UNITED STATES BANKRUPTCY COURT
SOUTHERN DISTRICT OF NEW YORK

In re: **SPIRIT AVIATION HOLDINGS, INC., et al.,** Debtors.	Chapter 11 Case No. 25-11897 (SHL) (Jointly Administered)

DECLARATION OF ANDREW COLOMBI

I, Andrew Colombi, declare as follows pursuant to 28 U.S.C. § 1746:

1. I am a co-founder of Tonic AI, Inc. ("Tonic") and its Chief Technology Officer.

2. I hold a Ph.D. in computer science from the University of Illinois at Urbana-Champaign, which I earned in 2008. Over the course of my eighteen-year career in the technology sector, I have gained extensive experience in the fields of data analytics, software engineering, and de-identification (defined below).

3. From 2008 until 2017, I worked at Palantir Technologies. From my hire date until approximately 2013, I was an "FDE" or "forward deployed engineer." In that role, I was embedded inside customers' organizations to customize Palantir's products to meet their specific data analytics needs. I also collaborated with clients to connect their data to those systems and draw conclusions from it. In or about 2013, I transitioned my focus to Palantir's product team, where I worked on the development of Palantir's product called "Foundry"—an ontology-powered enterprise data operating system which integrates various elements of a client's business to provide automated decision-making solutions in complex business settings. In this role, with another team leader, I oversaw a team of fifty individuals, including developers, designers, quality assurance staff, and project managers. Starting in or around 2016, following Foundry's launch, I resumed a

prototyping role, in which I was responsible for developing monitoring tools to oversee Palantir's increasingly extensive portfolio of system deployments and user accounts.

4. After almost a decade at Palantir, I co-founded Tonic in 2018 and continue to work here today as its Chief Technology Officer. I am generally familiar with Tonic's day-to-day operations, business, and technological capabilities. As a founder, I have been both an architect and implementer of many key technologies, including many of the methods Tonic uses to de-identify data. I also work directly with Tonic's customers to implement de-identification in their companies.

5. I submit this declaration in connection with the Debtors' request for approval by this Court of the sale of the Deidentified Data[1] to Google LLC (the "Buyer"), and the related engagement by the Debtors of Tonic to perform certain services in connection therewith.

6. Tonic is a privately owned company. Neither the Debtors nor the Buyer has any equity interests in, or management or governance rights with respect to, Tonic. Tonic maintains ordinary-course technology partnerships with each of the major cloud providers, including Google Cloud, and uses their products in its own operations. However, no product or service of the Buyer or its affiliates will be used to process or host the Spirit Data (defined below).

7. I am over the age of 18 years and am authorized to submit this Declaration. Except where stated otherwise, the facts below are based on my personal knowledge, my review of the transaction documents, court filings, and information provided by Tonic personnel working alongside me or under my supervision. If called to testify, I could and would testify competently to the facts set forth in this Declaration.

[1] The term "Deidentified Data" as used herein shall have the meaning ascribed to it in the Bill of Sale between the Debtors and Google LLC (the "Sale Agreement"). ECF No. 1802, Ex. B.

I. Overview of Tonic.

8. Tonic was founded in 2018 as a data de-identification and privacy company. For purposes of this Declaration, "de-identification" refers to the process of removing or transforming the elements of a dataset that identify the individuals it describes, so that the resulting data cannot reasonably be used to infer information about, or otherwise be linked to, a particular individual or household.

9. Tonic is recognized as a leader in data de-identification. In 2024 and 2026, Tonic was named to Rising in Cyber,[2] a list of thirty of the most promising private cybersecurity companies, chosen through nominations and a vote by more than one hundred Chief Information Security Officers and other security leaders. In addition, Fast Company named Tonic one of the World's Most Innovative Companies in 2021,[3] and Amazon Web Services has awarded Tonic its Generative AI Competency,[4] a partner designation AWS grants after technically validating a partner's solution.

10. Since its inception, hundreds of businesses in regulated industries have used Tonic's products to keep sensitive data out of software development, analytics, and AI training workflows by using the de-identification process. Enterprises in highly regulated industries, including online retailers, commercial insurers, investment advisors, health insurers, and telecom and internet service providers use Tonic's software to de-identify their data. More and more, companies use Tonic for the same purpose in AI development. For example, healthcare

---

2 https://risingincyber.com/ (last visited November 12, 2026).

3 https://www.tonic.ai/press-releases/tonic-ai-named-to-fast-companys-annual-list-of-the-worlds-most-innovative-companies-for-2021 (last visited November 12, 2026).

4 https://www.tonic.ai/press-releases/tonic-achieves-aws-generative-ai-competency (last visited November 12, 2026).

companies use Tonic's products to supply training data to AI developers who rely on Tonic to de-identify that data before any model is trained on it.

11. Tonic's products help customers use synthetic and de-identified data to overcome data-access barriers, and accelerate software and AI development, while protecting sensitive information. For AI training in particular, a digital healthcare company uses Tonic's products to address limited access to realistic testing and training data for generative AI. Another one of Tonic's clients—a provider of AI-powered solutions that automate private market contracts, compliance, and entity management—uses Tonic's products to support development involving commercial large-language models and proprietary models.

II. De-Identification Process.

12. Tonic works with many data formats, including, as relevant here, "structured data," "semi-structured data," and "unstructured data." The process for de-identification depends on whether the data to be de-identified (the "Relevant Information") is structured, semi-structured, or unstructured.

13. Structured data is information that is organized into a fixed, predefined format such as database tables. The data may be modeled in rows and columns, or in other forms. Irrespective of how the structured data is modeled, structured data is composed of "Records," where each Record has a predefined set of labeled data elements, and each label is called a "Field." Fields are attributes (e.g., Payment Type, Amount, Date, Month, Year), and the Record is the specific thing being recorded (e.g., a payment sent from a vendor). Therefore, in tabular data, the rows are Records, and the columns are the Fields. Other examples of structured data include data stored by applications such as Salesforce or QuickBooks.

14. Sometimes data is referred to as "semi-structured." Semi-structured data does not rigidly conform to one format, and instead the Fields differ from Record to Record. Semi-structured data commonly exists when data has evolved over time to include additional Fields, or includes some mandatory and some optional Fields. For example, data from an online retailer's order system may always require a customer's shipping address, but only include additional Fields where applicable, such as gift-wrapping options or a gift message, which other orders may not have.

15. "Unstructured data" means information that does not live in a predefined data model such as tables or similar forms. Common examples of unstructured data are e-mails, online chats, or scanned paper records. Unstructured data also frequently exists within an otherwise structured data set, for example, a "notes" column in an Excel file or a "notes" Field in a database.

III. Tonic's Engagement With Spirit.

16. The Buyer and the Debtors have entered into the Sale Agreement for the purchase of the Assets (as defined in the Sale Agreement), including all or a portion of the data set forth in Exhibit A to the Sale Agreement (the "Spirit Data").

17. The Debtors have entered into a Master De-Identification Services Agreement (the "MSA") and related Statement of Work ("SOW")[5] with Tonic under which Tonic will serve as the Deidentification Agent (the "Engagement").[6] In accordance with the Sale Agreement, the Spirit Data will not be transmitted to the Buyer directly but to a Deidentification Agent (as defined in the Sale Agreement) designated by the Debtors and reasonably acceptable to the Buyer, working

[5] *See* ECF No. 1802, Ex. C.

[6] To the extent there is any purported conflict between the MSA or SOW (as applicable) and the statements made herein, the terms of the MSA or SOW (as applicable) shall govern. Additional details regarding the workstream for the exclusion of confidential information are discussed below.

under the Debtors' supervision. Tonic has been designated by the Debtors as the Deidentification Agent.

18. Tonic's fee is a fixed amount subject to adjustment per the terms of the MSA and SOW, to be paid by the Buyer in accordance with the Sale Agreement and a separate payment agreement between Tonic and the Buyer. Tonic's fee does not depend on the outcome of the Engagement.

19. The Engagement contemplates two workstreams: (i) the de-identification of personal information in the Spirit Data before its transfer to the Buyer (the "De-identification Process"); and (ii) the identification and exclusion of the Excluded Data (as defined in the Sale Agreement) designated by the Debtors (the "Exclusion Process").[7] The protocols for the two workstreams, Exhibits C and D to the SOW, are the "Data Treatment Procedures" attached to the Proposed Sale Order[8] as Schedule 1.

20. The MSA requires Tonic to take measures, as described in the applicable SOW, to ensure that the Services Output Data (as defined in the MSA) "cannot reasonably be used to infer information about, or otherwise be linked to, any particular identifiable natural person or household, including by applying the de-identification standard reflected in Cal. Civ. Code § 1798.140(m) and comparable state privacy laws and, where applicable, the de-identification standards set forth under 45 C.F.R. §164.514 (the "De-identification Standard")."[9]

---

[7] MSA § 1(a).

[8] For purposes of this Declaration, the "Proposed Sale Order" refers to the proposed *Order (I) Authorizing The Transaction With Respect to the Deidentified Data Free and Clear of Liens, Claims, Interests and Encumbrances, (II) Authorizing the Debtors to Enter into and Perform Their Obligations Under the Sale Agreement and (III) Granting Related Relief*, ECF No. 1802, Ex. A

[9] MSA § 1(a).

21. Before Tonic releases any Spirit Data to the Buyer, it must have "(i) completed the agreed de-identification and targeted exclusion processing in accordance with the Data Treatment Procedures, (ii) documented validation and quality-control testing, (iii) confirmed, in its professional judgment, that the Services Output Data meets the De-Identification Standard and the Excluded Information Standard pursuant to the applicable SOW, (iv) issued the Certification required under the applicable SOW, and (v) confirmed that the Sale Order has been entered and is in effect" and any fees then due have been paid, and the Debtors have authorized the release in writing.[10]

22. Tonic, as an independent third party, is required to "exercise independent professional judgment in selecting and applying the procedures necessary" to carry out the De-identification Process and Exclusion Process, and may "escalate to [the Debtors] for direction any content that in Tonic's professional judgment cannot reliably be de-identified to the De-Identification Standard and/or excluded to the Excluded Information Standard."[11]

A. <u>De-Identification of Structured and Semi-Structured Data</u>.

23. ***Step 1 – Categorization.*** With respect to structured data and semi-structured data, the process for de-identification begins by categorizing each Field. Tonic's technology is capable of searching for and identifying categories of Relevant Information ("<u>Categories</u>"), which include names, email addresses, physical addresses, account numbers, payment information, and notes made within a Field. Tonic's technology is also capable of finding Relevant Information even if the Field is not labeled in a way that obviously suggests the Field contains Relevant Information.

[10] MSA § 1(e); SOW §§ 4, 7; *id*. Ex. E § 5.

[11] MSA § 2(g).

For example, Tonic's technology can recognize personally identifiable information ("PII") such as a name or date of birth, even if it appears in a column named "test_column."

24. The results of categorization are verified through the sampling and measurement process described in paragraphs 38 through 41.

25. ***Step 2 – Field Modification.*** After categorization, sensitive Fields are modified to strip the identifying data from the Records. The mechanism for modifying Fields depends on the nature of the Relevant Information being changed.

26. Tonic's technology replaces PII with synthetic values. This approach reduces re-identification risk because synthesis (rather than redaction) makes it more difficult to distinguish a residual (non-synthesized) identifier from the synthetic values applied to the entity type (*e.g.*, name, home address, Social Security Number); and synthesis introduces uncertainty in confirming which values are real versus synthesized.

27. Financial account identifiers (*e.g.*, bank accounts, credit card numbers) are replaced consistently with synthetic values drawn from ranges that cannot correspond to an issued card or account (or removed where synthetic values are not possible).

28. Indirect identifiers (*e.g.*, pieces of information about an individual that, if aggregated, could identify the individual)[12] are generalized, meaning replaced with a coarser value (for example, an age range or a salary band), to the extent required to meet the De-identification Standard, as measured under SOW Exhibit D item 10.

29. Unstructured data embedded in structured data (*e.g.*, "Notes" columns) are run through the process for de-identifying unstructured data described below.

[12] For example, a person's age, zip code, job title, and marital status, if triangulated, could result in the identification of a particular individual.

B. <u>De-Identification of Unstructured Data</u>.

30. ***Step 1 – Scanning.*** Unstructured data is scanned to identify the location and Category of Relevant Information. The process of locating and categorizing Relevant Information in unstructured data is different from the process of locating and categorizing structured data. That is because in structured data, the Fields are already known (*i.e.*, the columns in the Excel file are already there, with a label), whereas unstructured data, by its nature, does not typically have "Fields."

31. Tonic will use proprietary and third-party models to identify Relevant Information in unstructured data. Critically, these models use context to categorize data. For example, "Washington" may be a last name or a city depending on the context: "Good evening, Mr. Washington," as opposed to "My sister lives in Washington."

32. ***Step 2 – Modification.*** After categorization, the Relevant Information (as described below) is replaced using a mechanism specific to the category of data.

33. PII, financial account information, and indirect identifiers are replaced using the same approach for structured data described in paragraphs 26 through 28 above. For example, a person's name in an e-mail is replaced with a new value, which is then applied to each mention of that person's name across the unstructured data set.

34. Other categories of Relevant Information such as photographs and handwritten signatures, and payment card numbers, are treated as described in paragraph 36.

35. The De-identification Process Tonic will employ for the Spirit Data will generally follow Tonic's standard procedures for de-identification of structured, semi-structured, and unstructured data. However, Tonic has built, and may continue to build, certain technologies

tailored to the needs of the Engagement. For example, models may be fine-tuned to the detection needs of the Engagement, or adapters to new data formats may be developed.[13]

36. Exhibit D to the SOW sets out the de-identification procedures that Tonic will apply to the Spirit Data, summarized as follows:

- ***Direct identifiers*** — such as names, contact details, government-issued identification numbers, employee and account numbers, usernames, and email addresses — will be detected using named-entity recognition and pattern-matching technology, whether they appear in structured data fields or free text. Tonic will then link all references to the same individual across the dataset so that each person is assigned a single, consistent synthetic identity, and will replace the real identifying information with realistic but fictitious synthetic values that preserve the format and internal consistency of the data.

- ***Indirect identifiers*** — such as dates of birth, other personal dates, ages, job titles, work locations and base assignments, seniority numbers, and compensation information — will be generalized into ranges or bands (for example, converting an exact date of birth to a birth month and year, or converting an exact salary to a salary range) to the extent required to meet the applicable de-identification standard.

- ***Structured and semi-structured records*** — such as database tables and spreadsheets — will be reviewed field by field, with each field classified as a direct identifier, an indirect identifier, non-identifying information, or free text. Fields will then be further generalized or suppressed as necessary to meet the De-identification Standard.

- ***Free text***[14] — including email, chat messages, documents, help-desk tickets, source code, and content in languages other than English — will be run through the same detection, linking, and synthesis process.

- ***Payment card numbers and similar financial account identifiers*** will be replaced with synthetic values that cannot correspond to any real issued card or account. Where synthetic replacement is not technically feasible, the data will be removed entirely.

[13] Any such engagement-specific models remain within the processing environment, will not be used for any other customer or product, and will be deleted at the end of the Engagement together with the other processing artifacts. Processing artifacts may include custom entity types for identifiers specific to an airline's data, such as employee and crew identifiers, passenger record locators, and internal system identifiers; the use of the Debtors' identity rosters and system extracts to link references to the same individual across systems; and detection coverage for message headers, document properties, version histories, and embedded images.

[14] "Free text" refers to unstructured data.

- ***Photographs of identifiable individuals and handwritten signatures*** will be obscured or excluded from the output data.

- ***Scanned and image-based documents*** will be processed using optical character recognition (OCR) technology that captures the position of text within the image. Identifying information will then be replaced in place, and the resulting documents will be flattened.

- ***Health information***, including protected health information and consumer health-related information, will be de-identified in accordance with the De-Identification Protocol so that no raw health information is included in the output data delivered to the Buyer.

- ***Content that cannot reliably be treated*** by any of the foregoing methods will be suppressed at the specific element or region level where suppression alone cannot reliably meet the de-identification standard (for example, video or other unsupported formats), the entire unit will be excluded or escalated to Spirit for direction. In all cases, Tonic will use the least restrictive treatment that, in its reasonable professional judgment, reliably meets the standard.

37. Before Tonic runs the De-identification Process, Tonic will configure and test the De-identification Process and establish recall and precision thresholds.

38. ***Configuration and Testing.*** Tonic will configure and test the De-identification Process by using human annotators to label Relevant Information in a statistically significant set of samples of the Spirit Data. Human annotators will manually label the Relevant Information in each sample document to establish a "ground truth."[15] Tonic's de-identification model will then be run across the ground truth sample set annotated by the human annotators.[16] The model's output

[15] A "ground truth" is an accurate, verified data set that represents the standard against which the De-identification Process is measured.

[16] SOW Ex. D item 10; *id.* Ex. E §§ 2, 4.

will be compared against the human annotators' labels, so that recall[17] and precision[18] can be measured.

39. ***Recall and Precision Thresholds.*** Tonic evaluates recall and precision of its de-identification process by using two distinct but complementary measures: "per individual" and "per mention." The "per individual" metric determines whether all identifying references to a specific person — across the entirety of a dataset — have been successfully located and de-identified, ensuring complete coverage at the individual level. The "per mention" metric examines each discrete occurrence of identifying information to verify that it was correctly detected and appropriately transformed or removed. This dual-metric approach is more protective than a purely "per mention" approach, because a system may successfully process most mentions of a person's information while still missing isolated references that, if left in place, could allow that individual to be re-identified. By measuring accuracy at both levels simultaneously, Tonic's technology is designed to minimize that risk and provide a more reliable de-identification outcome.

40. Following receipt and review of the Spirit Data, Tonic will establish recall and precision thresholds, against which Tonic will measure the Services Output Data before delivery to the Buyer.[19] The thresholds will be numeric, set separately for each entity type and class of data, and measured both per mention and per individual.[20] Tonic will set these thresholds in its independent professional judgment, after consultation with the Debtors and the Buyer (but not

---

[17] "Recall" refers to the ratio of correct detections (true positives) against total occurrences (true positives *plus* false negatives). For example, if a document contains 100 names, and 95 names are detected correctly, the recall would be 95%.

[18] "Precision" refers to the ratio of correct detections (true positives) against total detections (true positives *plus* false positives). For example, if a document contains 100 names, and 95 names are detected correctly (true positives), and 2 non-names are detected as names (false positives), the precision would be 95 / 97, or 97.94%.

[19] SOW Ex. E § 2.

[20] SOW Ex. D § 10.

subject to the Buyer's consent), and Tonic will document its rationale for such thresholds.[21] These recall and precision thresholds and their rationale will be fixed and documented before full-corpus processing begins, as Item 10 of Exhibit D to the SOW requires.

41. Starting in Phase 2 (configuration and sample iteration), and concluding in Phase 4 (quality assurance), Tonic will sample processed content as it completes processing and, as set forth more fully in Exhibit E to the SOW, correct any gaps during processing.[22]

42. In addition to the recall and precision thresholds described above, at the conclusion of Phase 4, Tonic will produce the QA Assessment required by paragraph 9 of the Proposed Sale Order. The QA Assessment will be a quantitative assessment of the probability of re-identification of consumers or employees of the Debtors, applying the risk-based framework of ISO/IEC 27559:2022 (context assessment, data and attack modelling, quantified identifiability and documentation) (the "ISO 27559 Standard") and taking into account the Buyer Assumptions described below. The QA Assessment will address adversarial re-identification risk through attack modelling and quantitative assessment of identifiability under clauses 7.3 and 8.2.2 of the ISO 27559 Standard, which will also constitute the adversarial review contemplated by Item 10 of Exhibit D and Phase 4 of Exhibit E to the SOW. For identifiable groups formed by certain indirect identifiers in the Services Output Data, the benchmark threshold will be an average identifiability of 0.1 (the "ISO Benchmark Threshold"), which is the benchmark for a non-public release with a "low-med[ium] possibility of attack, low medium impact" under Annex B, Table B.2 of the ISO 27559 Standard.

[21] SOW Ex. D § 10; *id*. Ex. E § 2; MSA § 2(g).

[22] SOW Ex. E § 4.

43. In addition, an independent Verification Agent designated by the Debtors under paragraph 9 of the Proposed Sale Order will verify, based on reasonable samples of Tonic's proposed form of the Deidentified Data, that the QA Assessment was performed in accordance with its stated methodology and that its results are accurately reported.[23] The Verification Agent will not replicate or replace the Data Treatment Procedures, re-run the QA Assessment, apply a re-identification methodology that deviates from the methodology used to produce the QA Assessment, or alter the assumptions, thresholds or criteria fixed under the Data Treatment Procedures.

44. Delivery will not occur, until measured performance meets the recall and precision thresholds described in paragraph 40 and the ISO Benchmark Threshold; the QA Assessment is completed; the Verification Agent has completed its review of the QA Assessment in accordance with the Proposed Sale Order; and Tonic delivers the Certification in accordance with the MSA, SOW and Proposed Sale Order.[24]

C. The Exclusion Process.

45. The Exclusion Process will use a subset of the same tools that Tonic will use for the De-identification Process. Exclusion will occur in two phases. *First,* the Debtors will exclude the systems designated "Not Included" on the Asset Schedule before the Spirit Data is made available to Tonic, and will also exclude Privileged Materials as set forth in Section 1(b) of the Sale Agreement. *Second*, Tonic will apply search criteria to identify confidential third-party information ("Excluded Third-Party Data," as defined in the Sale Agreement) and residual data that otherwise falls within the categories of employee records that Exhibit A to the Sale Agreement

[23] Proposed Sale Order ¶ 9.

[24] SOW Ex. E § 5; Proposed Sale Order ¶ 9.

defines as Excluded Employee Data (the "Residual Excluded Employee Data"),[25] by applying the search terms that Specified Third Parties may propose under paragraphs 20-22 of the Proposed Sale Order and the search terms that Union Parties (as defined in the Sale Agreement) may propose under paragraphs 23–25 of the Proposed Sale Order.

46. Exhibit C to the SOW sets forth the details of the Exclusion Process. The Debtors will exclude certain designated systems, folders, or categories of records before Tonic receives any Spirit Data.[26] The Debtors will designate the criteria for identifying Excluded Third-Party Data and Residual Excluded Employee Data (the "Search Criteria")[27] within a fixed period after entry of the Proposed Sale Order.[28] The Search Criteria will be expressed in searchable terms including, for example, designated counterparties and their email domains; designated search terms applied to the full text of documents and their attachments; systems; modules; custodians; program names; and checksums[29] of known files.[30] Tonic will apply the Search Criteria across the Spirit Data, as agreed with the Debtors or determined by the Court if disputed.[31]

47. Tonic will use the following deterministic and contextual processes to identify Excluded Data:

- ***Deterministic processes*** use designated conditions provided by the Debtors to locate confidential data. These conditions include: third-party domains and addresses, search terms for file names, the checksums of known sensitive files, search terms for the contents of emails, chat messages and documents, and other conditions as described more fully in the SOW.

---

[25] The Excluded Employee Data is defined in paragraph C of Exhibit A to the Sale Agreement, and includes the Excluded Union Data defined in footnote 2 to that Exhibit.

[26] SOW Ex. C § 1.

[27] SOW Ex. C § 2; *Id.* Ex. B.

[28] Proposed Sale Order ¶¶ 20-25.

[29] A "checksum" is a unique digital fingerprint assigned to a file, allowing identical copies of that file to be identified wherever they appear in the data, even if the file has been renamed or moved.

[30] SOW Ex. B.

[31] Proposed Sale Order ¶¶ 11, 14, 20-25.

- ***Contextual processes*** surface, for the Debtors' review, data that is similar to the data identified deterministically but was not identified using purely deterministic processes. Contextual processes will be developed using, as training data, the data located via the deterministic process noted above. Data identified by contextual processes will be presented to the Debtors for review before being designated confidential third-party information. Any additional files designated confidential this way will be added to the list of checksums to be used in the deterministic processes above.

48. A match on the Search Criteria identifies a unit of data (*e.g.*, an email with its attachments, a chat conversation, a file, or a record in a structured system) (the "Responsive Material"). Correspondence with a designated counterparty and agreements and other designated files are always removed as whole units. In other content, where the Responsive Material is reasonably severable (the boundary is apparent or confirmed by the Debtors) and redaction or field-based exclusion is feasible without leaving content from which the excluded material could be reasonably inferred, Tonic will redact or suppress the severable portion (*e.g.*, an attachment without the cover message, a slide, section, sheet or column in a workbook). Otherwise, the unit is removed as a whole unit. Responsive Material whose treatment under the Search Criteria cannot be determined is quarantined and is not delivered. Every exclusion will be recorded in the exclusion manifest with the criterion under which it was made. After the Exclusion Process is completed, Tonic will re-run the Search Criteria against the data set to be delivered to the Buyer to identify any residual Responsive Material for removal prior to delivery.[32] Incidental mentions of Excluded Employee Data in other content that cannot reasonably be removed are de-identified through the De-identification Process.[33]

[32] SOW Ex. C §§ 2–4 and concluding paragraph.

[33] SOW Ex. C § 4 and concluding paragraph.

49. Tonic will record every excluded item in an exclusion manifest, and then reconcile file and record counts from ingest to delivery to account for each item excluded from delivery to the Buyer. The exclusion manifest will be delivered to the Debtors,[34] and to a Specified Third Party or Union Party in accordance with the terms of the Proposed Sale Order.[35]

D. Re-Identification Risk Mitigation.

50. The protection that the Data Treatment Procedures will afford here depends in part on how the Spirit Data is held and used after transfer. Tonic's calibration of the recall and precision thresholds (discussed above in paragraphs 379 through 41) and its application of the Exclusion Process rest on the assumptions set forth below.

51. ***Buyer Commitments and Security.*** Tonic will calibrate its recall and precision thresholds and apply the Exclusion Process based on the assumptions that: (i) the Buyer will maintain and use the Deidentified Data only in de-identified form not associated with any consumer or employee of Spirit and will not intentionally or knowingly reidentify it, or attempt to reidentify it (including, for the avoidance of doubt, to profile consumers or employees of the Debtors); (ii) the Buyer will treat the data repositories of Services Output Data delivered to it by Tonic as confidential and will maintain such data repositories of Services Output Data in a secure environment, in each case, consistent with how the Buyer treats and maintains its own confidential information of a like nature; and (iii) the Buyer will contractually obligate any party to which it transfers the Services Output Data to the same commitments, and any such party will be further bound automatically to such commitments by means of the Sale Order, with the result that the Services Output Data will not be made available to the public (collectively, the "Buyer

[34] SOW § 7.

[35] Proposed Sale Order ¶ 27.

Assumptions"). Clauses (i) through (iii) are consistent with the commitments in Section 1(c) of the Sale Agreement and paragraph 8 of the Proposed Sale Order. The Buyer Assumptions are the basis on which the recall and precision thresholds described in paragraph 40 will be calibrated. Nonconformities with the Buyer Assumptions will require recalibration.

52. ***Real-to-Synthetic Value Mapping.*** The mapping between real and synthetic values, and the entity graph that links references to the same individual, will not be delivered to the Buyer, so the Services Output Data carries no key by which its synthetic values could be reversed. Tonic will retain the mapping and the entity graph in the processing environment only until the post-delivery remediation period provided in the SOW closes, and will then destroy them, with written confirmation, and retain no copy.[36] If the Debtors or the Court determine that a mapping should survive for any purpose, it will be deposited with the Court or with an escrow agent the Court approves, not held by Tonic.[37] Tonic's copy of the Spirit Data will be securely deleted at the conclusion of the Engagement, with written confirmation.[38] Tonic may retain, as the engagement record provided for in the MSA and for seven years, the versioned configuration it applied (entity definitions, thresholds, treatment rules, and exclusion criteria), together with the quality-assurance measurements, the exclusion manifest, and the completeness reconciliation.[39] The engagement record is held under the confidentiality and security obligations of the MSA and does not include the Spirit Data, the mappings, or the entity graph, which are destroyed as described above.[40]

---

[36] SOW § 10; *id*. Ex. E § 5.

[37] SOW § 10.

[38] MSA § 6; SOW § 10; *id*. Ex. E § 5.

[39] MSA § 6.

[40] MSA § 6.

E. Deliverables and Certification.

53. At completion, in addition to the Services Output Data, Tonic will deliver to the Debtors: (i) a summary report of the configuration and rules applied; the QA methodology and measured results; (ii) an exclusion manifest identifying each item excluded pursuant to the Exclusion Protocol; (iii) a reconciliation tying record and file counts from ingest through delivery, including documented exclusions; and (iv) the written certification required by Section 3(c) of the Sale Agreement (the "Certification").[41]

54. The Certification will state that "(i) the procedures described in this SOW were performed; (ii) the Targeted Exclusion Protocol was applied to the Excluded Information using the criteria in effect for the processing run and validated as described in Exhibit C such that, in Tonic's professional opinion, the Services Output Data meets the Excluded Information Standard; and (iii) in Tonic's professional opinion, the De-Identification Protocol was applied such that the Services Output Data meets the De-Identification Standard. The Certification is a statement of procedures performed, results measured and Tonic's professional opinion as of its date, as provided in Section 9 of the Agreement and paragraph [19] of the Proposed Sale Order."[42]

IV. Conclusion.

55. The Data Treatment Procedures, together with the review provided for in the Proposed Sale Order, go well beyond the de-identification process that Tonic applies in the ordinary course of its business, including for the regulated organizations in healthcare, insurance and financial services that Tonic typically serves. Here, Tonic will perform its work under a court-approved protocol that adds a Debtor-designated exclusion layer applied before de-identification,

[41] SOW § 7.

[42] *Id.*

quality-assurance thresholds measured per mention and per individual and fixed with documented rationale before full-corpus processing, a quantitative re-identification assessment, independent verification, an exclusion manifest and completeness reconciliation, and ongoing supervision by this Court.

56. Tonic has agreed to implement the Data Treatment Procedures as they are described above, as set forth in the MSA, SOW and the Proposed Sale Order, before any Spirit Data is transferred to the Buyer.

57. I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

Executed on November 12, 2026
San Francisco, California

/s/ Andrew Colombi
Andrew Colombi
Co-founder and Chief Technology Officer
Tonic AI, Inc.