[Counsel on signature page]

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN JOSE DIVISION

CONCORD MUSIC GROUP, INC., ET AL., Plaintiffs, v. ANTHROPIC PBC, Defendant.	Case No. 5:24-cv-03811-EKL-SVK **[PROPOSED] PROTOCOL FOR PRODUCING DOCUMENTS AND ELECTRONICALLY STORED INFORMATION** Hon. Eumi K. Lee Magistrate Judge Susan van Keulen

[PROPOSED] PROTOCOL FOR PRODUCING DOCUMENTS AND ELECTRONICALLY STORED INFORMATION

Plaintiffs Concord Music Group, Inc., Capitol CMG, Inc., Universal Music Corp., Songs of Universal, Inc., Universal Music – MGB NA LLC, Polygram Publishing, Inc., Universal Music – Z Tunes LLC, and ABKCO Music, Inc. (collectively, "Publishers") and Defendant Anthropic PBC ("Anthropic") (together, "the Parties") hereby acknowledge, agree, and stipulate as follows:

PURPOSE

1. The Protocol for Producing Documents and Electronically Stored Information ("ESI Protocol") shall govern the production of documents and electronically stored information ("ESI") by the Parties in the above-captioned litigation. The ESI Protocol shall govern all productions made by a Party, including, without limitation, any production of material received from a third party who is subpoenaed in this action. The Parties each reserve the right to seek exceptions, amendments, or modifications to this Order by agreement, or from the Court for good cause shown.

DEFINITIONS

2. "Document" is defined to be synonymous in meaning and equal in scope to the usage of the term in Rule 34(a) of the Federal Rules of Civil Procedure and includes ESI existing in any medium from which information can be translated into reasonably usable form, including

but not limited to email and attachments, word processing documents, spreadsheets, graphics, presentations, images, text files, databases, instant messages, transaction logs, audio and video files, voicemail, internet data, computer logs, text messages, and backup materials. The term "Document(s)" shall include Hard Copy Documents, Electronic Documents, and Electronically Stored Information (ESI) as defined herein.

3. "Electronic Document or Data" means Documents or Data existing in electronic form at the time of collection, including but not limited to email or other electronic communications, word processing files (*e.g.*, Microsoft Word), computer presentations (*e.g.*, PowerPoint slides), spreadsheets (*e.g.*, Excel), and image files (*e.g.*, PDF).

4. "Electronically stored information" or "ESI" is information that is stored electronically as files, documents, or other data on computers, servers, mobile devices, online repositories, disks, USB drives, tape or other real or virtualized devices or digital media.

5. "Hard Copy Document" means Documents existing in paper form at the time of collection.

6. "Hash Value" is a numerical identifier that can be determined from a file, a group of files, or a portion of a file, based on a standard mathematical algorithm that calculates a value for a given set of data, serving as a digital fingerprint, and representing the binary content of the data to assist in subsequently ensuring that data has not been modified and to facilitate duplicate identification. Unless otherwise specified, hash values shall be calculated using the MD5 hash algorithm.

7. "Inaccessible Data" is any relevant ESI not reasonably accessible within the meaning of Fed. R. Civ. P. 26(b)(2)(B). "Presumptively Inaccessible Data" includes the following data sources and document types:

a. Back-up tapes, cold or offline storage, or other long-term storage media, including any tapes or storage system that was created strictly for use as a data back-up or disaster recovery medium.

b. Temporary data stored in a computer's random-access memory (RAM), or other operating system files.

c. Deleted, fragmented, unallocated space or other data only accessible by forensics.

d. Online access data such as temporary internet files, history, cache, cookies, and the like.

e. Data remaining from systems no longer in use that is unintelligible on systems in use.

f. Ephemeral data automatically deleted by its nature or design by the application, settings, or operating systems.

8. "Load File" is an electronic file containing information identifying a set of paper scanned (static) images or processed ESI and indicating where individual pages or files belong together as documents, including attachments, and where each document begins and ends. Load Files also contain data relevant to individual Documents, including extracted and user-created Metadata, coded data, as well as OCR or Extracted Text. A load file linking corresponding images is used for productions of static images (*e.g.*, TIFFs).

9. "Metadata" is the term used to describe the structural information of a file that contains data about the file, as opposed to describing the content of a file.

a. "Native Format" means the file format associated with the original creating application and as collected from custodians. For example, the native format of an Excel workbook is an .xls or .xlsx file.

b. "Optical Character Recognition" or "OCR" means a technology process that captures text from an image for the purpose of creating an ancillary text file that can be associated with the image and searched in a database. OCR software evaluates scanned data for shapes it recognizes as letters or numerals.

c. "Searchable Text" means the native text extracted from an Electronic Document or, when extraction is infeasible, by Optical Character Recognition text ("OCR text") generated from a Hard Copy Document or electronic image.

PRESERVATION & SCOPE OF DISCOVERY

10. Presumptively Inaccessible Data. The Parties agree that there is no need to preserve

Presumptively Inaccessible Data.

11. Meet and Confer Obligation. The Parties agree to meet and confer in an attempt to reach an agreement on the scope of relevant and proportional discovery within the meaning of Rule 26(b)(1), including: date ranges; file types; custodians; search terms; noncustodial sources of ESI; and/or information located in applications or databases.

12. Technology Assisted Review ("TAR"). A Party who intends to use TAR will describe a TAR Protocol that will include: (a) the TAR system to be used; (b) the criteria to be used to identify the universe of documents to which TAR is to be applied (the "TAR Universe"); (c) its methodology for training the TAR model and identifying responsive documents; (d) the methodology for validation testing to be used; (e) the subject-matter expert who will oversee the implementation of the TAR Protocol; and (f) the process by which documents excluded as not conducive to categorization (e.g., multimedia files, primarily numerical spreadsheets, database files) will be reviewed for production.

13. No Responsiveness Presumption. The fact that a document is captured by a search pursuant to an agreed-upon protocol does not mean that such document is responsive to a discovery request, relevant to this litigation, or will be produced.

14. Noncustodial Repositories & Business Records Management. Documents and categories of documents that are relevant to this action and responsive to a Party's document requests, and that are regularly maintained in a known location, or in a location that is knowable upon reasonable inquiry of those with knowledge about a Party's document management, may be collected and reviewed for responsiveness and privilege without the use of search terms or other agreed-upon advanced search methodology (*e.g.*, analytics, predictive coding, technology-assisted-review).

15. Deduplicating Documents. The Parties may deduplicate globally (*i.e.*, across all custodians). This will result in the Producing Party needing to produce only a single copy of responsive duplicate ESI. The Parties shall deduplicate stand-alone documents against stand-alone documents and shall deduplicate top-level email documents against top-level email documents. Deduplication shall not break apart families.

16. Most Inclusive Email Productions. Email threads are email communications that contain prior or lesser-included email communications that also may exist separately in the Party's electronic files. A most inclusive email thread is one that contains all the prior or lesser-included emails, including attachments, for that branch of the email thread. The Parties agree that removal of wholly included, prior-in-time, or lesser-included versions from potential production will reduce all Parties' costs of document review, production, and litigation-support hosting. For the avoidance of doubt, only email messages for which the parent document and all attachments are contained in the more inclusive email message will be considered less inclusive email messages that need not be produced; if the later message contains different text (such as where the later message adds in-line comments to the body of the earlier message), or does not include an attachment that was part of the earlier message, the earlier message must be produced. Use of email threading may not serve to obscure whether a recipient received an attachment.

17. Production of Short Message Data. Electronic messages exchanged between users on communication software such as Microsoft Teams and Slack shall be produced in a searchable format that preserves the conversational relationship and presentational features of the original messages, such as emojis, images, video files, and animations. Electronic messages must not be converted to unitized files that contain less than a 24-hour period of conversation. Redactions may be applied to privileged or non-responsive portions of a conversation. To the extent electronic messages cannot be produced in a reasonably usable format, the Parties will meet and confer to address the identification, production, and production format of short message data.

18. Documents Containing Internal Hyperlinks. The Parties agree that documents containing internal hyperlinks pointing to documents on a system within a Producing Party's custody, possession, or control (e.g., such as an internal document management system, Google Drive, Google Docs, SharePoint, Office365, or similar document hosting and collaboration service) do not need to be produced in the first instance as part of the same family group as the document residing at the location at which that hyperlink points. If there are particular documents containing hyperlinks where the document to which that hyperlink points cannot be located in the production, the Receiving Party may submit a list of such documents by Bates number to the

Producing Party, and the Producing Party will engage in reasonable efforts to locate the document at that pointed located and either identify it by Bates number or provide it if not already produced and not privileged.

19. Search and Collection Cutoff. The Parties have agreed in general to a March 22, 2024 cutoff date for document collections and searches for purposes of this ESI protocol. This agreement is subject to the Parties' duty to supplement under the Federal Rules, the need to address subsequent developments in the case as appropriate, and any good-faith compromises reached during discovery, and does not preclude Parties from seeking additional discovery as necessary.

PRODUCTION FORMATS

20. Document Image Format

a. The Parties will produce all Documents in Group IV single-page TIFF format, black and white, 300 dpi, unless specified otherwise below.

b. The Parties will produce all documents with both load files, as specified below, and metadata files, as specified in Addendum A.

c. Hard Copy Documents. The Parties shall scan all Hard Copy Documents using best efforts to have their vendors unitize documents correctly. The Parties commit to address situations where there are improperly unitized documents. The Parties agree to provide the following objective coding to scanned Hard Copy Documents, if applicable and/or available: beginning Bates number; ending Bates number; file name (beginning Bates number with .tif file extension); beginning attachment Bates number; ending attachment Bates number; page count; and source location/custodian.

d. Color Documents. Color documents (*e.g.*, color photographs or graphical representation in color) shall be produced in black and white except the Receiving Party may request higher resolution TIFF images or color images to render the image legible, understandable, or more usable. If color images are requested, the files shall be delivered in single page, JPEG format or Native format, at the discretion of the Producing Party.

e. Emails will be produced with the CC and BCC line displayed in the image.

f. If ESI in commercial or proprietary database formats can be produced in an existing and reasonably usable, delimited report format, the Parties will produce the information in *.csv format. If an existing report format is not reasonably available or usable, the Parties will meet and confer to attempt to identify a mutually agreeable form of production based on the specific needs and the content and format of data within such structured data source.

g. The Parties will provide full extracted text in the format of a single *.txt file for each file (*e.g.*, not one *.txt per *.tif image). Where ESI contains text that has been redacted under assertion of privilege or other protection from disclosure, the redacted *.tif image will be OCR'd and file-level OCR text will be produced in lieu of extracted text. Searchable text will be produced as file-level multi-page UTF-8 text files with the text file named to match the beginning Bates number of the file. The full path of the text file must be provided in the *.dat data load file.

21. Native File Production. The Parties will produce the following ESI in Native Formats with the metadata specified in Addendum A rather than document image format: spreadsheets (*e.g.*, *.xls, *.xlsx, *.csv), presentation files (*e.g.*, *.ppt, *.pptx, *.odp), audio or audiovisual files (*e.g.*, *.mp4, *.avi, *.mov, *.m4a, *.mp3), short message files, and ESI in commercial or proprietary database formats as specified in paragraph 17(f) above. Redacted ESI may be redacted natively, as feasible, or produced as redacted TIFFs with applicable, non-privileged metadata and OCR searchable text. The Parties will meet and confer regarding any good faith request for the production of other files or file types in native file format.

22. Document Unitization. If a document that contains an attachment(s) is responsive, the document and the attachment(s) will be produced. The Parties shall take reasonable steps to ensure that parent-child relationships within a document family (the association between an attachment and its parent document) are preserved. The child document(s) should be consecutively produced immediately after the parent document. For further clarification, this shall not require a Party to produce documents merely referenced in responsive documents; provided, however, that documents sent via a link within an email should be produced to the extent available. Extracted images from emails where the extracted image is present in the body of the email such as logos or

screenshots need not be produced as child documents.

23. Load Files. There will be two Load/Unitization files accompanying all productions. One will be the image load file and the other will be the metadata load file. Fielded data should be exchanged via a document-level-database load file in one of two delimited formats: either standard Concordance (DAT) or comma delimited (CSV).

a. Image Load File.

i. All image data should be delivered with a corresponding image load file in one of three formats: standard IPro (LFP), Opticon (OPT) or Summation (DII).

ii. Every document referenced in the product load file shall have all corresponding images, text, and data logically grouped together in a directory structure with a common key to properly load the data.

iii. Documents shall be produced in only one image load file throughout the productions, unless that document is noted as being a replacement document in the Replacement field of the data load file.

iv. The name of the image load file shall mirror the name of the delivery volume, and should have a lfp., opt, or .dii* extension (*e.g.*, ABC00l.lfp). The volume names shall be consecutive (*i.e.*, ABC001, ABC002, *et. seq.*) *If dii file is produced, the accompanying metadata load file shall be separate from the .dii file and not contained within the .dii file.

v. The load file shall contain one row per TIFF image.

vi. Every image in the delivery volume shall be contained in the image load file.

vii. The image key shall be named the same as the Bates number of the page. Load files shall not span across media (*e.g.*, CDs, DVDs, Hard Drives, etc.). A separate volume shall be created for each piece of media delivered.

b. Metadata Load File. The metadata fields associated with each Electronic Document or Data or ESI, to the extent they are available, will be produced as specified in the attached Addendum A.

PRIVILEGE LOG AND TREATMENT OF PRIVILEGED MATERIALS

24. Consistent with the Federal Rules of Civil Procedure, the Parties agree to serve a privilege log providing information regarding all documents withheld or redacted under a claim of privilege and/or work product. The Parties will meet and confer regarding the contents and format

of the privilege log, consistent with the initial agreement outlined below in paragraph 27.

25. The Parties agree that certain privileged communications or documents need not be included on a privilege log, including:

a. Communications regarding litigation holds or preservation, collection, or review to the extent the communication is a privileged communication related to this litigation;

b. Any communication or document that post-dates the filing of this complaint with outside counsel and in-house counsel to the extent the communication is privileged communication related to this litigation.; and

c. Work product of counsel and Parties that post-dates the filing of this complaint.

26. The Parties agree to negotiate a reasonable time within which to exchange privilege logs, but no later than 45 days after a Party's last production prior to the close of fact discovery.

27. Privilege Log Contents. For each document withheld or redacted, the privilege log shall contain the following information: a. The name and job title or capacity of the author; b. The name and job title or capacity of each recipient; c. The date the document was prepared and, if different, the date(s) on which it was sent to or shared with persons other than its author(s); d. The title and description of the document; e. The subject matter addressed in the document; f. The purpose(s) for which it was prepared or communicated; g. The specific basis for the claim that it is privileged; and h. For redacted documents only, the bates numbers corresponding to the first and last page of any document redacted.

28. Protocols for Logging Email Chains. Any email chain (i.e., a series of emails linked together by email responses and forwarding) that is withheld or redacted on the grounds of privilege, immunity or any similar claim shall be logged as one document and shall be identified by the top-most email in the chain that is withheld or redacted. The Parties shall not be required to log identical copies of an email that is included in a chain that has been logged in accordance with this Paragraph.

29. Protocols for Logging "Families." Each member of a family that is withheld or redacted on the grounds of privilege, immunity or any similar claim shall be identified on the log

separately.

30. Contesting Claim of Privilege or Work Product Protection. Nothing in this Order shall limit the Receiving Party's right to challenge (on grounds unrelated to the fact or circumstances of the disclosure) the Producing Party's claim that disclosed information is protected from disclosure by the attorney-client privilege or work product doctrine. If, after undertaking an appropriate meet-and-confer process, the Parties are unable to resolve any dispute they have concerning the protection of documents for which a claim of disclosure has been asserted, the Producing Party may file the appropriate motion or application as provided by the Court's procedures to compel production of such material. Any Protected Information submitted to the Court in connection with a challenge to the disclosing Party's claim of attorney-client privilege or work product protection shall not be filed in the public record, but rather shall be redacted, filed under seal, or submitted for in camera review.

DOCUMENTS PROTECTED FROM DISCOVERY

31. Pursuant to Fed. R. Evid. 502(d) and the Stipulated Protective Order (Dkt. 293), the production of a privileged or work-product-protected document, whether inadvertent or otherwise, is not a waiver of privilege or protection from discovery in this case or in any other federal or state proceeding. For example, the mere production of privileged or work-product-protected documents in this case as part of a mass production is not itself a waiver in this case or in any other federal or state proceeding. Nothing contained herein, however, is intended to limit a Party's right to conduct a review of ESI for relevance, responsiveness and/or privilege or other protection from discovery.

32. Communications involving trial counsel that post-date the filing of the complaint need not be placed on a privilege log. Communications may be identified on a privilege log by category, rather than individually, if appropriate.

33. Activities undertaken in compliance with the duty to preserve information and at the direction of counsel are protected from discovery pursuant to Fed. R. Civ. P. 26(b)(3)(A) and (B).

34. Nothing in this Order shall be interpreted to require disclosure of irrelevant information or relevant information protected by the attorney-client privilege, work-product

doctrine, or any other applicable privilege or immunity. The Parties do not waive any objections to the production, discoverability, admissibility, or confidentiality of documents and ESI.

SECURITY

35. The Parties will make reasonable efforts to ensure that any productions made are free from viruses and provided on encrypted media or by secure file transfer protocol ("FTP").

MISCELLANEOUS

36. Nothing herein shall preclude the Parties from agreeing in writing to amend or waive the terms of this ESI Protocol or to agree in writing to proceed differently for any given instance. Nor shall anything herein preclude any Party from moving the Court to amend the terms of this ESI Protocol for good cause shown, provided, however, that no Party may seek relief from the Court concerning compliance with the Order until it (i) has met and conferred in good faith with any Parties involved in the dispute to resolve or narrow the area of disagreement, and (ii) has given any such Parties a reasonable opportunity to cure (if cure is possible) any claimed deficiency in compliance with this ESI Protocol.

37. Nothing herein shall be construed to affect the admissibility of documents and ESI. All objections to the discoverability or admissibility of any documents and ESI are preserved and may be asserted at any time.

38. Nothing herein shall affect, in any way, a Producing Party's right to seek reimbursement for costs associated with collection, review, and/or production of documents in response to disproportionate ESI production requests, pursuant to Federal Rule of Civil Procedure 26.

39. Further, nothing herein is intended to prevent either Party from complying with the requirements of any applicable country or state's data privacy laws.

PURSUANT TO STIPULATION, IT IS SO ORDERED.

Dated: _____, 2025

HON. SUSAN VAN KEULEN
UNITED STATES MAGISTRATE JUDGE

Dated: February 12, 2025

/s/ Joseph R. Wetzel

Joseph R. Wetzel
Andrew M. Gass
Brittany N. Lovejoy
LATHAM & WATKINS LLP
505 Montgomery Street, Suite 2000
San Francisco, CA 94111
Telephone: 415.391.0600
joe.wetzel@lw.com
andrew.gass@lw.com

Allison L. Stillman
(admitted *pro hac vice*)
LATHAM & WATKINS LLP
1271 Avenue of the Americas
New York, NY 10020
Telephone: 212.751.4864
alli.stillman@lw.com

Sarang V. Damle
(admitted *pro hac vice*)
LATHAM & WATKINS LLP
555 Eleventh Street, NW, Suite 1000
Washington, D.C. 20004
Telephone: 202.637.2200
sy.damle@lw.com

Attorneys for Defendant

Respectfully submitted,

/s/ Timothy Chung

Matthew J. Oppenheim (admitted *pro hac vice*)
Nicholas C. Hailey (admitted *pro hac vice*)
Audrey L. Adu-Appiah
(admitted *pro hac vice*)
OPPENHEIM + ZEBRAK, LLP
4530 Wisconsin Ave., NW, 5th Floor
Washington, DC 20016
Telephone: (202) 480-2999
matt@oandzlaw.com
nick@oandzlaw.com
aadu-appiah@oandzlaw.com

Jennifer Pariser (admitted *pro hac vice*)
Andrew Guerra (admitted *pro hac vice*)
Timothy Chung
(admitted *pro hac vice*)
OPPENHEIM + ZEBRAK, LLP
461 5th Avenue, 19th Floor
New York, NY 10017
Telephone: (212) 951-1156
jpariser@oandzlaw.com
andrew@oandzlaw.com
tchung@oandzlaw.com

Richard S. Mandel (admitted *pro hac vice*)
Jonathan Z. King (admitted *pro hac vice*)
Richard Dannay (admitted *pro hac vice*)
COWAN, LIEBOWITZ & LATMAN, P.C.
114 West 47th Street
New York, NY 10036-1525
Telephone: (212) 790-9200
rsm@cll.com
jzk@cll.com
rxd@cll.com

Attorneys for Plaintiffs

SIGNATURE ATTESTATION PURSUANT TO CIVIL L.R. 5-1(h)

Pursuant to Civil L.R. 5-1(h), I hereby attest that concurrence in the filing of the document was obtained from the document's signatories. I declare under penalty of perjury that the foregoing is true and correct.

Dated: February 12, 2025

/s/ Timothy Chung
Timothy Chung

ADDENDUM A – METADATA

Field Name	Field Description	Required for Email	Required for Non-Email and Non-Messaging ESI	Required for Instant Message ESI
BegBates	Beginning Bates number (including Prefix)	x	x	x
EndBates	Ending Bates number (including Prefix)	x	x	x
ParentBates	Beginning Bates number for the parent email of a family (will not be populated for documents that are not part of a family)	x	x	x
BegAttach	Beginning Bates number of the first document in an attachment range (only in documents with attachments)	x	x	x
EndAttach	Ending Bates number of the last document in attachment range (only in documents with attachments)	x	x	x
AttachCount	Number of attachments to a document	x	x	x
AttachRange	Bates range beginning with BatesBeg and ending with BatesEndAttach (will not be populated for documents that are not part of a family)	x	x	x
CustodiansAll	Name of all custodian(s) of email(s) or file(s) produced (Last Name, First Name format) as available	x	x	x
Confidentiality Designation	Confidentiality Designation	x	x	x
Redacted	Descriptor for documents that have been redacted. “Yes” for redacted documents; “No” for unredacted documents.	x	x	x

Field Name	Field Description	Required for Email	Required for Non-Email and Non-Messaging ESI	Required for Instant Message ESI
DateCreated	Date that a non-email file was created (mm/dd/yyyy format)		x	
DateLastmodified	The application recorded time on which the document was last modified		x	
DateSent	Sent date of an email message (mm/dd/yyyy format) (a given email will have either a DateSent or Date Recvd, but not both)	x		
TimeSent	Time email was sent (hh:mm:ss format) (a given email will have either a TimeSent or Time Recvd, but not both)	x		
DateRcvd	Received date of an email message (mm/dd/yyyy format) (a given email will have either a DateSent or Date Recvd, but not both)	x		
TimeRcvd	Time email was received (hh:mm:ss format) (a given email will have either a TimeSent or Time Recvd, but not both)	x		
Author	Author field extracted from the metadata of a non-email document		x	
From	From field extracted from an email message	x		
To	To or Recipient field extracted from an email message	x		
Bcc	Blind Carbon Copy (“Bcc”) field extracted from an email message	x		
Cc	Carbon Copy (“Cc”) field extracted from an email message	x		
Title	Title field extracted from the metadata of a non-email document		x	

Field Name	Field Description	Required for Email	Required for Non-Email and Non-Messaging ESI	Required for Instant Message ESI
Filename	File name — Original name of file as appeared in original location		x	
EmailSubject	Subject line extracted from an email message	x		
FilePath	File path storage location of the document or email if the source application allows for the creation of folders	x	x	
FileExtension	File extension of native file	x	x	x
HashValue	MD5 or SHA-1 hash value used to deduplicate the data and uniquely identify each Native document	x	x	x
MsgID	Unique Message ID field	x		
NativeLink	Relative path to any files produced in native format	x	x	x
HiddenContent	Denotes presence of Tracked Changes/ Hidden Content/ Embedded Objects in item(s) (Y/N)	x		
TextLink	Relative path to any OCR/extracted text files in the production set	x	x	x
ChatSource	Messaging service name (Slack, Google Chat, Microsoft Team, etc.)			x
ChatName	Name of message channel/group			x
ChatParticipants	List of message channel/group participants			x

Field Name	Field Description	Required for Email	Required for Non-Email and Non-Messaging ESI	Required for Instant Message ESI
ChatThreadID	Unique Chat ID field			x
ChatStartDate	Date that messages started (mm/dd/yyyy format) (based on parsing)			x
ChatEndDate	Date that messages ended (mm/dd/yyyy format) (based on parsing)			x
ChatStartTime	Time messages started (hh:mm:ss format) (based on parsing)			x
ChatEndTime	Time messages ended (hh:mm:ss format) (based on parsing)			x