

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TREND MICRO, INC.

v.

WEBROOT, INC.

Petition for *Inter Partes* Review of

U.S. Patent No. 11,409,869

DECLARATION OF INGRID HSIEH-YEE, PH.D.

TABLE OF CONTENTS

I.	Introduction	3
A.	Qualifications and Professional Experience	3
B.	Scope of This Declaration.....	11
C.	Person of Ordinary Skills in the Arts.....	12
D.	Evidence Considered in Forming My Opinions	13
II.	Authenticity and Public Availability of <i>Tahan</i> (EX1005)	15
A.	Authentication.....	15
B.	Internet Archive Records	16
C.	Usage Record.....	19
D.	Summary of My Opinion on <i>Tahan</i>	19
III.	Authenticity and Public Availability of <i>Menahem</i> (EX1006).....	20
A.	Authentication.....	20
B.	Library of Congress Records	23
C.	Internet Archive Records	27
D.	Internet Archive Affidavit (EX1010).....	28
E.	Citation Record	29
F.	Summary of My Opinion on <i>Menahem</i>	29
IV.	Conclusion.....	30

I. Introduction

I, Ingrid Hsieh-Yee, Ph.D., do hereby declare as follows:

1. I have personal knowledge of all of the matters about which I testify in this declaration.

2. I have been retained as an independent expert witness on behalf of Trend Micro, Inc. (“TREND”) for proceedings before the United States Patent and Trademark Office (“PTO”) regarding the authenticity and public availability of certain prior art reference in relation to US Patent No. 11,409,869 (“the ’869 patent”).

3. I am being compensated for my work in this matter at my customary hourly rate. I am also being reimbursed for any reasonable expenses associated with my work and testimony in this investigation. My compensation is not contingent on the results of my study, the substance of my opinions, or the outcome of this matter.

A. Qualifications and Professional Experience

4. My complete qualifications and professional experience are described in my academic curriculum vitae (**Appendix A**). The following is a brief summary of my relevant qualifications and professional experience.

5. In 2022 I was conferred the title of Professor Emerita after 32 years as a professor in the Department of Library and Information Science of the Catholic University of America. I have experience working in an academic library, a medical library, and a legislative library. I hold a Ph.D. in Library and Information Studies

from the University of Wisconsin-Madison and a Masters in Library and Information Studies from the University of Wisconsin-Madison.

6. I am an expert on library cataloging and classification and have published two editions on this subject, *Organizing Audiovisual and Electronic Resources for Access: A Cataloging Guide* (2000 and 2006). I taught a variety of courses, including Cataloging and Classification, Advanced Cataloging and Classification, Organization of Internet Resources, Organization of Information, Digital Content Creation and Management, Metadata, Internet Searches and Web Design, Information Literacy Instruction, Advanced Information Retrieval and Analysis Strategies, and The Information Professions in Society. I am familiar with metadata schema design and implementation. In my teaching I covered the design and implementation of metadata in libraries, databases, search engines, digital repositories, digital libraries, and digital archives. I also covered how information organization affects the discovery and access to digital resources on the Internet. My research interests cover cataloging and classification, information organization, metadata, information retrieval, information architecture, digital collections, scholarly communication, social media, user interaction with information systems, and others.

7. I am very familiar with a library cataloging encoding standard known as the “Machine-Readable Cataloging” standard, also known as “MARC,” which

became the national standard for sharing bibliographic data in the United States by 1971 and the international standard by 1973. MARC is the primary communications protocol for the transfer and storage of bibliographic metadata in libraries. Experts in my field reasonably rely upon MARC records when forming their opinions.

8. A MARC record consists of several fields, each of which contains specific data about the work. Each field is identified by a standardized, unique, three-digit code corresponding to the type of data that follows. **Appendix B** is a true and correct copy of Parts VII to X of “Understanding MARC Bibliographic: Machine-Readable Cataloging” (<http://www.loc.gov/marc/umb/um07to10.html>) from the Library of Congress. *Understanding MARC Bibliographic* was originally published by the Follett Software Co. and copyrighted in 1988. Editions in 1989, 1990, 1994, 1998, 2000, 2003 and 2009 include changes. The latest edition was copyrighted in 2009. This document explains commonly-used MARC fields. For example, the personal author of the work is recorded in Field 100, the title is recorded in Field 245, publisher information is recorded in Field 260, the physical volume and characteristics of a publication are recorded in Field 300, and topical subjects are recorded in Fields 650.

9. The Online Computer Library Center (OCLC) is the largest bibliographic network of the world, with more than 534 million records and thousands of member institutions (many of which are libraries of some type) in more

than 100 countries. OCLC was founded in 1967 to promote and support library cooperation. According to the “Third Article, Amended Articles of Incorporation of OCLC Online Computer Library Center, Inc.,” OCLC was created “to establish, maintain and operate a computerized library network and to promote the evolution of library use, of libraries themselves, and of librarianship, and to provide processes and products for the benefit of library users and libraries, including such objectives as increasing availability of library resources to individual library patrons and reducing the rate of rise of library per-unit costs, all for the fundamental public purpose of furthering ease of access to and use of the ever-expanding body of worldwide scientific, literary and educational knowledge and information” (source: <https://www.oclc.org/content/dam/oclc/membership/articles-of-incorporation.pdf>).

10. OCLC members can contribute original cataloging records in MARC to the system or derive cataloging records from existing records, an activity referred to as “copy cataloging.” When an OCLC participating institution acquires a work, it can create an original MARC record for the work in OCLC’s Connexion system (a system for catalogers to create and share MARC records), and the system will automatically generate a code for the date of record creation in the yymmdd format, and the creating library’s OCLC symbol is recorded in subfield “a” of the 040 field. Once the MARC record is in Connexion, it becomes available to other OCLC members for adoption to their local online catalogs (*i.e.*, copy cataloging).

11. After a MARC record is created in Connexion, it also becomes searchable and viewable on WorldCat, which is a free web portal for users to explore more than 10,000 libraries worldwide. The record in WorldCat, however, is not presented in MARC fields. Instead, the data elements are labeled to help users interpret the record. Thus, the information stored in MARC records in Connexion is available to the interested public through the user-friendly WorldCat web portal.

12. WorldCat (<http://www.worldcat.org>) is “the world’s largest network of library content and services” and its features are summarized in “What is WorldCat” (<https://www.worldcat.org/about>). WorldCat began operation in January 1998 as a free web portal to more than 10,000 libraries worldwide. Through WorldCat, users can search for information resources in their local libraries and libraries around the world. WorldCat allows users to search for books, CDs, videos, and many new types of digital content, such as audiobooks, in many languages. Users can also retrieve research materials and article citations with links to their full text. After an item is retrieved, WorldCat helps users identify a library nearby that holds the item or all the libraries that hold the item. WorldCat is an efficient way to explore the content held by more than 10,000 libraries around the world. WorldCat is publicly accessible, and there is no fee to perform a search on WorldCat.

13. Library online catalogs gained acceptance in the early 1980s and many libraries migrated their systems to the World Wide Web in the mid-1990s. Library

online catalogs are based on MARC records that represent their collections in order to help the public understand what materials are publicly accessible in those libraries. Most libraries with online catalogs have made their catalogs freely available on the Web. These online catalogs offer user-friendly search interfaces. Strong user interest in keyword searches and the popularity of Google have led to the “googlization” of library search systems. As a result, many library catalogs now provide a single search box for users to conduct keyword searches, with additional support for searches by author, title, subject terms, and other data elements such as ISBN (International Standard Book Number). Library catalogs these days also offer features for users to narrow their search results by language, year, format, and other elements. Many libraries display MARC records on their online catalogs with labels for the data elements to help the public interpret MARC records. Many libraries also offer the option to display MARC records in MARC fields.

14. Libraries create MARC records for works they acquire, including books, serials, motion pictures, and publications in other formats. Monograph cataloging is fairly common in libraries, and most libraries make a newly cataloged monograph available to the public soon after the cataloging work is completed, usually within a week. Libraries can create original cataloging records or use an existing record in OCLC to create a copy cataloging record. As soon as the cataloging record is completed, it is added to the library’s online catalog for users.

If the record is an original record, it is also entered into the OCLC system. If it is a copy cataloging record, the library's holding symbol is attached to the existing original record in the OCLC system to facilitate searching and interlibrary loan.

15. The cataloging of serials and the serial check-in process are discussed here to show how libraries usually provide access to newly received serial issues. According to the glossary of the *RDA: Resource Description and Access* cataloging standard, a serial is “a mode of issuance of a manifestation issued in successive parts, usually bearing numbering, that has no predetermined conclusion. A serial includes a periodical, monographic series, newspaper, etc.” Because the publisher of a serial makes new issues of the serial available successively, a customary cataloging practice is to create one bibliographic record for the serial, and the serial record (encoded in MARC) typically provides information on the beginning date and frequency of the serial, not the dates of individual issues. In other words, libraries typically do not create MARC records for individual issues of a serial. Instead, they rely on a serial check-in system to track the receipt of new issues. A common check-in practice is to date stamp a new issue when it arrives. This practice has become automated since the late 1990s, and libraries now vary in how they share the receipt date of a new serial issue with the public. Some libraries use a date stamp, some affix a label to indicate the receipt date, some pencil in the receipt date, and some do not provide the information to the public.

16. The serial check-in process usually takes less than an hour, and one of the steps involves placing a date stamp on the new issue to document the date the issue is checked in. After that, the holdings information of the serial is updated in the library's catalog so that users know which issues are available for request or access. After serial check-in is completed, the new issue is placed on the shelf with the previous issues of the serial. Libraries with a public periodical room typically place new issues in the periodical room for easy user access. Because information presented in serials often reflects latest discovery, a general practice of libraries is to make new issues of serials available for user access soon after they are checked in, usually within a week.

17. I am personally familiar with many online catalogs, databases, and search engines. In preparing for this declaration, I used the following authoritative information systems to search for records:

- Google Scholar (<https://scholar.google.com>)
- Internet Archive (<https://archive.org>)
- Online catalog of the British Library (<http://explore.bl.uk/>)
- Online catalog of the Library of Congress (<https://catalog.loc.gov>)
- WorldCat (<https://www.worldcat.org>)

B. Scope of This Declaration

18. I have been asked to offer opinions on the authenticity and public availability date of the following documents:

Tahan, G., Rokach, L., & Shahr, Y., (2012, April), Mal-id: Automatic malware detection using common segment analysis and meta-features, *Journal of Machine Learning Research*, 13(spring), pp. 949-979, **EX1005** (“Tahan”);

Menahem, E., Shabtai, A., Rokach, L., & Elovici, Y., (2009), Improving malware detection by applying multi-inducer ensemble, *Computational Statistics & Data Analysis*, vol. 53, no. 4, pp. 1483-1494, **EX1006** (“Menahem”).

19. I am rendering my expert opinion on the authenticity of the document referenced herein and on whether it was published. I will provide my opinion on when and how this document was publicly disseminated or otherwise made publicly available to the extent that persons interested and ordinarily skilled in the subject matter or art, exercising reasonable diligence, could have located and retrieved this document.

20. I am informed by counsel that an item is considered authentic if there is sufficient evidence to support a finding that the item is what it is claimed to be. I am also informed that authenticity can be established based on the contents of the

documents themselves, such as the appearance, content, substance, internal patterns, protocols followed or other distinctive characteristics of the item.

21. I am informed by counsel that a given reference is “published” if it is publicly accessible upon a satisfactory showing that such a document has been disseminated or otherwise made publicly available to the extent that persons interested and ordinarily skilled in the subject matter or art could locate it exercising reasonable diligence and obtain the document. I have also been informed by counsel that materials available in a library constitute “printed publications” if they are cataloged and indexed according to general library practices and protocols that make the references available and accessible to members of the interested public.

22. I have been informed by counsel that the ’869 patent was filed on February 14, 2020 in the United States, and the earliest priority date of the ’869 is on May 12, 2015. I have assessed public accessibility of the documents referenced below as of this claimed priority date.

C. Person of Ordinary Skills in the Arts

23. I am told by counsel that the subject matter of this proceeding relates to malware detection systems.

24. I have been informed by counsel that a “person of ordinary skill in the art at the time of the inventions” (POSA) is a hypothetical person who is presumed to be familiar with the relevant field and its literature at the time of the inventions.

This hypothetical person is also a person of ordinary creativity, capable of understanding the scientific principles and literature applicable to the pertinent field.

25. I am told by counsel that persons of ordinary skill in this subject matter or art would have had a bachelor's degree in computer science, or a related field, or an equivalent technical degree or equivalent work experience, and an additional one to two years of education or experience in the field of computer security topics such as malware/exploit detection and/or network forensics. More education can supplement practical experience and vice versa.

26. It is my opinion that such a person would have been actively engaged in learning about the field, possibly through formal instruction through the bibliographic resources. By at least May 12, 2015, such a person would have had access to a vast array of print and electronic resources, including at least the documents referenced below.

D. Evidence Considered in Forming My Opinions

27. In the preparation of this declaration, I have reviewed the documents referenced below and any other documents I reference herein, and each of these is a type of material that experts in my field would reasonably rely upon when forming their opinions:

- (1) The document referenced above in Section I.B;
- (2) "JMLR Volume 13" website, archived by the Internet Archive on

- June 13, 2013, available at <https://web.archive.org/web/20130613002437/http://jmlr.org/papers/v13/>, accessed and on February 16, 2023, **Appendix 1005A**;
- (3) Abstract and full text of *Tahan*, archived by the Internet Archive on September 21, 2013, available at <https://web.archive.org/web/20130921035310/http://jmlr.org/papers/v13/tahan12a.html> and <https://web.archive.org/web/20130921100447/http://jmlr.org/papers/volume13/tahan12a/tahan12a.pdf>, accessed and on February 16, 2023, **Appendix 1005B**;
- (4) Early citations to *Tahan*, obtained from Google Scholar, **Appendix 1005C**;
- (5) Library copy of *Menahem*, obtained from the Library of Congress, **Appendix 1006A**;
- (6) Bibliographic and MARC records for *Computational Statistics & Data Analysis* that contains *Menahem*, available at <https://lcn.loc.gov/84642334> from the online catalog of the Library of Congress, accessed and obtained on February 17, 2023, **Appendix 1006B**;
- (7) Library copy of the front matter of the issue of *Computational Statistics & Data Analysis* that contains *Menahem*, obtained from the British Library, **Appendix 1006C**;
- (8) Copy of *Menahem*, archived by the Internet Archive on December 3, 2011, available at <http://web.archive.org/web/20111203120533/https://www.ise.bgu.>

[ac.il/faculty/liorr/malens.pdf](https://www.tau.ac.il/faculty/liorr/malens.pdf), accessed and obtained on February 17, 2023, **Appendix 1006D**;

- (9) Selected citations to *Menahem*, obtained from Google Scholar, **Appendix 1006E**;
- (10) Affidavit of Mr. Nathaniel E. Frank-White of the Internet Archive, received from counsel, **EX1010**.

These records are identified and discussed in this declaration. Experts in the field would reasonably rely on the data described herein to form their opinions.

II. Authenticity and Public Availability of *Tahan* (EX1005)

A. Authentication

28. **EX1005** is a true and correct copy of “Mal-id: Automatic malware detection using common segment analysis and meta-features,” (“*Tahan*”), *Journal of Machine Learning Research*, 13(spring), pp. 949-979, that I obtained from the journal website. When I began preparing this declaration, I searched Google Scholar for *Tahan* and a link from the article entry led me to the JMLR copy at <https://www.jmlr.org/papers/volume13/tahan12a/tahan12a.pdf> . This publisher copy is presented as **EX1005** in this declaration.

29. **EX1005** is a true and correct copy of *Tahan* obtained from the journal website. Page 1 shows the title, authors, abstract and keywords (“computer security, malware detection, common segment analysis, supervised learning”) of this article. The top shows the manuscript was submitted on August 11 and revised on November

11 of 2011, and the article was published on April 12 of 2012 in “Journal of Machine Learning Research 13 (2012) 949-979.” The bottom of the page shows a “2012” copyright date, and the authors are the copyright holders. **EX1005** shows *Tahan* has a total of 31 pages, including eight tables, 10 figures (most of them color), and 41 references.

30. The JMLR copy of *Tahan* (**EX1005**) is in a condition that creates no suspicion about its authenticity. Specifically, this copy is not missing any intermediate pages, the text on each page appears to flow seamlessly from one page to the next, and there are no visible alterations to the document. Moreover, **EX1005** was found in the custody of the journal website, a place where, if authentic, it would likely be found. Accordingly, I see no reason to question the authenticity of this copy of *Tahan* (**EX1005**).

B. Internet Archive Records

31. To determine when *Tahan* became publicly accessible on the Internet, I searched the Internet Archive for records. **Appendix 1005A** is a true and correct copy of the JRML volume 13 website archived by the Internet Archive. I personally located, identified, and obtained this archived file. This is the type of material experts in my field would reasonably rely upon when forming their opinions.

32. Page 1 of **Appendix 1005A** is a screenshot of a Wayback Machine calendar to show the URL of the preserved website is <http://jmlr.org/papers/v13/> and

it was archived 116 times from “10 Jun 2013” to “6 Dec 2022.” Page 1 also shows this particular copy was archived on “JUN 13 2013” and is available at <https://web.archive.org/web/20130613002437/http://jmlr.org/papers/v13/>. Internet Archive is a digital library that archives scholarly materials, cultural artifacts, and other web resources on the Internet and provides the public with free access to its collections. I have used the Internet Archive in the past and am familiar with their practice of assigning an extended URL to an archived resource in the format of [http://www.archive.org/web/\[Year in yyyy\]\[Month in mm\]\[Day in dd\]\[Time code in hh:mm:ss\]/\[Archived URL\]](http://www.archive.org/web/[Year in yyyy][Month in mm][Day in dd][Time code in hh:mm:ss]/[Archived URL]) to show the point of time the archived file was captured. The URL of this archived copy shows this website was archived on “20130613002437” (*i.e.*, June 13, 2013, 00 a.m., 24 minutes, 37 second). Because Internet Archive does not archive web resources as soon as they appear on the Internet, the archive date is not the first public availability date of an archived resource. Furthermore, a web resource will need to be available on the Internet before it can be archived. For these reasons, the archive date in the extended URL means that an archived copy of a web resource was available before, and definitely no later than, the archive date. It is therefore my opinion that this archived copy of this website would have been publicly available by June 13, 2013.

33. Pages 2 to 12 of **Appendix 1005A** are the PDF version of the archived website. Each entry represents a paper, including the title, authors, issue date,

pagination, year of publication, and links to the abstract and the PDF of the paper. Page 4 shows the *Tahan* article was published in “13(Apr):949-979, 2012” and links to the abstract and the PDF version are provided.

34. **Appendix 1005B** is a true and correct copy of the abstract and the full text of *Tahan* archived by Internet Archive. I personally located, identified, and obtained this archived file. This is the type of material experts in my field would reasonably rely upon when forming their opinions.

35. Page 1 of **Appendix 1005B** is a screenshot of a Wayback Machine calendar to show the URL of the archived abstract of *Tahan* is <http://jmlr.org/papers/v13/tahan12a.html>, and the abstract was archived 14 times from “21 Sep 2013” to “29 Oct 2020.” The calendar shows this particular copy was archived on SEP 21 2013” and is available at <http://web.archive.org/web/20130921035310/http://jmlr.org/papers/v13/tahan12a.html> Jul 2013.” The URL of this archived copy shows this abstract was archived on “20130921035310” (*i.e.*, September 21, 2013, at 3 a.m., 53 minutes, 10 seconds).

36. Page 2 of **Appendix 1005B** is a Wayback Machine calendar to shows the URL of the archived article is <http://jmlr.org/papers/volume13/tahan12a/tahan12a.pdf> and the article was archived 21 times from “21 Sep 2013” to “25 Sep 2022.” The calendar also shows this particular copy was archived on “SEP 21 2013” and is available at

<http://web.archive.org/web/20130921100447/http://jmlr.org/papers/volume13/tahan12a/tahan12a.pdf>. The “20130921100447” means this article was archived on September 21, 2013, at 10 a.m., 4 minutes, 47 seconds. Pages 3 to 33 of **Appendix 1005B** is the PDF version of *Tahan*. I have closely compared this copy with the publisher copy (**EX1005**) and found them to be the same.

37. These archived records from the Internet Archive show that volume 13 of JMLR made *Tahan* accessible from the journal website by June 13, 2013, when that website was archived, and the abstract and full text of *Tahan* became publicly accessible by September 21, 2013, when they were archived. It is therefore my opinion that the general public would have been able to access *Tahan* by June 13, 2013, and definitely no later than September 21, 2013, on the Internet.

C. Usage Record

38. Actual usage of a publication is reflected by the papers that make reference to it. The citation history on Google Scholar shows that *Tahan* has been cited 106 times, and eight citations were published in 2013. **Appendix 1005C** is a list I compiled from Google Scholar to demonstrate early usage. The earliest citations appeared in January 2013.

D. Summary of My Opinion on *Tahan*

39. The publisher copy of *Tahan* (**EX1005**) is an authentic copy, the archived website of volume 13 of JRML (**Appendix 1005A**) shows that by June 13,

2013, visitors to the website would have learned that *Tahan* was published in April 2012, and they could follow the links to access the abstract and full text of this article. The archived copies of the abstract and full text of *Tahan* (**Appendix 1005B**) show that by September 21, 2013, *Tahan* was publicly accessible. Based on these records, it is my opinion that *Tahan* was publicly accessible as early as June 13, 2013, and no later than September 21, 2013, on the Internet. Citation history on Google Scholar shows eight citations were published in 2013 and the earliest citation appeared in January 2013.

III. Authenticity and Public Availability of *Menahem* (EX1006)

A. Authentication

40. **EX1006** is a true and correct copy of “Improving malware detection by applying multi-inducer ensemble,” (“*Menahem*”), by Menahem et al., *Computational Statistics & Data Analysis*, vol. 53, no. 4, pp. 1483-1494, that I obtained from the publisher Elsevier. When I began preparing this declaration I searched Google Scholar by the title of *Menahem* for records and a link from the entry led me to the publisher’s webpage for *Menahem* at <https://www.sciencedirect.com/science/article/abs/pii/S0167947308004763>. I then followed the PDF link in the record to obtain a copy. This publisher copy is presented as **EX1006** in this declaration.

41. **EX1006** is a true and correct copy of *Menahem* obtained from the

publisher. Page 1 shows the title, authors, and abstract of this article, and also shows the manuscript was received “23 July 2008,” accepted “4 October 2008,” and became available online on “19 October 2008.” This page shows *Menahem* appears in *Computational Statistics and Data Analysis* 53 (2009) 1483-1494, and is available on ScienceDirect, which is a platform of Elsevier to provide full text articles in science, technology, and other fields. The bottom of the page shows *Menahem* appears in a journal whose ISSN (International Standard Serial Number) is “0167-9473” and the article has a “2008” copyright date, with Elsevier B.V. as the copyright holder, and the doi (digital object identifier) of *Menahem* is “10.1016/j.csda.2008.10.015.” **EX1006** shows *Menahem* has a total of 12 pages, including eight equations, three figures (some color), ten tables, and 35 references.

42. To determine if *Menahem* was available at libraries, I searched WorldCat by the journal title and located a record at <https://worldcat.org/title/962364724> that shows the Library of Congress and the British Library are two of the holdings libraries. I then checked the online catalog of the Library of Congress, and the search results confirmed the holding information, so I obtained a scanned copy of *Menahem* via Wisconsin TechSearch (WTS), a document delivery service based at the University of Wisconsin-Madison. This library copy is presented as **Appendix 1006A** in this declaration.

43. **Appendix 1006A** is a true and correct copy of *Menahem* obtained from

the Library of Congress. Page 1 shows the covers and spine of volume 53 of *Computational Statistics & Data Analysis* that contains “Feb 15-Mar 15 2009, 4-5” and has a call number of “QA 276.A1 C665 Set 1.” Page 2 is the cover of the journal that identifies this issue as “Volume 53, Issue 4, 15 February 2009” and the ISSN as “0167-9473.” The bottom of the page shows this journal is the official journal of iasc of isi, and iasc is the International Association for Statistical Computer, a section of the International Statistical Institute, and the journal is available online at ScienceDirect at www.sciencedirect.com. Page 3 is the issue cover that shows Elsevier of Amsterdam is the publisher. This page also carries a date stamp, but enlarged copies on pages 4 and 5 show that the stamp is not legible. Page 6 is the copyright page that shows a “2009” copyright date for this issue, and B.V. as the copyright holder. Pages 7 to 10 are the table of contents that shows *Menahem* begins on page 1483 and ends on page 1494. Pages 11 to 22 (internal pages 1483 to 1494) are the full text of *Menahem*. I have closely compared this copy of *Menahem* (**Appendix 1006A**) with the publisher copy (**EX1006**) and found them to be the same.

44. The Library of Congress copy of *Menahem* (**Appendix 1006A**) is in a condition that creates no suspicion about its authenticity. Specifically, this copy is not missing any intermediate pages, the text on each page appears to flow seamlessly from one page to the next, and there are no visible alternations to the document. In addition, I have found this copy to be the same as the publisher copy. Moreover,

Appendix 1006A was found in the custody of a national library, a place where, if authentic, it would likely be found. Accordingly, I see no reason to question the authenticity of this copy of *Menahem*.

B. Library of Congress Records

45. **Appendix 1006B** is a true and correct copy of the bibliographic and MARC records for *Computational Statistics and Data Analysis* that contains *Menahem*. I personally located, identified, and obtained these records from the online catalog of the Library of Congress. These are the type of records experts in my field would reasonably rely upon when forming their opinions.

46. The top of **Appendix 1006B** is the bibliographic record that identifies this publication as “periodical or newspaper” and shows the title, publishers, description, ISSN, and other bibliographic details. The “Item Availability” area shows the call number of this journal is “QA276.A1 C665 Set 1” and users can request access at the Jefferson or Adams Building Reading Rooms. The status of the journal issues is “Not Charged” (meaning available for user access), and the “Older receipts” field shows volume 53, no. 4 is included in the library’s holdings. The record also shows unbound issues with the call number “QA276.A1 C665 Copy 1 Unbound issues” can be requested at the Newspaper & Current Periodical Reading Room in Madison Building “LM133”.

47. Pages 6 to 10 are the MARC record. The first six digits of Field 008

show the MARC record was created on “830415” (*i.e.*, April 15, 1983), and the “c19839999” code after the creation date indicates this publication is a continuing resource that began publication in 1983 and has no end date of publication. Subfields “a” and “c” of Field 040 show “LUU” created the original record and subfield “d” shows “DLC” modified the original record to create a copy cataloging record later. “LUU” is the OCLC code for the library of Louisiana State University, and “DLC” is the OCLC library symbol for the Library of Congress, according to the *Directory of OCLC Members* (<https://www.oclc.org/en/contacts/libraries.html>). The serial MARC record has made *Computational Statistics & Data Analysis* that contains *Menahem* discoverable in the catalog of the Library of Congress.

48. Field 022 shows the ISSN of this journal is “0167-9473,” Field 245 shows the journal title, Field 246 shows a variant title, and Fields 260 show the publishers. Field 362 shows the journal began publication in March 1983, and Fields 321 and Field 310 show the publication frequency changed over time. Field 530 indicates the journal is also “issued online via the World Wide Web” and Field 550 shows this journal has been “[t]he Official journal of International Association for Statistical Computer” since January 1994. Field 710 shows the parent organization is included in the record to help users discover this journal.

49. Subjects of this journal are represented in Field 050 by a Library of Congress Classification (LCC) number of “QA276.A1” that represents journals on

“Mathematical statistics” and in Field 082 by a Dewey Decimal Classification number of “519.5/05” that represents journal on “Statistical mathematics.” Subjects are also represented in Fields 650 by three Library of Congress subject heading, with “Statistics,” “Mathematical statistics,” and “Mathematical statistics |x Data processing” (“|x” indicates a topical subdivision) as the main topics, each followed by “Periodicals,” a form subdivision encoded in subfield “v”, to indicate the main topics are treated in a journal.

50. This serial MARC record has made *Computational Statistics & Data Analysis* discoverable in the library’s catalog. Interested users would have been able to conduct subject searches by the LCC number, the DDC number, and the Library of Congress subject headings to locate this journal. They would also have been able to search for it by its ISSN, title, variant title, and the parent organization of the journal.

51. Because the date stamp on the Library of Congress copy of *Menahem* is not legible, I obtained from the British Library the front matter of the journal issue containing *Menahem* to determine a date of public availability. **Appendix 1006C** is the copy obtained from the British Library. I personally located, identified and obtained this copy. This is the type of material experts in my field would reasonably rely upon when forming their opinions.

52. **Appendix 1006C** is a true and correct copy of the front matter of

volume 53, no. 4 of *Computational Statistics & Data Analysis*. Page 1 is the cover that is identical to the cover included in the Library of Congress copy of this issue (**Appendix 1006A**). The cover carries a date stamp indicating the receipt date is “24/02/09” (*i.e.*, February 24, 2009), and the call number of this journal is “3390.625000.” Page 2 is the copyright page, and pages 3 to 6 are the table of contents. The British Library’s copy of the front matter of vol. 53, issue 4 of this journal is identical to the front matter in the Library of Congress copy (**Appendix 1006A**). The date stamp in the British Library copy informs me that the library received this issue on February 24, 2009. As discussed above, because the publisher of a serial makes new issues of the serial available successively, a customary library cataloging practice is not to create MARC records for individual issues of a serial. Instead, libraries rely on a serial check-in system to track the receipt of new issues. The serial check-in process usually takes less than an hour, and one of the steps involves placing a date stamp on the new issue to document the date the issue is checked in. After that, the holdings information of the serial is updated in the library’s catalog so that users know which issues are available for request or access. After serial check-in is completed, the new issue is placed on the shelf with the previous issues of the serial. Because information presented in serials often reflects latest discovery, a general practice of libraries is to make new issues of serials available for user access soon after they are checked in, usually within a week. The date stamp of

on the British Library copy (**Appendix 1006C**) informs my opinion that vol. 53, issue 4 of *Computational Statistics & Data Analysis* that contains *Menahem* was received by the library on February 24, 2009, and after the issue went through the serial check-in process, it would have been publicly accessible no later than March 3, 2009, at the library.

C. Internet Archive Records

53. To determine if *Menahem* was available on the Internet, I searched the Internet Archive for records. **Appendix 1006D** is a true and correct copy of *Menahem*. I personally located, identified, and obtained this copy from the Internet Archive. This is the type of material experts in my field would reasonably rely upon when forming their opinions.

54. **Appendix 1006D** is a true and correct copy of *Menahem*. Page 1 is a screenshot of a Wayback Machine calendar to show the URL of the archived article is <https://www.ise.bgu.ac.il/faculty/liorr/malens.pdf> and the article was archived 12 times from “3 Dec 2011” to “2 Jan 2023.” The calendar also shows this particular copy was archived on “DEC 03 2011” and is available at <http://web.archive.org/web/20111203120533/https://www.ise.bgu.ac.il/faculty/liorr/malens.pdf>. As discussed above, the archive time in this extended URL means this copy was archived on “20111203120533” (*i.e.*, December 3, 2011, at 12 p.m., 5 minutes, 33 seconds).

55. Pages 2 to 22 are the PDF version of *Menahem*. I have closely compared the archived copy (**Appendix 1006D**) with the publisher copy (**EX1006**) and found them to have the same content. The differences are that the archived copy shows five keywords provided by the authors, does not carry page numbers, does not show the corresponding author note, the ISSN of the journal or the doi of *Menahem*. In addition, the references are numbered in the archived copy.

56. The archived copy of *Menahem* (**Appendix 1006D**) informs my opinion that a copy of *Menahem* would have been publicly accessible no later than December 3, 2011, on the Internet.

D. Internet Archive Affidavit (EX1010)

57. My analysis of the archived copy of *Menahem* (**Appendix 1006D**) above is supported by the Declaration of Mr. Nathaniel E. Frank-White of the Internet Archive (**EX1010**) that I obtained from counsel. Mr. Frank-White states that “[t]he Internet Archive is a website that provides access to a digital library of Internet sites and other cultural artifacts in digital form. Like a paper library, we provide free access to researchers, historians, scholars, and the general public” (Page 1, ¶ 2). He further explains that “[t]he Internet Archive assigns a URL on its site to the archived files in the format `http://www.archive.org/web/[Year in yyyy][Month in mm][Day in dd][Time code in hh:mm:ss]/[Archived URL]` aka an ‘extended URL’ ... The date indicated by an extended URL applies to a preserved instance of a file for a given

URL” (Page 1, ¶ 5). He then states that Exhibit B to his declaration “are true and accurate copies of the Internet Archive’s records of the archived files for the URLs and the dates specified in the attached coversheet of each printout” (Page 2, ¶ 7).

58. Mr. Frank-White’s Exhibit B (**EX1010**, Pages 12 to 33) shows the URL and the archived copy of *Menahem* are the same as the URL and the content in the archived copy I obtained and discussed above (**Appendix 1006D**). Exhibit B thus confirms that this archived copy includes a “true and accurate” copy “of the Internet Archive’s records of the archived files” (**EX1010**, Page 1, ¶ 6) and supports my analysis that *Menahem* was indeed archived on December 3, 2011, by the Internet Archive.

E. Citation Record

59. Actual usage of a publication is reflected by the papers that make reference to it. The citation history on Google Scholar shows *Menahem* has been cited at least 174 times. **Appendix 1006E** presents citations from 2009 to 2011 to demonstrate early usage. The earliest citation appeared in 2009.

F. Summary of My Opinion on *Menahem*

60. The “24/02/09” date stamp on the British Library copy of the front matter of the journal issue containing *Menahem* (**Appendix 1006C**), the Library of Congress copy of *Menahem* (**Appendix 1006A**), the bibliographic and MARC records for *Computational Statistics & Data Analysis* that contains *Menahem*

(**Appendix 1006B**), and the Internet Archive archived copy of *Menahem* (**Appendix 1006D**) that shows this article was archived on December 3, 2011, together with my knowledge of customary library cataloging and processing of serials, inform my opinion that the journal issue *containing Menahem* has been publicly accessible at the British Library no later than March 3, 2009, and the archived copy of *Menahem* would have been publicly accessible no later than December 3, 2011. Citation history shows the earliest citation was published in 2009.

IV. Conclusion

61. In signing this declaration, I recognize that the declaration will be filed as evidence in a Petition for *Inter Partes* Review before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to deposition in the case. If deposition is required of me, I will appear at a reasonable time and place to be agreed upon.

62. I hereby declare that all statements made herein on my own knowledge are true and that all statements made on information and belief are believed to be true, and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Date: 8-15-2023

Executed: Ingrid Hsieh-Yee
Ingrid Hsieh-Yee, Ph.D.