instructed not to send a detailed written report to the Board of Directors, but instead convey his findings orally, at a high level only.⁶⁶ Mudge found the request unusual, but as a new team member, complied. With the benefit of hindsight, Mudge now interprets this instruction as an overt act in furtherance of an ongoing effort to restrict critical information and defraud the Board of Directors and Twitter shareholders.⁶⁷

53. Executive action: As Twitter's Security Lead, Mudge was responsible not just for identifying the problems, but also for fixing them. And over the course of 2021, he designed and implemented a long-term strategy for reform. Among other things, Mudge:

- a. Stood up a world-class **Privacy Engineering team**, recruited some of the best leadership talent in the world, quantified the problem for Twitter for the

⁶⁶ On multiple occasions, executive team members shared that they believed that the best type of Board was one that was uninformed so as to keep them very hands off and mostly out of Twitter's business. Note that Dorsey always encouraged Mudge to be direct, unfiltered, honest, and transparent with the Board of Directors.

⁶⁷ On information and belief, with respect to this episode, the particular Twitter staff lawyer was acting on instructions from Twitter General Counsel [REDACTED]. Mudge prefers not to provide the name of the staff lawyer that conveyed the instructions, on the grounds that he (Mudge) has no reason to suspect that particular staff lawyer of harboring fraudulent intent.

-
- first time,⁶⁸ and achieved more progress in only 8 months than had been made over several years prior⁶⁹;
- b. Solicited independent report by the [REDACTED] to formally identify **platform integrity (manipulation, disinformation, and spam) capabilities and gaps** for the first time;
 - c. Procured resources and head count to enable significant growth of the **InfoSec team** for the purpose of enabling reform and accountability;
 - d. Created the **#Protect Initiative**, a formal 3-year Board-reported objective to address critical privacy, security, and platform manipulation issues;
 - e. Oversaw **improvements on user safety cases** that drove down a backlog⁷⁰ of more than 1M cases to approximately 200K, and placed Twitter on its way to running within internal support service level agreements (“SLAs,” in which Twitter defines the level of service that it needs to meet in order to responsibly serve its customers) in 2022 for the first time ever;⁷¹
 - f. Demanded **data-driven metrics, and accountable ownership** of every process;
 - g. Began **aggressively recruiting diverse top talent** from across the industry.

54. Disengaged CEO: CEO Jack Dorsey had recruited Mudge personally. They got along well, and Mudge has never suspected Dorsey of harboring bad intent. But Dorsey, the high-profile CEO of one of the most prominent companies on earth, was experiencing a drastic loss of focus in 2021. Dorsey attended meetings

⁶⁸ General Counsel [REDACTED] did not only provide legal advice, but supervised operational (non-legal) staff on non-legal matters. With respect to one operational (non-legal) privacy matter, after Gadde was shown quantified data for the first time, she stated approximately “...so this proves that we [Twitter] haven’t made any progress over the past 4 years.”

⁶⁹ Mudge notes that Twitter engineers worked very hard in the prior years in good faith, but without leadership having domain knowledge and expertise to direct them to measure the problem and correctly direct the effort and solutions the underlying problem grew larger, not smaller.

⁷⁰ Backlogs included items such as harassment, violations of various rules, and reported accounts and tweets, problems with accounts, etc. It was historically the norm that cases in backlogs would eventually become so old that they would be silently closed, which most would agree is inappropriate support.

⁷¹ In or around October 2021, Mudge learned that even the @TwitterSupport account was historically unmanned. Through new leadership, brought on by Mudge, other overlooked fundamental issues such as language support and staffing safety and abuse agents to match timezones when issues were being reported were identified and improved.

sporadically, and when he did, he was extremely disengaged.⁷² In some meetings—even after he was briefed on complex corporate issues—Dorsey *did not speak a word*. Mudge heard from his colleagues that Dorsey would remain silent for days or weeks. Worried about Dorsey’s health, the senior team mostly tried to cover up for him,⁷³ but even mid- and lower-level staff could tell that the ship was rudderless.

55. Lack of Support: Whatever the cause, Dorsey’s absent behavior was anomalous and unhelpful in summoning the herculean effort needed to fix Twitter’s problems. In theory, Dorsey supported Mudge, and delegated him a huge amount of responsibility. But in deed, Mudge was getting little to no actual support for his task of fundamentally changing the risky behaviors of over 8,000 employees, and the entire corporate culture. Other senior executives took advantage of Dorsey’s absence to stay in their separate silos, pursuing their separate interests without interference. Unsurprisingly, this dynamic had negative consequences.

56. Cascading data center problems: In or around the spring of 2021, Twitter’s primary data center began to experience problems from a runaway engineering process, requiring the company to move operations to other systems outside of this datacenter. But, the other systems could not handle these rapid changes and also began experiencing problems. Engineers flagged the catastrophic danger that all the data centers might go offline simultaneously. A couple months earlier in February, Mudge had flagged this precise risk to the Board because Twitter data centers were fragile, and Twitter lacked plans and processes to “cold boot.” That meant that if all the centers went offline simultaneously, even briefly, Twitter was unsure if they could bring the service back up. Downtime estimates ranged from weeks of round-the-clock work, to permanent irreparable failure.

57. “Black Swan” existential threat: In fact, in or about Spring of 2021, just such an event was underway, and shutdown looked imminent. Hundreds of engineers nervously watched the data centers struggle to stay running. The senior executive

⁷² Over the course of 12 months, Mudge had no more than 6 one-on-one telephone calls with Dorsey, each lasting less than 30 minutes and almost all at the request of Mudge. During these calls, Dorsey cumulatively spoke perhaps fifty words. The total set of their electronic communications, again predominantly initiated by Mudge, came to no more than a couple dozen text messages.

⁷³ One executive team member bragged to Mudge about trying to get Dorsey to break his silence by prodding and aggravating him (Dorsey).

who supervised the Head of Engineering, aware that the incident was on the verge of taking Twitter offline for weeks, months or permanently, insisted the Board of Directors be informed of an impending catastrophic “Black Swan” event. Board Member ██████ responded with words to the effect of “Isn’t this exactly what Mudge warned us about?” Mudge told ██████ that he was correct. In the end, Twitter engineers working around the clock were narrowly able to stabilize the problem before the whole platform shut down.⁷⁴

58. Software Development Life Cycle (SDLC): An SDLC is a uniform process to develop and test software, and a basic best practice for engineering development at commercial companies. Twitter’s need to implement an SDLC was more than a best practice, it had been required since the 2011 FTC Consent Order⁷⁵ and reported regularly to the Board of Directors.⁷⁶ In or around May 2021, Mudge instructed that the Board Risk Committee receive accurate data showing that the company only had a *template* for the SDLC, not even a functioning process, and by Q2 2021 that template had only been rolled out for roughly 8 to 12% of projects.

59. Board Misled on SDLC: Board ██████ became incensed and noted that for years the board had been hearing “the (SDLC) effort was getting closer to being complete.” ██████ realized he and the Board had been misled, and was not happy. After the meeting, an executive called Mudge to state that he and Agrawal were upset with Mudge for providing accurate information to the Risk Committee, and that he and Agrawal deserved credit for their efforts. The call was a turning point for Mudge. He realized that for years, Agrawal and other executives had been misleading the board by **reporting their efforts, not actual results.**

⁷⁴ This is one of the specific items on which Agrawal had challenged Mudge earlier that year, which Mudge interpreted as defensiveness over problems that had developed on Agrawal’s watch as Chief Technology Officer, see discussion above.

⁷⁵ See FTC 2011 Decision & Order (requiring Twitter to implement technical safeguards appropriate for its size and complexity, the nature and scope of its activities, and the sensitivity of the nonpublic consumer information). U.S. Fed. Trade Comm’n, *In the Matter of Twitter, Inc.*, (No. C-4316), Decision & Order (Mar. 2, 2011), available at: <https://www.ftc.gov/sites/default/files/documents/cases/2011/03/110311twitterdo.pdf>.

⁷⁶ See FTC 2011 Decision & Order requiring Twitter to implement technical safeguards appropriate for its size and complexity, the nature and scope of its activities, and the sensitivity of the nonpublic consumer information. U.S. Fed. Trade Comm’n, *In the Matter of Twitter, Inc.*, (No. C-4316), Decision & Order (Mar. 2, 2011), available at: <https://www.ftc.gov/sites/default/files/documents/cases/2011/03/110311twitterdo.pdf>.

60. Attempts to engage Agrawal: On many of these deficiencies, the Chief Technology Officer behaved defensively. Beginning mid-2021, Mudge initiated bi-weekly one-on-one meetings with Agrawal to flag issues with him first, and tried to get Agrawal's buy-in for reforms. For a period of time, these regular meetings seemed to improve their working relationship.

61. Anomalous Handling of Report on Platform Integrity: With authority to engage consultants, Mudge hired [REDACTED] to produce a report on Twitter's capacity to combat mis- and dis-information, fight spam and hostile actors, and promote overall platform integrity. In or around May or June of 2021, [REDACTED] reported its initial findings, which were devastating. Word got out to other senior executives, who became concerned about the impact on Twitter's reputation were the findings to become publicly known. Without notifying Mudge, others on the executive team approached [REDACTED] and ordered them to open a separate contract with an outside law firm. [REDACTED] was told to send their report there first; the external law firm was responsible for removing factual information that would be especially embarrassing for Twitter, and return to [REDACTED] a "clean" version to present to Mudge. When Mudge learned of this, he wondered whether this process was illegal or unethical. Further, despite the fact that the report did not touch on legal or compliance issues in any way, lawyers applied the erroneous label "Privileged and Confidential / Attorney Work Product" to the report.⁷⁷ Twitter counsel explicitly told Mudge that **this was intended to hide the findings and prevent them from becoming known internally or externally.**⁷⁸

62. Perverse bonus structure: In or around July 2021, Twitter announced the "Value Creation Award,"⁷⁹ a new bonus structure in which top executives could individually earn over \$10 million for generating short-term growth of mDAU ("monetizable daily active users," see description above Section II). No bonus was provided for

⁷⁷ See Exhibit 2; After consulting with independent "filter counsel," we have concluded that this document is not in fact subject to attorney-client privilege. (1) The report does not contain or discuss legal advice or exposure, nor does it discuss legal or regulatory options or contain legal citations. (2) It was written by non-lawyers at [REDACTED] for Mudge, a non-lawyer executive tasked with security, privacy, and content moderation at scale.

⁷⁸ *Id.*; As described above, we have obtained an independent legal opinion that this document is not in fact subject to Attorney-Client Privilege, and also that the Work Product Doctrine has no application in the context of voluntary, protected disclosures under the Dodd-Frank Act.

⁷⁹ See page 62, SEC Schedule 14A, available at https://www.sec.gov/Archives/edgar/data/1418091/000114036122012589/ny20001921x1_pre14a.htm#tDC

improving platform privacy, security or integrity. Mudge came to believe that short-sighted incentives like this were an important cause of Twitter's egregious ongoing deficiencies.

63. **Failed “stemming” for hateful ad targeting:** Twitter maintains a list of hateful terms and slurs that cannot be used for ad targeting. But Mudge learned that the list was not “stemming” properly, meaning that even minor variations on slurs were able to be used for targeting for an unknown period (Twitter SIM 154).
64. **Failed logins:** In or around August 2021, Mudge notified then-CTO Agrawal and others that the login system for Twitter's engineers was registering, on average, between 1500 and 3000 failed logins every day, a huge red flag. Agrawal acknowledged that no one knew that, and never assigned anyone to diagnose why this was happening or how to fix it.
65. **No employee computer backups:** In or around Q3 or Q4 2021, Mudge learned that no Twitter employee computers were being backed up at all. Supposedly, Twitter's IT department had managed a backup system for years, but it had never been tested and Mudge learned that it was not functioning correctly. Obviously this raised fundamental risks for corporate data integrity, including financial data, for any information needing to be recovered that was located exclusively on employee laptops. (To the extent that financial staff's data was at risk, it could constitute material weaknesses in internal financial controls required under SEC regulations.) Other Twitter executives—aware that the company was chronically out of compliance with most government requests for information—tried to look on the bright side, noting that going forward Twitter would have a valid excuse for not responding to regulator queries about which data particular employees had access to on which days. They explicitly decided not to replace or fix the employee backup system, but instead discontinued the service entirely.
66. Knowing that employees often had actual data from production systems on their laptops several executives and leaders commented to the effect of “this is actually a good thing because it means we [Twitter] cannot comply with [legal requests] and have less exposure”.

67. At the end of 2021, the Head of Privacy Engineering and the Chief Privacy Officer reported accurately to the Board that:⁸⁰

Every new employee has access to data they do not need to have access to for the purpose of their role. Until we have implemented a mature centrally owned and operated system to manage access to data (e.g., entitlements and review, Role Based Access Controls, audits, etc) we are at risk of inappropriate access or use of data. Our inability to delete data compounds that risk, as we retain data that we should not have and which is therefore accessible by people who do not need to have access to this data.

68. **Deficient moderation for “Spaces”:** In December 2021, an executive incorrectly told staff and Board members that Twitter’s “Spaces” product was being appropriately moderated. But Mudge researched and discovered that about half of “Spaces” content flagged for review was in a language that the moderators did not speak, and that there was little to no moderation happening.

69. **Log4j:** In December 2021, the world discovered that “Log4j,” a very common piece of software deployed in hundreds of applications across hundreds of millions (or billions) of computers worldwide, contained a previously-unrecognized (“zero-day”) security vulnerability. Overnight, a huge number of computers around the globe needed patching, or else they would be easy for adversaries to exploit. Left unaddressed, Log4j lets hackers break into systems, steal passwords and login information, extract data, and infect networks with malicious software. Log4j was *already* actively being exploited to compromise computers worldwide by criminals and governments alike. As “the most severe computer vulnerability in years,”⁸¹ the FTC instructed companies to pursue remediation, and that they could request detailed explanation and data on a company’s Log4j remediation efforts.⁸² In January 2022, Mudge determined and reported to the executive team that (because of poor engineering architecture decisions that preceded Mudge’s employment) Twitter had over 300 corporate systems and upwards of 10,000 services that might still be affected, but Twitter was unable to thoroughly assess its exposure to Log4j,

⁸⁰ See Exhibit 1

⁸¹ “What the Log4j vulnerability is, who is affected - NCSC.GOV.UK.”
<https://www.ncsc.gov.uk/information/log4j-vulnerability-what-everyone-needs-to-know>.

⁸² “FTC warns companies to remediate Log4j security vulnerability.” 4 Jan. 2022,
<https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2022/01/ftc-warns-companies-remediate-log4j-security-vulnerability>.

and did not have capacity, if pressed in a formal investigation, to show to the FTC that the company had properly remediated the problem.⁸³

70. Unlicensed machine learning materials for core algorithms: In January 2022, in the days before he was terminated, Mudge learned that Twitter had never acquired proper legal rights to training material used to build Twitter's key Machine Learning models.⁸⁴ The Machine Learning models at issue were some of the core models running the company's most basic products, like which Tweets to show each user. Days before Mudge was fired in January 2022, Mudge learned that Twitter executives had been informed of this glaring deficiency several times over the past years, yet they never took remedial action.

71. Misleading regulators in multiple countries: When, years earlier, the FTC had asked questions about the training material used to build Twitter's machine learning models, Twitter realized that truthful answers would implicate the company in extensive copyright / intellectual property rights violations. Twitter's strategy, which executives explicitly acknowledged was deceptive, was to decline to provide the FTC with the requested training material, and instead pointed the FTC towards particular models that would not expose Twitter's failure to acquire appropriate IP rights. In early 2022, the Irish-DPC and French-CNIL were expected to ask similar questions, and a senior privacy employee told Mudge that Twitter was going to attempt the same deception. Unless circumstances have changed since Mudge was fired in January, then Twitter's continued operation of many of its basic products is most likely unlawful and could be subject to an injunction, which could take down most or all of the Twitter platform. Before Mudge could dig deeper into this issue he was terminated.

72. Penetration by Foreign Intelligence & Threats to Democracy: Over the course of 2021, Mudge became aware of multiple episodes suggesting that Twitter had been

⁸³ The head of the Detection and Remediation Team had told Mudge that the lack of visibility across the Twitter system meant that there was no way to determine whether log4j was successfully remediated and that upwards of 10,000 instances of the vulnerability could still be running and would not be able to be identified. Similarly, multiple teams were reporting conflicting internal numbers of systems needing remediation or evaluation or that had been fixed.

⁸⁴ See Exhibit 39

penetrated by foreign intelligence agencies and/or was complicit in threats to democratic governance, including:⁸⁵

- a. **The Indian government** forced Twitter to hire specific individual(s) who were government agents, who (because of Twitter's basic architectural flaws) would have access to vast amounts of Twitter sensitive data. Twitter's transparency reports purported to quantify the number of government data requests from the Indian government, but the company did not in fact disclose to users that it was believed by the executive team that the Indian government had succeeded in placing agents on the company payroll. By knowingly permitting an Indian government agent direct unsupervised access to the company's systems and user data, Twitter executives violated the company's articulated commitments to its users.⁸⁶
- b. Twitter executives opted to allow Twitter to become more dependent upon revenue coming from **Chinese entities** even though the Twitter service is blocked in China. After Chinese entities paid money to Twitter, there were concerns within Twitter that the information the Chinese entities could receive would allow them to identify and learn sensitive information about Chinese users who successfully circumvented the block, and other users around the world. Twitter executives knew that accepting Chinese money risked endangering users in China (where employing VPNs or other circumvention technologies to access the platform is prohibited) and elsewhere. Twitter executives understood this constituted a major ethical compromise. Mr. Zatko was told that Twitter was too dependent upon the revenue stream at this point to do anything other than attempt to increase it.
- c. The **Nigerian government** blocked Twitter in June 2021, then falsely claimed to be in negotiations with Twitter executives; Twitter's failure to correct the false record on many reported non-existent discussions with the Nigerian government permitted Nigeria to negotiate unilaterally through media and

⁸⁵ Mudge has submitted a separate disclosure including details and documentation of these incidents to the Counterintelligence and Export Controls Section within the National Security Division of the U.S. Department of Justice, and to the Senate Select Committee on Intelligence.

⁸⁶ "India - Twitter Transparency Center." <https://transparency.twitter.com/en/reports/countries/in.html>.

dictate unfavorable terms for final resolution.⁸⁷ Twitter's deliberate decision to refrain from correcting misinformation about Twitter's proposed negotiations with the Nigerian government directly harmed Twitter shareholders. Permitting the Nigerian government to impose various conditions on the platform harmed free expression rights and democratic accountability for Nigerian citizens;⁸⁸

- d. A few months before CTO Parag Agrawal was promoted to CEO, Agrawal suggested to Mudge that Twitter should **consider ceding to the Russian Federation's** censorship and surveillance demands as a way to grow users in Russia. Although Mr. Agrawal's suggestion was never pursued or implemented, the fact that Twitter's current CEO even suggested Twitter become complicit with the Putin regime is cause for concern about Twitter's effects on U.S. national security. This was a strong departure from the message Mr. Dorsey had conveyed to Mr. Zatko. This interaction was notable because Mr. Zatko was already directing teams to prepare for possible Russian incursions into Ukraine;
- e. Shortly before Mudge was [REDACTED] terminated, Twitter received specific information from a U.S. government source that **one or more particular company employees were working on behalf of another particular foreign intelligence agency.**

73. In none of these cases did Twitter choose to focus on the long term health of the platform and company. Mudge's inference from these and other episodes was that some senior executives were intent on hiding bad news, or even misrepresenting it, instead of trying to fix it. This was possibly because (a) executives had personal financial incentives to grow mDAU / active users; or (b) they didn't know any better; or (c) some of them had built the broken system in the first place.

74. **Squeezing Local Staff:** Countries where Twitter had a physical presence, including actual full time employees (FTEs), and particularly where Twitter had official offices, represented heightened risk to Twitter and the Twitter platform. In addition to the risk exposed by Twitter's fundamental lack of information security and privacy

⁸⁷Ruth Maclean, *Nigeria Lifts 7-Month Ban on Twitter*, New York Times, Jan. 13, 2022, <https://www.nytimes.com/2022/01/13/world/africa/nigeria-lifts-twitter-ban.html>

⁸⁸ See Exhibit 40

control, described in other disclosures, there was the physical safety of the employees to consider. The threat of harm to Twitter employees was sufficient to cause Twitter to seriously consider complying with foreign government requests that Twitter would otherwise fundamentally oppose. The governments of **India, Nigeria, [REDACTED] and Russia** sought, with varying success, to force Twitter to hire local FTEs that could be used as leverage.

75. As of the date of Mudge's termination, January 19, 2022, Twitter remained out of compliance in multiple respects with the 2011 FTC Consent Order, which has the force of law following Twitter's consent to its terms. While the company had made progress on privacy because of Mudge's leadership, it was not on track to ever achieve compliance for other important items, especially in the area of information security (in which Mudge's reform efforts had been repeatedly and unreasonably blocked). Twitter's non-compliance constitutes violations of the Federal Trade Commission Act, 15 U.S. Code §§ 41-58.

[Disclosure continues next page]

V. New CEO Enables Fraud

76. **Dorsey Out, Agrawal In:** In November 2021, Twitter announced that Dorsey was stepping down, and would be replaced by Agrawal, effective November 29.

77. [REDACTED]

78. [REDACTED] The full Board of Directors was set to meet on December 9, 2021, followed by a Board Risk Committee meeting on December 16. And Mudge was becoming increasingly concerned about the accuracy of the information that Board members would receive.

79. **Inaccurate Board Materials:** [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]. Upon review of the materials, Mudge determined that the materials contained false and misleading statements about the state of Twitter's information security and privacy.⁸⁹

80. Mudge's concerns are described in greater detail below, and in the exhibits to this disclosure. At a high level, Mudge flagged four issues which were improperly omitted or presented in a misleading way:

⁸⁹ Some of the emails evidencing deliberate misrepresentations to the Board in December 2021 constitute privileged attorney-client communications. Therefore, although Mudge is unable to share those particular communications, he can nevertheless describe the underlying facts. "[T]he protection of the [attorney-client] privilege extends only to communications, and not to facts. A fact is one thing and a communication concerning that fact is an entirely different thing." *Upjohn Co. v. United States*, 449 U.S. 383 at 395-96 (1981) (internal citations omitted).

-
- a. **Basic security protections on software (software and system patching):** The materials reported a misleading statistic that 92% of employee computers had security software installed, implying those computers were secure. In fact, that software's most basic function was to determine whether that particular computer's software and settings met basic security standards—and the software was reporting that one-third of the systems were critically insecure. A full 30% of employee systems were reporting that they had disabled critical safety settings such as software updates. Other critical flaws reported by the software brought this number closer to 50%. This crucial context was not included for the Board.
- b. **Access control to systems and data:** On the issue that led to the FTC complaint in 2011 and the July 2020 hack,⁹⁰ the materials did not contain the overall numbers, which were getting worse. Instead, the materials took a small, cherry-picked subset of data that could be made to look like a positive trend, and turned that into a graph with a downward trajectory. The graph misleadingly suggesting that Twitter was making significant progress in reducing access to production systems. Mudge knew that the actual underlying data showed that at the end of 2021, 51% of the ~11 thousand full-time employees had privileged access to Twitter's production systems, a 5% increase from the 46% of total employees in February of 2021 that Mudge had shared in his initial findings delivered to the Board in early 2021.
- c. **Volume and frequency of security incidents:** A graphic in the document showed only a subset of security incidents, presented as if to encompass all security incidents.⁹¹ Twitter's actual total number of security incidents in 2021 was closer to 60, a marked difference from what was being implied. The misleading graphic also attributed only 7% of incidents to access control,

⁹⁰ As Mudge noted in his February 2021 oral presentation to the Board, Twitter engineers have privileged access to Twitter's production systems, which meant 46% of full-time employees had such access to Twitter's live data. While the deck inaccurately reported improvements in this area, the problem was actually getting worse. Data showed that at the end of 2021, 51% of full-time employees had privileged access to Twitter's production systems, a 5% increase from Mudge's February 2021 report to the Board.

⁹¹ Twitter's actual total number of security incidents in 2021 was closer to 60, an alarmingly high number and particularly concerning, given that Twitter faced scrutiny from the FTC and similar international regulatory agencies. Because the document miscategorized various incidents, the graphic misattributed only 7% of incidents to access control, when in reality the root causes of 60% of incidents were actually access control issues, a problem that plagued Twitter's security and had not been properly reported to the Board historically.

when in reality access control was the root cause of 60% of security incidents.

- d. **Lack of Software Development Life Cycle (or related processes and compliance)**, was presented as largely completed, instead of still in the initial phases of planning, as discussed above in Section IV.

81 [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

82. **December 9:** At the December 9 full Board meeting, the inaccurate materials were not shared only because of Mudge's strenuous efforts. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] Mudge then briefed the Board on his 3-year #Protect Initiative that would provide Board-level quantified visibility into the areas of privacy, information security, and scope and remediation efforts around malicious actors and activity on the platform.

83.

[REDACTED]

[REDACTED]

[REDACTED]

84. [REDACTED]

85. **December 14:** Agrawal [REDACTED] He said that the materials as they stood would be presented to the Board Risk Committee on December 16, despite Mudge’s informing Agrawal that the information in the materials were materially misleading.

86. [REDACTED]

[Disclosure continues next page]

87. [REDACTED]

[REDACTED]

94 [REDACTED]

95 After Mudge’s signature in the email there were detailed explanations of a subset of items Mudge felt were critical so as to ensure there was no chance for Parag to be confused as to why there were concerns.

88. **December 16:** The Risk Committee received the misleading document before the meeting, [REDACTED]

[REDACTED] This screenshot⁹⁶ of unsolicited, real-time instant messages during the Board Risk Committee meeting confirms that other participants recognized untrue [REDACTED]

[REDACTED] 2:34 PM Employee 1
This is not accurate.

[REDACTED] 2:34 PM Employee 2
[REDACTED] talking about the system access plan only
like ssh stuff

[REDACTED] 2:34 PM Employee 1
wildly different from the overall.

[REDACTED] 2:35 PM Employee 2
are you going to clarify?
because I do not want to be trying to re-explain this next quarter
this is "how many have endpoint software" not how many are in a
good state

[REDACTED] 2:36 PM Employee 1
[REDACTED] It's not good.

89. For those without a technical background, these messages are discussing "endpoint software," meaning software on individual employee machines (as opposed to, e.g., server software running in Twitter's data centers) and "access

⁹⁶ See Exhibit 9

control” to Twitter’s production systems and data using only the “ssh” encryption protocol. [REDACTED] misleading in multiple respects:

- a. First, [REDACTED] stated that 92% of Twitter employee computers had this security software installed, implying a high level of security. But in truth, the endpoint software did not actually provide or constitute security on its own. Rather, the endpoint software’s primary function was to evaluate whether the employee computer had basic security configurations enabled. And most importantly, the software on these endpoint computers was reporting dire problems. Over 30% of the more than 10,000 employee computers were lacking the most basic security settings, such as enabling software updates.
- b. Second, [REDACTED] a broad discussion of the company’s overall access control issues, [REDACTED] cited progress on a very small subset of the problem referred to as “system access plan” (which only applied to a small percentage of relevant users) and only the “ssh” communication method (which was merely one method among many). Therefore the cherry-picked numbers misrepresented the problem.
- c. Board members were not given enough information to draw these elementary distinctions.

90. [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

Further Redacted for Congress



*Report government and corporate lawbreaking.
Without breaking the law.*

91. In the e-mail Mudge re-articulated some of the most concerning items

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

⁹⁷ See Exhibit 9



Report government and corporate lawbreaking.
Without breaking the law.

92. [Redacted text block]

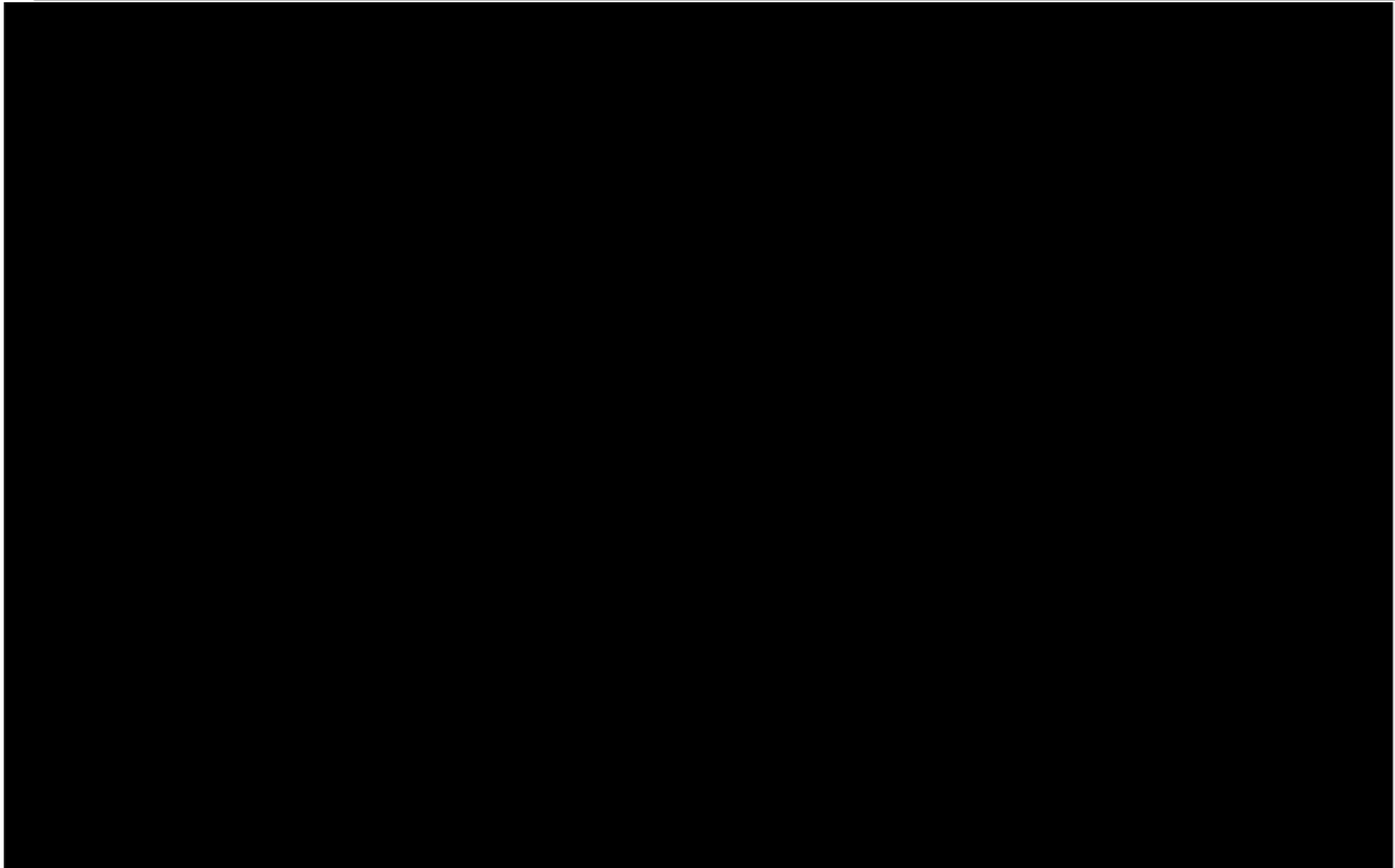
93. [Redacted text block]

94. [Redacted text block]

Further Redacted for Congress



Report government and corporate lawbreaking.
Without breaking the law.



[Disclosure continues next page]

VI. 2022: Termination in the New Year

95. On **January 4, 2022**, [REDACTED] [REDACTED]
[REDACTED] Mudge sent
the following email describing the December 16 meeting as “at worst fraudulent”:⁹⁹

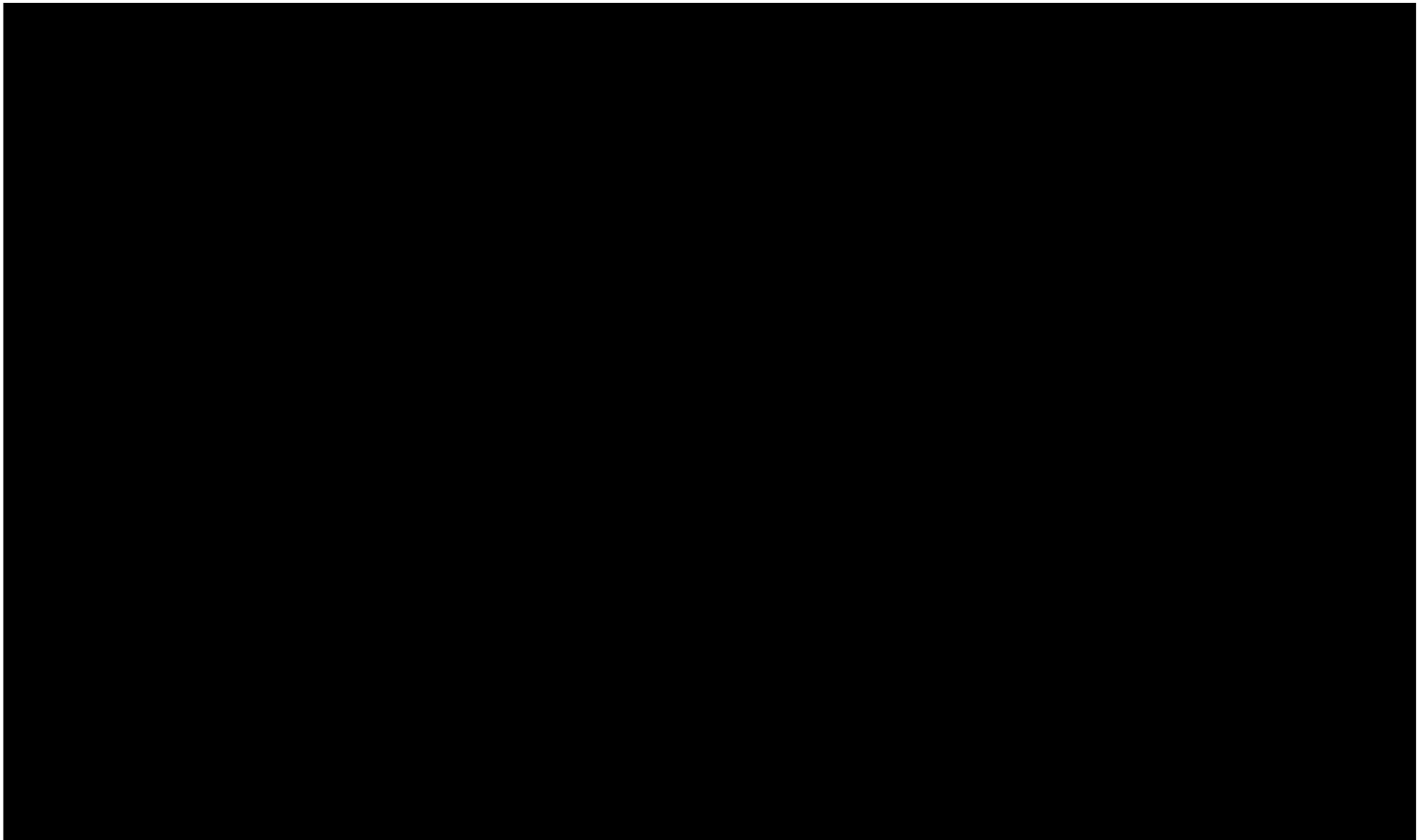
[Disclosure continues next page]

⁹⁹ Exhibit 11, p. 1-2

Further Redacted for Congress



*Report government and corporate lawbreaking.
Without breaking the law.*



— PROTECTED & SENSITIVE WHISTLEBLOWER DISCLOSURE —

96.

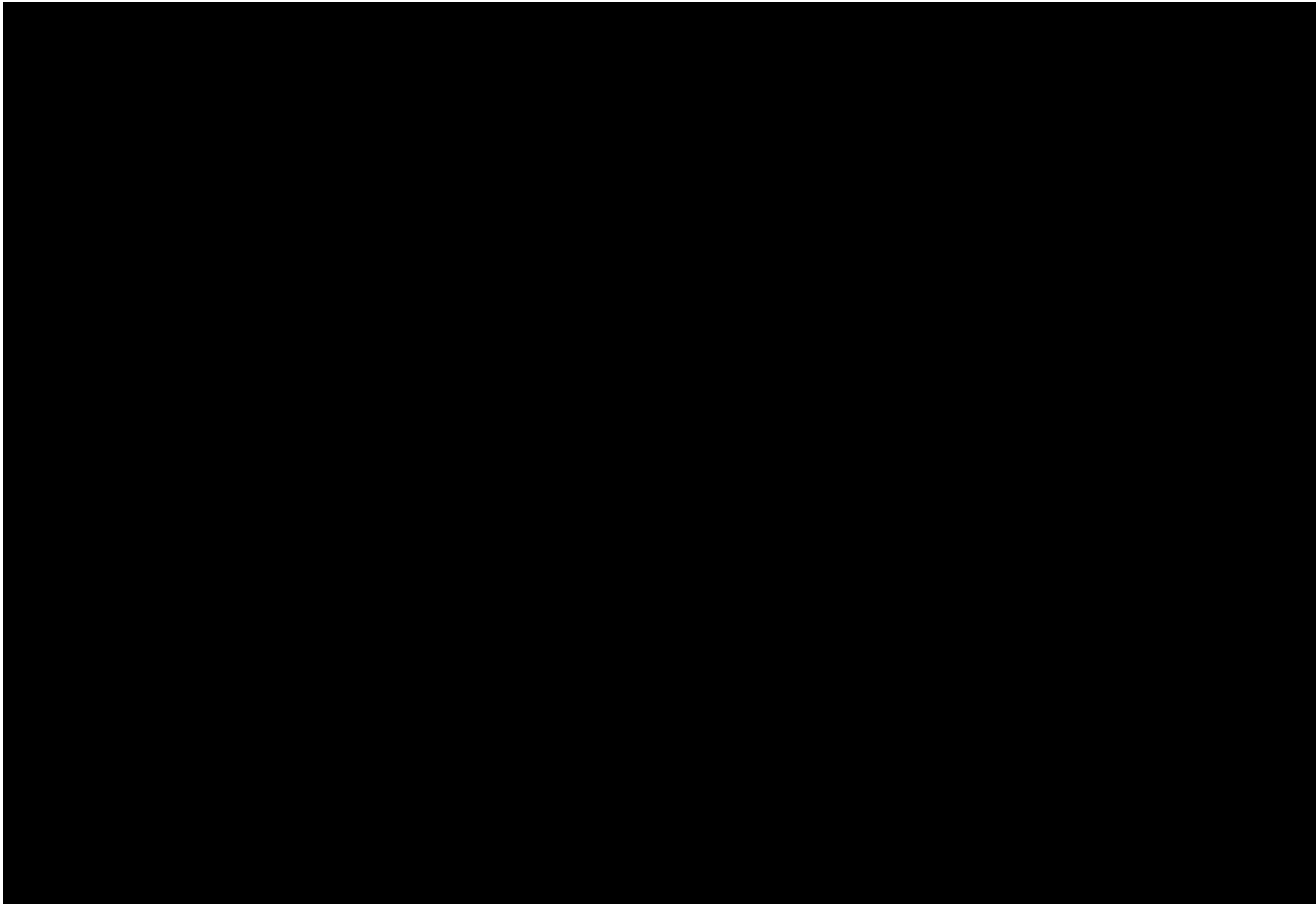
97. **January 11:** Following Mudge’s reference to “fraud,” Twitter [REDACTED] initiated an internal investigation and interviewed Mudge on January 11, 2022. [REDACTED] [REDACTED] agreed that the information delivered to the Risk Committee was inappropriate and inaccurate, and [REDACTED] that he (Mudge) should write a report correcting the misrepresentations. Mudge immediately began drafting a corrective report for the Board as agreed upon. [REDACTED]

98. **Memorandum for the record:** Mudge documented his concerns again on January 12, [REDACTED]

Further Redacted for Congress



*Report government and corporate lawbreaking.
Without breaking the law.*



— PROTECTED & SENSITIVE WHISTLEBLOWER DISCLOSURE —

Further Redacted for Congress

99. **January 18, 11:16am:** Following a request from Twitter’s Chief Compliance Officer, Mudge sent an email confirming he planned to provide corrected materials for the Board “by the end of this week,” as part of a fraud investigation [REDACTED] Less than two hours later, [REDACTED] emailed Mudge, and surprised him with a request to do a call 45 minutes later with [REDACTED] [REDACTED] [REDACTED] Twitter’s Board of Directors’ Risk Committee.¹⁰²

100. [REDACTED], Mudge stated he was close to finishing corrective materials for the Board as agreed upon [REDACTED]

¹⁰¹ On information and belief, [REDACTED] one of the Board members who had advocated strongly for the decision to promote Agrawal to the CEO position, [REDACTED]. [REDACTED]
[REDACTED] Twitter's recent SEC filing also shows that [REDACTED] holds [REDACTED] shares of Twitter stock, then worth [REDACTED]. [REDACTED] therefore has a substantial financial interest in hiding the Company's dire cybersecurity condition.

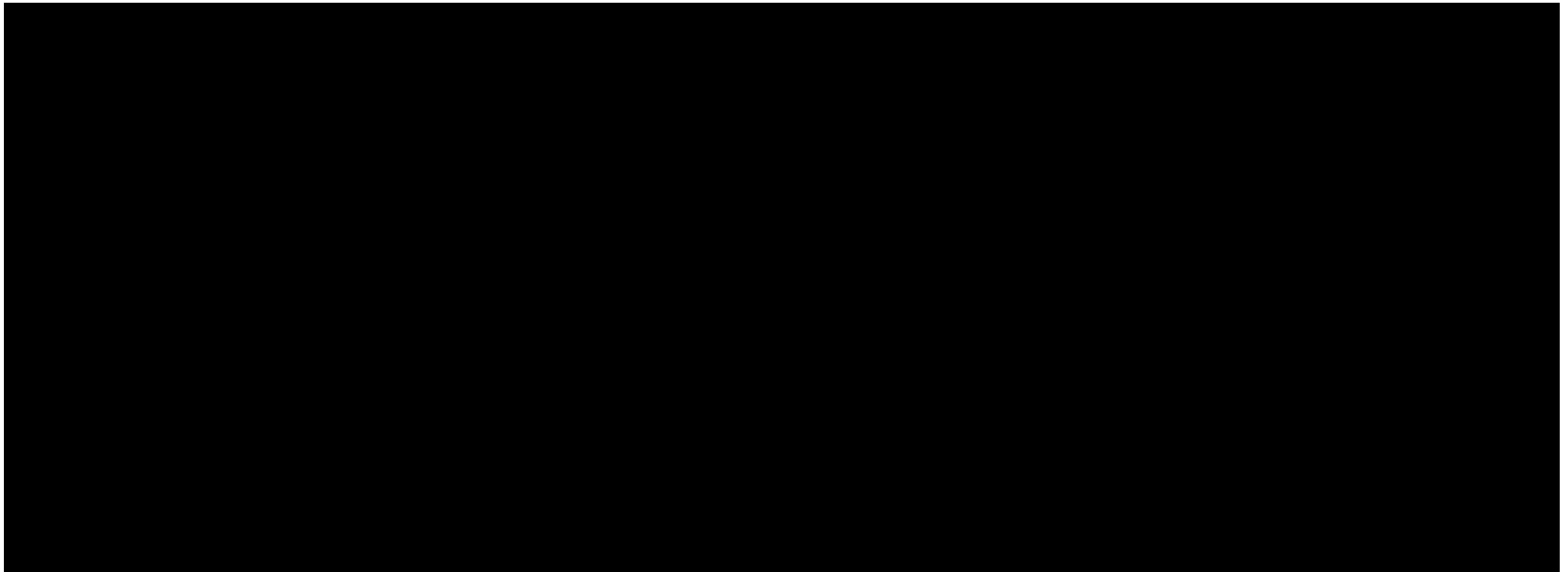
¹⁰² See Exhibit 5

Further Redacted for Congress



*Report government and corporate lawbreaking.
Without breaking the law.*

██ When Mudge attempted to correct this false statement, ██████████
interrupted and refused to let Mudge continue. Mudge sent the following email after the call:¹⁰³



¹⁰³ See *Id.* Although this email contains the words “Privileged and Confidential,” and Twitter counsel was copied, the email is not subject to a claim of attorney-client privilege, because it was a message from one non-lawyer, Mudge, to another non-lawyer, Agrawal, and it was not sent for the purpose of seeking legal advice. We have redacted any arguably privileged content.

101. On January 19, 2022, [REDACTED] [REDACTED]
[REDACTED] called Mudge and terminated his employment.¹⁰⁴

[REDACTED]

¹⁰⁴ See Exhibit 36

102.

103. Based on statements made by Mudge during the termination meeting, later on that same day, January 19, Twitter's [REDACTED] began emailing Mudge, requesting that he continue working to produce corrected materials for the Board of Directors. [REDACTED]

¹⁰⁵ Exhibit 27. No attorney-client privilege attaches to post-termination communications between Mudge and Twitter counsel, because any attorney-client relationship was severed the moment he was terminated and was no longer a "constituent" of Twitter counsel's client, the corporation. Courts have occasionally upheld privilege claims for post-separation communications between corporate counsel and former employees, but those have occurred pursuant to a formal Joint Defense agreement, or when the corporation and the former employee had a shared interest, e.g. in defending tort claims. But Twitter counsel's post-termination communications with Mudge were not pursuant to a Joint Defense Agreement nor a shared interest; in fact the parties were by that point adverse since Mudge had raised concerns during his termination meeting that his termination was retaliation for disclosing violations of law. We are not aware of any authority suggesting that Twitter counsel can assert a valid privilege claim in these circumstances. Further, [REDACTED] never indicated these communications were privileged. [REDACTED] acting in an operational, non-legal capacity.

¹⁰⁶ From the moment Mudge was fired, any attorney-client relationship between himself and Twitter counsel was also terminated. Therefore no attorney-client privilege could attach to his post-termination communications with Twitter counsel or staff.

104. [REDACTED]

[REDACTED]

105. [REDACTED] Twitter staff
inferred (incorrectly) that Mudge was implicated in serious misconduct:¹⁰⁸

107 [REDACTED]
108 [REDACTED]

Further Redacted for Congress



Report government and corporate lawbreaking.
Without breaking the law.



Further Redacted for Congress



*Report government and corporate lawbreaking.
Without breaking the law.*

106. [REDACTED] Twitter revoked Mudge's access to his Twitter devices and Twitter systems. Nevertheless, Twitter [REDACTED] requesting he work **without salary, entirely from memory**, to produce corrective materials for the Board. [REDACTED] tone grew increasingly urgent:¹⁰⁹

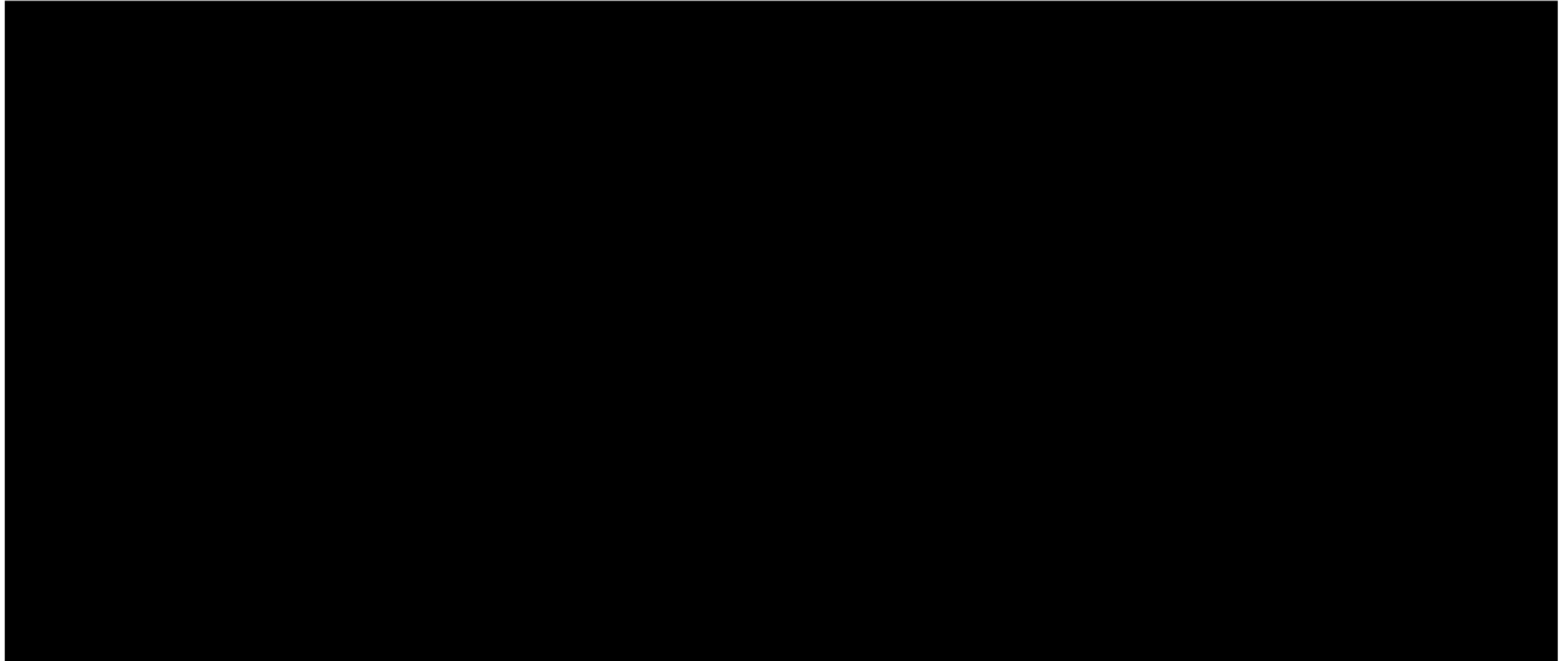
A large, solid black rectangular box covers the majority of the page, redacting the content of the document. It starts below the paragraph and extends nearly to the bottom of the page.

¹⁰⁹ See Exhibit 16

Further Redacted for Congress



*Report government and corporate lawbreaking.
Without breaking the law.*



Further Redacted for Congress



Report government and corporate lawbreaking.
Without breaking the law.

-
107. On February 2, 2022, Mudge sent an email directly to Board member [REDACTED] laying out many of his concerns.
108. On or about February 14, 2022, after at least 150 hours of unpaid work on his personal computer, Mudge sent a 27-page corrective report to Twitter [REDACTED] and two Board Directors. The report is included as an exhibit to this filing.¹¹⁰
109. On or around February 16, on information and belief, the Board Risk Committee held an emergency meeting to consider Mudge's disclosures. Then on February 18, Agrawal publicly announced he was taking unexpected "paternity leave."¹¹¹ Agrawal returned to work a few weeks later, before his child was born.

[Disclosure continues next page]

¹¹⁰ See Exhibit 1

¹¹¹ "Twitter CEO Parag Agrawal will take paternity leave after three" 16 Feb. 2022, <https://www.washingtonpost.com/technology/2022/02/16/twitter-ceo-parag-agrawal-paternity-leave/>.

VII. Material Misrepresentations and Omissions, and Other Legal Violations

110. Under SEC rules, it is unlawful for a publicly-traded company to “make any untrue statement of a material fact or to omit to state a material fact” in connection with the purchase or sale of a security.¹¹² Misrepresentations create liability if they occur in SEC filings, or in any other public communications available to investors. Further, it is unlawful “[t]o employ any device, scheme, or artifice to defraud,”¹¹³ or “[t]o engage in any act, practice, or course of business which operates or would operate as a fraud or deceit upon any person”¹¹⁴ that is “in connection with the purchase or sale of any security.”
111. For years, across many public statements and SEC filings, Twitter has made material misrepresentations and omissions, and engaged in acts and practices operating as deceit upon its users and shareholders, regarding security, privacy and integrity.
112. Twitter’s misrepresentations are especially impactful, given that they are directly at issue in Elon Musk’s contemplated takeover of the company. For example, Twitter filed SEC Schedule 14A on May 17, 2022, which noted that:¹¹⁵

The obligations of Parent and Acquisition Sub [*n.b., the acquiring companies controlled by Mr. Musk*] to consummate the merger are subject to the satisfaction or waiver of each of the following additional conditions, any of which may be waived by Parent:

- Twitter having performed and complied in all material respects with the obligations required by the merger agreement to be performed or complied with by it on or prior to the closing of the merger;

¹¹² 17 C.F.R. § 240.10b-5(b).

¹¹³ 17 C.F.R. § 240.10b-5(a).

¹¹⁴ 17 C.F.R. § 240.10b-5(c).

¹¹⁵ Page 14, SEC Schedule 14A, available at

<https://www.sec.gov/Archives/edgar/data/0001418091/000119312522152250/d283119dprem14a.htm>

- the accuracy of the representations and warranties of Twitter in the merger agreement, subject to applicable materiality or other qualifiers, as of the effective time of the merger or the date in respect of which such representation or warranty was specifically made;
- and the absence of any Company Material Adverse Effect (as defined in the section of this proxy statement captioned “The Merger Agreement—Representations and Warranties”) having occurred that is continuing.

113. But as part of the same SEC filing, in the *Agreement and Plan of Merger - Execution Version*, Twitter made fraudulent misrepresentations:¹¹⁶

ARTICLE IV — REPRESENTATIONS AND WARRANTIES OF THE
COMPANY [n.b. Twitter, Inc.] ...

[T]he Company hereby represents and warrants to Parent and Acquisition Sub as follows:...

Section 4.5 ... Compliance With Laws. ...Neither the Company nor any of its Subsidiaries is in default or violation of any Law applicable to the Company, any of its Subsidiaries or by which any of their respective properties or assets are bound, except for any such defaults or violations that would not have a Company Material Adverse Effect. Notwithstanding the foregoing, no representation or warranty in Section 4.5(a) or this Section 4.5(b) is made with respect to Company SEC Documents or financial statements, “disclosure controls and procedures” or “internal control over financial reporting,” employee benefits matters, Intellectual Property Rights matters, Tax matters, which are addressed exclusively in Section 4.6 (Company SEC Documents; Financial Statements), Section 4.8 (Disclosure Controls and Procedures), Section 4.12 (Employee Benefit Plans), Section 4.14 (Intellectual Property Rights), Section 4.15 (Taxes), respectively. ...

¹¹⁶ Pages A-20-24, Exhibit A to SEC Schedule 14A, available at <https://www.sec.gov/Archives/edgar/data/0001418091/000119312522152250/d283119dprem14a.htm>

Section 4.6 Company SEC Documents: ... [T]he Company SEC Documents complied in all material respects with the requirements of the Securities Act and the Exchange Act, as the case may be, and the applicable rules and regulations promulgated thereunder, and none of the Company SEC Documents at the time it was filed (or, if amended or supplemented, as of the date of the last amendment or supplement) contained any untrue statement of a material fact or omitted to state any material fact required to be stated therein or necessary to make the statements therein, in light of the circumstances under which they were made, or are to be made, not misleading. ...

Section 4.7 Information Supplied None of the information supplied or to be supplied by or on behalf of the Company or any of its Subsidiaries expressly for inclusion or incorporation by reference in the proxy statement relating to the matters to be submitted to the Company's stockholders at the Company Stockholders' Meeting (such proxy statement and any amendments or supplements thereto, the "Proxy Statement") shall, at the time the Proxy Statement is first mailed to the Company's stockholders and at the time of the Company Stockholders' Meeting to be held in connection with the Merger, contain any untrue statement of material fact or omit to state any material fact required to be stated therein or necessary to make the statements therein, in light of the circumstances under which they were made, not misleading at such applicable time...

Section 4.14 Intellectual Property. (a) Except as would not have a Company Material Adverse Effect, the Company and its Subsidiaries solely and exclusively own all patents, trademarks, trade names, copyrights, Internet domain names, service marks, trade secrets and other intellectual property rights (the "Intellectual Property Rights") purported to be owned by the Company and its Subsidiaries (the "Company Intellectual Property"), free and clear of all Liens, except Permitted Liens.

(b) To the Knowledge of the Company, the conduct of the business of the Company and its Subsidiaries as currently conducted does not infringe, misappropriate or otherwise violate any Intellectual Property Rights of any other Person, except for any such infringement, misappropriation or other violation that would not have a Company Material Adverse Effect. To the Knowledge of the Company, no other

Person is infringing, misappropriating or otherwise violating any Company Intellectual Property, except for any such infringement, misappropriation or other violation as would not have a Company Material Adverse Effect.

The Company and its Subsidiaries are in compliance with all applicable Laws, Contracts to which the Company or its Subsidiaries are bound, and internal- and external-facing policies of the Company or its Subsidiaries, in each case, relating to privacy, data protection, and the collection and use of information that constitutes “personal information” under applicable Laws (“Personal Information”) collected, used or held for use by the Company or its Subsidiaries, except where the failure to be in compliance would not have a Company Material Adverse Effect.

(e)¹¹⁷ Neither the Company nor any of its Subsidiaries has experienced any unauthorized access to the information technology systems owned or used by the Company or its Subsidiaries or Personal Information collected, used, held for use or otherwise processed by the Company or its Subsidiaries, except as would not have a Company Material Adverse Effect.

114. Intellectual Property: While several of Twitter’s representations and warranties are untrue, in particular note that the “intellectual property” statements are egregious lies. In fact, Twitter senior leadership have known for years that the company has never held proper licenses to the data sets and/or software used to build some of the key Machine Learning models used to run the service. Litigation by the true owners of the relevant IP could force Twitter to pay massive monetary damages, and/or obtain an injunction putting an end to Twitter’s entire Responsible Machine Learning program and all products derived from it. Either of these scenarios would constitute a “Material Adverse Effect” on the company.

115. Twitter has consistently misrepresented in SEC filings its capacity to recover from even a brief outage of only a few data centers, with misrepresentations *increasing* after the Spring 2021 “Black Swan” event that threatened the platform’s survival:

¹¹⁷ *Sic.* It appears that these sections were mislabelled (a), (b), (e).

a. *“We have implemented a disaster recovery program, which allows us to move production to a back-up data center in the event of a catastrophe... [T]his program is functional...”*¹¹⁸

b. *“We have implemented a disaster recovery program ... this program is functional.”*¹¹⁹

c. These claims are fundamental to any proper valuation of Twitter’s business. And they are significantly misleading; some would consider them plainly false. While Twitter is expanding its data centers with another location, and has been working on migrating non-operational infrastructure to a cloud,¹²⁰ no current program is functional enough to mitigate, or more importantly reliably recover from, the type of “Black Swan” event our client warned Twitter of in his 2021 Q1 board presentation.¹²¹ Furthermore, Twitter failed to report the actual Spring 2021, “Black Swan” event, that threatened Twitter’s survival to the extent that it was the subject of an emergency disclosure to the Board of Directors.

116. Twitter made multiple misrepresentations in its SEC Form 10-K for the fiscal year ended December 31, 2021:¹²²

a. *“We focus on ... measures to help protect the privacy of people on Twitter.”*

i. But when the FTC asked Twitter whether it fully deleted the data of users who left the service, Twitter deliberately misled the FTC by stating those accounts were “deactivated,” even when the data was not fully deleted. And in late 2021, Mudge sent memos to executive team members arguing that, in light of the egregious and ongoing misrepresentations to the FTC, French and Irish regulators, plus the very real possibility of multi-billion dollar fines or even bans from major markets, Privacy should become

¹¹⁸ SEC Form 10-Q for the quarter ended September 30, 2021, available at <https://www.sec.gov/ix?doc=/Archives/edgar/data/0001418091/000141809121000209/twtr-20210930.htm>

¹¹⁹ SEC Form 10-K, Twitter, Inc., Fiscal Year ended December 31, 2021, available at <https://www.sec.gov/Archives/edgar/data/0001418091/000141809122000029/twtr-20211231.htm>

¹²⁰ See Exhibit 3; see also “Partly Cloudy: The start of a journey into the cloud - Twitter Blog.” 8 Apr. 2019, https://blog.twitter.com/engineering/en_us/topics/infrastructure/2019/the-start-of-a-journey-into-the-cloud.

¹²¹ *Id.*

¹²² SEC Form 10-K, Twitter, Inc., Fiscal Year ended December 31, 2021, available at <https://www.sec.gov/Archives/edgar/data/0001418091/000141809122000029/twtr-20211231.htm>

Twitter's #1 priority. But Mr. Agrawal's executive team did not even reply to Mudge's e-mail and they left critical privacy compliance initiatives out of the company's top five priorities.

- ii. Twitter internally knew that the vast majority of data in their systems, data that was supposed to be registered and tagged to enable appropriate privacy handling, was not compliant with privacy requirements under the FTC consent decree and international regulations. The executives also knew that compounding this problem was the speed at which more non-compliant data was being created - thereby growing the body of data for which Twitter was out of compliance.

b. "Our prioritization of the long-term health of our service may adversely impact our short-term operating results."

- i. Contrary to this assertion, Twitter's unwillingness to accept short-term costs is jeopardizing its future. Further, Twitter's bonus structure provided executives with cash incentives for quickly growing the user count, which could only happen if the company ignored or deprioritized privacy, security and platform integrity.
- ii. The 2011 FTC Consent Order and the 2020 FTC Draft Complaint both identified protection of sensitive user data as crucial problems to be addressed. But in the decade since then, things actually got meaningfully worse, with sensitive customer information like emails and phone numbers improperly used for marketing, simultaneously while the company negotiated a new settlement with the FTC in 2020 and 2021.
- iii. As directly stated to the Board by a member of the executive team, it was a conscious choice by Twitter to deprioritize health and integrity of the service in order to direct resources towards mDAU growth. Twitter product release managers were given the authority to override security and privacy issues when choosing whether to make changes or ship a product, and in fact were encouraged to do so. This was reported up to Mudge numerous times, but his efforts to stop it were rebuffed. A good example is the Fleets service,¹²³ which supposedly permitted sending Tweets that would automatically disappear (similar to Snapchat). The Fleets product avoided undergoing security and privacy reviews before

¹²³ https://blog.twitter.com/en_us/topics/product/2021/goodbye-fleets

launch. When serious issues were identified at the last minute, Fleets was launched without addressing them. Shortly after the product launch security and privacy issues were found in the service causing teams to scramble to patch and fix overlooked issues when they impacted real users.¹²⁴

- c. **Omission:** Twitter has made no disclosure at all about the material information security problem that was the basis for the original 2011 FTC complaint, namely (1) the company's inability to limit employee access control to its production systems and (2) its highly anomalous practice of permitting engineers to build, test and deploy new code directly in live-production systems. This omission is material as a reasonable investor would want to know about Twitter's lack of basic engineering hygiene and security exposure.

117. **Misrepresenting the 2020 Hack:** Following the July 2020 hack by teenagers, Twitter provided updates via unsigned blog entries.¹²⁵ Broadly speaking, Twitter drastically overstated the sophistication of the hack, and misrepresented the sophistication of its own defenses. But in several cases, the blog posts went beyond misleading, and constitute outright falsehoods:

¹²⁴<https://www.vice.com/en/article/n7vnab/psa-twitter-doesnt-automatically-delete-your-fleets-after-24-hours>

¹²⁵ "An update on our security incident - Twitter Blog." 30 Jul. 2020, https://blog.twitter.com/en_us/topics/company/2020/an-update-on-our-security-incident.

Further Redacted for Congress



Report government and corporate lawbreaking.
Without breaking the law.

the world that help with account support. Our teams use proprietary tools to help with a variety of support issues as well as to review content in line with The Twitter Rules (<http://twitter.com/rules>) and respond to reports. Access to these tools is strictly limited and is only granted for valid business reasons. We have zero tolerance for misuse of credentials or tools, actively monitor for misuse, regularly audit permissions, and take immediate action if anyone accesses account information without a valid business reason. While these tools, controls, and processes are constantly being updated and improved, we are taking a hard look at how we can make them even more sophisticated.

Since the attack, we've significantly limited access to our internal tools and systems to ensure ongoing account security while we complete our investigation. As a result, some features (namely, accessing the Your

We're always investing in increased security protocols, techniques and mechanisms – it's how we work to stay ahead of threats as they evolve. Going forward, we're accelerating several of our pre-existing security workstreams and improvements to our tools. We are also improving our methods for detecting and preventing inappropriate access to our internal systems and prioritizing security work across many of our teams. We will continue to organize ongoing company-wide phishing exercises throughout the year.

118. In particular, based on evidence provided in this disclosure and interpreting these statements the way an investor or user would read them, it is not true that:

- a. “[a]ccess to these tools is strictly limited”
- b. “[w]e have zero tolerance for misuse of credentials or tools”
- c. “[w]e ... regularly audit permissions”
- d. “[w]e...take immediate action if anyone access account information without a valid business reason”
- e. “These tools, controls and processes are constantly being updated and improved”

-
- f. “Since the attack, we’ve significantly limited access to our internal tools and systems”
 - g. “We’re always investing in increased security protocols, techniques and mechanisms”
 - h. “We are also improving our methods for detecting and preventing inappropriate access to our internal systems”
 - i. “We will continue to organize ongoing company-wide phishing exercises throughout the year.” (No phishing exercise was ever reported to Mudge as Security Lead, nor were any new policies or practices implemented based on any phishing exercise; on information and belief, the company never organized a phishing exercise as described here.)
119. **False assurances on security:** A September 24, 2020 blog post by Parag Agrawal and Damien Kiernan also included multiple false assertions:¹²⁶

Our continued work to keep Twitter secure

By
[Parag Agrawal](#)
and
[Damien](#)
Thursday, 24 September 2020

¹²⁶ "Our continued work to keep Twitter secure." 24 Sep. 2020, https://blog.twitter.com/en_us/topics/company/2020/our-continued-work-to-keep-twitter-secure.

Further Redacted for Congress



Report government and corporate lawbreaking.
Without breaking the law.

To further secure our internal tools from potential misuse, we have been strengthening the rigorous checks that team members with access must undergo. This also helps reduce the potential for an unauthorized person to get access to our systems. We have strict principles around who is allowed access to which tools and at what time, and require specific justifications for customer data to be accessed.

behavior on your account to help you keep it secure, we have internal detection and monitoring tools that help alert us of unusual behavior or possible unauthorized attempts to access our internal tools. These tools are constantly being improved, even since the July incident, to include things like expanding our detection and response efforts to include suspicious authentication and access activity.

Our teams have also been investing in additional penetration testing and scenario planning to help secure Twitter from a range of possible threats, including in the context of the upcoming 2020 US elections. Specifically, over a five month period from March 1 to August 1, Twitter's cross-functional elections team conducted tabletop exercises internally on specific election scenarios. Some of the topics included:

courses, we've also enhanced training content on secure coding, threat modeling, privacy impact assessments, and privacy by design (https://blog.twitter.com/en_us/topics/company/2019/privacy_data_protection.html) so privacy is integrated into everything we design and build by default.

Finally, we continue to invest in and scale the processes in place to review products for security and privacy concerns before they launch. If a project could have significant privacy impacts, we conduct a detailed impact assessment to make sure we're taking appropriate measures before we launch it. We've significantly increased the number of privacy reviews and impact assessments the past few years. Specifically, in

120. In particular, the evidence provided in this disclosure demonstrates that many of these statements are simply false:

-
- a. Twitter did not have “rigorous checks that team members with access must undergo”. Twitter did not perform meaningful vetting for employees with privileged access, nor were team members with access to production systems and data evaluated differently from other employees;
 - b. Twitter did not have “strict principles around who is allowed access to which tools and at what time, and require specific justifications for customer data to be accessed”; according to expert quantification and analysis in January 2022, over half of Twitter’s 8,000-person staff was authorized to access the live production environment and sensitive user data. Twitter lacked the ability to know who accessed systems or data or what they did with it in much of their environment. The ranks of those given access to Twitter's production environment, systems, and services which contained user data continued to increase through the end of 2021. For the vast majority of methods that staff and contractors accessed sensitive data, including user data, there were no time-based limits on access; such limits applied only to a small subset of tools.¹²⁷

¹²⁷ See Exhibit 1, p. 21

At the beginning of 2021, 46% of all FTEs had privileged access to production systems and data. By Q4 2021 this number was 51% of employees. Twitter has grown meaningfully in its number of employees. The percentage of employees with privileged access has increased on top of this.



Access to Twitter's Datacenter Production Environment¹¹

- i. Dec 2020 46% of employees (2,763 out of 5917)
- ii. Dec 2021 51% of employees (3,995 out of 7714)
- (*) The dip was an unintended (internal) incident

- c. Twitter did not have “internal detection and monitoring tools that help alert us of unusual behavior or possible unauthorized attempts to access our internal tools”; on information and belief they had no meaningful detection or monitoring tools;¹²⁸
- d. It is not true that Twitter “teams have also been investing in additional penetration testing and scenario planning”; Twitter had neither an internal red team nor a third party engaged to do meaningful internal penetration testing within the InfoSec organization, at least none ever reported to Mudge as Security Lead. An excerpt of Mudge’s analysis:¹²⁹

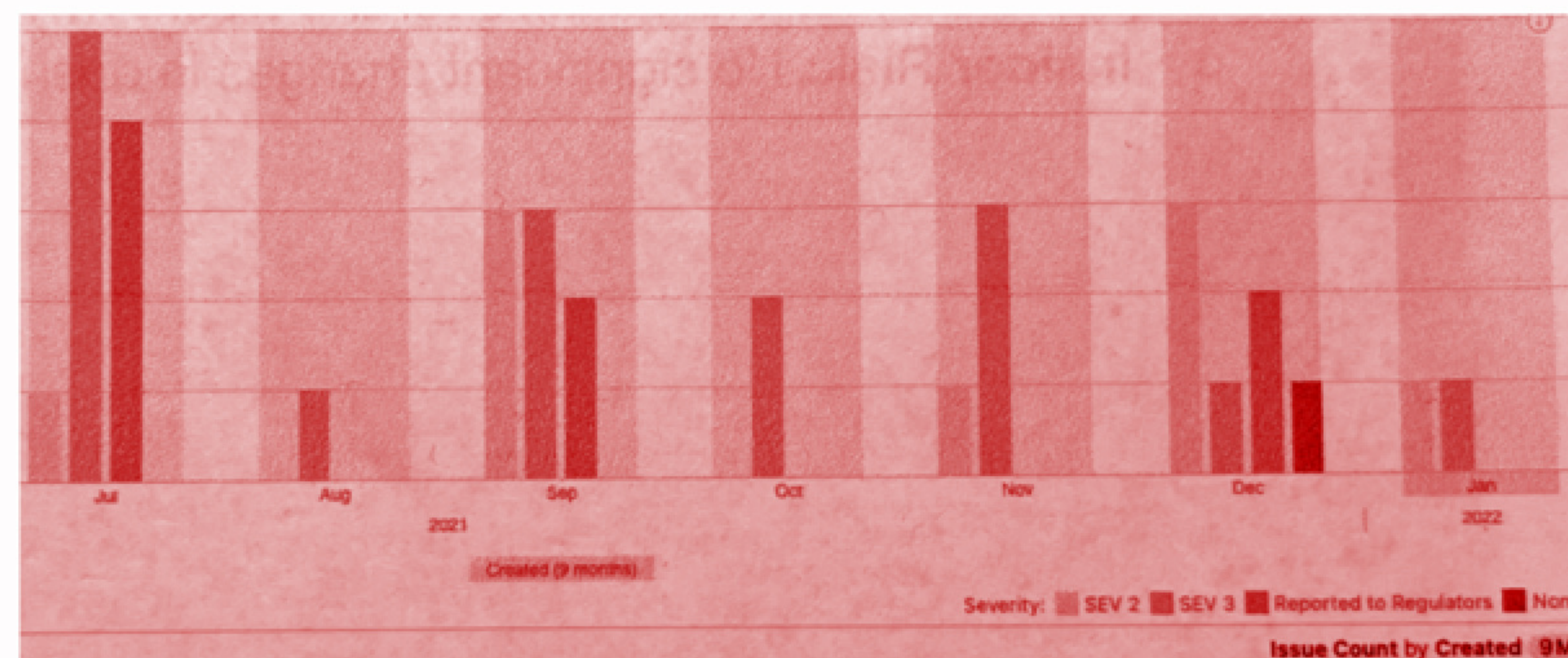
¹²⁸ *Id.*

¹²⁹ *Id.*, p. 16

Further Redacted for Congress

Twitter has an unacceptable, and near continuous, number of security and privacy incidents. I estimate there were more than 50 Incidents in 2021¹⁵; approximately an incident per week. Based on my professional experience, peer companies do not have this magnitude or volume of incidents.

H2 2021 had 11 Incidents that were required to be reported to regulators, 5 of which happened in Q4.



The Incidents were predominantly related to areas where Twitter has systemic, long lived, problems: 'Access Control' and 'Security Configuration and Bugs'. Together these problems account for more than 80% of the Incidents.

¹⁵ My notes capture 48 incidents in the period of April - November 2021. 50+ is an extrapolation to include January-March, and December, at the same Incident rate, as my data sources were taken away before this document could be completed.

- e. Twitter had not “enhanced training content on secure coding, threat modeling, privacy impact assessments, and privacy by design so privacy is integrated into everything we design and build by default.” Twitter neither followed a mature Software Development Life Cycle nor had one been rolled out across engineering and existing projects and programs. If and when the InfoSec or Privacy teams learned about a project whether from the project manager or through the grapevine, security and privacy reviews often had to be forced into projects. It was further noted that very few of the products submitted for security / privacy review included threat modeling on how the products could

be abused by bad actors. The omission of threat modeling, indicated that engineers had not considered the question of vulnerabilities;¹³⁰

Twitter does not have an industry-appropriate Software Development Life Cycle (SDLC) and Twitter has thus far operated largely without one at all. If it were not for an FTC consent decree, it is possible that Twitter would not be working to put together and deploy an SDLC. This is very atypical in the industry and is a significant risk to the company.

Due to this deficiency, it is inappropriate to label any SDLC progress as “Compliant” to the Committee, as was done in the Information Security documents sent to the Risk Committee. Doing so misrepresents Twitter’s situation as it would be seen by regulators and auditors.

- f. “If a project could have significant privacy impacts, we conduct a detailed impact assessment to make sure we’re taking appropriate measures before we launch it”; but until 2021 Twitter did not employ trained Privacy Engineers. Instead Twitter relied on regular engineers to implement privacy measures without the benefit of guidance from senior Privacy Engineering leadership or people with appropriate domain expertise.

121. Securities Violations: Based on the foregoing, it is likely that Twitter and Mr. Agrawal’s actions constitute violations of multiple SEC rules and regulations:

- a. 15 U.S. Code § 7142 on corporate responsibility for SEC reporting;
- b. 15 U.S. Code § 7262 on management assessment of internal controls.
- c. 18 U.S. Code § 1350(c):

whoever...willfully certifies any statement ...[filed for a public company with the SEC] knowing that the periodic report accompanying the statement does not comport with all the requirements set forth in this section shall be fined not more than \$5,000,000, or imprisoned not more than 20 years, or both;

¹³⁰ *Id.*, p. 15



Report government and corporate lawbreaking.
Without breaking the law.

d. Twitter CEO Parag Agrawal signed the company’s 2021 Annual Report:¹³¹

Pursuant to the requirements of the Securities Exchange Act of 1934, this report has been signed below by the following persons on behalf of the registrant and in the capacities and on the dates indicated:		
Signature	Title	Date
<u>/s/ Parag Agrawal</u> Parag Agrawal	Chief Executive Officer and Director (Principal Executive Officer)	February 16, 2022

[Disclosure continues next page]

¹³¹ "twtr-20211231 - SEC.gov."
<https://www.sec.gov/Archives/edgar/data/0001418091/000141809122000029/twtr-20211231.htm>.

VIII. Conclusion

116. For the foregoing reasons, please open an investigation into legal violations by Twitter, Inc.
117. As a senior executive, Mudge was awarded Twitter stock, for which he previously created and has followed without deviation an Automatic Securities Disposition Plan pursuant to SEC rules codified at 17 C.F.R. § 240.10b5-1(c).
118. Whistleblower Aid is a non-profit legal organization that helps workers report their concerns about violations of the law safely, lawfully, and responsibly. We respectfully request the SEC's assistance ensuring that our client never faces retaliation.

Sincerely,



John N. Tye, Founder & Chief Disclosure Officer
Whistleblower Aid

[Redacted]

[Redacted]

[Redacted]



Andrew P. Bakaj, Senior Counsel
Whistleblower Aid

[Redacted]

[Redacted]



Report government and corporate lawbreaking.
Without breaking the law.

A handwritten signature in black ink that reads "Kyle Gardiner".

Kyle Gardiner, Associate Counsel
Whistleblower Aid

[Redacted]

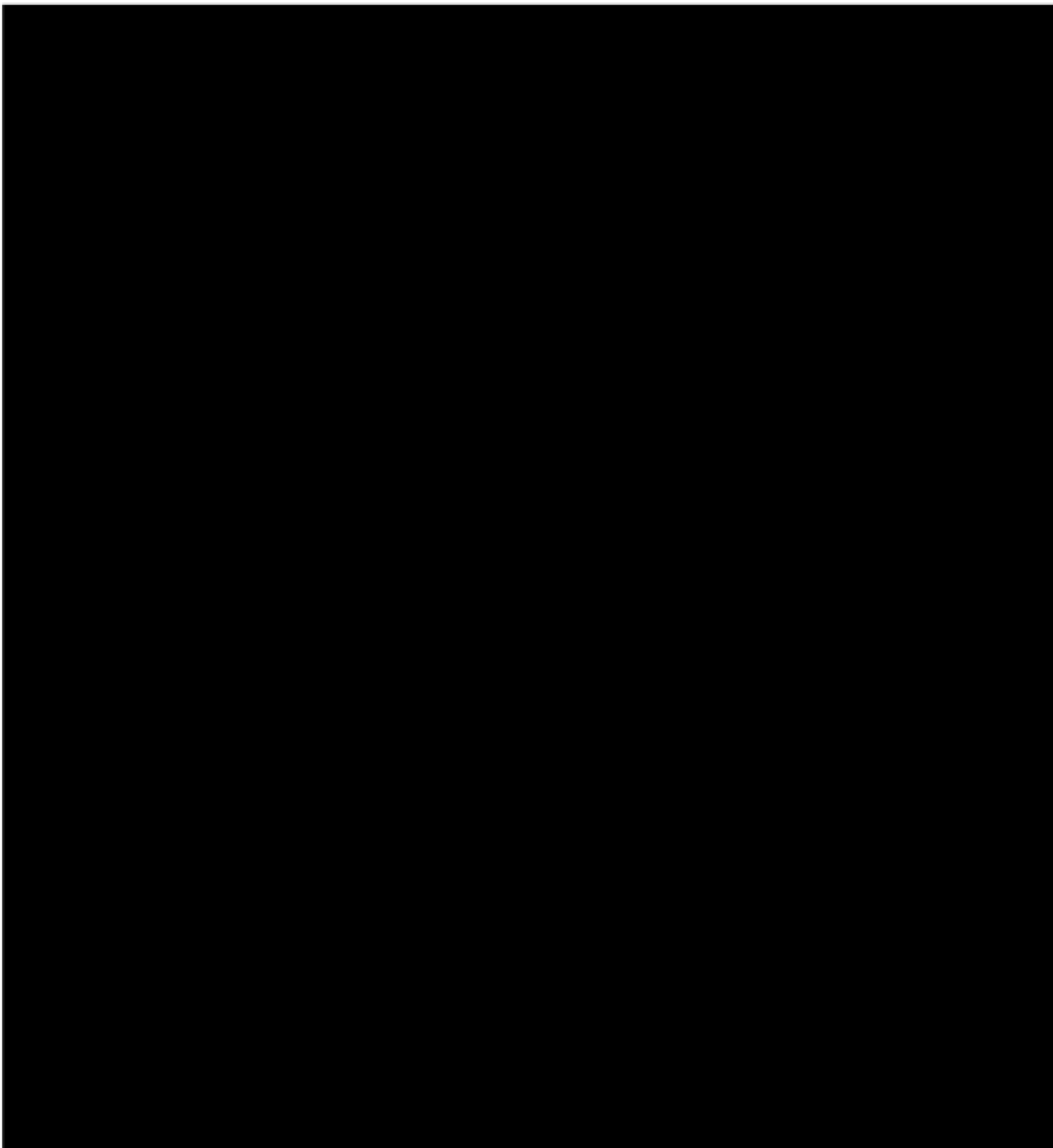
[Redacted]

Whistleblower Aid

[Redacted]

[Redacted]

Washington, DC [Redacted]



Exhibits Enclosed:

Exhibit Label	Exhibit Link
1	20220214_Twitter Q4 2021 Risk Committee Issues_redacted.pdf
2	2021xxxx [Redacted] Privileged Confidential Twitter_Dis-Mis-information-assessment_2021_Q1_redacted
3	202103xx_2021 Q1 board voice.pdf
4	Exhibit 4_202112xx_DRAFT_2021 Q4 Privacy and Data Protection Report
5	20220118_Email_to_Parag_following_Omid_meeting_Jan18_2022_redacted

Further Redacted for Congress



Report government and corporate lawbreaking.
Without breaking the law.

6	20220202_Mudge_2_ [REDACTED] Feb 2 2022_redacted
7	[REDACTED]
8	202111xx_Protect 2022 Strategy
9	20211217_Gmail_action items from Mudge to Parag 1 of 2_redacted
10	[REDACTED]_redacted
11	20210112_Parag_e-mails_Memorandum for the record_Jan_12_2022_redacted
13	20220118_Gmail - Re_Privileged and Confidential - Priority Meeting Request_Jan_18_redacted
14	Exhibit 14_202112xx_DRAFT NOT DELIVERED_2021 Q4 Information Security Report
15	[REDACTED]_redacted
16	[REDACTED]_redacted
17	202112xx_Copy of access_control_actual.jpg202112xx_Access_control_actual
19	[REDACTED]
20	[REDACTED]
21	2021XXXX_Snapshot of data center security system deficiencies_redacted
22	[REDACTED]
23	202106xx_Q2 board meeting - Mudge Board voice.pdf
24	[REDACTED]_redacted
25	20220127_email draft Marianne 6 weeks v 1 [was_Re_Follow-Up]_redacted

Further Redacted for Congress



Report government and corporate lawbreaking.
Without breaking the law.

26	<div>redacted</div> <div>redacted</div>
27	20220119_email_Follow-Up_redacted
28	<div>redacted</div> <div>redacted</div>
29	<div>redacted</div> <div>redacted</div>
30	<div>redacted</div> <div>redacted</div>
31	<div>redacted</div> <div>redacted</div>
32	<div>redacted</div> <div>redacted</div>
33	<div>redacted</div> <div>redacted</div>
34	20220214_Fwd_Q4_2021_Risk_Committee_Issues_redacted
35	<div>redacted</div> <div>redacted</div>
36	<div>redacted</div> <div>redacted</div>
37	<div>redacted</div> <div>redacted</div>
38	202201xx_Q4_Risk_Committee_Information_Security_Corrective_Document_and_Timeline.docx_redacted
39	Excerpt on Licensing
40	20211001_Mudge_Twitter_Nigeria_Notes_redacted
41	<div>redacted</div>
42	202112xx_NEW_Corrected_Risk_Information.docx

— END OF DISCLOSURE —

— PROTECTED & SENSITIVE WHISTLEBLOWER DISCLOSURE —