**UNITED STATES DISTRICT COURT**
**SOUTHERN DISTRICT OF NEW YORK**

| | |
|---|---|
| MOJO NICHOLS, SUSAN BREWSTER, DUANE DEA, MARYANNE DERACLEO, KAREN KELLY, REBECCA RICHARDS, JENNIFER SELLERS, and STACY SPENCER,<br><br>Individually and on Behalf of All Others Similarly Situated,<br><br>Plaintiffs,<br><br>v.<br><br>NOOM, INC., ARTEM PETAKOV, and JOHN DOES 1 TO 5,<br><br>Defendants. | Case No. 20 Civ. 3677 (LGS) (KHP)<br><br>**DECLARATION OF DOUGLAS E. FORREST** |

DOUGLAS E. FORREST, of full age, hereby declares under penalty of perjury as follows:

1. I am the Vice President, eDiscovery Analytics & Strategy, at International Litigation Services ("ILS"), in Irvine, California. I have been retained as a consultant for Plaintiffs in this action. I am fully familiar with the facts contained herein based upon my personal knowledge.
2. This is my second declaration in this matter; my prior declaration was dated October 8, 2020, ECF No. 65-3.
3. I provide this declaration in support of Plaintiffs' "extenuating circumstances" motion to explain why (i) collecting linked documents contained in Defendant Noom's email, Google drive documents, and Slack messages is neither difficult nor expensive, and (ii) only Defendant Noom has the ability to provide the metadata to link separately produced documents to their parent document.

4. In preparation for this declaration, I reviewed, *inter alia*:

   a. ECF No. 214 – Plaintiffs' extenuating circumstances letter motion;

   b. ECF No. 226 – Defendants' opposition ("Noom's Opposition to Plaintiffs' Letter Motion");

   c. ECF No. 226-2 – Declaration of Stephen Gresch ("Gresch Declaration"); and

   d. ECF No. 225 – the parties' March 3, 2021 joint pre-conference submission to the Honorable Katharine H. Parker ("March 3 Joint Pre-Conference Letter").

## I. QUALIFICATIONS

5. I am a graduate of Stanford Law School, where I was a Note Editor of the Law Review. I was admitted to the bar in 1977, and, after practicing law at Breed, Abbott & Morgan and Cravath, Swaine & Moore, where I worked extensively on discovery matters, I developed significant expertise in computer technology design and implementation, both generally and with respect to litigation support and e-discovery.

6. As an attorney at Cravath, I relied on Aquarius, the first large-scale implementation of computerized litigation support, which was implemented on the IBM antitrust cases.

7. As Director of Litigation Services at Legal Information Technology, Inc., I was instrumental in introducing imaging, coding and search technology for discovery to law firms, and pioneered in integrating imaging with legacy search systems such as BRS.

8. As a systems architect, application designer and programmer, I created case management, litigation support and document repository systems (including WIDE, and LIT CaseWorks for Lotus Notes), SaaS (Software as a Service) knowledge management applications (including LexisNexis Total Alerts and LexisNexis Clipper), and e-discovery and production operation systems. As the Chief Technology Officer of Ozmosys, I

designed Knowledge Management (KM) based systems, including Alerts Manager, Clipper, and Open Alerts.

9. At ILS, I direct e-discovery analytics and strategy, and design, direct and oversee programming, *inter alia*, to collect and process ESI from cloud-based sources.

10. I have advised, consulted or acted as a declarant or affiant with respect to ESI in many cases, including:

    *a.* *In re: Ethiopian Airlines Flight ET 302*, Lead Case: 1:19-cv-02170 (N.D. Ill.) (Boeing 737 Max Crashes);

    b. *In Re: 3M Combat Arms Earplug Product Liability Litigation*, MDL No. 2885 (N.D. Fla.) ("3M");

    c. *In Re: Volkswagen "Clean Diesel" Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2672 (N.D. Cal.) ("Volkswagen");

    *d.* *In Re: Intel Corp, CPU Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2828 (D. Or.);

    *e.* *In Re: Chrysler-Dodge-Jeep EcoDiesel Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2777 (N.D. Cal.);

    *f.* *Lafferty v. Alex Jones* (Conn. Super. Ct.) (Sandy Hook parents);

    *g.* *Soto v. Bushmaster Firearms International* (Conn. Super. Ct.) (Sandy Hook parents);

    *h.* *Small v. UMC*, Case No. 2:13-cv-0298-APG-PAL (D. Nev.) (widely commented on spoliation case, *see, e.g.*, https://e-discoveryteam.com/2018/09/16/the-importance-of-witness-interviews-what-happens-in-vegas-shouldnt-stay-in-vegas/ (last accessed on March 8, 2021)

    i. *In Re: Takata Airbags Product Liability Litigation*, MDL No. 2599 (S.D. Fla.) (exploding airbags);

    j. *In Re: Testosterone Replacement Therapy Products Liability Litigation*, MDL No. 2545 (N.D. Ill.);

    k. *In Re JCCP 4771, Zoloft Birth Defect Cases*, (Cal. Super. Ct.);

    l. *Da Silva Moore v. Publicis Groupe & MSL Group,* No. 11 Civ. 1279 (ALC)(AJP) (S.D.N.Y.) (seminal TAR case);

m. *In re Actos (Pioglitazone) Products Liability Litigation*, MDL No. 6:11-md-2299 (W.D. La.) (initial award of $9 billion for spoliation); and

n. *In Re Xarelto (Rivaroxaban) Products Liability Litigation*, MDL No. 2592 (E.D. La.).

11. I have served as a speaker or panelist on many CLE webinars including West LegalEdcenter's *Top ESI Mistakes Made in Mass Tort Disputes* (Sept. 14, 2017), *Lessons Learned from Recent eDiscovery Disasters* (Feb. 26, 2018), and *The State of E-Discovery in 2018: Analysis & Review* (Sept. 27, 2018), and HarrisMartin's MDL Conferences on *The Mass Tort Litigation Landscape – A Critical Analysis Agenda* (Sept. 26, 2018), *Current Mass Torts from E-Discovery Through Exit Strategies – Navigating "Game-Changing" Dynamics* (Nov. 26, 2018), and *An Analysis of Today's Mass Tort Landscape Agenda* (Mar. 27, 2019).

12. I am a member of the Sedona Conference Working Group 1: Electronic Document Retention and Production, and a member of EDRM's newly formed Analytics & Machine Learning Project.

## II. DISCUSSION

13. Noom asserts that (i) "there is no practical means of producing all hyperlinks in Gmail or Google drive documents," (ii) MetaSpike's Forensic Evidence Collector ("FEC") "tool is completely unworkable for Google drive and **does not even offer the capability of downloading links from Google drive documents**," and (iii) "the only way to satisfy Plaintiffs' request would be to re-perform Noom's entire collection and attempt to manually download every link in the collection." Noom's Opposition to Plaintiffs' Letter Motion at 3 (emphasis in original). All three of Noom's assertions are incorrect and only serve to muddy the waters.

14. Noom's emphasized assertion that FEC does not offer the capability to download links from Google drive documents is simply irrelevant. While Plaintiffs had suggested FEC as a superior tool to collect **Gmail** as, unlike Google Vault, it can collect Google drive documents attached to emails via links, Plaintiffs have not suggested that it be used to collect Google drive documents attached to other Google drive documents via links. The straightforward fix for collecting such documents is discussed in paragraphs 23 and 24, *infra*.

15. Noom's assertions that there is no "practical means" to collect "all hyperlinks in Gmail or Google drive documents" and that such links must therefore be "manually" collected are both overstated and fundamentally incorrect.

16. Overstated for two reasons. First, Plaintiffs are not requesting that Noom collect and produce "all" links. Plaintiffs can access *publicly accessible* ESI and documents linked to by Gmail and Google drive documents and accordingly have not requested that Noom do so. What Plaintiffs *cannot* access are documents in Noom's private Google drive which are attached to Gmail or other Google drive documents via internal links, and are accessible only by Noom.

17. Moreover, even if Noom separately and independently produces Google drive documents that at some point were also attached to Gmail and/or other Google drive documents via links, there will be no metadata linking the "child" to the "parent" document, even though those linkages are substantive evidence in and of themselves, and Plaintiffs will be unable to determine or establish those linkages.

18. Furthermore, by not collecting Google drive documents attached via links, and searching those Google drive documents as family members of their parent documents, parents

such as basic cover memorandums, *e.g.*, "see attached," which do not contain search terms themselves will not be collected, concealing strong direct ties between the Google drive documents attached via links and the recipients of the cover memorandums.

19. Second, Noom's assertions are also overstated because Noom's claim that "there is no practical means of producing all hyperlinks **in Gmail**" (emphasis added) has been directly contradicted by Plaintiffs' prior suggestion of FEC, which is a widely used[1] tool that can collect Google drive documents **attached to Gmail** via links, and moreover, is a tool upon which **Noom itself, by its own admission, is relying upon and using to collect other Google workspace ESI**.[2] March 3 Joint Pre-Conference Letter at 3, Noom's response to Plaintiffs' Agenda Item 1(b).

20. Noom's assertions that there are no "practical means" to collect "all hyperlinks in . . . Google drive documents" and that such links must therefore be "manually" collected are silly. It is no different than claiming that there are no "practical means" to travel from California to New York because, by horseback, it is too far and would take far too long.

21. Google drive is a cloud document service provided by Google, a company which was and is built on the collection of documents and linked documents from URLs.

---

[1] Prominent FEC users, including users in government and law enforcement, and in prominent firms in the Am Law 100, the accounting Big 3, and litigation support and forensic services, include Paul Weiss, PriceWaterhouseCoopers, the Office of the California Attorney General, FTI Consulting, Deloitte, Grant Thornton, Clifford Chance, Stroz Friedberg, Winston & Strawn, the Federal Trade Commission, the Norwegian Tax Administration, and the Australian Federal Police. (Battle-Tested Software, https://www.metaspike.com/forensic-email-collector/#customers, accessed on March 7, 2021).

[2] Despite Noom's adoption of FEC to collect Google Calendar ESI, the Gresch Declaration denigrates FEC, stating "TLS' testing at the time of the collection showed the results to be significantly problematic and requiring additional supervision and quality control checks." Gresch Decl. ¶ 8. However, it is impossible to credit, or even evaluate, such claims in the absence of any specific supporting facts or details, and none are given. Moreover, given the acceptance of FEC by prominent and demanding organizations such as those set out in the prior footnote, and ILS's own experience with the tool, perhaps any issues that Mr. Gresch or his colleagues at TransPerfect encountered were not fairly attributable to FEC itself.

22. Google provides a well-known and well-established API (Application Programming Interface) to enable its clients to interface with Google drive documents and folders by creating programs that "read, write, and sync files in Google Drive." Google Drive API, (https://developers.google.com/drive, accessed on March 7, 2021).

23. Creating a program to utilize that API here—to extract links to Google drive documents from other Google drive documents, retrieve those documents, and produce them as attachments—is not rocket science, and is well within the capabilities of a first or second year computer sciences student, let alone the capabilities of the alpha programmers at TransPerfect and its Digital Reef eDiscovery division,[3] which is described as "a leader in global e-discovery and legal technology" (Digital Reef, https://www.digitalreefinc.com/, accessed on March 7, 2021),[4] or even at Noom itself.[5]

24. The basic logic and execution of such a program is simple:

    a. Parse parent documents for link URLs to Google drive documents.

---

[3] The TransPerfect Legal Solutions ("TLS") web page (https://www.transperfectlegal.com/technology/digital_reef, accessed on March 7, 2021) points to the Digital Reef website (https://www.digitalreefinc.com/) under the heading "Visit our dedicated site for more detailed information."

[4] TransPerfect's website currently lists 39 open Software Development positions. TransPerfect, *Search for Jobs* (https://transperfect.wd5.myworkdayjobs.com/transperfect/5/refreshFacet/318c8bb6f553100021d223d9780d30be, accessed on March 7, 2021). One Digital Reef job offering posted just 7 days ago on LinkedIn has already garnered 79 applications for a "passionate mid-level to Senior Software Engineer to join our top-notch engineering team in our Maynard, MA office," with required experience and qualifications including, at a minimum, a BS in Computer Science or a related discipline, and 5 or more years in software development. LinkedIn, *Digital Reef Software Engineer* (https://www.linkedin.com/jobs/view/2376164747/?refId=5io8rSOVqQ5Uj3w%2FhqND2w%3D%3D, accessed on March 7, 2021).

[5] There are currently 12 open positions under software development listed for Noom on the Indeed.com jobs site. *See* Indeed, *Noom Jobs* (https://www.indeed.com/cmp/Noom/jobs?c=techsoftware, accessed on March 7, 2021).

b. Parse each link URL to extract the Google drive "file id" of each linked document.

c. Retrieve each linked document using the Google Drive API methods "file.export," for downloading Google drive documents (Google Docs, Sheets, Slides, and so on), or "file.get," for other document types, *e.g.*, Word, etc., stored in Google drive. *See* Google Drive API > v3, *Download Files* (https://developers.google.com/drive/api/v3/manage-downloads, accessed on March 7, 2021).

d. Associate each downloaded document with its parent.

Each of these steps is relatively simple to execute, and the same program logic and most of the code would also be applicable even to parent documents which are Slack documents or Gmails.

25. I estimate the time required for a reasonably experienced programmer to develop such a program to be less than one to two weeks, at a fraction of the cost put forth by Noom.

26. Commercial solutions to provide equivalent functionality may also exist. *See, e.g.*, Onna Help Center, *How to collect from G Suite* (https://support.onna.com/en/articles/2440761-how-to-collect-from-g-suite, accessed on March 7, 2021) ("Links will be collected from individual drives and team drives."); *see also* Hanzo, *Hanzo Hold's New Follow-the-Link Capability Preserves Full Text of Linked Documents in Google Drive* (https://www.hanzo.co/blog/hanzo-holds-new-follow-the-link-capability-preserves-full-text-of-linked-documents-in-google-drive, accessed on March 7, 2021) (Slack document links).

27. The Gresch Declaration also states:

> There is an inherent problem in associating hyperlinked documents as 'attachments' in the course of reviewing and producing documents. Traditionally, a 'parent item' is considered an email or electronic document that has an additional file embedded directly in it, for example an email attachment or an embedded image in a presentation slide. FEC treats the hyperlinked documents the same way, even though the hyperlink is a reference to a document that exists independently outside of the 'parent' item.

Gresch Decl. ¶ 9.

28. However, "traditionally" as used in the Gresch Declaration harkens back to the world before cloud platforms such as Google drive had a significant corporate market share, and attachments were always physical.

29. Times and corporate practices have changed, especially for cloud-based services such as Noom's weight-loss service. The way that Noom apparently and routinely uses links to Google drive documents instead of physically attaching them is just such a change, and requires a corresponding recognition that a corporation's decision to adapt such technology does not exempt it from its discovery obligations.[6]

---

[6] A producing party may not "shield itself from discovery by utilizing a system of recordkeeping which conceals rather than discloses relevant records, or makes it unduly difficult to identify or locate them, thus rendering the production of documents an excessively burdensome and costly expenditure." *Wesley v. Muhammad*, No. 05 Civ. 5833 (GEL) (MHD), 2008 WL 4386871, at *5 (S.D.N.Y. 2008) (quotation marks and citation omitted); *see also Kozlowski v. Sears Roebuck & Co.*, 73 F.R.D. 73, 76 (D. Mass. 1976) (same and observing that "[t]o allow a defendant whose business generates massive records to frustrate discovery by creating an inadequate filing system, and then claiming undue burden, would defeat the purposes of the discovery rules."); *Pom Wonderful LLC v. Coca-Cola Co.*, No. 08 Civ. 86237 (SJO) (FMO), 2009 WL 10655335, at *3 (C.D. Cal. Nov. 30, 2009) (rejecting argument that defendant "'has no method to automatically re-link emails with their alleged "missing" attachments[]' and that requiring it to do so would 'employ a tedious manual process,'" because a defendant "cannot seek to preclude plaintiff from pursuing discovery based on a record-keeping system that is plainly inadequate"); *Lou v. Ma Labs, Inc.*, No. 12 Civ. 5409 (WHA) (NC), 2013 WL 12328278, at *2 (N.D. Cal. Mar. 28, 2013) ("The fact that a corporation has an unwieldy record keeping system which requires it to incur heavy expenditures of time and effort to produce requested documents is an insufficient reason to prevent disclosure of otherwise discoverable information.") (quotation marks and citation omitted).

30. Similarly, contrary to Noom's claim in footnote 3 of its Opposition to Plaintiffs' Letter Motion, Plaintiffs' inclusion of document stubs in the definition of family groups in the ESI protocol (ECF No. 146 at 13) was intended to cover this exact situation, with "stub" being used in a generic sense to mean any circumstances where direct inclusion of an internal document was bypassed for a link to its internal location.

31. The Gresch Declaration's language in paragraph 9 about a linked document existing "independently outside of the 'parent' item" carries an implication that such links are unreliable because a linked to document may have changed after a document which links to it was created. But that is exactly why programs such as FEC include options to collect the versions of linked-to internal documents that existed as of the time an email was sent, obviating concerns of later alterations.

32. The Gersch Declaration at paragraph 10 raises concerns about the ability of deduplicating functionality to miscalculate when evaluating FEC-collected emails which properly treat links to internal documents as attachments against identical emails which were not so collected, as the MD5 hash values would differ.

33. Two points in response. First, the family groups of such emails would not in fact be identical and shouldn't be treated as if they were. Second, if a need arises to calculate the MD5 hash value of a parent email containing Google drive documents attached via links, the MD5 value of the parent email sans family can be readily calculated for comparison purposes.

I declare under penalty of perjury that the foregoing is true and correct.

Dated: March 8, 2021

/s/ Douglas Forrest
Douglas Forrest