

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

ROKU, INC.,
Petitioner,

v.

ANONYMOUS MEDIA RESEARCH HOLDINGS, LLC,
Patent Owner

Case IPR2024-01054
U.S. Patent No. 8,510,768

**DECLARATION OF DR. SAMUEL H. RUSS IN SUPPORT OF PETITION
FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 8,510,768**

Mail Stop "PATENT BOARD"

Patent Trial and Appeal Board
U.S. Patent & Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

Roku EX1003
IPR2024-01054
U.S. Patent No. 8,510,768

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	MY BACKGROUND AND QUALIFICATIONS	2
III.	MATERIALS CONSIDERED.....	5
IV.	RELEVANT LEGAL STANDARDS	7
	A. Obviousness.....	8
	B. Level of Skill in the Art	10
	C. Motivation to Combine and Hindsight Bias	11
	D. Claim Construction	12
V.	TECHNOLOGY BACKGROUND.....	13
	A. Content Identification Was Well-Known	13
	B. Generating Play Streams Was Well-Known.....	14
	C. Scrubbing Data Was Well-Known	16
VI.	OVERVIEW OF THE '768 PATENT	21
	A. Summary of the Disclosure	21
	B. Prosecution History Summary	41
	C. Claim Construction	45
VII.	SUMMARY OF THE APPLIED REFERENCES.....	46
	A. OVERVIEW OF U.S. PATENT APPLICATION PUBLICATION NO. 2004/0260682 TO HERLEY ET AL. ("HERLEY682") (EX1005).....	46
	B. OVERVIEW OF U.S. PATENT APPLICATION PUBLICATION NO. 2005/0193016 TO SEET ET. AL ("SEET") (EX1006)	49
	1. U.S. Provisional Patent Application 60/545,681, ("'681 provisional application") (EX1019) provides support for at least claim 9 of Seet.	54
	2. The '681 provisional application describes the portions of Seet relied upon as prior art.	57
	C. OVERVIEW OF U.S. PATENT NO. 7,333,864 TO HERLEY ("HERLEY864") (EX1007).....	61

D.	OVERVIEW OF U.S. PATENT APPLICATION PUBLICATION NO. 2002/0083060 TO WANG ET AL. (“WANG060”) (EX1008).	72
VIII.	GROUND 1: CLAIMS 1-3, 5, 6, 8-11, 13, AND 14 OF THE ’768 PATENT ARE UNPATENTABLE UNDER 35 U.S.C. § 103 UNDER HERLEY682 AND SEET.....	77
A.	Motivation to Combine Herley682 and Seet	78
B.	Independent Claim 1	81
1.	Preamble [1.P]	81
2.	Element [1.1]	83
3.	Element [1.2]	93
4.	Element [1.3]	97
5.	Element [1.4]	101
6.	Element [1.5]	103
7.	Element [1.6]	106
8.	Element [1.7]	109
9.	Element [1.8]	118
C.	Dependent Claim 2	119
D.	Dependent Claim 3	120
E.	Dependent Claim 5	124
F.	Dependent Claim 6	127
G.	Dependent Claim 8	128
H.	Independent Claim 9	130
1.	Preamble [9.P]	130
2.	Element [9.1]	133
3.	Element [9.2]	133
4.	Element [9.3]	134
5.	Element [9.4]	134
6.	Element [9.5]	135
7.	Element [9.6]	135
8.	Element [9.7]	136
I.	Dependent Claim 10	136

J.	Dependent Claim 11	136
K.	Dependent Claim 13	137
L.	Dependent Claim 14	137
IX.	GROUND 2: CLAIMS 1-6 AND 8-14 OF THE '768 PATENT ARE UNPATENTABLE UNDER 35 U.S.C. § 103 UNDER HERLEY682 IN VIEW OF SEET AND HERLEY864.....	137
A.	Motivation to Combine Herley864 with Herley682 and Seet	137
B.	Independent Claims 1 and 9	140
C.	Dependent Claims 4 and 12	143
D.	Dependent Claims 2, 3, 5, 6, 8, 10, 11, 13, and 14	146
X.	GROUND 3: CLAIMS 7 AND 15 ARE OBVIOUS IN VIEW OF HERLEY682 AND SEET (WITH OR WITHOUT HERLEY864) AND WANG060.....	146
A.	Claims 7 and 15	146
XI.	SECONDARY CONSIDERATIONS OF NON-OBVIOUSNESS.....	150
XII.	CONCLUSION	151

I, Dr. Samuel Russ, declare as follows:

I. INTRODUCTION

1. I have been engaged by Sterne, Kessler, Goldstein & Fox P.L.L.C. on behalf of Roku, Inc. (“Roku”) to provide this Declaration concerning technical subject matter relevant to the *inter partes* review of U.S. Patent No. 8,510,768 (EX1001, “’768 patent”) which is entitled “MEDIA USAGE MONITORING AND MEASUREMENT SYSTEM AND METHOD,” to Jonathan S. Steuer, et al., which is owned by Patent Owner Anonymous Media Research, LLC.

2. I am over 18 years of age. I have personal knowledge of the facts stated in this Declaration.

3. I have reviewed, and am familiar with, the specification, claims, and prosecution history of the ’768 patent. I will cite to the specification of a U.S. Patent using the following format: Column:Line Number(s). For example, the citation to EX1001, 1:1-10, points to the ’768 patent specification at column 1, lines 1-10.

4. In forming my opinions expressed in this Declaration, I have reviewed and, in some cases, relied upon the list of materials provided in Section III. To the best of my knowledge, those exhibits listed in Section III are true and accurate copies of what they purport to be. An expert in the field would

reasonably rely on them to formulate opinions such as those set forth in this Declaration.

5. I am being compensated at my customary consulting rate by Roku for my time spent in connection with this proceeding. My compensation is not dependent on, and in no way affects, the substance of my opinions. Nor is my compensation dependent on the outcome of this proceeding. My findings are based on my education, experience, and background in the fields discussed below.

II. MY BACKGROUND AND QUALIFICATIONS

6. My academic and professional pursuits are closely related to the subject matter of the '768 patent. Based on my education and work experience, I am well qualified to serve as a technical expert in this matter.

7. A summary of my educational and professional experience is set forth in my *curriculum vitae* ("CV"), which is submitted as Exhibit 1004. Exhibit 1004 also includes a list of my publications and the cases in which I have testified at deposition, hearing, or trial. Some of the relevant points in my CV are described below.

8. I received a Bachelor's degree in Electrical Engineering from the Georgia Institute of Technology ("Georgia Tech") in 1986 and a Ph.D. in Electrical Engineering from Georgia Tech in 1991.

9. From 2007 to the present, I have been a member of the faculty of the University of South Alabama as an Assistant and Associate Professor in the Department of Electrical and Computer Engineering. During that time, I have won awards for excellent teaching and have been actively publishing research in home networking and digital video recording (DVR) technologies.

10. I am active in the Institute of Electrical and Electronic Engineers (IEEE) and am a Distinguished Lecturer for the IEEE Consumer Electronics Society. As a consultant, I have conducted briefings for members of the financial community on technology trends in the cable, satellite, and IPTV sectors.

11. From 2000 to 2007, I worked for Scientific-Atlanta (now Cisco's Service Provider Video Tech. Group), where I managed a cable set-top box (STB) design group that designed four STB models, including the Explorer 4200 (non-DVR) and 8300 (DVR) models. Both models sold several million units. As design group manager, I was responsible for managing the design and prototyping activities of the group and for interfacing with other groups (especially integrated-circuit design, procurement, software developers, the factory where prototypes were built, and product managers) and for maintaining the hardware and mechanical development schedule. Since the products were produced in extremely high volumes, the projects had very high visibility in the company, and therefore carried a great deal of responsibility.

12. Also, while at Scientific-Atlanta, I became a staff expert in home networking, conducting demonstrations of wireless video technology and managing a group that developed a new coaxial home networking system. The coaxial system won a Technology and Engineering Emmy® Award in 2013. I became a staff expert in DVR reliability, and led a team that improved the software, hardware, repair, and manufacturing processes. I am a named inventor on over fifty patent applications that were filed while I was at Scientific-Atlanta, twenty-nine of which have issued as U.S. patents, as of the writing of this report. Many have also issued as European and Canadian patents.

13. From 1999 to 2000, I was a Staff Electrical Engineer and then Matrix Manager at IVI Checkmate (now Ingenico), where I managed the hardware design team that completed the design of the eN-Touch 1000 payment terminal. This terminal was in widespread use, for example, at the self-checkout at Home Depot.

14. I also served on the faculty of Mississippi State University from 1994 to 1999 as an Assistant Professor in the Department of Electrical & Computer Engineering, where I taught circuit board design and two-way interactive video classes, among other things.

15. I have also authored at least 15 journal articles and at least 31 conference papers, including a conference paper on digital video recording that

won second place in a “best paper” competition at the 2011 International Conference on Consumer Electronics in Las Vegas, NV.

III. MATERIALS CONSIDERED

16. In formulating my opinions, I have relied upon my training, knowledge, and experience that are relevant to the '768 patent. Furthermore, I have specifically considered the following documents listed below, in addition to other sources cited in this Declaration. I understand that the references are true and accurate copies of what they appear to be.

17. I have been asked to provide technical review, analysis, insights, and opinions regarding the references cited in the Petition and discussed below. My opinions are not only based on these references but also based on my over 30 years of experience as a computer architect, computer system designer, personal computer graphics designer, educator, and executive in the electronics industry.

<i>Exhibit #</i>	<i>Description</i>
1001	U.S. Patent No. 8,510,768 to Steuer et al. (“’768 patent”)
1002	File History of the ’768 Patent (“’768 Patent File History”)
1004	<i>Curriculum Vitae</i> of Dr. Samuel Russ
1005	U.S. Patent Application No. 2004/0260682 to Herley et al. (“Herley682”)
1006	U.S. Patent Application No. 2005/0193016 to Seet et al. (“Seet”)
1007	U.S. Patent No. 7,333,864 to Herley (“Herley864”)

<i>Exhibit #</i>	<i>Description</i>
1008	U.S. Patent Application No. 2002/0083060 to Wang et al. (“Wang060”)
1009	Ingunn Myrtveit, Erik Stensrud, and Ulf H. Olsson “Analyzing Data Sets with Missing Data: An Empirical Evaluation of Imputation Methods and Likelihood-Based Methods,” in IEEE Transactions on Software Engineering, vol. 27, no. 11, 999-1013, November 2001. (“Myrtveit”)
1010	Mong Li Lee, W. Hsu and Vijay Kothari, “Cleaning the spurious links in data,” in IEEE Intelligent Systems, vol. 19, no. 2, 28-33, March-April 2004. (“Lee”)
1011	U.S. Patent No. 5,581,658 to O’Hagan et al. (“O’Hagan”)
1012	Maurico Hernandez and Salvatore Stolfo, “Real-world Data is Dirty: Data Cleansing and The Merge/Purge Problem,” in Data Mining and Knowledge Discovery, 2, 9-37 (1998). (“Hernandez”)
1013	Kamakshi Lakshminarayan, Steven A. Harp, Tariq Samid, “Imputation of Missing Data in Industrial Databases,” Applied Intelligence 11, 259-275, (1999) (“Lakshminarayan”)
1014	Eloi Batlle, Jamue Masip, Enric Gaus, “Automatic Song Identification in Noisy Broadcast Audio,” Audiovisual Institute and Department of Technology, Barcelona, Spain, (2002) (“Batlle”)
1015	“Signature,” Merriam-Webster Dictionary, accessed at https://www.merriam-webster.com/dictionary/signature
1019	U.S. Provisional Application No. 60/545,681, filed February 17, 2004 (“Seet provisional”)

18. This Declaration represents the opinions I have formed to date. I may consider additional documents as they become available or other documents that

are necessary to form my opinions. I reserve the right to revise, supplement, or amend my opinions based on new information.

19. In light of my review of the exhibits, I recognize that the specification of the '768 patent purports to “provide a media measurement system and method that enhances recognition (e.g., in terms of accuracy or efficiency) of the content of a media sample by analyzing information exogenous to the sample.” EX1001, 1:60-64. In particular, the '768 patent purports to “analyze sequence data” to “enhance content identification.” EX1001, 2:1-3.

20. I understand that the '768 patent purports to claim priority to U.S. Application No. 11/139,330, filed May 26, 2005, which claims the benefit of Provisional Application No. 60/574,836, filed May 27, 2004. I am familiar with the technology described in the '768 patent and the state of the art as of the earliest possible priority date, May 27, 2004. I have reviewed the exhibits discussed herein with the May 27, 2004 priority date in mind. My opinions would remain unchanged should the '768 patent not be entitled to the May 27, 2004 priority date and instead be entitled to the later date of May 26, 2005.

IV. RELEVANT LEGAL STANDARDS

21. When considering the '768 patent and stating my opinions, I am relying on legal principles that have been explained to me by counsel.

22. I understand that for a claim to be found patentable, the claims must be, among other requirements, novel and nonobvious from what was known at the time of the invention.

23. I understand that the information that is used to evaluate whether a claim is novel and nonobvious is referred to as prior art.

24. I understand that in this proceeding, Petitioner has the burden of proving that each element of the challenged claims is disclosed by the prior art or would have been obvious based on the prior art.

A. Obviousness

25. I understand that a patent claim is invalid if the claims would have been obvious to a person of ordinary skill in the art (“POSA”) at the time of the invention. I understand that the obviousness inquiry should not be done in hindsight, but from the perspective of a POSA as of the time of invention of the patent claim.

26. I understand that to obtain a patent, the claims must have, as of the time of the invention, been nonobvious in view of the prior art.

27. I understand that a claim is obvious when the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious to a POSA at the time the invention.

28. I understand that to prove that the prior art reference or a combination of prior art references renders a patent obvious, it is necessary to: (1) identify the particular references that, singly or in combination, make the patent obvious; (2) specifically identify which elements of the patent claim appear in each of the asserted references; and (3) explain how a POSA could have combined the prior art references to create the claimed invention and what the combination would be.

29. I understand that to support a conclusion that a prior art reference or a combination of prior art references renders a patent obvious, there must be some documentary evidence. Mere statements about what is basic knowledge or common sense, i.e., common knowledge as a replacement for documentary evidence, is insufficient to support a conclusion of obviousness.

30. I understand that certain objective indicia can be important evidence regarding whether a patent is obvious. Such indicia include: industry acceptance, commercial success of products covered by the patent claims; long-felt need for the invention; failed attempts by others to make the invention; copying of the invention by others in the field; unexpected results achieved by the invention as compared to the closest prior art; praise of the invention by the infringer or others in the field; taking of licenses under the patent by others; expressions of surprise

by experts and those skilled in the art at the making of the invention; and the patentee proceeded contrary to the accepted wisdom of the prior art.

B. Level of Skill in the Art

31. I have been asked to consider the level of ordinary skill in the field that someone would have had at the time the alleged invention. In deciding the level of ordinary skill, I considered the following:

- The educational level of the inventor;
- the types of problems encountered in the field;
- prior art solutions to those problems;
- the rapidity with which innovations are made;
- the sophistication of the technology; and
- the educational level of active workers in the field.

32. In my opinion, a POSA at the time of the invention would have had a bachelor's degree in electrical engineering, computer engineering, computer science, or equivalent degree with at least two years of relevant industry experience, including in digital media content delivery.

33. Throughout my Declaration, even if I discuss my analysis in the present tense, I am always making my determinations based on what a POSA would have known at the time of the invention. Additionally, throughout my

Declaration, even if I discuss something stating “I,” I am referring to a POSA’s understanding.

C. Motivation to Combine and Hindsight Bias

34. I understand that obviousness can be established by combining multiple prior art references to meet each and every claim element, but I also understand that a proposed combination of references can be susceptible to hindsight bias.

35. I understand that references are more likely to be combinable if the nature of the problem to be solved is the same.

36. I understand that if the combination of references results in the references being unsatisfactory for their intended purposes or the combination changes the references’ principle of operation, a POSA would not have a motivation to combine the references.

37. I understand that obviousness may be defeated if the prior art indicates that the invention would not have worked for its intended purpose or otherwise teaches away from the invention. I further understand that a reference teaches away when a POSA, upon reading the reference, would be discouraged from following the path set out in the reference, or would be led in a direction divergent from the path that was taken in the claim. I also understand that teaching away, e.g., discouragement, is strong evidence that the references are

not combinable and/or that a POSA would not be motivated to combine them. I also understand that a disclosure of more than one alternative does not necessarily constitute a teaching away. I understand that the combination does not need to result in the most desirable embodiment.

D. Claim Construction

38. I understand that the claims of a patent are read in light of the specification of the patent as understood by POSA at the time of the invention. I understand that “the time of the invention” refers to the effective filing date of the patent in question, which I am informed by counsel is no earlier than May 27, 2004, for the purposes of this proceeding. Whether or not explicitly stated below, my analysis of the prior art is made as of the time the invention was made.

39. I understand that claim terms are given their plain and ordinary meaning as would be understood by a POSA as of the effective filing date. I understand that an inventor may provide a special meaning for a term, but that she must do so in the specification with precision and clarity.

40. In construing claims, I understand that evidence that is intrinsic to the patent and prosecution history is given more weight than extrinsic evidence. For example, if there are specific statements in the specification that define the invention, those statements are strong evidence of a definition for a term. In this Declaration, I have understood the claims in view of the general principles

outlined above. Unless I indicate otherwise below, I have interpreted each claim term discussed herein to have its plain and ordinary meaning.

V. TECHNOLOGY BACKGROUND

41. The '768 patent is directed to providing a “media measurement system and method that enhances recognition (e.g. in terms of accuracy or efficiency) of the content of a media sample by analyzing information exogenous to the sample.” EX1001, 1:60-64. In particular, the '768 patent purports to “analyze sample sequence data” to “enhance content identification.” EX1001, 2:1-3.

A. Content Identification Was Well-Known

42. The '768 patent admits that content identification systems, including methods and algorithms for content identification, were well known. The '768 patent explains that:

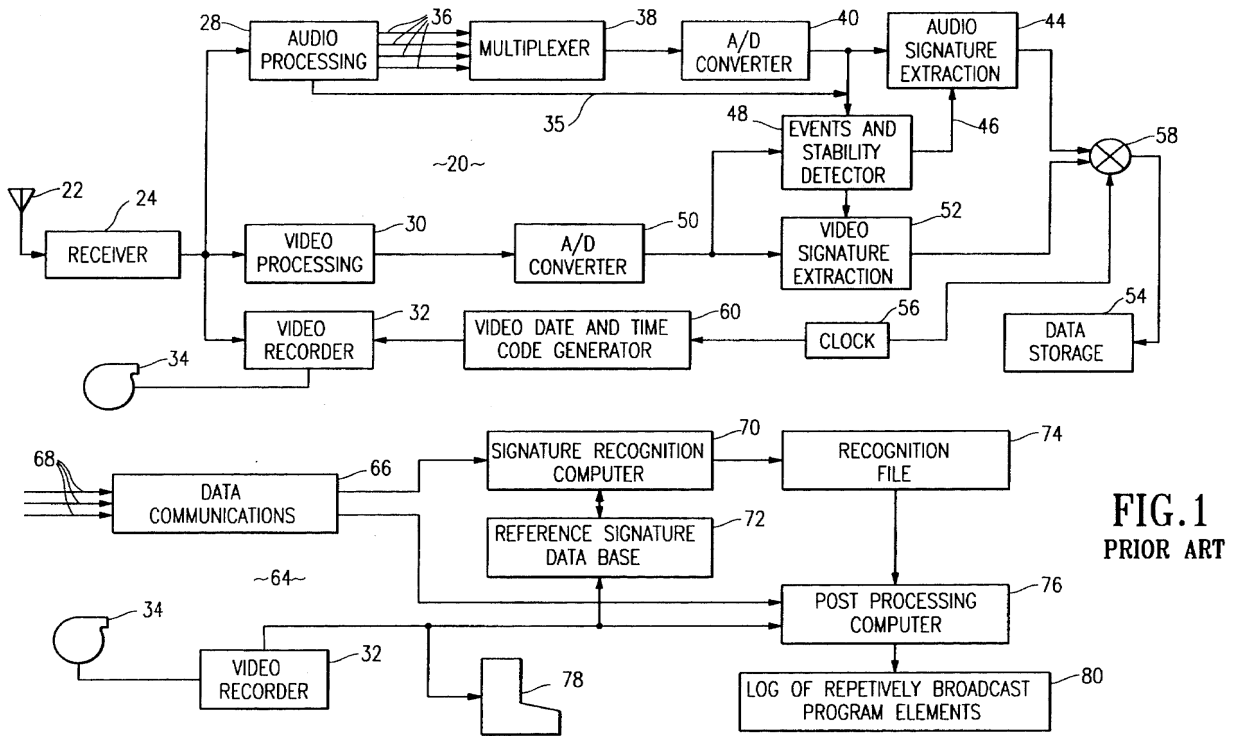
Content recognition system 124 is adapted to implement one or more *known content recognition methods that identify content* by extracting parameters from a media signal and applying an algorithm to those parameters to search for a content match. Those skilled in the art will recognize that *many such methods and algorithms exist*. One such example is described in U.S. patent application Ser. No. 09/839,476 entitled “System and Methods for Recognizing Sound and Music Signals in High Noise and Distortion” by Wang et al. and published Jun. 27, 2002 with publication number US2002/0083060. Aspects of other such examples are described in the following published PCT

applications by Konin-Klijke Philips Electronics N.V.: “Fingerprinting Multimedia Contents” (WO2004/044820, published May 27, 2004), “Fingerprint Extraction” (WO2004/030341 published Apr. 8, 2004), and “Improvements in and Relating to Fingerprint Searching” (WO2004/040475 published May 13, 2004).

EX1001, 5:42-62 (emphasis added).

B. Generating Play Streams Was Well-Known

43. The ‘768 patent explains that a “raw play stream” functionally relates “a sequence of samples to content identification results.” EX1001, 6:14-16, 8:1-3. By this definition, the generation of play streams in content identification systems was well known by a POSA in the art by the time of the ‘768 disclosure. For example, U.S. Patent No. 5,581,658 to O’Hagan (“O’Hagan,” EX1011), which was filed in 1993, and issued in 1996 (almost two decades prior to the ‘768 patent), describes the generation of play streams as being known to a POSA at that time – over three decades ago. O’Hagan describes play stream generation in relation to Figure 1, reproduced below, and which O’Hagan identifies as prior art. EX1011, 4:64-6:19.



EX1011, FIG. 1.

44. O'Hagan describes the automatic identification of broadcast performances, such as music and commercials. EX1011, 1:9-12. Station monitoring unit 20 receives radio or television signals at antenna 22 and processes the signals to extract audio or video signatures representative of the received signals. EX1011, 4:65-5:42. The received audio signals pass through audio processing circuitry 28, multiplexer 38, A/D converter 40, and audio signature extraction circuit 44, which outputs an audio signature. EX1011, 5:10-21. The received video signals pass through video processing circuitry 30, A/D converter 50, and video signature extraction circuit 52, which outputs a video

signature. EX1011, 5:21-30. The extracted audio and video signatures are stored along with their timestamps in a storage medium 54. EX1011, 5:30-36. The extracted audio and/or video signatures are then identified by comparison to a library of reference signatures. EX1011, 5:40-51. O'Hagan's prior art utilizes a signature recognition computer 70 to compare extracted signatures to a library of reference signatures stored in memory 72, and "searches for instances of successive recognition of extracted signatures to identify a particular program" and then classifies and stores the signatures in recognition file 74. EX1011, 5:52-64. At a minimum, recognition file 74 comprises a play stream consistent with the '768 patent's description. Further, O'Hagan discloses that the recognition file 74 is further provided to a post-processing computer 76 which, in turn, produces a file or log 80 of identified programs along with their start times and durations if successive recognized extracted signatures are detected. EX1011, 5:64-6:3. Additionally, post processing computer 76 assigns a code to non-recognized or unmatched signatures for later manual review and identification. EX1011, 6:7-10.

45. O'Hagan evidences that the generation of play streams in the content identification art was well known three decades ago – in the mid-1990s.

C. Scrubbing Data Was Well-Known

46. The '768 patent defines "scrubbing" as "reviewing for missing or incorrect data." EX1001, 8:34-35. This concept was notoriously well known at

the time of the '768 patent. By this definition, scrubbing data was well known by a POSA in the art decades prior to the time of the '768 patent disclosure, and a POSA would have been familiar with these well-known data error correction and data imputation techniques, inasmuch as these concepts were well-published and well-understood. For example, the publications to Lee (EX1010) and Myrtveit (EX1009) describe examples of known data imputation techniques. Stated differently, references like these make clear that scrubbing data was a well-known technique that would have been within the knowledge and skillset of a POSA.

47. Lee notes that data quality problems can arise from a variety of sources. EX1010, 28. “When mining these data warehouses for decision making information, making logical and well-informed decisions becomes problematic if the data’s quality is doubtful.” EX1010, 28. Lee teaches that “clean data is crucial” for many database implementations, including data mining. Lee notes “[d]ata cleaning refers to processes for detecting and removing errors and inconsistencies from data.” EX1010, 28. “Research in this area as included outlier detection, noise handling for classification, and duplicate elimination.” EX1010, 28. “Given the ‘garbage in, garbage out’ principle, clean data is crucial for database integration, data warehousing, and data mining, thus leading to data cleaning’s strong association with these research areas.” EX1010, 28.

48. Lee further explains that “[d]ata cleaning refers to processes for detecting and removing errors and inconsistencies from data.” EX1010, 28. Lee notes a type of erroneous data is known as ‘spurious links’ and Lee identifies a “context-based solution” to identify data that needs correcting. EX1010, 28.

49. Lee describes automated software-based methods for identifying and correcting data. EX1010, 30. “We implemented our algorithms in Java and ran the experiments on a Pentium 4 1.6-GHz system with 256 Mbytes of RAM running Windows XP Professional. We used IntelliClean4 on the three data sets to find record sets with high similarity in the spurious attribute, and we used DM2-CBA7 to generate association rules to determine the context attributes.” EX1010, 30.

50. Myrtveit discusses automated software-based imputation methods for data sets containing missing data. EX1009, 999. As described in Myrtveit, “imputation-based techniques” are one type of “missing data techniques (MDTs)” to handle missing data. EX1009, 1001. One such imputation technique is “similar response pattern imputation (SRPI)” in which the most similar entry without missing data is determined, and those values are copied into the missing data. EX1009, 1002. In SRPI, “the idea is to identify the most similar project without missing [data] and to copy the values of this project to fill the [data] holes in the project with the missing [data].” EX1009, 1002. Myrtviet describes another technique known as “[h]ot deck imputation (HDI)” which “replaces

missing values by drawing from an *estimated* distribution for each missing value.” EX1009, 1002. Myrtviet teaches “[i]t is common [in hot deck imputation] to use the *sample* distribution [...] as the distribution (or deck) to draw from.” EX1009, 1002, 1003. Continuing, Myrtviet teaches “SRPI belongs to the hot deck class,” and “picks the nearest neighbor rather than drawing from the sample at random.” EX1009, 1003. Myrtveit articulates clearly that “[i]mputation-based methods replace missing values by suitable estimates.” EX1009, 1001. In my opinion, a POSA would have immediately understood the desirability and advantages of replacing missing data values with suitable data estimates in order to enhance the integrity and reliability of the data, given the known problem of “dirty data” and known desirability for “cleaned” or “scrubbed” data.

51. Looking back in time, the article to Hernandez, entitled “Real-world Data is Dirty: Data Cleansing and the Merge/Purge Problem” and published in 1998 (more than two decades prior to the ’768 patent) also describes the known problem of dirty data. Hernandez notes that “[e]rrors due to data entry mistakes, faulty sensor readings or more malicious activities, provide scores of erroneous data sets that propagate errors in each successive generation of data.” EX1012, 10. Hernandez describes a basic data cleansing solution utilizing a sorted neighbor method (considering neighbor data values to resolve data errors). EX1012, 11-18.

52. The Lakshminarayan article, entitled “Imputation of Missing data in Industrial Databases” and published in 1999, further evidences the well-known problem of dirty data and well-known data scrubbing techniques. The Lakshminarayan article discusses well-known approaches for data imputation in incomplete data structures. Lakshminarayan notes “[i]ncomplete data is an unavoidable problem in dealing with most of the real world data sources, and various methods for dealing with such data have been developed.” EX1013, 261. In the article, Lakshminarayan concludes that the mining of databases is often hampered by missing and erroneous data entries, and notes reasons for completing incomplete databases and summarizes well-known statistical approaches for missing data imputation. EX1013, 274. Broadly, Lakshminarayan provides a brief overview of existing methods for dealing with missing data, as roughly grouped following known taxonomy as 1) ignoring and discarding incomplete records and attributes, 2) parameter estimation in the presence of missing data, and 3) imputation procedures. EX1013, 261-262.

53. Finally, I note the 2002 article by Batlle entitled “Automatic Song Identification in Noisy Broadcast Audio” as exemplifying that it was well known at the time of the ’768 patent that a problem with broadcast audio data is that the data is dirty. Batlle notes “[o]ne of the main problems with broadcast audio is that the received audio suffers several transformations before reaching the listener []

and, therefore, the original and the broadcast songs are very different from the signal point of view.” EX1014, Abstract. This “dirty” data can impede proper automatic song identification. Although the Batlle reference pursues a different solution to the problem of dirty data, Batlle evidences a POSA knew that data cleansing was, at the time of the ’768 patent, known as a necessary component in audio content identification systems, especially when analyzing broadcast audio.

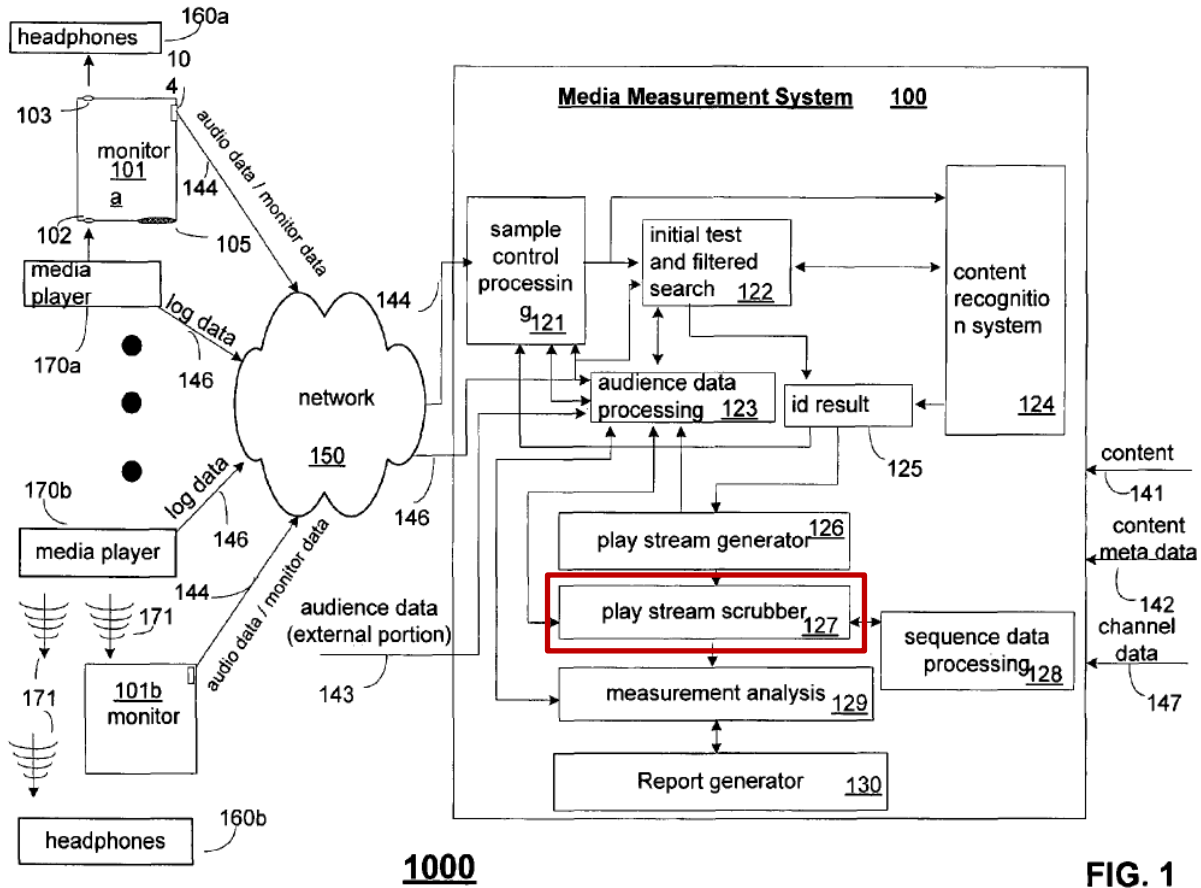
VI. OVERVIEW OF THE ’768 PATENT

54. With the background discussed in Section V in mind, I now turn to the ’768 patent, described below.

A. Summary of the Disclosure

55. The ’768 patent purports to provide a media measurement system and method that enhances recognition (e.g. in terms of accuracy or efficiency) of the content of a media sample by analyzing information exogenous to the sample. EX1001, 1:60-64. In particular, the ’768 patent purports to “analyze sample sequence data” to “enhance content identification.” EX1001, 2:1-3.

56. Figure 1 illustrates a media usage monitoring and measurement system 100 and monitoring devices 101. EX1001, 2:66-3:3, FIG. 1.



EX1001, FIG. 1 (annotated).

57. The '768 patent explains that a “monitoring device 101a includes a microphone 105, a media player port 102, a headphone port 103, and a data upload port 104.” EX1001, 3:4-6. Audio data captured by the monitor 101a, along with associated monitor data (e.g. device ID) (collectively, “audio data/monitor data” 144), is uploaded to system 100. EX1001, 3:38-40. System 100 is also adapted to receive known content 141, which may include media files associated with a known piece of media content (e.g. a song, movie, commercial,

television show, video game, etc.) and corresponding content meta data 142.

EX1001, 4:13-16.

58. Media measurement system 100 includes sample control processing module 121 that receives the audio data and monitor data (e.g. device ID), divides the audio data into samples if the audio data was not previously divided, and determines which samples should be submitted for further content recognition testing. EX1001, 4:49-57. “Initial test and filtered search module 122 conducts one or more initial identification attempts.” EX1001, 4:63-64. “After completing the initial identification attempt, filtered search module 122 passes the identification result to ID result module 125.” EX1001, 5:29-31.

59. “If an identification has been made, the result is passed to play stream generator 126 and to sample control processing module 121.” EX1001, 5:31-33. “Play stream generator 126 generates a raw play stream relating a sequence of samples to content identification results.” EX1001, 6:14-16. The raw play stream is then processed by play stream scrubber 127, which analyzes the sequence of sample identification results utilizing sequence data processing module 128. EX1001, 6:16-20. Based on the analysis, play stream scrubber 127 generates a clean play stream and clean play list for use by measurement analysis module 129. EX1001, 6:20-25. Results of the media measurement analysis can then be obtained via report generator 130. EX1001, 6:25-27.

60. Figure 2 illustrates a media measurement method 200 carried out by system 100 of Figure 1. EX1001, 6:62-65. The '768 patent explains that audio data is uploaded to the system 100 by monitoring device 101a, which includes a microphone 105 and a media player port 102 to capture audio energy when an audience member is consuming media from media player 170a. EX1001, 3:4-37, FIG. 2.

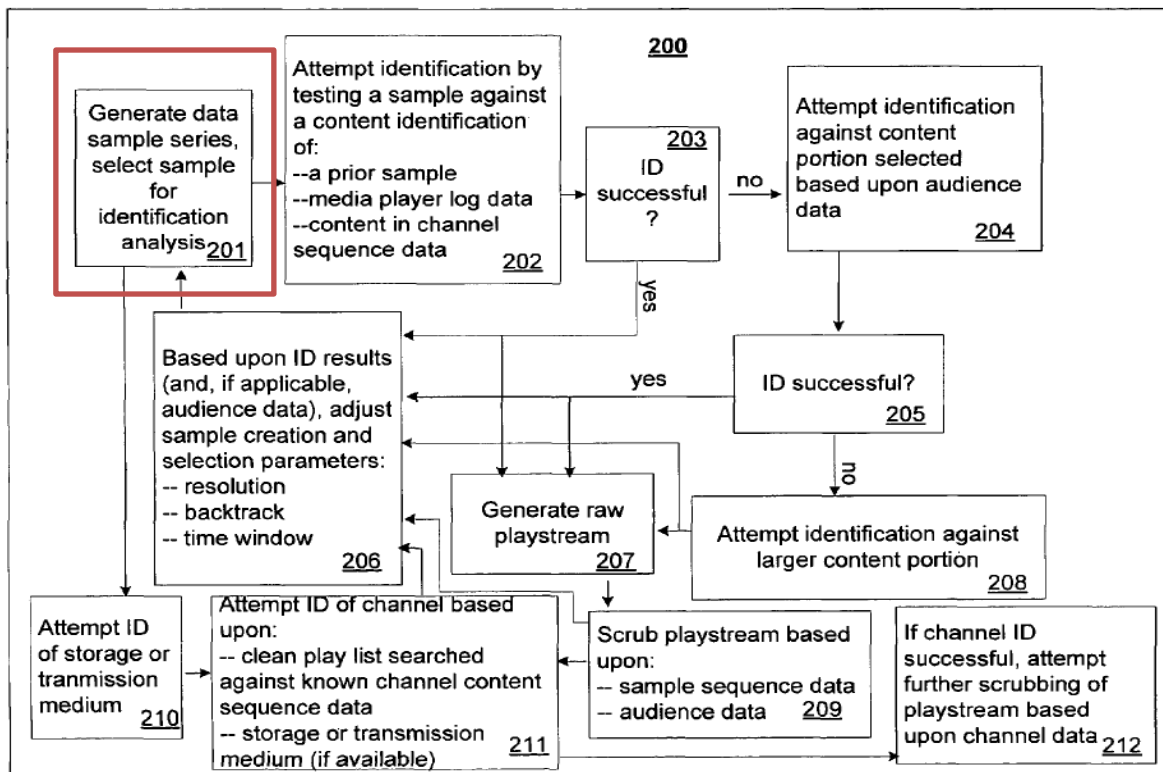


FIG. 2

EX1001, FIG. 2 (annotated).

61. At step 201, audio data is received from a monitoring device. At step 202, the initial test and filtered search module 122 attempts to identify a sample in a series of samples via initial test methods by determining whether the audio

data in the tested sample matches data associated with the content ID

corresponding to the most recently identified sample. EX1001, 7:17-21.

62. If the sample is identified at step 202, the result is provided to audio sample adjustment module (step 206) and play stream generator 126 (step 207). EX1001, 7:24-27. If not, then step 204 searches for a content match against targeted content selected based at least in part upon additional audience data (step 208). EX1001, 7:27-33.

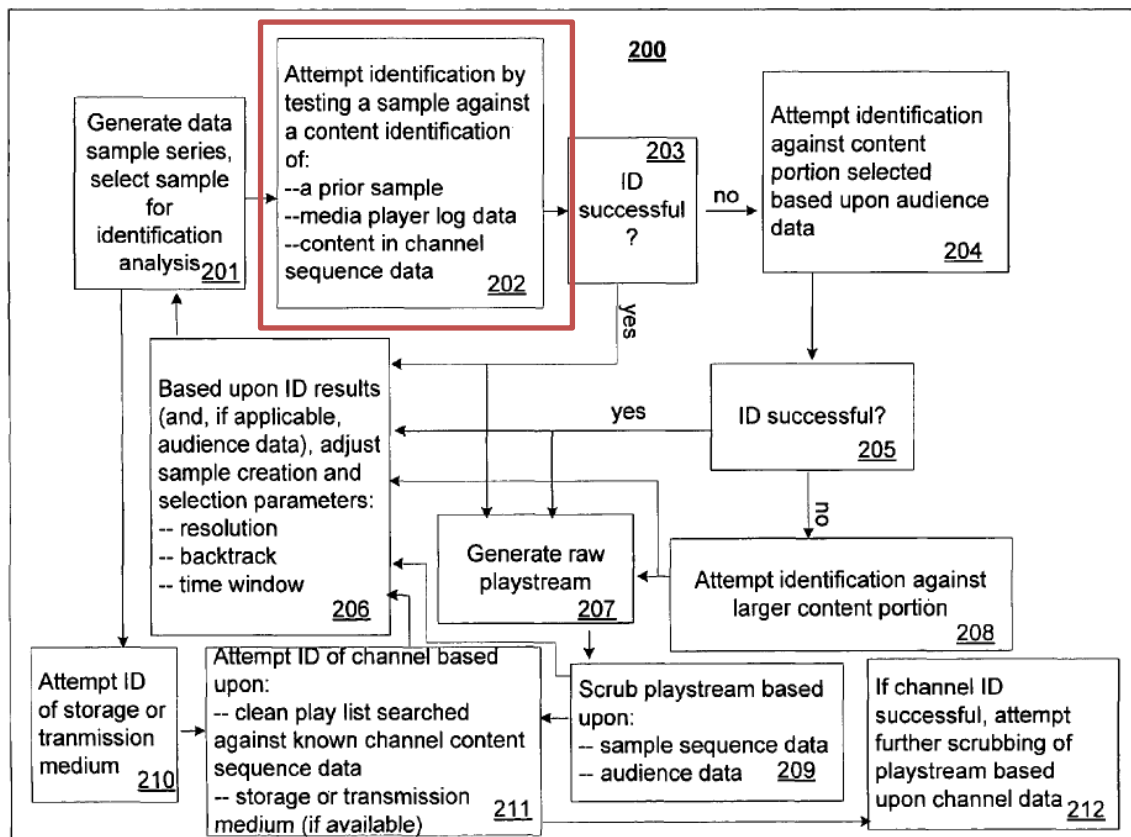


FIG. 2

EX1001, FIG. 2 (annotated).

63. If the sample is identified at 204, the result is provided to audio sample adjustment module (step 206) and play stream generator 126 (step 207). EX1001, 7:40-41. If not, then step 208 searches for a content match against a larger portion of the content. EX1001, 7:41-42.

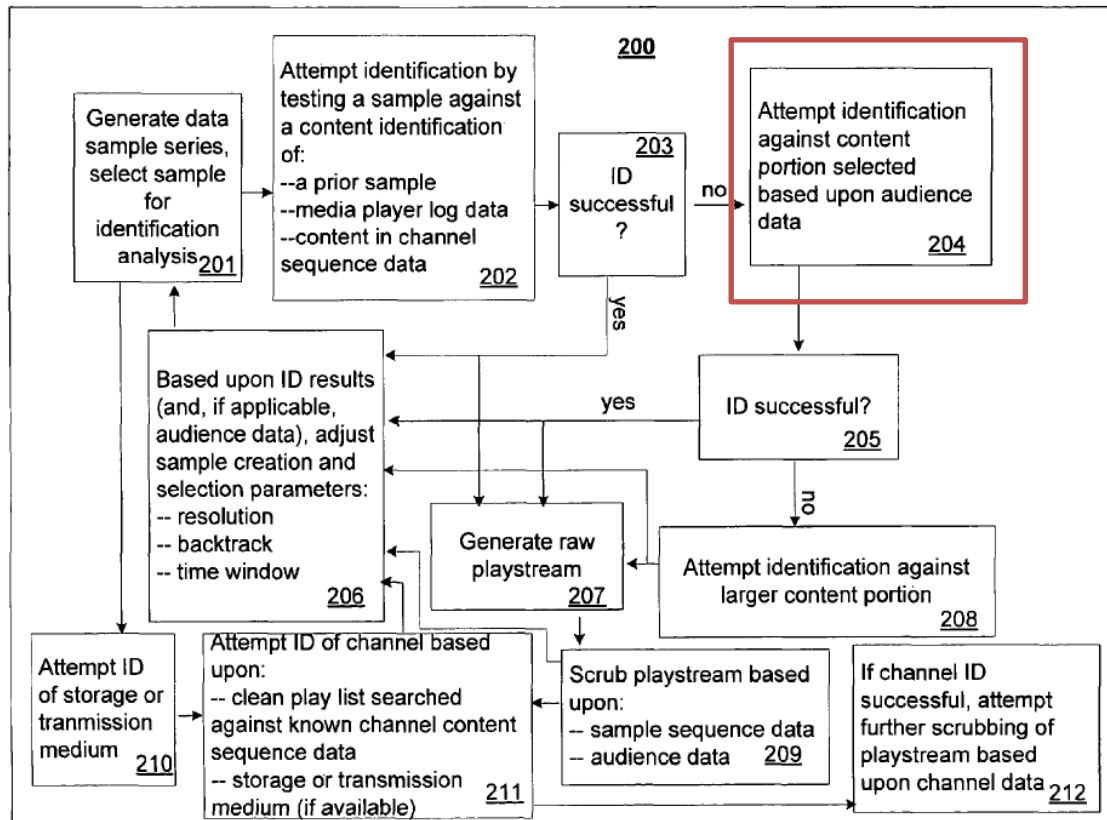


FIG. 2

EX1001, FIG. 2 (annotated).

64. The '768 patent explains that there are two possible results of steps 202, 204, and 208: an identification has been made or no identification has been made. EX1001, 7:42-45. Whatever the result is at step 208 (identification or no identification) that result is passed to steps 207 and 206. EX1001, 7:37-38.

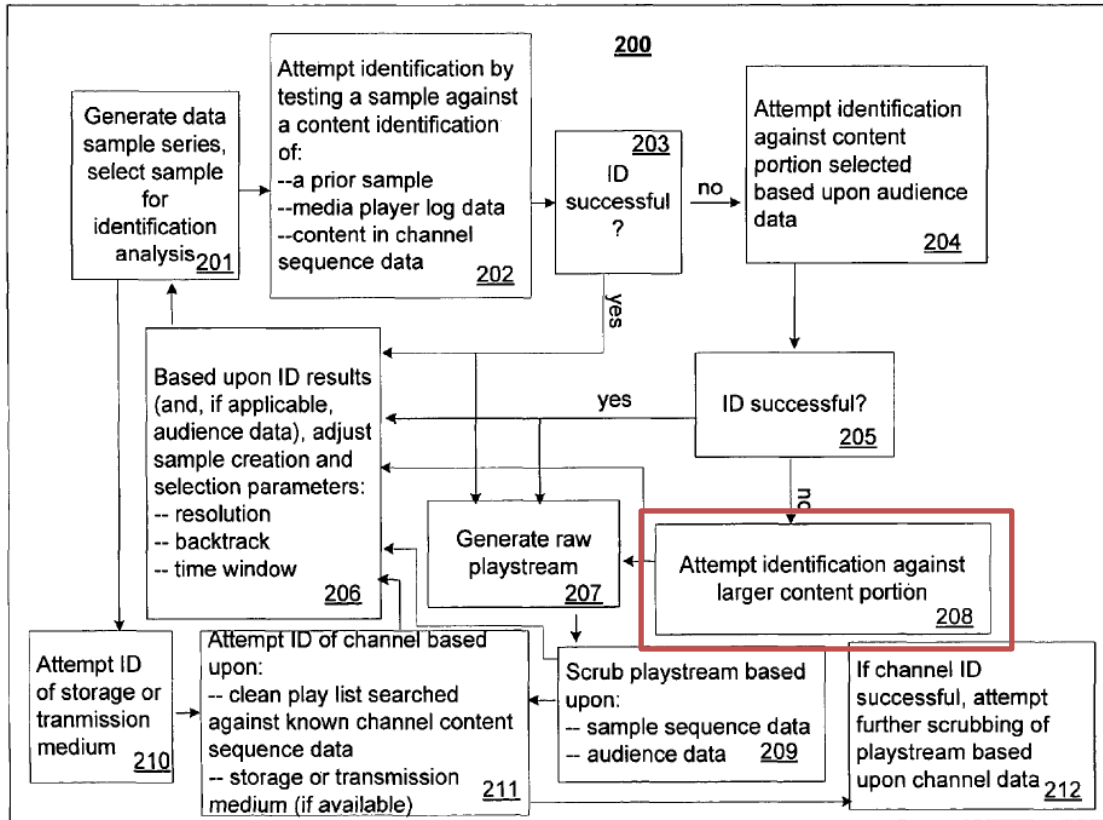


FIG. 2

EX1001, FIG. 2 (annotated).

65. Steps 202, 204, and 208 illustrate three attempts of sample identification and comprise content recognition system 124. The '768 patent explains, however, that the content recognition system 124 did not implement a new or novel method or technique of identifying content but instead simply “implement[s] one or more known content recognition methods that identify content by extracting parameters from a media signal and applying an algorithm to those parameters to search for a content match.” EX1001, 5:38-42. Indeed, the '768 patent explains that:

Those skilled in the art will recognize that many such methods and algorithms exist. One such example is described in U.S. patent application Ser. No. 09/839,476 entitled “System and Methods for Recognizing Sound and Music Signals in High Noise and Distortion” by Wang et al. and published Jun. 27, 2002 with publication number US2002/0083060. Aspects of other such examples are described in the following published PCT applications by Konin-Klijke Philips Electronics N.V.: “Fingerprinting Multimedia Contents” (WO2004/044820, published May 27, 2004), “Fingerprint Extraction” (WO2004/030341 published Apr. 8, 2004), and “Improvements in and Relating to Fingerprint Searching” (WO2004/040475 published May 13, 2004). Such distinct methods and algorithms may have varying degrees of efficiency, accuracy, and applicability to certain types of content (or different settings of a particular algorithm might have different efficiency depending on content type). Thus, in some embodiments, a system such as system 100 may intelligently select among multiple such algorithms (or multiple settings of a single algorithm) based on expected content characteristics.

EX1001, 5:42-62.

66. As I previously noted, regardless of the whether sample identification is made or not made in steps 202, 204, and 208, the results of the identification

attempts are passed to audio sample adjustment module 206 and raw play stream generator 207. At step 207, a raw play stream is generated indicating a series of samples and corresponding identification results. EX1001, 6:14-16, 8:1-3, FIG. 2.

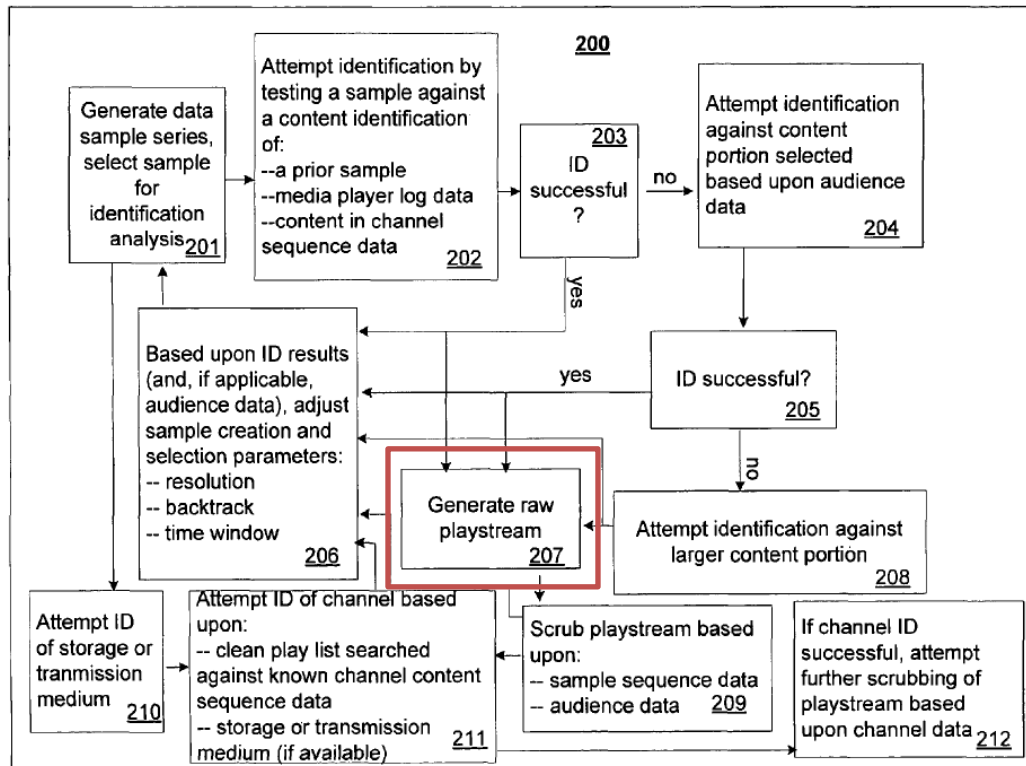


FIG. 2

EX1001, FIG. 2 (annotated).

67. A POSA would understand that Figure 4 illustrates a raw play stream 410, generated at step 207, which includes a sequence of analyzed samples ('row' column; e.g., 1), log times ('log time' column; e.g., t_0), and corresponding identification results ('content ID' column; e.g., song 51,) and referred to as sample sequence data ('ssd')." EX1001, 9:12-31.

68. I note that the '768 patent indicates that the data representations in the raw play stream may be expressed in different terms. For example, “the times in list 410 may be relative to an arbitrary value (e.g., ‘ t_0 ’) rather than representing an absolute time; however, in many instances, a log time will in fact indicate the actual time the sample started to be captured.” EX1001, 9:27-31.

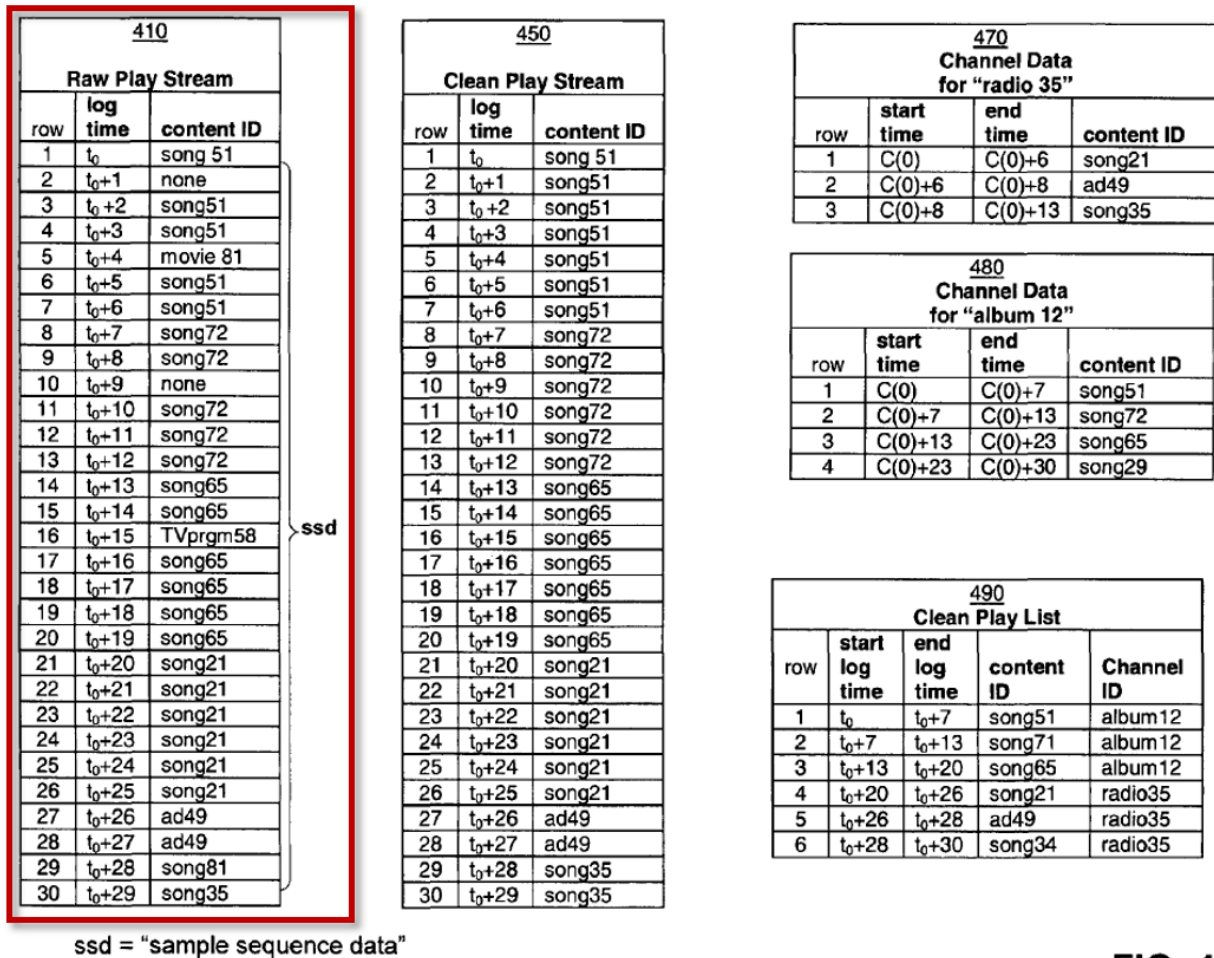


FIG. 4

EX1001, FIG. 4 (annotated).

69. Turning to step 209, play stream scrubber 127 scrubs the raw play stream based upon the sample sequence data and audience data, and creates a clean play stream. EX1001, 6:20-25, 8:3-6, FIG. 2.

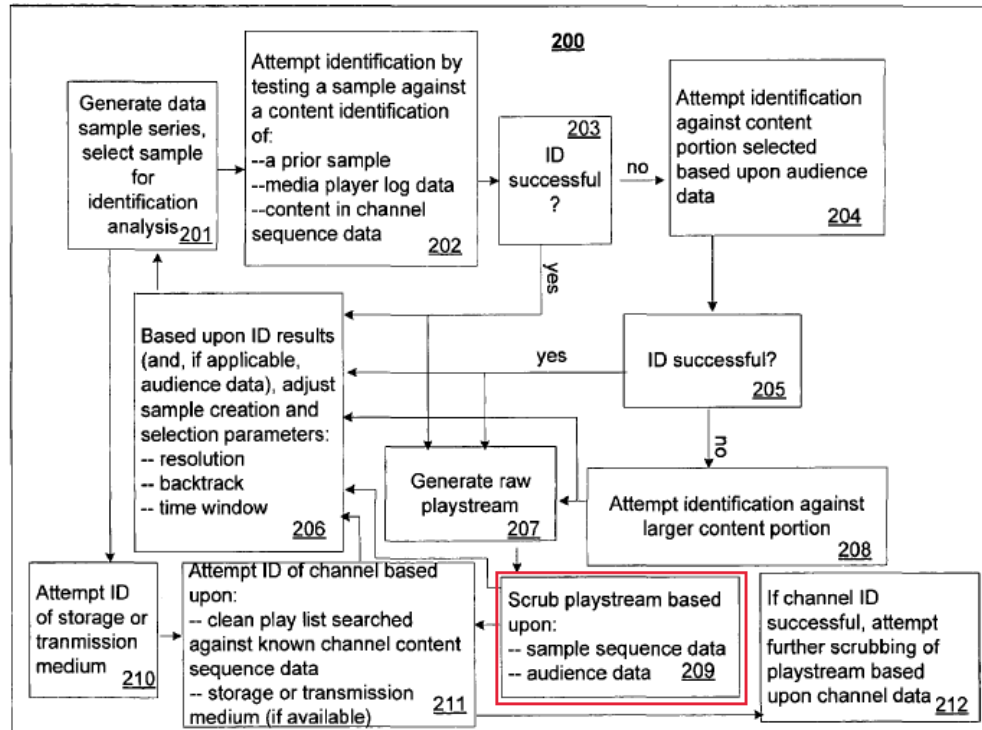


FIG. 2

EX1001, FIG. 2 (annotated).

70. During this scrubbing process, the raw play stream 410 is reviewed for missing or incorrect data in order to generate a clean play stream 450.

EX1001, 8:32-35, 9:12-51, FIG. 4. I note that the '768 patent provides no additional details about how the “scrubbing” is achieved. The '768 patent does not provide details regarding whether the scrubbing is manually performed by a

user or whether any reviewing for missing or incorrect data happens with or without user intervention.

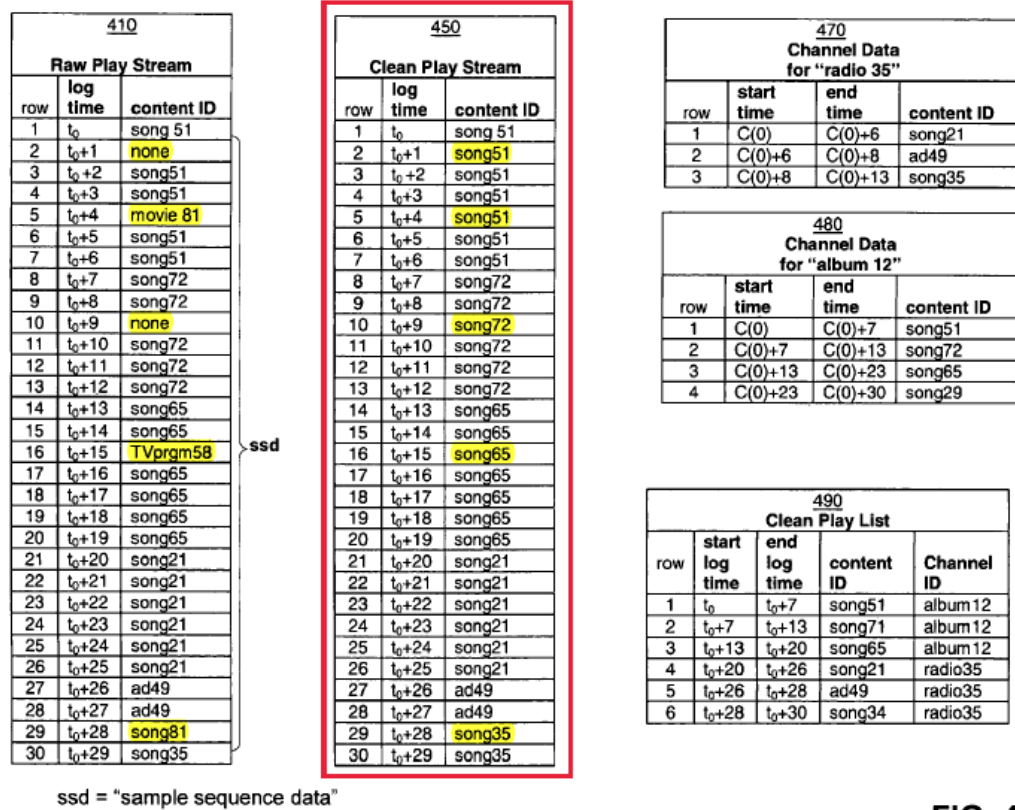


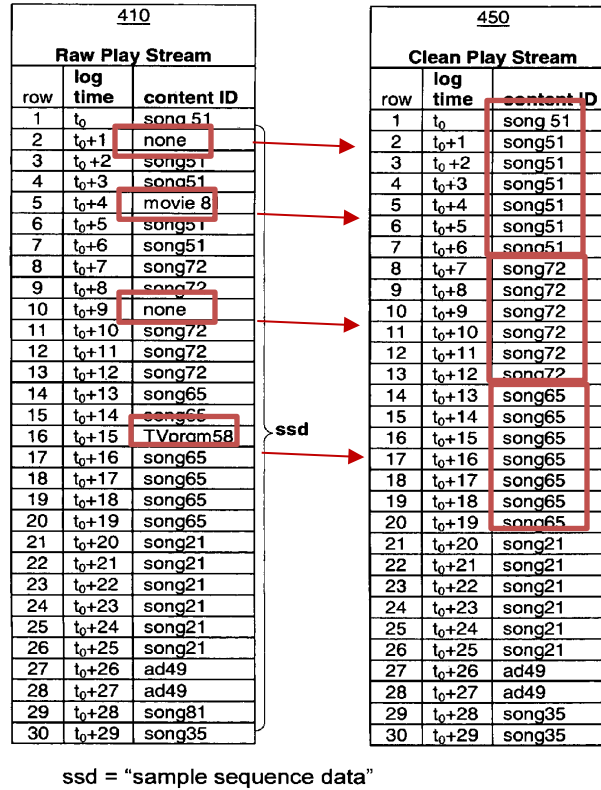
FIG. 4

EX1001, FIG. 4 (annotated).

71. The '768 patent describes that "[s]ample sequence data may be used to "scrub" Raw Play Stream 410." EX1001, 9:32-33. "By analyzing the sequence of data on Raw Play Stream, it is possible to locate and correct any identifications that are likely to be incorrect or to supply missing identifications where the system has not otherwise been able to provide an identification." EX1001, 9:33-37.

72. Continuing, I note the disclosure that “[s]ample sequence data can be used to derive portions of Clean Play Stream 450 from Raw Play List 410 in the following manner.” EX1001, 9:38-40. “Referring to Raw Play Stream 410, the content IDs associated with samples 1, 3, 4, 6, and 7 all indicate the same media content referenced as ‘song 51.’” EX1001, 9:40-42. “The content ID value associate[d] with sample 2 indicates that no content has yet been successfully associated with that sample and the content ID value associated with sample 5 indicates ‘movie81.’” EX1001, 9:42-44. “Given this pattern of sample sequence data, it is reasonable to assume that the person using the corresponding monitor listened to song51 through the time period associated with the audio data of samples 1-7, and thus, the ‘scrubbed’ version of the play stream, i.e. Clean Play Stream 450, indicates song51 for each of samples 1-7.” EX1001, 9:45-51, FIG. 4.

73. Thus, as described, the ’768 patent’s “scrubbing” process analyzes the raw play stream 410 for missing or suspect data entries, and replaces the identified missing or suspect data entries with data that is based upon a reasonable assumption of what the data should be, given a particular sample sequence of data. This describes known data imputation techniques.



EX1001, FIG. 4 (annotated).

74. Continuing, the '768 patent describes "[i]nformation about the sequence of content in a play stream can be used to generate a play list and identify a channel, and the identified channel data can also potentially be used to help further 'scrub' a play stream." EX1001, 9:52-55. "With respect to further 'scrubbing' the Raw Play Stream 410, in this example, channel data may be used to correct the apparent misidentification of sample 29." EX1001, 10:5-7. "Note that the sample sequence data alone in this example does not provide a clear basis for correcting sample 29." EX1001, 10:8-9. "Sample 28 is identified as 'ad49,' sample 29 as 'song81,' and sample 30 as 'song35.'" EX1001 10:9-10. "However,

given the sequence ‘song21, ad49, and song35’ is associated with a known sequence of the radio broadcast referenced as ‘radio 35,’ there is a basis in channel data 470 for believing that the proper content ID for sample 29 is ‘song35’ rather than ‘song81.’” EX1001, 10:11-15, FIG. 4. I note that the sequence of “song21, ad49, and song35” is illustrated in Figure 4 as Channel Data for “radio 35” in table 470. As I noted above, this passage of the ’768 patent also describes known data imputation techniques.

75. The ’768 patent describes another embodiment where data from yet another data source may be utilized to “scrub” the raw data stream. “Audience data may also be used in the ‘scrubbing’ process.” EX1001, 10:28-29. “For example, if ‘song81’ is radically different from the user’s known consumption habits and ‘song35’ is within the user’s known consumption habits, another basis might be provided for suspecting that ‘song35’ rather than ‘song81’ is the correct content ID for sample 29 of Play Stream 410.” EX1001, 10:29-33, FIG. 4. It is my opinion that these two examples, when viewed as alternatives, suggest to a POSA that the particular source or origin of the data utilized to “scrub” the raw play stream is not critical to the scrubbing process.

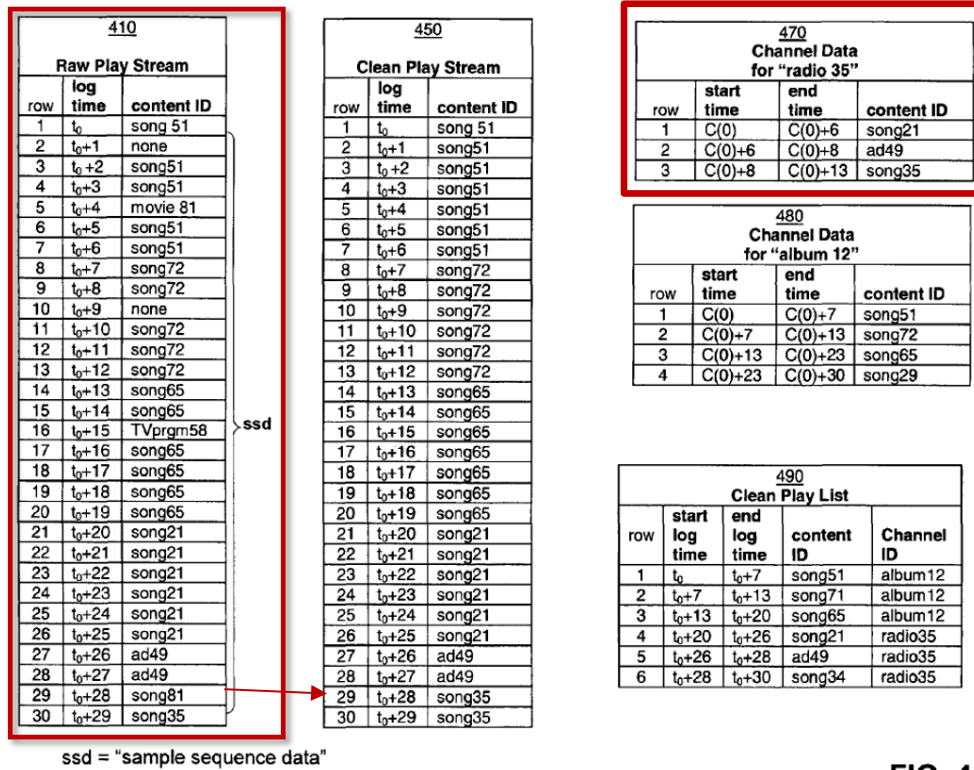


FIG. 4

EX1001, FIG. 4 (annotated).

76. Further, I note that there may be multiple “scrubbing” passes to generate the clean play stream and that additional “scrubbing” can be utilized to generate clean play list 490, illustrated in Figure 4. “Clean play list 490 is generated by a scrubbing step such as step 209 of FIG. 2, or steps 301, 302, or 305 of FIG. 3; Channel data 470 and Channel data [480].” EX1001, 9:12-16, FIG. 4.

77. I note that there appear to be several typographical errors related to Figure 4. The ’768 patent states “, “[p]lay stream generator 126 generates a raw play stream relating a sequence of samples to content identification results (for an example of a portion of a raw play stream, see FIG. 3 and accompanying text).

EX1001, 6:14-17. I presume the reference should have been to Figure 4 which shows “raw play stream 410.” The ’768 patent’s reference to 490 in 9:16 is presumed to be a typographical error and should be 480. Further, based on the context of Figure 4, in Clean Play List 490, it appears that “song71” should be “song72” and “song34” should be “song35”. These are highlighted in my annotated FIG. 4 below.

410 Raw Play Stream			450 Clean Play Stream			470 Channel Data for "radio 35"				480 Channel Data for "album 12"				490 Clean Play List				
row	log time	content ID	row	log time	content ID	row	start time	end time	content ID	row	start time	end time	content ID	row	start log time	end log time	content ID	Channel ID
1	t ₀	song 51	1	t ₀	song 51	1	C(0)	C(0)+6	song21	1	t ₀	t ₀ +7	song51	1	t ₀	t ₀ +7	song51	album 12
2	t ₀ +1	none	2	t ₀ +1	song51	2	C(0)+6	C(0)+8	ad49	2	t ₀ +7	t ₀ +13	song71	2	t ₀ +7	t ₀ +13	song71	album 12
3	t ₀ +2	song51	3	t ₀ +2	song51	3	C(0)+8	C(0)+13	song35	3	t ₀ +13	t ₀ +20	song65	3	t ₀ +13	t ₀ +20	song65	album 12
4	t ₀ +3	song51	4	t ₀ +3	song51					4	t ₀ +20	t ₀ +26	song21	4	t ₀ +20	t ₀ +26	song21	radio35
5	t ₀ +4	movie 81	5	t ₀ +4	song51					5	t ₀ +26	t ₀ +28	ad49	5	t ₀ +26	t ₀ +28	ad49	radio35
6	t ₀ +5	song51	6	t ₀ +5	song51					6	t ₀ +28	t ₀ +30	song29	6	t ₀ +28	t ₀ +30	song34	radio35
7	t ₀ +6	song51	7	t ₀ +6	song51													
8	t ₀ +7	song72	8	t ₀ +7	song72													
9	t ₀ +8	song72	9	t ₀ +8	song72													
10	t ₀ +9	none	10	t ₀ +9	song72													
11	t ₀ +10	song72	11	t ₀ +10	song72													
12	t ₀ +11	song72	12	t ₀ +11	song72													
13	t ₀ +12	song72	13	t ₀ +12	song72													
14	t ₀ +13	song65	14	t ₀ +13	song65													
15	t ₀ +14	song65	15	t ₀ +14	song65													
16	t ₀ +15	TVprgm58	16	t ₀ +15	song65													
17	t ₀ +16	song65	17	t ₀ +16	song65													
18	t ₀ +17	song65	18	t ₀ +17	song65													
19	t ₀ +18	song65	19	t ₀ +18	song65													
20	t ₀ +19	song65	20	t ₀ +19	song65													
21	t ₀ +20	song21	21	t ₀ +20	song21													
22	t ₀ +21	song21	22	t ₀ +21	song21													
23	t ₀ +22	song21	23	t ₀ +22	song21													
24	t ₀ +23	song21	24	t ₀ +23	song21													
25	t ₀ +24	song21	25	t ₀ +24	song21													
26	t ₀ +25	song21	26	t ₀ +25	song21													
27	t ₀ +26	ad49	27	t ₀ +26	ad49													
28	t ₀ +27	ad49	28	t ₀ +27	ad49													
29	t ₀ +28	song81	29	t ₀ +28	song35													
30	t ₀ +29	song35	30	t ₀ +29	song35													

ssd = "sample sequence data"

FIG. 4

EX1001, FIG. 4 (annotated).

78. My annotations also highlight the ’768 patents conversion of clean play stream 450 into a clean play list 490. Clean play list 490 appears in Figure 4 as a listing of the content capture along with an indication of the order in which it

was captured. A POSA would understand from Figure 4 that clean play list 490 is a condensed version of clean play stream 450. In fact, the '768 patent describes “[t]he resulting clean play stream may be converted to a play list listing the content captured and indicating the order in which it was captured.” EX1001, 8:6-8.

79. The '768 patent also makes clear that the purported invention is not limited to the particular organization or content of data described in Figure 4 and 5. The '768 patent explains:

It will be understood by those skilled in the art that in the context of data sets that might, in alternative examples, make up a "raw play stream," "a clean play stream," or a "clean play list," the particular organization, division, and content of the data may vary considerably from that illustrated in FIG. 4 (as well as that illustrated in FIG. 5 below) without departing from the spirit and scope of the present invention. To cite but one example, such play streams and/or play lists might include several data fields in addition to those illustrated. Moreover, the illustrated data sets might not be collected in the same table but, depending upon database architecture, might be spread out in different tables and/or relational structures. Furthermore, the data sets similar to those illustrated might not be stored at all in a database but rather might be generated on the fly and then discarded as mere

intermediate steps to generating a particular report that the system is asked to produce. These and other variations will be readily apparent to those skilled in the art. The examples shown are chosen primarily for usefulness in illustrating aspects of a particular embodiment. While they are useful in illustrating certain underlying principles of certain aspects of the present invention, they should not be considered, in and of themselves, to limit the scope of the invention.

EX1001, 10:34-56.

80. I note that the '768 patent includes several descriptions that are relevant to the plain and ordinary meaning of claim terms in the context of the '768 patent specification.

81. I note that the '768 patent explains that a “raw play stream” functionally relates “a sequence of samples to content identification results” EX1001, 6:14-16 and 8:1-3. I have reviewed the prior art documents with this description of “raw play stream” in mind.

82. I note that the '768 patent describes “scrubbing” as “reviewing for missing or incorrect data.” EX1001, 9:34-35. I have reviewed the prior art documents with this description of “scrubbing” in mind.

83. I note that '768 disclosure fails to provide meaning to “an expected pattern of sample sequence data” as the only appearance of “expected pattern”

occurs in the '768 claims. However, “expected pattern” may be inferred from the disclosure given in relation to the raw data stream 410 of Figure 4 and the disclosure “Given this pattern of sample sequence data, it is reasonable to assume that the person using the corresponding monitor is listening to song51 through the time period associated with the audio data of samples 1-7, and thus the ‘scrubbed’ version of the play stream, i.e. Clean Play Stream 450, indicates song51 for each of samples 1-7.” EX1001, 9:45-51. I have reviewed the prior art documents with this inference of “expected pattern of sample sequence data” in mind.

84. I note that the '768 patent explains that signal landmarks are a part of the stored versions of the media content. In particular, the '768 patent explains “audio signatures (sometimes referred to as audio fingerprints) and landmarks or other values useful in the content recognition process,” EX1001, 3:38-54, and “content may be stored signatures and landmarks derivable from a song’s audio, but not necessarily an audio file useable to play the song itself.” EX1001, 4:23-24. I have reviewed the prior art documents with this description of “signal landmark” in mind. I also note that the specification expressly regards “signatures” and “fingerprints” as overlapping, if not synonymous, concepts.

B. Prosecution History Summary

85. The '768 patent claims priority to U.S. Pat. Appl. No. 60/574,836, filed May 27, 2004, and is a divisional of U.S. Pat. Appl. No. 11/139,330, filed May 26, 2005. EX1001, (62), (60). Applicant filed the '768 patent application on September 11, 2008. EX1001, (22). The file wrapper for the '768 patent is attached as Exhibit 1002.

86. As summarized below, the Applicant obtained the '768 patent by amending the claims and arguing the cited prior art references failed to disclose all the limitations recited in the amended claims. On January 3, 2011, the USPTO mailed the first Office Action on the merits. EX1002, 331. In that Action, the Examiner rejected claims 1-5 under 35 U.S.C. § 103 as being unpatentable over Lamb (US 5,437,050) in view of Wang (US 2002/0083060) and Hansen (US 2002/0038456). EX1002, 333.

87. On March 31, 2011, the Applicant conducted an interview with the Examiner. EX1002, 321. During the interview, the Applicant proposed amendments that clarified that the “sample sequence data” corresponds to the sequence of the generated “raw play stream of content identification results.” EX1002, 323.

88. On May 3, 2011, the Applicant submitted an Amendment and Reply in response to the Office Action. EX1002, 308. In the Amendment and Reply, the

Applicant amended claim 1 to “make clear” that the raw play stream includes “a sequence of content identification results,” and that “scrubbing” is done by:

analyzing sample sequence data including the sequence of content identification results, wherein analyzing includes:

identifying results that are apparently correct in view of their sequence relative to other results in the sequence of content identification results; and

identifying results, if any, that are apparently incorrect in view of their sequence relative to other results in the sequence of content identification results. EX1002, 309, 312.

89. On August 5, 2011, the USPTO mailed a Final Office Action.

EX1002, 139. In the Final Office Action, the Examiner rejected claims 1-5 under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. EX1002, 141. In particular, the Examiner asserted that the limitations “identifying results that are apparently correct” and “identifying results, if any, that are apparently incorrect” are “vague because the recited language does not provide a description that one of ordinary skill in the art would understand based on the teachings in the Applicant’s specification.” EX1002, 142.

90. The Examiner also rejected claim 1 under 35 U.S.C. § 103 as being unpatentable over Deng (WO 2005/035877) in view of Hansen, rejected claim 2 under 35 U.S.C. § 103 as being unpatentable over Deng in view of Hansen, and further in view of Lee (WO 2005/006768), rejected claim 3 under 35 U.S.C. §

103 as being unpatentable over Deng in view of Hansen, and further in view of Ellis (US 5,436,653), claim 4 under 35 U.S.C. § 103 as being unpatentable over Deng in view of Hansen, and further in view of Barton (US 2002/0072982), and claim 5 under 35 U.S.C. § 103 as being unpatentable over Deng in view of Hansen, and further in view of Herley682 (US 2004/001161). EX1002, 142-147.

91. On February 6, 2012, the Applicant submitted an Amendment and Reply, along with a Request for Continued Examination, in response to the Final Office Action. EX1002, 125-130, 122. In the Amendment and Reply, the Applicant amended claim 1 to make clear that “any determinations of correctness or incorrectness of results are made ‘in view of their sequence relative to other results in the sequence of content identification results.’” EX1002, 128. The Applicant also amended claim 1 to recite that the “generating the raw play stream” step comprises:

- (i) receiving the plurality of audio data samples at one or more computers;
- (ii) with respect to an electronic database including a plurality of audio data representations, each of the plurality of audio data representations being associated with a content identifier, using one or more computers to query the electronic database to obtain a sequence of content identification results by searching the electronic database to obtain respective likely matches between respective audio data samples of the plurality of audio data samples and respective audio data representations of the plurality of audio data representations, the

respective audio data representations
corresponding to respective content identifiers; and
(iii) using the respective content identifiers to
generate the sequence of content identification
results in the raw play stream;

EX1002, 126. The Applicant also added new claims 6-8. EX1002, 127.

92. On October 2, 2012, The USPTO issued another Non-Final Office Action. EX1002, 82. In the Office Action, the Examiner rejected claims 1, 5, 6, and 8 under 35 U.S.C. § 103 as being unpatentable over Iyoku (US 2004/0055445) in view of Deng and Hansen, rejected claim 2 under 35 U.S.C. § 103 as being unpatentable over Iyoku in view of Deng and Hansen, and further in view of Lee, rejected claim 3 under 35 U.S.C. § 103 as being unpatentable over Iyoku in view of Deng and Hansen, and further in view of Ellis, claim 4 under 35 U.S.C. § 103 as being unpatentable over Iyoku in view of Deng and Hansen, and further in view of Barton, and claim 7 under 35 U.S.C. § 103 as being unpatentable over Iyoku in view of Deng and Hansen, and further in view of Wang (US 2002/0083060). EX1002, 86-92.

93. On March 28, 2013, the Applicant conducted an interview with the Examiner. EX1002, 54. During the interview, the Applicant presented a new claim for the Examiner's consideration. EX1002, 54. The Examiner agreed that the new claim was in condition for allowance. EX1002, 54. On April 1, 2013, the Applicant submitted an Amendment and Reply in response to the Office Action

issued on October 2, 2012. EX1002, 57. In the Amendment and Reply, the Applicant cancelled claims 1-8 and added new claims 9-23. EX1002, 58-61. Claim 9 is identical to the claim proposed during the Examiner interview. EX1002, 58-59, 54.

94. The Applicant further noted that it would be clear to one skilled in the art that the limitation “data for a sequence of audio data samples,” as recited by new claims 9 and 17 is “broad enough to cover implementations in which the data has already been divided into samples prior to being submitted over the network to one or more computers and implementations in which the data is to be divided into samples after being received at the one or more computers.” EX1002, 62-63.

95. The USPTO issued a Notice of Allowance on April 24, 2013 allowing claims 9-23. EX1002, 27. In allowing these amended claims, the Examiner found that the prior art does not disclose or suggest the combined elements/steps of independent claim 9. EX1002, 33-34. While the Examiner cited Wang060, the Examiner was unaware, however, that the allowed features are disclosed in prior art to Herley682, Seet, Herley864, and Wang060.

C. Claim Construction

96. I do not believe that any particular construction of any particular claim term is required for purposes of this proceeding. I have therefore interpreted the

claim terms recited in the '768 patent consistent with their plain and customary meaning to a person of ordinary skill in the art, as informed by the specification.

97. I reserve the right to offer claim construction opinions, however, in the event that Patent Owner proposes, or the Board adopts, particular constructions for claim terms.

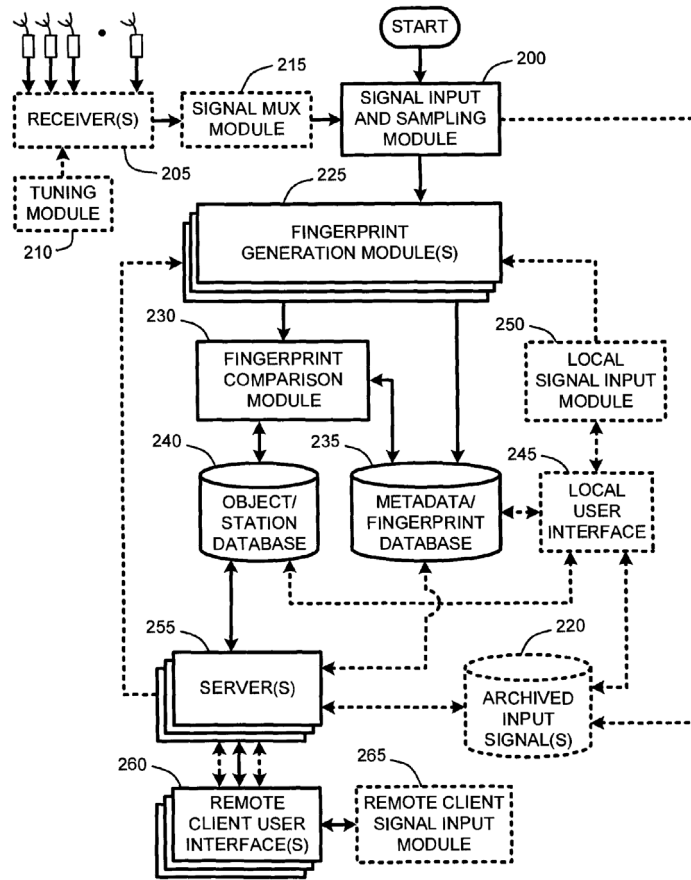
VII. SUMMARY OF THE APPLIED REFERENCES

A. OVERVIEW OF U.S. PATENT APPLICATION PUBLICATION NO. 2004/0260682 TO HERLEY ET AL. (“HERLEY682”) (EX1005)

98. Herley682 (EX1005) was published on December 23, 2004, and filed on June 19, 2003. I understand Herley682 is therefore prior art under 35 U.S.C. § 102(e) (pre-AIA) regardless of the '768 patent's priority date.

99. Herley682 is directed to “a system for identifying content of a signal, and in particular, to a system and method for sampling one or more channels of a broadcast spectrum, such as a radio frequency spectrum, identifying and storing content information for each sampled channel, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1005, ¶2.

100. Figure 2 of Herley682, reproduced below, illustrates an exemplary architectural diagram of an implementation of an embodiment thereof. EX1005, ¶38.



EX1005, FIG. 2.

101. “Sampling module 200 [to] receive[s] and sample[s] one or more signals, such as, for example, audio broadcast signals, video broadcast signals, etc.” EX1005, ¶75. “[S]ignal samples are continuously provided in real-time to one or more fingerprint generation modules 225 for extracting fingerprints from samples of the input signals.” EX1005, ¶84.

102. Extracted fingerprints are provided to fingerprint comparison module 230. EX1005, ¶86. Fingerprint comparison module 230 searches a metadata/fingerprint database 235 to determine if there is a match between the

extracted fingerprint and pre-computed fingerprints stored in database 235.

EX1005, ¶86. “If a match is identified ..., then the sampled portion of the input signal from which the fingerprint was extracted is identified as belonging to an object of interest corresponding to the fingerprint and metadata in the metadata/fingerprint database 235.” EX1005, ¶86. “This information, including the metadata, is stored in an object/station database 240.” EX1005, ¶86.

103. The object/station database 240 includes statistical information (e.g., the time that the object was identified, the station, channel, frequency, etc., where the input signal was monitored) and any metadata for each identified object of interest. EX1005, ¶86. “Metadata can include any identifying information that it the user desires to have associated with particular objects or object types.” EX1005, ¶68. For a song, the metadata may include, for example, the song’s name, artist, album title, copyright year, length, music genre, etc. EX1005, ¶¶26, 28, 68.

104. “[I]f the fingerprint extracted from the sampled input signal does not match any entries in the metadata/fingerprint database 235, that fingerprint is stored to the metadata/fingerprint database [235] as an ‘unknown object’ entry.” EX1005, ¶88. A copy of the sample and any statistical information available for the sample also are stored in database 235. EX1005, ¶88.

105. Metadata/fingerprint database 235 allows for unknown objects to be manually identified by a user. EX1005, ¶89. “[W]henver any metadata in the metadata/fingerprint database 235 is edited or modified, any corresponding entry in the object/station database 240 is automatically updated to reflect the changes in the metadata. In the fashion, the two databases are kept synchronized with respect to metadata content.” EX1005, ¶89. “Alternately such [unknown objects] metadata can be added by automatically querying a more authoritative and up-to-date fingerprint database.” EX1005, ¶17.

B. OVERVIEW OF U.S. PATENT APPLICATION PUBLICATION NO. 2005/0193016 TO SEET ET. AL (“SEET”) (EX1006)

106. Seet (EX1006) published on September 1, 2005, was filed on February 17, 2005, and claims priority to provisional application No. 60/545,681 (“the ’681 application”) filed on February 17, 2004. I understand that Seet is therefore prior art under 35 U.S.C. § 102(e) (pre-AIA) as of February 17, 2004 because it is entitled to its provisional filing date for prior art purposes, for the reasons explained below in Section VII.B.1.

107. Seet is directed to “Generation of a Media Content Database by Correlating Repeating Media Content in Media Streams.” EX1006, Title. In Seet, “media” is defined as including at least audio and video data. EX1006, ¶¶ 26, 50. Figure 1 of Seet, reproduced below, illustrates a system for creating and maintaining a media content database. EX1006, ¶15.

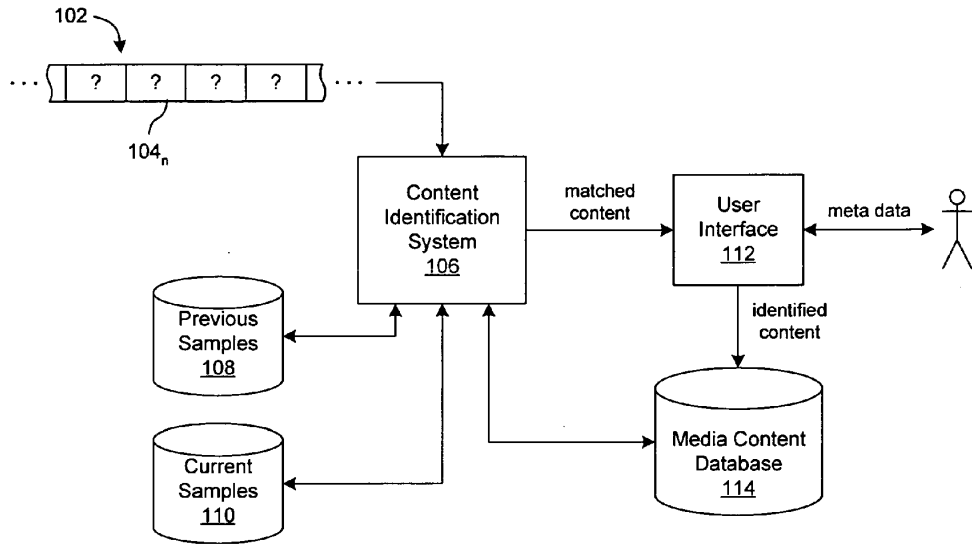


FIG. 1

EX1006, FIG. 1.

108. Content identification system 106 samples portions of media items 104_n from an incoming media data stream 102 to extract characteristic information and store that information in the previous samples database 108.

EX1006, ¶19. The content identification system 106 attempts to match samples in the current samples database 108 with samples in the previous samples database 110. EX1006, ¶19. Matched samples are made accessible by a review through a user interface 112 which allows the reviewer to observe the matched media content and provide identifying meta-data for that content. EX1006, ¶19. The identified media content is then added to the media content database 114, where it can be used by the content identification system 106 at a later time in an

application that requires media content recognition. EX1006, ABSTRACT, ¶¶14, 19.

109. As illustrated in Figure 2, reproduced below, the general process for creating and maintaining a media content databases 114 includes extracting a set of samples 202 from a first period of media stream 102. EX1006, ¶23. Samples 104_n (unknown media) are subsequently extracted. EX1006, ¶23. With the two sets of samples extracted, the system attempts to match 206 at least some of the samples in the second set of samples to the first set of samples. EX1006, ¶24. Matches occur between the two samples if their characteristic information differs within a predetermined tolerance. EX1006, ¶24. Matched content is then provided to a reviewer, from which meta-data are received 208. EX1006, ¶24. The metadata are then stored 210 in the media content database 114, along with the characteristic information of their associated samples, for use in the media content recognition application. EX1006, ¶24. The characteristic information (i.e., the fingerprint) and a marker (such as a timestamp) are associated with each other. EX1006, ¶27.

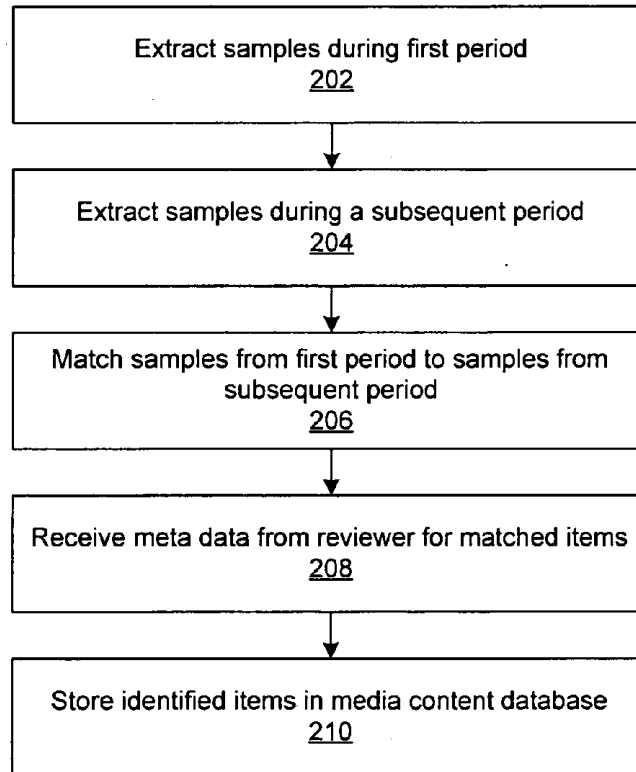


FIG. 2

EX1006, FIG. 2.

110. Because the samples typically obtained in noisy conditions, they may not match exactly (or within a predetermined tolerance) even where they are of the same content. EX1006, ¶31. Figure 4, reproduced below, illustrates a process in which a match for an unmatched sample is inferred based on its temporal location to other match samples. EX1006, ¶18.

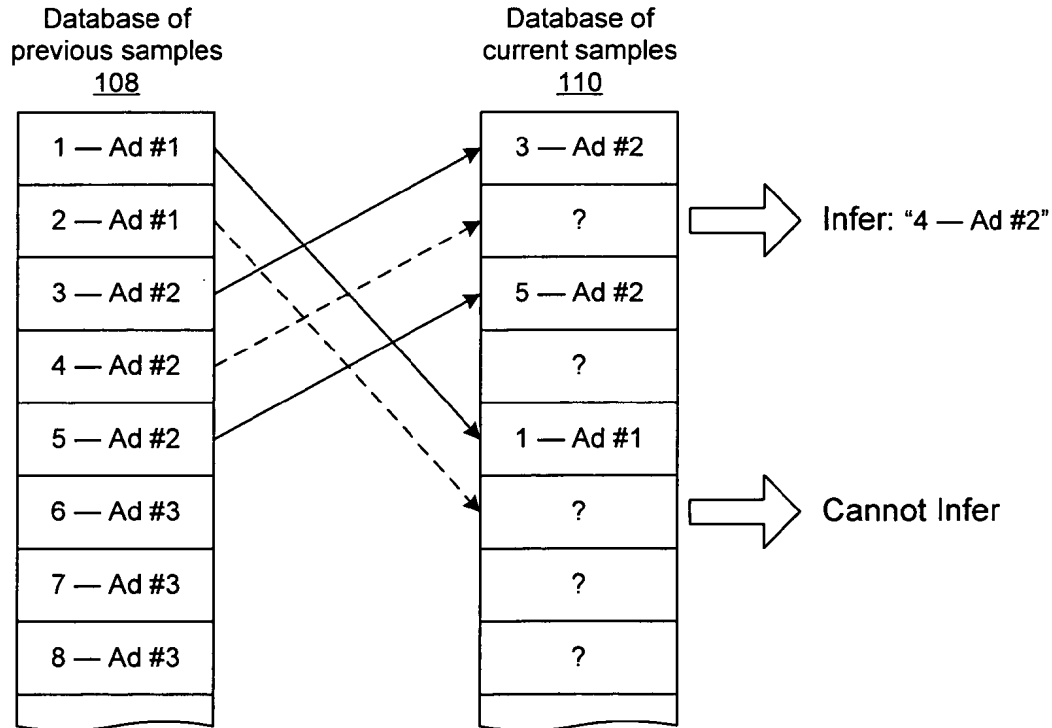


FIG. 4

EX1006, FIG. 4.

111. System 106 may infer additional matches. EX1006, ¶32. A matching inference may be made for some samples based on their position in media stream 102 relative to other match samples. EX1006, ¶32. Specifically if a match is not made for a particular sample, its surrounding samples may still be in the correct sequence and timing so that a match for the unmatched sample can be made. EX1006, ¶32. In figure 4, the second sample in the current samples database 110 was not matched to any sample in the previous samples database. EX1006, ¶32. However, because its previous sample was matched to samples 3 and its following sample was matched to sample 5, the system can infer that the second

sample matches sample 4 of the previous samples database. EX1006, ¶32. The temporal consistency of the bounding samples allows one to infer that they are of the same medial item 104_n (in this case, an advertisement labeled #2); therefore, the middle sample is also of the same media item 104_n. EX1006, ¶32.

1. U.S. Provisional Patent Application 60/545,681, (“’681 provisional application”) (EX1019) provides support for at least claim 9 of Seet.

112. Seet claims priority to U.S. Provisional Patent Application 60/545,681, (“’681 provisional application”) (EX1019) filed on February 17, 2004. The ’681 provisional application is entitled “Media Content Database Creation And Maintenance Using Time-Delayed Auto-Correlation Of Frequency Characteristics” and defines “media” as any streaming audio source in any form, be it radio, television, cable, satellite, internet streams or the like.” EX1019, 4.

113. It is my understanding from counsel that a nonprovisional patent application publication is prior art as of the earlier filing date of a provisional application if the disclosure of the provisional provides §112 support for at least one claim in the publication, and the provisional describes the subject matter relied upon as prior art.

114. I have reviewed the ’681 provisional application document, and conclude that it provides § 112 support for at least independent claim 9 as set forth in the table below.

Seet Claim 9	'681 provisional application (EX1019)
9. A method for generating entries for a media content database, the method comprising:	“Media Content Database Creation and Maintenance Using Time-Delayed Auto=correlation of Frequency Characteristics,” EX1019, 4, Title. “The invention described herein attempts to solve the issue of content acquisition by utilizing the signal itself as the data source, and discusses a method by which this database can be created in an automated fashion.” EX1019, 4, Background
receiving a first media stream;	“Instead of using a high-quality original master recording (say a CD, or digital video tape, etc), we suggest using the same audio signal that is to be identified, but from a previous period. In the case of radio or television monitoring, we propose that the previous period (i.e. day or day-period) for that particular station (or genre, or group of similar stations), be fingerprinted, and inserted into a finger print database.” EX1019, 4. “DB of Previous Day’s Samples,” EX1019, 5, Figure 1, EX1019, 6, Figure 2.
receiving a second media stream;	“Instead of using a high-quality original master recording (say a CD, or digital video tape, etc), we suggest using the same audio signal that is to be identified, but from a previous period. In the case of radio or television monitoring, we propose that

Seet Claim 9	'681 provisional application (EX1019)
	the previous period (i.e. day or day-period) for that particular station (or genre, or group of similar stations), be fingerprinted, and inserted into a finger print database.” EX1019, 4. “During the subsequent period, matches are then preformed against this database of previous samples, and if matched, then the identifier (or identifiers if multiple matches are found) and a time stamp are recorded.” EX1019, 5. “Next Day’s Samples’ Results Are Matched,” EX1019, 5, Figure 1. “Next Day’s Samples’ Results,” EX1019, 6, Figure 2.
matching at least some of the second media stream to the first media stream to determine repeating content between the streams; and	“During the subsequent period, matches are then preformed against this database of previous samples, and if matched, then the identifier (or identifiers if multiple matches are found) and a time stamp are recorded.” EX1019, 5. “Now the relationship has been established between the previous period’s database and the current period’s samples, it can be asserted that portions of the previous database that were matched are repeated content like advertisements, motions, songs, or the like.” EX1019, 5.

Seet Claim 9	'681 provisional application (EX1019)
	“Next Day’s Samples’ Results Are Matched,” EX1019, 5, Figure 1. “Matched content to be reviewed by human to add content description EX1019, 5, Figure 1.
presenting the repeating content to a reviewer for identification.	“Now a human reviewer can manually listen to just the series of identified samples and add metadata (like add name, or artist, and title information) to the series of samples.” EX1019, 5 “Matched content to be reviewed by human to add content description,” EX1019, 5, Figure 1. “Finally, an elegant user interface could present the sets of matched samples to the user in such a way as the user can just listen to as much of the sequence of samples as he or she likes, and enter in the content meta-data before jumping to the next sequence.” EX1019, 6.

2. The '681 provisional application describes the portions of Seet relied upon as prior art.

115. I also conclude that the '681 provisional application also describes the subject matter relied upon as prior art, as I indicate throughout my analysis

below with parallel citations to both EX1006 and EX1019. I highlight some particular portions of the '681 provisional here.

116. Reviewing the '681 provisional application, I note in particular that '681 provisional application's Figure 2, reproduced below, appears equivalent to Seet's Figure 4 (reproduced above) and illustrates the "scrubbing" aspects relied upon below as prior art.

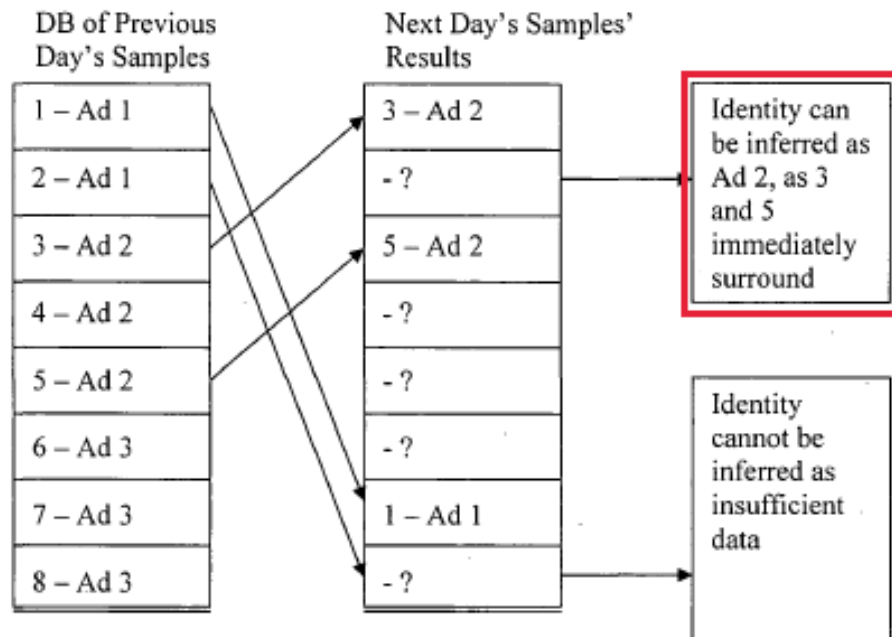


Figure 2. Auto-correlation Identity Inference Example
EX1019, Fig. 2 (annotated).

117. The '681 provisional application describes "[i]dentified content is left in the database for all future days until a set expiration date is reached since the last identification, and any samples from the current day that did not match

are now added into the database for the next Day's identification; the cycle repeats." EX1019, 5. Thus, the '681 provisional application makes clear that unidentified samples are stored in the database just as the identified data.

118. Continuing, the '681 provisional application describes "[a]dditionally, should the database itself contain duplicates, the post-processing algorithm should include measures to group matches into sets correspond the duplicate sample sets." EX1019, 6. "This requires that the underlying fingerprinting software returns all matches, not just the first match." EX1019, 6. "Further, if some matches are not made, but surrounding samples are in correct sequence and timing to infer the identity of the missing sample, that sample can thus be identified." EX1019, 6.

119. I also note that the '681 provisional application includes program code. EX1019, 7-10. The program code referenced below establishes how the inferred value is determined.

```
for(my $n = 1; $n < $num_play; $n++) #leave the first one
{
    #If first and third are rationally spaced, set second to expected
    value
    if($play[$n - 1]->{'media_id'} + 1 != $play[$n]->{'media_id'} &&
        $play[$n - 1]->{'media_id'} + 2 == $play[$n + 1]->{'media_id'})
    {
        $play[$n]->{'media_id'} = $play[$n - 1]->{'media_id'} + 1;
        $play[$n]->{'play_offset'} = 0;
        $play[$n]->{'play_confidence'} = 0;
    }

    #If first and fourth are rationally spaced, set second and third to
    expected value
    if($play[$n - 1]->{'media_id'} + 1 != $play[$n]->{'media_id'} &&
```

EX1019, 9.

120. Further, the program code referenced below establishes that after the program code determines the inferred value to substitute for the unknown value, the inferred value is written into the unknown data position in the play stream, thereby generating a new play stream with the inferred value.

```
for(my $n = 0; $n < $num_play; $n++)
{
    print STDERR "$n: $play[$n]->{'play_id'}, $play[$n]->{'media_id'},
    $play[$n]->{'play_offset'}, $play[$n]->{'play_confidence'}\n";
    $Broadcastid::play->{'play_id'} = $play[$n]->{'play_id'};

    &Broadcastid::update_play("media_id=$play[$n]-
    >{'media_id'},play_offset=$play[$n]-
    >{'play_offset'},play_confidence=$play[$n]-
    >{'play_confidence'},play_retrieved=NOW(),play_status=6");

    $Broadcastid::media->{'media_id'} = $play[$n]->{'media_id'};
    &Broadcastid::update_media("media_count=media_count+1");
}
```

EX1019, 10.

121. Thus, I conclude that the “Database of current samples 110” illustrated in Seet’s Figure 4 is equivalent to “Next Days’ Sample Results” as illustrated in the ’681 provisional application’s Figure 2. I also conclude that these data structures contained in the ’681 provisional application and Seet are equivalent to the “raw data stream” as described in the ’768 patent, and that a data structure containing Seet’s inferred data taught in the ’681 provisional application and in Seet are equivalent to the “clean play stream” described in the ’768 patent.

C. OVERVIEW OF U.S. PATENT NO. 7,333,864 TO HERLEY (“HERLEY864”) (EX1007)

122. Herley864 (EX1007) issued on February 19, 2008, and was filed on June 22, 2002. I understand that Herley864 is therefore prior art under 35 U.S.C. § 102(e) (pre-AIA) regardless of the ’768 patent’s priority date. Herley864 is directed to identification of repeating objects or music in an audio stream, and in particular, to a system and method for jointly segmenting and identifying repeating objects from one or more broadcast audio streams. EX1007, 1:16-19.

123. Herley864 describes the systems as a “repeating object extractor.” EX1007, 1:66. The Herley864 system “automatically segments and identifies audio objects such as songs, station identifiers, jingles, advertisements, emergency broadcast signals, etc., in an audio stream comprised of repeating and non-repeating audio objects.” EX1007, 1:66-2:12. “Unlike conventional schemes,

instead of trying to segment the audio stream into audio objects first, and then attempting to identify the segments, the repeating object extractor performs a joint segmentation and identification of the stream.” EX1007, 1:66-2:12. “This joint segmentation and identification provides a computationally efficient technique identifying and segmenting audio objects in an audio stream, even in an environment where the endpoints of audio objects are unknown or corrupted by voiceovers, cross-fading, or other noise.” EX1007, 1:66-2:12.

124. Like Herley682 described above, Herley864 is applicable to both audio streams and video streams, as Herley864 captures an “audio or audio/video broadcast via radio, television, the Internet, or other network.” EX1007, 2:13-15. “The [captured] audio stream, either buffered or read from a stored location, is then processed in segments, or windows.” EX1007, 2:19-21. “While discrete windows can be used, the use of overlapping or sliding windows to segment the audio stream for analysis provides better results.” EX1007, 2:21-23. “The segment is first filtered or processed to provide a vector of at least one representative parameter for characterizing the segment.” EX1007, 2:27-29.

125. “In particular, in processing the audio stream a set of at least one filter, which produces representative information about the window or data currently being processed, is applied to the audio stream.” EX1007, 2:47-58. “Thus, given a number of filters, an output vector including elements for each

filter is produced for one or more contiguous windows of the audio stream.”

EX1007, 2:47-58. “In other words, given one or more windows of the audio stream, the set of at least one filter produces a vector of outputs comprised of characteristic information which represents the audio data in those windows.”

EX1007, 2:47-58. “This output vector is then used to identify potential audio object matches.” EX1007, 2:47-58.

126. “Processing of the audio stream begins by applying a set of at least one filter to the current window of the stored or buffered audio stream.” EX1007, 10:30-45. “This set of filters provides a suite of algorithms for targeting particular aspects of the audio stream for producing representative information about the window or segment currently being processed.” EX1007, 10:30-45. “In other words, given K filters, there is vector of representative outputs $F_j = (f_{j0}, f_{j1}, f_{j2}, \dots, f_{jK-1})$, for the j^{th} window of the audio stream.” EX1007, 10:30-45. “Thus, for any given window or group of contiguous windows of the audio stream, the set of at least one filter produces a vector of outputs comprised of characteristic information that represents the audio data for that portion of the audio stream.” EX1007, 10:30-45. “This output vector is then used to identify potential audio object matches by checking the database for potential matches.” EX1007, 10:30-45.

127. “Once the output vector has been computed, it is passed to the database to determine whether any existing entries are possible matches.” EX1007, 3:1-15. “When potentially matching entries are located in the database, more detailed information regarding each potentially matching entry is provided for further comparison and analysis in order to help make a positive identification and segmentation of a portion of the audio stream.” EX1007, 3:1-15. “In particular, when the initial search identifies possible candidates that approximately match the output vector, a more detailed ‘match vector’ of information for each possibly matching audio object is returned for a comparison to the target portion of the audio stream.” EX1007, 3:1-15. “This match vector contains information such as, for example, a low-dimension version of the audio object, its duration, and other pertinent quantities, such as, for example, title and artist information.” EX1007, 3:1-15.

128. “Once a positive match has been found, the search terminates, the segmentation is completed, and the audio object is stored along with its meta-data.” EX1007, 3:43-49. “Alternately, the starting and ending points of the audio object within the audio stream is stored in a file or database so that the audio object can be accessed as desired, or so that statistics regarding audio object play can be maintained.” EX1007, 3:43-49.

129. Figure 2 of Herley864, reproduced below, illustrates an exemplary architectural diagram of an implementation of an embodiment thereof. EX1007, 4:15-17.

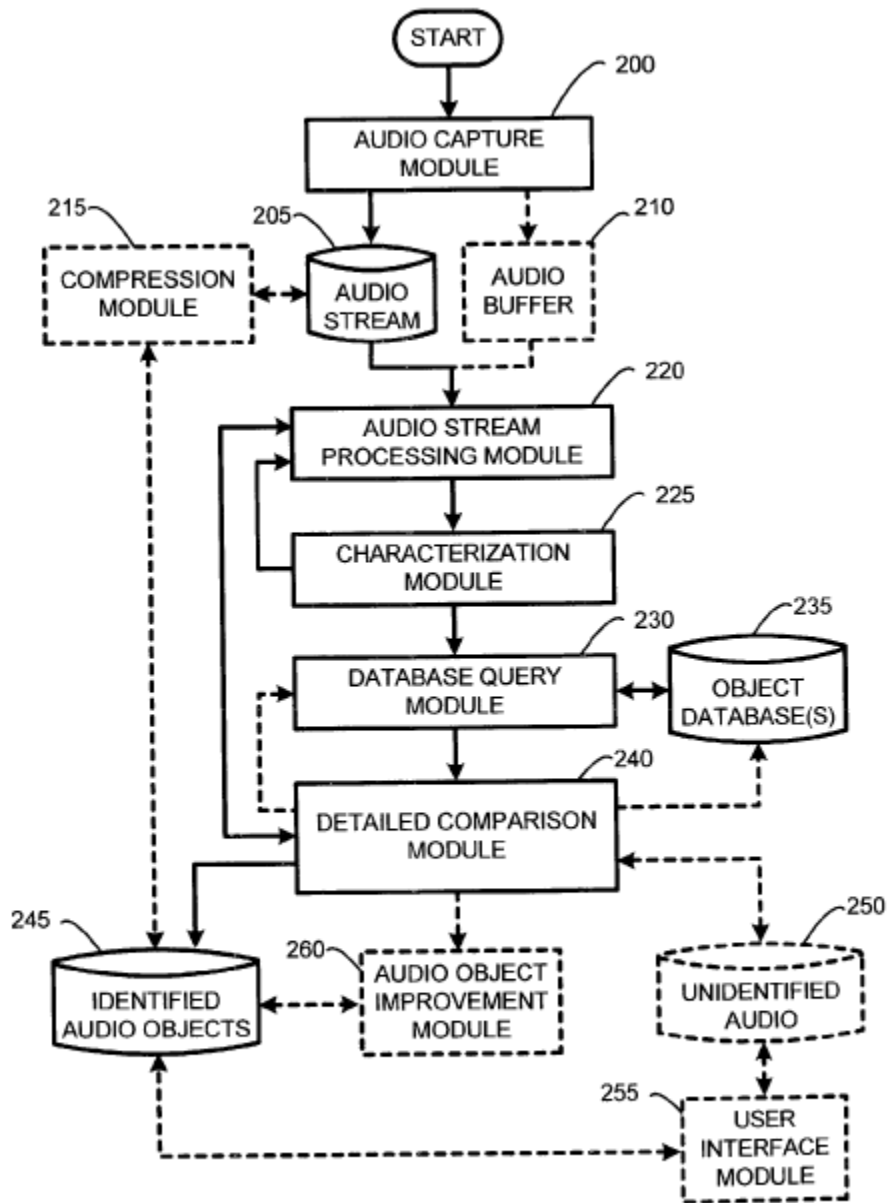


FIG. 2

EX1007, FIG. 2.

130. Audio capture module 200 captures one or more audio streams containing audio information, using well-known audio stream capture techniques. EX1007, 14:59-15:2. “[A]udio stream processing module 220 then begins examining the audio stream, either stored or buffered, 205 or 210, respectively.” EX1007, 15:11-13. In particular, audio stream processing module 220 provides segments or windows of the audio stream to a characterization module 225 for computing parametric information of the audio object. EX1007, 15:14-16. For example, if audio object represents a song, “the parametric information includes elements such as, for example, beats per minute (BPM), energy occupancy of various frequency bands, zero-crossings, etc., as estimated or calculated over each current window.” EX1007, 8:49-54.

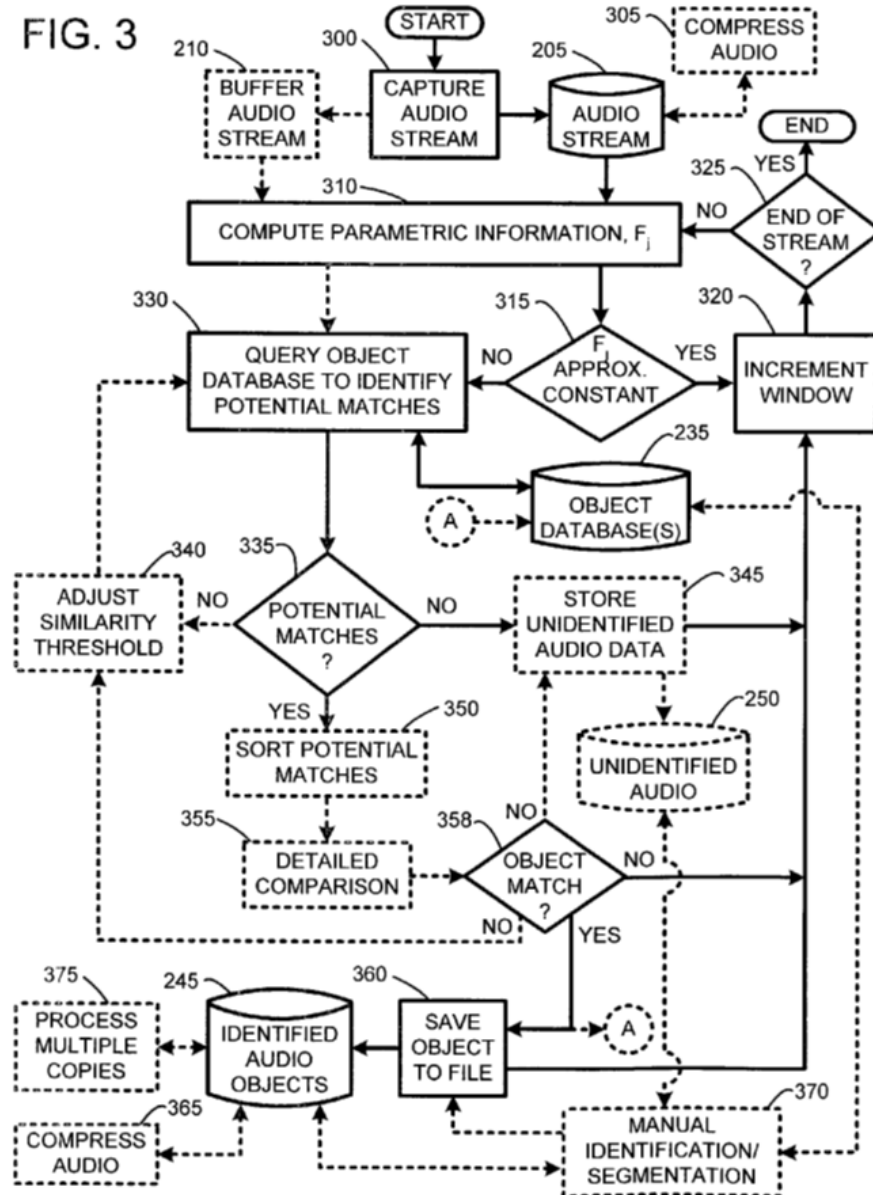
131. “[D]atabase query module 230 ... queries the object database 235 in order to identify potential matches to the portion of the audio stream being examined.” EX1007, 15:29-32. “[T]he object database contains parametric information for characterizing each known audio object.” EX1007, 8:39-40. Potential matches are identified, and the detailed comparison module 240 “perform[s] a detailed comparison between the portion of the audio stream being analyzed and one or more of the potentially matching object.” EX1007, 16:4-10. Once identified, detailed comparison module 240 “uses the duration information

[of the known audio object] to identify the nominal extents [(i.e., the endpoints)] of the identified audio object.” EX1007, 16:24-26.

132. If any portion of the audio stream cannot be identified, the “unidentified audio portions 250 are saved ... for later manual browsing, classification, or identification via a user interface module 255.” EX1007, 16:37-41. The “parametric and detailed information for characterizing the manually identified audio object is then made available to the object database for use in future automatic identifications of repeat instances of the manually identified audio object.” EX1007, 10:12-16. “[T]he extents of any such unknown or unidentified audio can be determined by identifying audio objects existing on either side of the unidentified audio within the stream.” EX1007, 13:26-28.

133. “FIG. 3 illustrates an exemplary system flow diagram for automatically identifying and segmenting audio objects in an audio stream.” EX1007, 4:18-20.

FIG. 3



EX1007, FIG. 3.

134. I now summarize the Herley864 system operation, with reference to Figure 3. Figure 3 begins “by capturing 300 an audio stream 205.” EX1007, 17:15-30. “Once the audio stream has been captured or buffered, 205 or 210, respectively, parametric information describing the audio stream is computed

310.” EX1007, 17:31-37. “In particular, starting with the beginning of the audio stream 205 or 210, segments or windows of the audio stream are analyzed using a set of K filters to generate a vector of representative outputs F_j , as described above.” EX1007, 17:31-37. “In one embodiment, so long as the output vector, F_j , is approximately constant 315 from one segment to the next, it is assumed that contiguous segments or windows of the audio stream belong to the same audio object and the window or segment is incremented 320 and F_j is recomputed 310 for the additional segment.” EX1007, 17:41-46. “In yet another embodiment, every segment of the audio stream, 205 or 210, is used to query 330 the object database 235. In this embodiment, as soon as the output vector, F_j , is computed 310 for each segment, the output vector, F_j , is provided for querying 330 the object database 235 to locate potential matches as described herein.” EX1007, 17:60-18:6. “However, this embodiment is computationally more expensive, as more database queries 330 are performed than with the aforementioned embodiment where the output vector, F_j , represents multiple segments or windows of the audio stream 205 or 210.” EX1007, 17:60-18:6. “In addition, while still useful, because shorter portions of the audio stream 205 or 210 are used in this embodiment, it is more likely that erroneous or incorrect matches will be identified.” EX1007, 17:60-18:6.

135. “In either case, as noted above, when querying 330 the object database 235, exact matches are not required, or even expected, in order to identify potential matches.” EX1007, 18:12-30. “Once one or more potential matches have been identified, the match vector, G_i is returned from the object database 235 for each potential match.” EX1007, 16:4-18. “A detailed comparison module 240 then uses the match vector, G_i , to perform a detailed comparison between the portion of the audio stream being analyzed and one or more of the potentially matching objects.” EX1007, 16:4-18. “This detailed comparison includes either a direct comparison of the current portions of the audio stream and the potential matches, or a comparison between a lower-dimensional version of the current portion of the audio stream and the potential matches.” EX1007, 16:4-18.

136. “If the detailed comparison shows that the current portion of the audio stream does not match any of the potential matches, it is identified as part of a new audio object in the database, and its position within the stream and parametric information is saved to the database as noted above.” EX1007, 13:18-28. “Alternately, as noted above, this unidentified audio is also made available to a user via a user interface for manual segmentation and identification.” EX1007, 13:18-28. “Note also that the extents of any such unknown or unidentified audio can be determined by identifying audio objects existing on either side of the

unidentified audio within the stream.” EX1007, 13:18-28. “In another embodiment, prior information, such as the most recently identified audio objects from the same audio stream, is used to modify match vector probabilities.” EX1007, 3:34-41. “For example, if a song of one genre was the most recently identified, it is more probable that a song of the same genre is being played.” EX1007, 3:34-41. “Consequently, a higher probability will be assigned to possible song matches of the same genre than to possible song matches of different genres.” EX1007, 3:34-41.

137. “In one embodiment, those portions of the audio stream for which no potential match is identified 335 are stored 345 in a file or database of unidentified audio data 250.” EX1007, 18:31-43. “As described herein, this unidentified audio data 250 can then be manually segmented, identified, or processed via a user interface.” EX1007, 18:31-43. “However, once one or more potential matches have been identified 335, the match vector, G_i , is returned from the object database 235 for each potential match.” EX1007, 18:31-43. “In another embodiment, these match vectors are sorted 350 in order of probability of match, such that those match vectors having a higher probability of match are compared to the portion of the audio stream being analyzed prior to those match vectors having a lower probability of match.” EX1007, 18:31-43.

138. Once the process concludes, Herley864's identified audio objects extracted from the audio stream are "saved 360 to a computer readable file or database 245 for later use." EX1007, 19:5-10.

D. OVERVIEW OF U.S. PATENT APPLICATION PUBLICATION NO. 2002/0083060 TO WANG ET AL. ("WANG060") (EX1008)

139. Wang060 (EX1008) was published on June 27, 2002. I understand that Wang060 is therefore prior art under 35 U.S.C. § 102(b) (pre-AIA) regardless of the '768 patent's priority date. Wang060 is directed to a "method for recognizing an audio sample locates an audio file that most closely matches the audio sample from a database indexing a large set of original recordings." EX1008, Abstract.

140. Wang060 notes a growing need for automatic recognition of music or other audio signals generated from a variety of sources. EX1008, ¶3. Wang060 notes problems with recognition methods due to "noisy" signals, required computational resources, unfeasible scalability, and strict sample length requirements. EX1008, ¶¶9-11. Wang060 seeks to alleviate these industry problems.

141. Figure 1 illustrates a method for recognizing a sound sample:

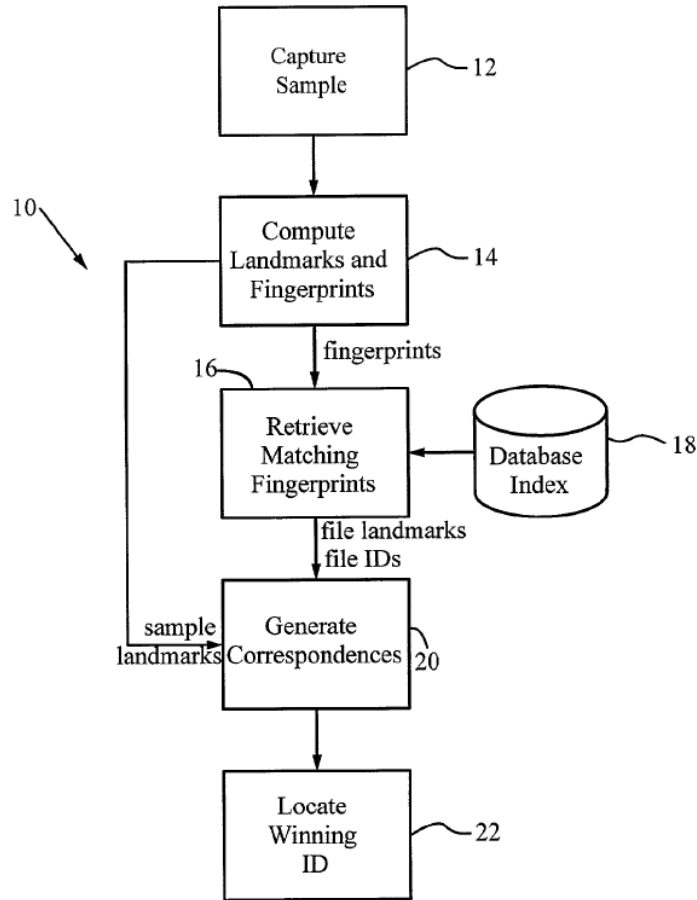


FIG. 1

EX1008, FIG. 1.

EX1008, ¶¶26, 39-24.

142. Wang060 explains that after the content is identified, a report for the identified sound samples, such as a log file of information about the identified sound recordings (e.g., title, artist, album cover art, lyrics, or purchase information), as well as “a list of music played from any sound stream (radio, internet streaming radio, hidden microphone, telephone call, etc.)” can be created.

EX1008, ¶107.

143. Wang060 describes a “method for recognizing a media sample, such as an audio sample, given a database index of a large number of known media files.” EX1008, ¶18. “The database index contains fingerprints representing features at particular locations of the indexed media files.” EX1008, ¶18. The unknown media sample is identified with a media file in the database (the winning media file) whose relative locations of fingerprints most closely match the relative locations of fingerprints of the sample. EX1008, ¶18. In the case of audio files, the time evolution of fingerprints of the winning file matches the time evolution of fingerprints in the sample. EX1008, ¶18.

144. Wang060 discloses “Preferably, the particular locations within the sample are reproducibly computed in dependence on the sample.” EX1008, ¶20. “The timepoints occur at, for example, local maxima of spectral L_p norms of the audio sample. EX1008, ¶21. “Fingerprints can be computed by any analysis of the audio sample, and are preferably invariant to time stretching of the sample.” EX1008, ¶21. “Examples of fingerprints include spectral slice fingerprints, multi-slice fingerprints, LPC coefficients, cepstral coefficients, and frequency components of spectrogram peaks.” EX1008, ¶21.

145. Wang060 explicitly recognizes that its “method is particularly useful for recognizing audio samples.” EX1008, ¶21.

146. Figures 3 and 4 of Wang060 illustrate database construction. “Before sound recognition can be performed, a searchable sound database index must be constructed.” EX1008, ¶47. “A block diagram of a preferred method 40 for constructing the searchable sound database index according to a first embodiment is shown in FIG. 3.” EX1008, ¶48. “In this embodiment, landmarks are first computed, and then fingerprints are computed at or near the landmarks.” EX1008, ¶48.

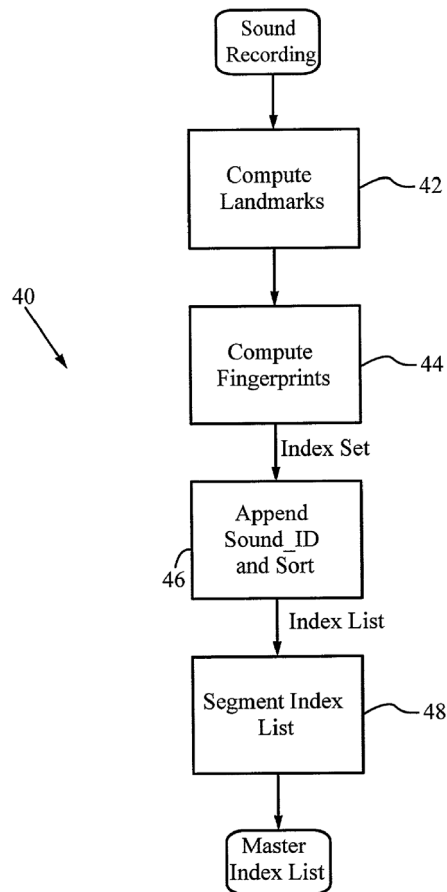


FIG. 3

EX1008, FIG. 3.

147. “To index the sound database, each recording in the collection is subjected to a landmarking and fingerprinting analysis that generates an index set for each audio file.” EX1008, ¶49. “FIG. 4 schematically illustrates a segment of a sound recording for which landmarks (LM) and fingerprints (FP) have been computed.” EX1008, ¶49. “Landmarks occur at specific timepoints of the sound and have values in time units offset from the beginning of the file, while fingerprints characterize the sound at or near a particular landmark.” EX1008, ¶49. “Thus, in this embodiment, each landmark for a particular file is unique, while the same fingerprint can occur numerous times within a single file or multiple files.” EX1008, ¶49.

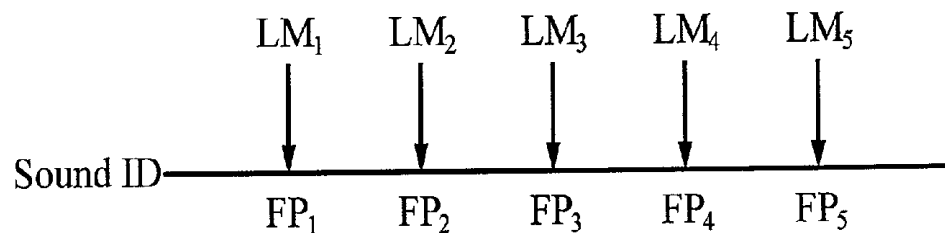


FIG. 4

EX1008, FIG. 4.

148. Wang060 discloses a variety of techniques for computing landmarks. EX1008, ¶¶51-54. “Once the landmarks have been computed, a fingerprint is computed at each landmark timepoint in the recording in step 44.” EX1008, ¶55.

The fingerprint is generally a value or set of values that summarizes a set of features in the recording at or near the timepoint.” EX1008, ¶55. “In a currently preferred embodiment, each fingerprint is a single numerical value that is a hashed function of multiple features.” EX1008, ¶55. “Possible types of fingerprints include spectral slice fingerprints, multi-slice fingerprints, LPC coefficients, and cepstral coefficients.” EX1008, ¶55. “Of course, any type of fingerprint that characterizes the signal or features of the signal near a landmark is within the scope of the present invention.” EX1008, ¶55. “Fingerprints can be computed by any type of digital signal processing or frequency analysis of the signal.” EX1008, ¶55.

VIII. GROUND 1: CLAIMS 1-3, 5, 6, 8-11, 13, AND 14 OF THE ’768 PATENT ARE UNPATENTABLE UNDER 35 U.S.C. § 103 UNDER HERLEY682 AND SEET

149. After reviewing the ’768 claims at issue, the Herley682 reference, and the Seet reference, I conclude that Herley682 and Seet disclose many of the same core features described and claimed in the ’768 patent. To the extent that there are differences, the differences would have been considered obvious by a POSA having knowledge of Seet and seeking to implement the Herley682 system. For the reasons outlined below, I conclude that claims 1-3, 5, 6, 8-11, 13, and 14 would have been obvious over Herley682 and Seet.

A. Motivation to Combine Herley682 and Seet

150. A POSA would have found it obvious to combine Herley682 and Seet for multiple reasons. Both Herley682 and Seet share many similarities, structures, goals, and methodologies. For example, both Herley682 and Seet relate to methods and systems of identifying objects of interest (e.g., songs, audio) within signals (e.g., a radio or television (media) broadcast). EX1005, ¶2; EX1006, ¶3; EX1019, 4. They both do so by receiving signal samples, extracting sample “fingerprints,” and comparing these to fingerprints of known objects in order to determine if there is a match and thereby identify the received signal. EX1005, ¶4; EX1006, ¶21; EX1019, 4. And both Herley682 and Seet describe how to handle the situation when there is not a match. EX1005, ¶¶16, 17, 71; EX1006, ¶32; EX1019, 6.

151. For example, Herley682 discloses a “system and method for sampling one or more channels of a broadcast spectrum, such as a radio frequency spectrum, identifying and storing content information for each sampled channel, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1005, ¶2. Herley682’s “interactive signal analyzer” among other things, “monitors one or more signals, identifies objects of interest within such signals, [and] stores statistical and metadata information describing such objects to an object database.” EX1005, ¶10. And

Herley682 teaches that these monitored signals may be, for example, audio signals, video signals, or combined audio/video signals. EX1005, ¶11.

152. In Herley682, unmatched objects are stored as “unknown objects” and can be manually identified by a user. EX1005, ¶¶16-17, 71. A POSA would understand that Herley682’s manual identification of unknown objects is a data imputation technique, similar to those described above in Section V.C.

153. Similarly, Seet teaches using system 106 to “infer additional matches” for unmatched (unknown) samples “based on their position in the media stream.” EX1006, ¶32, EX1019, 6 (“Further, if some matches are not made, but surrounding samples are in the correct sequence and timing to infer the identity of the missing sample, that sample can thus be identified.”). A POSA would understand that Seet’s “inference” technique is a data imputation technique, and data imputation techniques such as Seet’s inference and Herley682’s manual correction to correct missing, unknown, or incorrect data were well known to a POSA, as I described above in section V.C.

154. I conclude that a POSA would have understood and been motivated to use Seet’s teachings of inferring matches to identify Herley682’s unmatched objects as a POSA would recognize the efficiencies gained by replacing manual activity with an automated process as such inferences would reduce the number of manual identifications required in Herley682. This conclusion is further

supported by Seet's method providing provides "[s]ignificant operational and competitive advantages" and "improve[d] accuracy of matches," as well as "largely automat[ing]" "creating and maintaining the media content database." EX1006, ¶13, EX1019, 6 ("Since there is envisioned to be a great deal of database activity, efforts should be taken to reduce the search time"). In other words, a POSA, implementing the system of Herley682, would recognize the need for more automated handling of unmatched samples and would find, in Seet, a logical and easily implemented system for solving the issue. Further, a POSA would have further understood that using Seet's method to automatically correct the unknown objects in Herley682 would increase data reliability.

155. Because Herley682 and Seet are in the same field of endeavor, methods and systems of identifying objects of interest, a POSA would have been motivated to combine these references, and would have had a reasonable expectation of success in doing so. EX1005, ¶2; EX1006, ¶3; EX1019, 5. A POSA applying the teachings of Seet to the system and method taught in Herley682 would have viewed the combination as a simple combination of prior-art elements according to known methods to yield predictable results. Similarly, a POSA would have understood the combination would have merely been the application of known techniques (Seet's method of inferring additional matches for unmatched samples) to a known system/method ready for improvement

(Herley682's system/method, which includes identifying such unmatched samples) to yield the predictable result of efficiently identifying objects of interest in a broadcast signal. And because both Herley682 and Seet implement their media content identification methods in conventional database systems with conventional hardware/software, the combination would have been well within ordinary skill. *See* EX1005, ¶50; EX1006, ¶51.

B. Independent Claim 1

1. Preamble [1.P]

[1.P] A media measurement method comprising:

156. Initially, I note that the '768 patent explains generally that media measurements methods are known, and that traditional media measurement systems have focused on directly monitoring channels being utilized by audience members, noting however that such systems are limited in that “they can only track content that has been properly encoded.” EX1001, 1:34-36.

157. The '768 patent further explains “the present invention provides a media measurement system and method that enhances recognition (e.g., in terms of accuracy or efficiency) of the content of a media sample by analyzing information exogenous to the sample.” EX1001, 1:60-64.

158. To the extent the preamble of claim 1 is limiting, I conclude that Herley682 teaches the preamble. Herley682 discloses: “The invention is related to a system for identifying content of a signal, and in particular, to a system and method for sampling one or more channels of a broadcast spectrum, such as a radio frequency spectrum, identifying and storing content information for each sampled channel, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1005, ¶2.

159. Herley682 further discloses: “A system and method for providing automatic object identification and user interaction with respect to objects of interest within a signal, is referred to herein as an “interactive signal analyzer.” EX1005, ¶10. “The interactive signal analyzer monitors one or more signals, identifies objects of interest within such signals, stores statistical and metadata information describing such objects to an object database, and provides an interactive user interface to the object database for providing responses to user queries regarding objects or signals characterized in the object database.” EX1005, ¶10. “As described below, the interactive signal analyzer provides a number of advantages that makes it well suited for providing an interactive object database for viewing and interacting with information extracted from one or more signals.” EX1005, ¶10. “For example, in addition to providing a useful technique for gathering statistical information regarding objects within a signal such as, for

example, an audio media stream, automatic identification of objects within the media stream allows a user to interact with that statistical information either in real-time, or subsequent to signal transmission.” EX1005, ¶10.

160. Herley682 describes performing this analysis on “signals,” (e.g., “media”), which is disclosed as being “any type or dimensionality (audio, video, etc.) except where particular signal types are explicitly referred to.” EX1005, ¶11. “For example, such signals include an audio signal which is considered to be a one-dimensional signal; an image is which considered to be a two-dimensional signal; and video data which is considered to be a three-dimensional signal.” EX1005, ¶11. “Further, in the context of a combined audio/video signal, audio objects can be used to identify an associated video sequence, since the audio portion of a combined audio/video signal will typically remain approximately the same between repeating instances of the signal.” EX1005, ¶11. Thus, I conclude the teachings of Herley682 are equally applicable to audio or video signals.

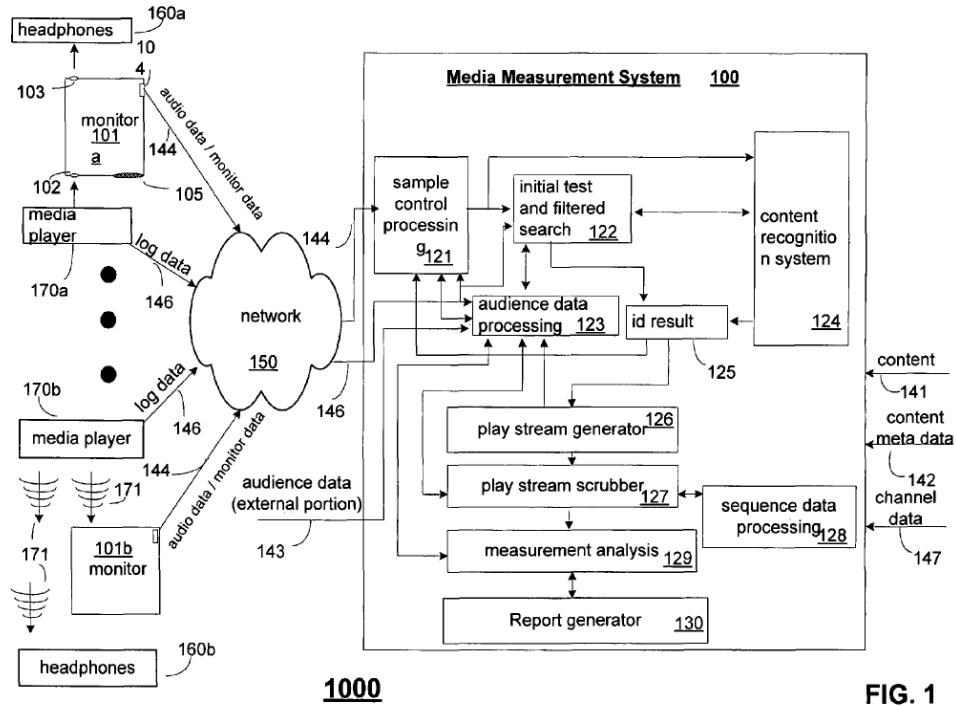
2. Element [1.1]

[1.1] (a) receiving over a network, at one or more computers, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network;

161. Herley682 teaches claim element [1.1].

162. First, I consider the disclosure of the '768 patent with respect to limitations recited in this claim element. The '768 patent explains, with reference to Figure 1, that audio data is captured by monitors 101 and uploaded to media measurement system 100, e.g., via network 150. EX1001, 3:38-43. "The audio data uploaded from a monitor 101 may be raw audio data, or may be audio data that has undergone one of a variety of levels of processing prior to upload." EX1001, 3:43-46. "The uploaded audio data may include parameters useable to calculate audio signatures (sometimes referred to as audio fingerprints) and landmarks or other values useful in the context recognition process." EX1001, 3:46-49. "Alternatively, the audio data may comprise pre-calculated signatures, marks, or other data useful for content recognition." EX1001, 3:49-51. "On the other hand, in some embodiments, the uploaded audio data may simply represent a raw signal corresponding to audio energy received at monitor 101." EX1001, 3:51-54. A POSA would understand this teaching to mean the uploaded audio data may be, for example, (1) a raw signal corresponding to audio energy received at monitor 101; (2) raw audio data that has undergone one of a variety of levels of processing prior to upload (e.g., data that "include[s] parameters useable to calculate audio signatures (sometimes referred to as audio fingerprints) and landmarks or other values useful in the content recognition process"); or (3) data that has undergone certain processing prior to upload (e.g., "the audio data may

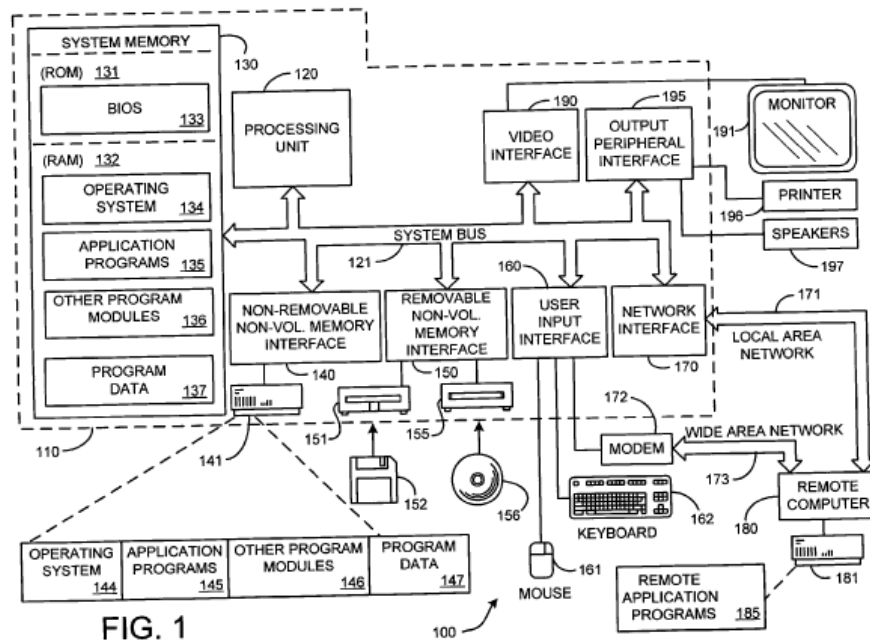
comprise pre-calculated signatures, landmarks, or other data useful for content recognition”). EX1001, 3:46-54.



EX1001, FIG. 1.

163. The '768 patent's sample control processing module 121 divides the audio data into samples (or receives pre-divided audio data) and determines which samples should be submitted for content recognition testing. EX1001, 4:53-62. Next, “[c]ontent recognition system 124 . . . implement[s] one or more known content recognition methods that identify content by extracting parameters from a media signal and applying an algorithm to those parameters to search for a content match.” EX1001, 5:38-42. The provided examples include various fingerprint searching techniques. *See* EX1001, 5:43-55.

164. Herley682 teaches the same features. Herley682 discloses that “FIG. 1 is a general system diagram depicting a general-purpose computing device constituting an exemplary system for automatically identifying and storing content information for sampled signals, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1005, ¶37.



EX1005, FIG. 1.

165. Herley682 discloses that “[t]he invention is operational with numerous other general purpose or special purpose computing system environments or configurations.” EX1005, ¶50. “Examples of well known computing systems, environments, and/or configurations that may be suitable for use with the invention include, but are not limited to, personal computers, server

computers, hand-held, laptop or mobile computer or communications devices such as cell phones and PDA's, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.” EX1005, ¶50.

166. A POSA would understand from these disclosures that Herley682 is operable “over a network” and Herley682’s system, or individual components, are implemented utilizing “one or more computers.” Further, using a network of computers in a distributed computing environment was a well-known technique for applying additional computing power to perform compute-intensive data analysis. Much of my research work at Mississippi State in the late 1990s was on using distributed computing.

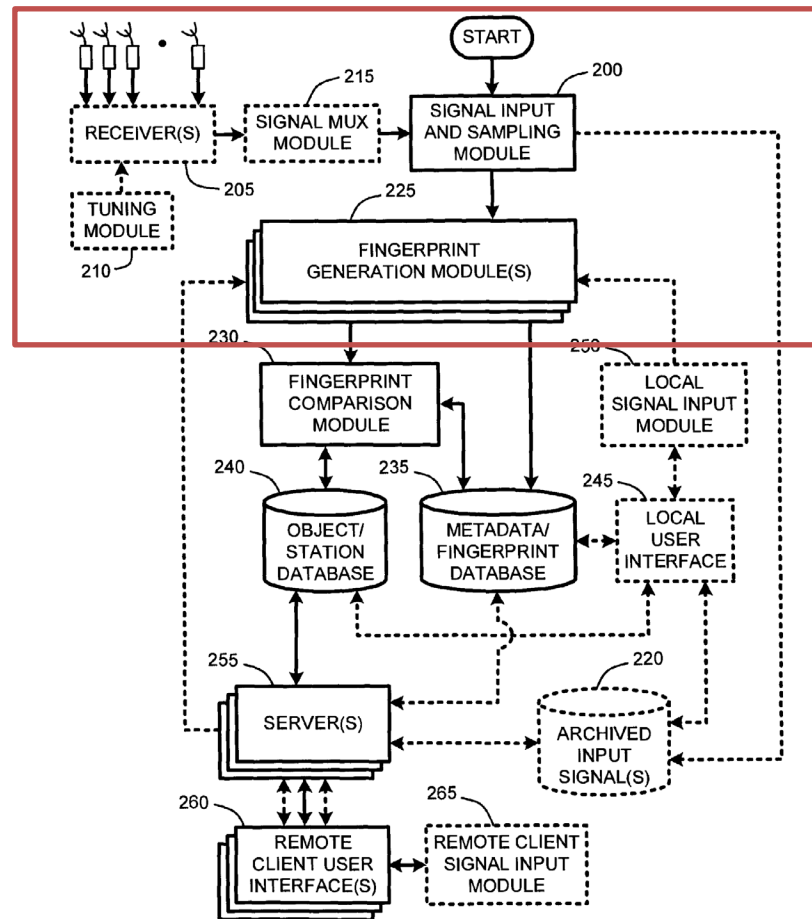
167. Herley682 further discloses that “[t]he invention may be described in the general context of computer-executable instructions, such as program modules, being executed by a computer.” EX1005, ¶51. “Generally, program modules include routines, programs, objects, components, data structures, etc., that perform particular tasks or implement particular abstract data types.” EX1005, ¶51. “The invention may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network.” EX1005, ¶51. “In a distributed

computing environment, program modules may be located in both local and remote computer storage media including memory storage devices.” EX1005, ¶51. “With reference to FIG. 1, an exemplary system for implementing the invention includes a general-purpose computing device in the form of a computer 110.” EX1005, ¶51.

168. Herley682 discloses that “[t]he computer 110 may operate in a networked environment using logical connections to one or more remote computers, such as a remote computer 180.” EX1005, ¶59. “The remote computer 180 may be a personal computer, a server, a router, a network PC, a peer device, or other common network node, and typically includes many or all of the elements described above relative to the computer 110, although only a memory storage device 181 has been illustrated in FIG. 1.” EX1005, ¶59. “The logical connections depicted in FIG. 1 include a local area network (LAN) 171 and a wide area network (WAN) 173, but may also include other networks. Such networking environments are commonplace in offices, enterprise-wide computer networks, intranets, and the Internet.” EX1005, ¶59.

169. Further, Herley682 discloses that “[i]n particular, as illustrated by FIG. 2, a system and method for automatically identifying and storing content information for sampled signals, and providing a user interface for allowing interactive user queries and display of the stored content information begins by

using a signal input and sampling module 200 to receive and sample one or more signals, such as, for example, audio broadcast signals, video broadcast signals, etc.” EX1005, ¶75. “Note that sampling frequency of the input signal is discussed in more detail in section 3.1.1.” EX1005, ¶75. “In one embodiment, radio or television broadcasts are captured using one or more receivers 205.” EX1005, ¶75.



EX1005, FIG. 2 (annotated).

170. Herley682 also discloses: “Once the signal input and sampling module 200 begins to receive one or more signals, or one or more multiplexed

signals, signal samples are continuously provided in real-time to one or more fingerprint generation modules 225 for extracting fingerprints from samples of the input signals (see Section 3.1.1).” EX1005, ¶84. Thus, the generated fingerprints are “representations of audio data captured at a media monitoring device” as claimed. “Once the fingerprint extraction module 225 has extracted a fingerprint from the sampled input signal, that fingerprint is provided to a fingerprint comparison module 230.” EX1005, ¶86.

171. Here, too, a POSA would understand that a “module” refers to program modules that may be stored and executed on one or more computers within the Herley682 disclosure. Herley682 further teaches that “[t]he invention may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network.” EX1005, ¶51; *see also* EX1005. ¶¶59-60 (discussing Herley682’s operation in a networked environment). In at least this instance, a POSA would have further understood that Herley682’s fingerprints (the representations of audio data captured) are provided to fingerprint comparison module 230 (one or more computers) over a network.

172. My conclusions are further supported by Herley682’s Figure 5, reproduced below. Figure 5 illustrates “an exemplary system flow diagram for

automatically identifying and storing content information for sampled signals.”

EX1005, ¶41.

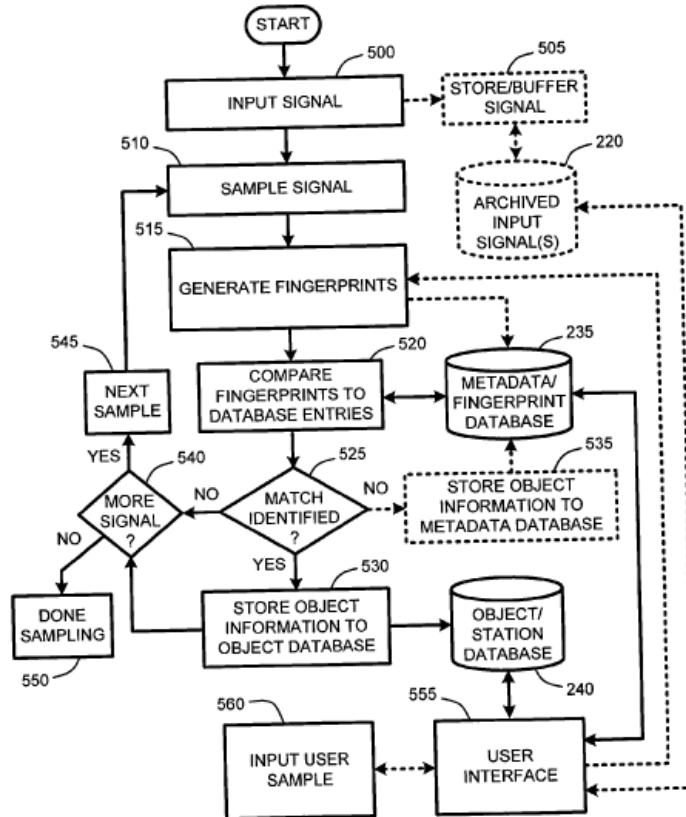


FIG. 5

EX1005, FIG. 5.

173. Now referring to both Figures 2 and 5, I will walk through Herley682’s teachings of the claimed features. Receiver 205 captures a radio or television broadcast. EX1005, ¶¶75, 160 (input signal 500). Herley682’s sampling module 200 receives and samples these broadcast signals, and sends the samples (*data for a sequence of audio data samples*) to fingerprint generation module(s) 225, which in turn extracts fingerprints from the samples. EX1005,

¶¶24, 75, 84, 161 (sample signals 510; generate fingerprints 515). Herley682 describes that such “fingerprints” are “used to uniquely identify the signal segments from which they are derived.” EX1005, ¶13. In other words, Herley682’s fingerprints are *representations of captured audio data* as claimed. “Once the fingerprint extraction module 225 has extracted a fingerprint from the sampled input signal, that fingerprint is provided to a fingerprint comparison module 230.” EX1005, ¶86.

174. And, as I explained above, a POSA would understand that a “module” (as in signal input and sample module 200, fingerprint generation module(s) 225, and fingerprint comparison module 230) refers to program modules that may be stored and executed on one or more computers within the Herley682 disclosure. Herley682 further teaches that “[t]he invention may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network.” EX1005, ¶51; *see also* EX1005 ¶¶59-60 (discussing Herley682’s operation in a networked environment). In at least this instance, a POSA would have further understood that Herley682’s broadcast signals, samples (the data for a sequence of audio data samples), and fingerprints (the representations of audio data captured) are provided respectively to signal input and sample module 200,

fingerprint generation module(s) 225, and fingerprint comparison module 230
(one or more computers) over a network.

3. Element [1.2]

[1.2] (b) using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers;

175. Herley682 teaches element [1.2].

176. Looking to the '768 patent for understanding, in the '768 patent, the “content recognition system 124” which, as described above, is “implement[ed as] one or more known content recognition methods that identify content by extracting parameters from a media signal and applying an algorithm to those parameters to search for a content match.” EX1001, 5:38-42. The provided examples include various fingerprint searching techniques. *See* EX1001, 5:43-55. Herley682 utilizes a fingerprint searching technique.

177. Specifically, Herley682 discloses that “[o]nce the fingerprint extraction module 225 has extracted a fingerprint from the sampled input signal, that fingerprint is provided to a fingerprint comparison module 230.” EX1005, ¶86, 161, FIG. 5, (step 520). “The fingerprint comparison module 230 then searches a metadata/fingerprint database 235 having pre-computed fingerprints and metadata describing objects represented by the pre-computed fingerprints.” EX1005, ¶86. “If a match is identified by the fingerprint comparison module 230,

then the sampled portion of the input signal from which the fingerprint was extracted is identified as belonging to an object of interest corresponding to the fingerprint and metadata in the metadata/fingerprint database 235.” EX1005, ¶86. “This information, including the metadata, is stored in an object/station database 240.” EX1005, ¶86. “Therefore, for each identified object of interest, the object/station database includes statistical information such as the time that the object was identified, the station, channel, frequency, etc., where the input signal was monitored, and any metadata stored in the metadata/fingerprint database 235 for that particular object.” EX1005, ¶86.

178. Herley682 also discloses by “[r]eferring [now] to FIG. 5 in combination with FIG. 2, in one embodiment, the process can be generally described as a system and method for identifying objects in one or more sampled signals, and providing an interactive user interface for interacting with statistical information and metadata describing objects identified within the sampled signals.” EX1005, ¶160. “In particular, as illustrated by FIG. 5, the interactive signal analyzer described herein begins by inputting one or more signals 500. In one embodiment, the input signals 500 are stored or buffered 505 to the archived input signal database 220.” EX1005, ¶160. “Whether or not the input signals 500 are stored or buffered 505, they are then sampled 510.” EX1005, ¶161. “The size and period of the sample are dependent on both the length of any objects in the

input signals 500, and the type of fingerprint engine being used.” EX1005, ¶161.

“For example, the DDA-based fingerprint engine described above computes trace fingerprints 515 every 186 ms over samples of the input signal 500.” EX1005,

¶161. “Once the trace fingerprint has been generated 515 for a particular sample, that trace fingerprint is compared 520 to the fingerprints of known audio objects that are stored in the aforementioned metadata/fingerprint database 235.”

EX1005, ¶161. “As discussed above, this metadata/fingerprint database 235 includes pre-computed fingerprints and metadata describing objects represented by the pre-computed fingerprints.” EX1005, ¶161.

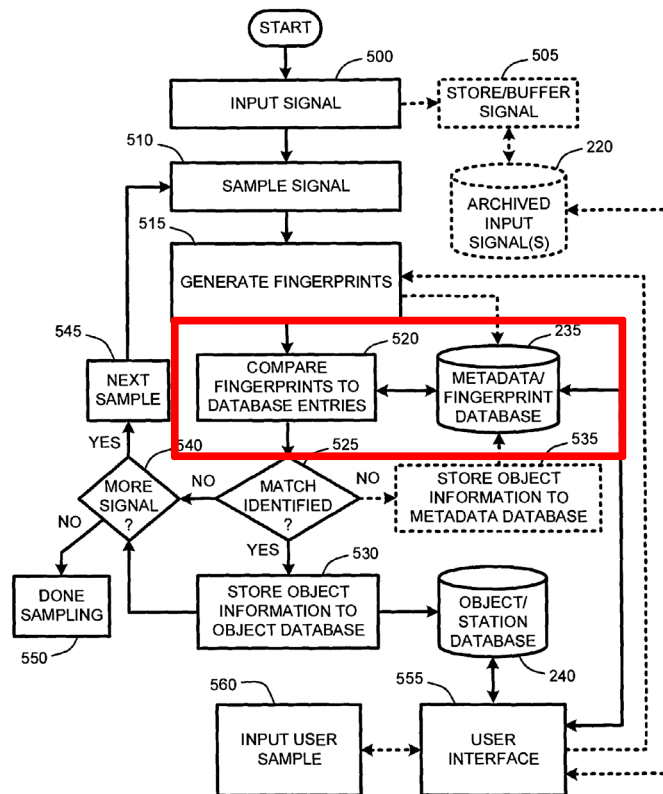


FIG. 5

Herley, FIG. 5 (annotated).

179. Herley682 further describes that, for a song, the metadata can include the name, artist, album title, copyright year, length, music genre or subgenre. EX1005, ¶¶26, 28, 68. This metadata corresponds to the claimed corresponding content identifiers, which is consistent with the '768 patent's description of "content ID" as "a reference corresponding to a particular piece of content," e.g., "song51." EX1001, 9:19-20, FIG. 4. Therefore, Herley682 teaches element 1.2.

4. Element [1.3]

[1.3] (c) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples; wherein:

180. The Herley862-Seet combination teaches or at least renders obvious element [1.3].

181. Initially, I note that the '768 patent explains that a “raw play stream” functionally relates “a sequence of samples to content identification results.” EX1001, 6:14-16. Thus, the claimed “raw play stream” encompasses generating “a sequence of content identification results corresponding to the sequence of audio data samples.” The '768 patent also discloses “step 207 generates a raw play stream indicating a series of samples and corresponding identification results.” EX1001, 8:1-3. Further, Figure 4 illustrates an example of such a “series of samples and corresponding identification results” as a table including columns for “row,” “log time,” and “contentID.” EX1001, 8:1-3, FIG. 4; *see also* EX1001, 10:34-56 (describing that the invention is not limited to any specific organization, division, or content of data). Because Herley682 stores such a sequence or series of content identification results corresponding to audio data samples, I conclude that Herley682 generates a “raw play stream,” for the reasons explained below.

182. Herley682 discloses that “[o]nce the fingerprint extraction module 225 has extracted a fingerprint from the sampled input signal, that fingerprint is

provided to a fingerprint comparison module 230.” EX1005, ¶¶86, 162. “The fingerprint comparison module 230 then searches a metadata/fingerprint database 235 having pre-computed fingerprints and metadata describing objects represented by the pre-computed fingerprints.” EX1005, ¶¶86, 162. “If a match is identified by the fingerprint comparison module 230, then the sampled portion of the input signal from which the fingerprint was extracted is identified as belonging to an object of interest corresponding to the fingerprint and metadata in the metadata/fingerprint database 235.” EX1005, ¶¶86, 162; *see also* FIG. 5 (step 525). “This information, including the metadata, is stored in an object/station database 240.” EX1005, ¶¶86, 162; *see also* FIG. 5 (step 530). “Therefore, for each identified object of interest, the object/station database includes statistical information such as the time that the object was identified, the station, channel, frequency, etc., where the input signal was monitored, and any metadata stored in the metadata/fingerprint database 235 for that particular object.” EX1005, ¶¶86, 162; *see also* FIG. 5 (step 530).

183. Herley682 discloses that “[i]f a matching fingerprint is identified 525 in the metadata/fingerprint database 235, then the metadata associated with that matching fingerprint is stored 530 to the aforementioned object database 240, along with statistical information regarding the sample, such as the time that the identified object appeared in the input signal 500, and other information, as

appropriate, such as the station, channel, frequency, etc. where the input signal was monitored.” EX1005, ¶162. “Further, in one embodiment, a copy of the sample extracted from the input signal 500 is also stored along with the statistical information and metadata.” EX1005, ¶162. “In a related embodiment, if a subsequent instance of a particular object is identified in the input signal, rather than creating a new entry in the object database 240, a counter representing the total number of identifications for that object of interest is simply incremented, and the statistical information documenting the time and source of the associated sample is stored to the object database.” EX1005, ¶162. From these teachings, a POSA would have understood that trace fingerprints are generated (step 515) and compared (step 520) for a plurality of samples taken from input signal 500 (step 510). EX1005, ¶161 (discussing sample size and period), ¶13 (“‘trace fingerprints’ ... are generated several times per second on an incoming signal”).

184. A POSA would have further understood from the above teachings that Herley682 stores in the object database 240, at minimum, a list of data relating sampled portions, metadata (e.g., content identification results), and the time(s) in which the identified object appeared in the input signal. A POSA would have understood that the collection of data stored in Herley682’s object database 240 teaches a *raw play stream*. Seet similarly teaches storing samples and corresponding identifiers in the database of current samples 110. EX1006,

¶30, FIG. 3; EX1019, 5, FIG. 1. This is consistent with the '768 patent's description that a "raw play stream relat[es] a sequence of samples to content identification results" noted above. EX1001, 6:14-16; *see also* EX1001, 8:1-3 ("[S]tep 207 generates a raw play stream indicating a series of samples and corresponding identification results."), 9:12-19 ("FIG. 4 illustrates Raw Play Stream 410 generated by a process such as step 207 of FIG. 2 Referring to FIG. 4, Raw Play Stream 410 itemizes a sequence of analyzed samples and relates that sequence to corresponding identification results in the content ID column."), FIG. 4 (raw play stream 410).

185. Further, it would have been obvious to store the data in object database 240 in sequential order, or in other words, as the data in the stream is processed. Herley682 also suggests the value in knowing the sequential, chronological order of the objects in the stream, e.g., the importance of maintaining the temporal integrity of the sample data. Herley682 discloses that in "one embodiment a user is given the ability to request the portions of a stream surrounding a particular object. For example, an advertiser might wish to see the context in which his/her commercial plays." EX1005, ¶81. This would suggest to a POSA, at a minimum, to sort the data in Herley682's object database 240 based on the time data to put it in sequential, chronological order so that such contextual information could be extracted from the data.

186. Further, Seet provides an express teaching of maintaining the temporal integrity, i.e., chronological order, of the sample data, confirming that a POSA would have been motivated to do so. In Seet, the list of match identifications is maintained in sequential order to allow for using information from temporally adjacent samples to identify unmatched samples. EX1006, ¶¶30, 32, FIGs. 3, 4; EX1019, 5-6, FIGs. 1, 2.

5. Element [1.4]

[1.4] the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database;

187. I conclude that Herley682 teaches this element.

188. As I discussed above for elements [1.2] and [1.3], Herley682's fingerprint comparison module 230 searches (e.g., queries, "querying") metadata/fingerprint database 235 (e.g., an electronic database) to compare the trace fingerprints (e.g., representations of audio data captured representing the audio data samples) to pre-computed fingerprints of known audio objects (e.g., respective audio data representations) in order to determine if there is a match (e.g., attempt to determine respective likely matches). EX1005, ¶¶86, 161, FIG. 5 (step 520). If a matching fingerprint is identified, the list of data relating sampled

portions, metadata (*content identification results*), and times are stored in the object database 240. EX1005, ¶¶86, 162, FIG. 5 (steps 525, 530).

189. Herley682 discloses: “Whether or not the input signals 500 are stored or buffered 505, they are then sampled 510.” EX1005, ¶161. “The size and period of the sample are dependent on both the length of any objects in the input signals 500, and the type of fingerprint engine being used.” EX1005, ¶161. “For example, the DDA-based fingerprint engine described above computes trace fingerprints 515 every 186 ms over samples of the input signal 500.” EX1005, ¶161. “Once the trace fingerprint has been generated 515 for a particular sample, that trace fingerprint is compared 520 to the fingerprints of known audio objects that are stored in the aforementioned metadata/fingerprint database 235.” EX1005, ¶161. “As discussed above, this metadata/fingerprint database 235 includes pre-computed fingerprints and metadata describing objects represented by the pre-computed fingerprints.” EX1005, ¶161.

190. In other words, a POSA would have understood that the data stored in Herley682’s sequence of content identification results in object database 240 were obtained by querying metadata/fingerprint database 235 (the electronic database) to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the metadata/fingerprint database 235 (the electronic

database). And, as explained above in for element [1.3], this process of searching, matching, and storing is completed for a series of samples from input signal 500, resulting in a sequence of content identifications results in object database 240 (stored in sequential order as taught by Seet). Therefore, a POSA would understand the claimed “sequence [...] is obtained by” the querying process just described. EX1005, ¶¶86, 160, 161; EX1006, ¶32, FIG. 4; EX1019, 6, FIG. 2.

6. Element [1.5]

[1.5] a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; and

191. I conclude that Herley682 teaches this element.

192. As I noted above with respect to elements [1.2] - [1.4], for a particular trace fingerprint (e.g., a representation of audio data captured representing the respective audio data sample), the metadata stored in the object store in Herley682’s object database 240 when a matching fingerprint is identified corresponds to the claimed content identification result.

193. Specifically, Herley682 discloses that the metadata comes from the pre-computed fingerprint of a known audio object found to be a likely match in

the metadata/fingerprint database 235. Herley682 discloses “Once the fingerprint extraction module 225 has extracted a fingerprint from the sampled input signal, that fingerprint is provided to a fingerprint comparison module 230.” EX1005, ¶86. “The fingerprint comparison module 230 then searches a metadata/fingerprint database 235 having pre-computed fingerprints and metadata describing objects represented by the pre-computed fingerprints.” EX1005, ¶86. “If a match is identified by the fingerprint comparison module 230, then the sampled portion of the input signal from which the fingerprint was extracted is identified as belonging to an object of interest corresponding to the fingerprint and metadata in the metadata/fingerprint database 235.” EX1005, ¶86. “This information, including the metadata, is stored in an object/station database 240.” EX1005, ¶86. “Therefore, for each identified object of interest, the object/station database includes statistical information such as the time that the object was identified, the station, channel, frequency, etc., where the input signal was monitored, and any metadata stored in the metadata/fingerprint database 235 for that particular object.” EX1005, ¶86.

194. This alone is sufficient to meet element [1.5] which recites parts (i) and (ii) in the alternative. However, Herley682 also teaches part (ii). Specifically, Herley682 discloses “[a]lternately, in another embodiment, if a matching fingerprint is not identified 525 in the metadata/fingerprint database 235, then

information characterizing the current sample, i.e., the trace fingerprint generated from that sample, and the time and source of the signal from which the sample was extracted, is stored to the metadata/fingerprint database 235 as an ‘unknown object’ entry, along with either a copy of the sample, or a pointer to a location in the input signal where the sample was taken.” EX1005, ¶164, FIG. 5 (step 535). “Further, any statistical information available for the sample, such as, for example, the broadcast time of the sample, and the source of the signal, is also stored to the metadata/fingerprint database 235.” EX1005, ¶164, FIG. 5 (step 535). “Therefore, any subsequent occurrences of an object having a matching fingerprint will be identified, as the metadata/fingerprint database 235 will include a fingerprint entry for that unknown object. In this case, both the first instance of the unknown object, and each subsequent instance, will then be added to the object database 240 along with any statistical information that is available for that object.” EX1005, ¶164, FIG. 5 (step 535). Thus, I conclude that the inclusion of the “unknown object” in object database 240 corresponds to *an indication of the absence of a respective likely match*, as recited in part (ii) or element [1.5]. EX1005, ¶164, FIG. 5 (step 535). I note that Seet further supports my conclusion inasmuch as Seet also discloses that when unable to identify a sample by using the sample’s neighbors, Seet acknowledges that the sample remains unidentified. EX1006, ¶32; EX1019, 6.

7. Element [1.6]

[1.6] the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database;

195. I conclude that the Herley682-Seet combination teaches or at least renders obvious element [1.6].

196. Herley682 discloses that “[g]iven a fingerprint engine such as the DDA-based fingerprint engine described above, identification of objects in one or more input signals serves as the basis for populating the aforementioned interactive object database.” EX1005, ¶120. “As noted above, when a match is identified via a fingerprint comparison, the metadata from the metadata database is attached to the portion of the stream then playing since it can be inferred that the object identified in the database is playing at that point in the stream.”

EX1005, ¶120. “Note that the term ‘object database’ is used interchangeably with the term ‘station database’ in the following discussion, as a tested embodiment of the object database includes object identification and metadata describing each object identified in one or more FM radio broadcast streams.” EX1005, ¶120.

“Further, as noted above, the discussion of FM radio audio streams is not intended to limit the interactive signal analyzer to use with radio signals. In fact,

the interactive signal analyzer may be used with any desired type of signal, and the discussion of the interactive signal analyzer in the context of FM radio signals is provided simply as one example of tested embodiment of the interactive signal analyzer.” EX1005, ¶120.

197. Continuing, Herley682 discloses “As noted above, if an object is detected more than once in an input signal, a counter for that object is incremented in the object database rather than creating a new entry for the already played object. Further, in one embodiment, the existing entry for that object in the D_s database is expanded to include the time and date of each new occurrence.” EX1005, ¶122. “Consequently, over time, the station database, D_s, will be populated with more and more information about the content that each monitored station plays.” EX1005, ¶122. “[S]tations will appear in the station database D_s with one or more entries, depending upon the number of objects played and identified.” EX1005, ¶123.

198. A POSA would have understood that radio stations typically play multiple songs. Indeed, Herley682 notes “it was observed that some stations play a limited collection of songs that are repeated fairly often, while others play larger collections.” EX1005, ¶123. Further, Herley682’s database supports data mining to determine, for example, “the top 100 most played songs on a particular radio station.” EX1005, ¶130. A POSA would also have understood that these

teachings indicate that Herley682's object database 240 would include multiple entries therein, e.g., at least one for each song played on the station. Further, a POSA would have understood or at a minimum found it obvious that, at least most of the time, these multiple entries would be comprised of either *different content identifiers*, identifying the different songs played on the station, or they would include *at least one content identifier* and an *indication of the absence of a respective likely match*, e.g., an "unknown object" as discussed above for element [1.5] part (ii). Therefore, a POSA would conclude that Herley682 discloses or renders obvious element [1.6].

199. I have considered Patent Owner's possible response to my conclusions, and to the extent Patent Owner argues this would not have been rendered obvious based on Herley682 alone, Seet expressly teaches a database of samples that includes multiple entries, both identified samples and unmatched samples. EX1006, ¶32, FIG. 4; EX1019, 6, FIG. 2. For example, Seet's database 110 teaches both *different content identifiers*, e.g., "Ad #2" and "Ad #1", and *at least one content identifier* and an *indication of the absence of a respective likely match*, e.g., "Ad #2" and "?" (which corresponds also to Herley862's "unknown object" as discussed above for element [1.5] part (ii)). EX1006, ¶32, FIG. 4; EX1019, 6, FIG. 2. Therefore, a POSA would conclude that the Herley682-Seet combination renders obvious element [1.6].

8. Element [1.7]

[1.7] (d) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and

200. The Herley682-Seet combination teaches or at least renders obvious element [1.7].

201. Initially, I note that as I discussed above in section V.C., data “scrubbing,” generally, is a well-known data cleansing and/or data error correction and/or data imputation technique.

202. The ’768 patent explains that a raw play stream is prepared “for further processing necessary for channel matching by scrubbing (i.e. reviewing for missing or incorrect data) the play stream using non-channel data.” EX1001, 8:32-35. The ’768 patent provides a contextual discussion of this “scrubbing” process in the discussion of Figure 4 describing “it is possible to locate and correct any identifications that are likely to be incorrect or to supply missing identifications where the system has not otherwise been able to provide an identification.” EX1001, 9:32-37. From this disclosure, a POSA would understand that the claimed “scrubbing” merely requires reviewing the data in the raw play stream for missing or incorrect data. However, the ’768 patent provides

no additional details about how the “scrubbing” is achieved. It does not provide details regarding whether the scrubbing is manually performed by a user or whether any reviewing for missing or incorrect data happens with or without user intervention. If *arguendo* the claim somehow required automatic scrubbing, it is nevertheless clear that if the only difference between the prior art (manual scrubbing) and the patent (automatic scrubbing with no algorithmic disclosure or special technique disclosed), then this is not a patentable distinction. It is merely using a conventional general-purpose computer to carry out an activity that can be performed by humans and the human mind.

203. I also note that the '768 patent describes the “sample sequence data” both in terms of its function (e.g., [...] “used to ‘scrub’ Raw Play Stream 410”, EX1001, 9:32-22; “derive portions of Clean Play Stream 450 from raw play list 410”, EX1001, 9:38-39) and in terms of its content (e.g., “includes content offset data,” EX1001, 9:26; “includes at least one of log times, content identifiers, and content offsets,” EX1001, 16:30-32, 15:8-10 (dependent claim 2)). Therefore, a POSA could conclude that the “sample sequence data” of claim 1 may include any of the disclosed content or perform any of the disclosed functions. *See also* EX1001, 10:34-56.

204. Further, I note that the '768 patent describes, with reference to Figure 4, that “Sample sequence data may be used to ‘scrub’ Raw Play Stream 410.”

EX1001, 9:32-37. “By analyzing the sequence of data on Raw Play Stream, it is possible to locate and correct any identifications that are likely to be incorrect or to supply missing identifications where the system has not otherwise been able to provide an identification.” EX1001, 9:32-37.

205. As best understood from the specification of the '768 patent, a POSA would have understood that an indication that an audio sample was not identified or is missing (e.g., the “none” entries in Figure 4 of the '768 patent) is a “pattern of the sample sequence data” that is not within “an expected pattern of sample sequence data” as recited in the claim, because an identification is expected.

206. Further, I note that '768 disclosure fails to provide meaning to “an expected pattern of sample sequence data” as the only appearance of “expected pattern” occurs in the '768 claims. However, “expected pattern” may be inferred from the disclosure given in relation to the raw data stream 410 of Figure 4 and the disclosure “Given this pattern of sample sequence data, it is reasonable to assume that the person using the corresponding monitor is listening to song51 through the time period associated with the audio data of samples 1-7, and thus, the ‘scrubbed’ version of the play stream, i.e. Clean Play Stream 450, indicates song51 for each of samples 1-7.” EX1001, 9:45-51. Thus, I have interpreted the '768 patent’s claimed “expected pattern of sample sequence data” as “reasonable assumptions from data points surrounding the missing or incorrect data.”

207. A POSA would have understood that an indication that an audio sample was not identified or is missing (e.g., “none” entries in Figure 4 of the ’768 patent) is “a pattern of the sample sequence data” that is not within an “expected pattern of sample sequence data” as recited in element [1.7] inasmuch as a POSA would know that a media source would generally be outputting known music or video content. Thus, as best understood in light of the specification of the ’768 patent, a POSA would have understood that “scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data” includes identifying and modifying data in “the raw play stream” that indicates a missing content identification result based on reasonable assumptions from data points (e.g., content identification results) near the missing data in the raw play stream. I conclude that Herley682 teaches or renders obvious claim element [1.7] for the following reasons.

208. As discussed above for elements [1.3] and [1.5], the collection of data stored in Herley682’s object database 240 teaches the claimed “raw play stream” and this data includes at least sampled portions, metadata (e.g., the ’768 patent’s “content identifiers”), and times (e.g., the ’768 patent’s “log time”).

Therefore, I conclude that Herley682 teaches the claimed “sample sequence data” as the data contained in object database 240.

209. In addition to teaching the claimed sample sequence data, Herley682 discloses a methodology for handling unidentified data samples within the sequence. *See* EX1005, ¶¶17, 18, 71, 88, 89, 164, 165. More particularly, a POSA would understand that Herley682 describes “unknown trace fingerprints” and includes storing the unknown data sample(s) and subsequent manual identification of the “unknown” data sample. Herley682 discloses that “[i]n a related embodiment, **if the fingerprint extracted from the sampled input signal does not match any entries in the metadata/ fingerprint database 235, that fingerprint is stored to the metadata/fingerprint database as an “unknown object” entry, along with either a copy of the sample, or a pointer to a location in the input signal where the sample was taken.**” EX1005, ¶88 (emphasis added). “Further, any statistical information available for the sample, such as, for example, the broadcast date and time of the sample, and the station, channel, frequency, etc. where the input signal was monitored.” EX1005, ¶88. “If any subsequent occurrences of an object having a matching fingerprint are identified, then both the first instance of the unknown object, and each subsequent instance, will then be added to the object/station database 240 along with any statistical information that is available for that object.” EX1005, ¶88.

210. Continuing, Herley682 discloses “[f]urther, in another related embodiment, the metadata/ fingerprint database 235 is open to user browsing and editing, so that unknown objects can be manually identified by a user via a local user interface 245, or alternately, via a remote client user interface 260.” EX1005, ¶89 (emphasis added). “In addition, whenever any metadata in the metadata/fingerprint database 235 is edited or modified, any corresponding entry in the object/station database 240 is automatically updated to reflect the changes in the metadata.” EX1005, ¶89. “In th[is] fashion, the two databases are kept synchronized with respect to metadata content.” EX1005, ¶89.

211. A POSA would have understood from this disclosure that Herley682’s “unknown” data entries correspond to the ’768 patent’s “none” data entries in the respective “raw play streams.” These unknown/unmatched entries are not within an “expected pattern of sample sequence data” because it would be expected to find a match to the sample. Because of this expectation (that is, to find a match to every sample), a POSA would have been motivated to fill in the metadata (e.g., Herley682’s “unknown data” or equivalent to the 768 patent’s “content identifiers”) in Herley682 because having a complete data set would be more useful for any subsequent statistical data analysis by Herley682’s customers. In fact, Herley682 teaches that these “unknown objects” results can and should be changed. EX1005, ¶¶17, 71, 89. Herley682 teaches that users can

“manually identify such unknown objects” in the metadata/fingerprint database.

EX1005, ¶¶18, 71, 89. In Herley682, when the metadata for the unknown objects is edited in the metadata/fingerprint database 235, “any corresponding entry in the object/station database 240 is automatically updated,” so the two databases are kept synchronized with respect to the metadata content. EX1005, ¶¶89, 165.

In this manner, a complete data set can be generated.

212. A POSA would understand from Herley682 that a complete data set is important. Herley682 teaches that a user may query the object/station database 240 with any number of different requests. *See* EX1005, ¶¶125-152. However, if the object/station database 240 contains incomplete or inaccurate data the query may return inaccurate results. A POSA would have thus been motivated to identify unknown/unmatched data entries in order to obtain a complete and accurate data set in the object/station database 240.

213. In summary, if the pattern of the sample sequence data in Herley682’s object/station database 240 does not correspond to an expected pattern of sample sequence data, for example because certain data is “unknown”/missing, it would have been obvious to a POSA to determine whether to change a result of the sequence of content identification results, based on other content identification results similar in time, resulting in an automatic update to the metadata in object database 240. EX1005, ¶¶89, 165.

214. From these teachings, a POSA would have been motivated to determine to change a result of the sequence of content identification results, if the pattern of the sample sequence data in Herley682's object/station database does not correspond to an expected pattern of sample sequence data because the data is unknown or missing. In other words, it would have been obvious, upon noticing missing data in object/station database 240 (i.e., "unknown" objects), to thereafter manually identify the unidentified audio, based on Herley682's express teaching of a user interface 245, 555 provided for at least that purpose. A POSA would have been motivated to do this in order to obtain a complete data set in the object/station database 240.

215. In addition, a POSA, aware of the Seet teachings, would be further motivated to identify Herley682's missing data because, as discussed above in VII.B, Seet explicitly teaches how to infer (identify) missing data. EX1006, ¶32; EX1019, 6, FIG. 4. A POSA would understand that Herley682's manual identification would be time intensive (slow), and Seet's inference technique of missing data imputation would be implemented quickly via software. Further, as discussed above in Section V.C, it was well known in the art at the time of the '768 patent to use other surrounding data in order to clean, or "scrub" a data set. Thus, a POSA would have looked to Seet's teachings of inferring matches as an improved way to identify Herley682's unmatched objects as a POSA would have

been motivated to automate Herley682's manual processing with Seet's inference technique, as such would be faster – a stated goal of Herley682. EX1005, ¶80.

216. More particularly, in Seet, “[a] matching inference may be made for some samples based on their position in a media stream 102 relative to other matched samples.” EX1006, ¶32; EX1019, 6. In particular, an unmatched sample may be identified using “surrounding samples [that are] in the correct sequence and timing so that a match for the unmatched sample can be made.” EX1006, ¶32; EX1019, 6. The example in Seet's Figure 4 identifies unmatched samples in the same way as disclosed in Figure 4 of the '768 patent. That is, to identify the unmatched second sample (*determine whether to change a result of the sequence of content identification results*), Seet looks to the previous sample and the following sample, and “[t]he temporal consistency of the bounding samples allows one to infer that they are of the same media item ... (in this case, an advertisement labeled #2); therefore, the middle sample is also of the same media item.” EX1006, ¶32; EX1019, 6. Seet, thus, uses *an expected pattern of sample sequence data*, to identify and correct missing data in the raw play stream.

217. Given the above interpretations and analysis, I conclude that the Herley682-Seet combination teaches or at least renders obvious all of the features of element [1.7].

9. Element [1.8]

[1.8] (e) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.

218. I conclude the Herley682-Seet combination teaches element [1.8].

219. Initially, I note that nothing the language of claim 1, or within the disclosure of the '768 patent, excludes manual changing of the data in the play stream. In fact, the '768 patent provides no guidance as to how any changes to the raw play stream are made. Herley682 teaches a user (via a user interface 245, 555) to manually identify the unidentified audio either directly in object database 240, or alternatively in meta/fingerprint database 235, resulting in a corresponding automatic updates being made to the metadata in object database 240. EX1005, ¶¶17, 18, 71, 89, 165. The updated collection of data in object database 240 corresponds to generating a clean play stream. This interpretation is consistent with the discussion in the '768 patent, in which the “none” entries in the raw play stream 410 are updated to include a content identifier in clean play stream 450. *See* EX1001, FIG. 4 (e.g., sample 2, sample 10).

220. Additionally, as discussed for element [1.7], Herley682 in view of Seet teaches identifying and updated the information for unidentified audio in the object database 240. EX1005, ¶¶17, 18, 71, 89, 165. A POSA would understand that the *updated* collection of data in object database 240 corresponds to the

clean play stream. Similarly, as described above in Sections VII.B, a POSA would understand that that a data structure containing Seet's inferred data taught in the '681 provisional application and in Seet are equivalent to the "*clean play stream*" described in the '768 patent.

221. In light of the above discussion, I conclude that a POSA would have found claim 1 of the '768 patent obvious in light of the disclosures of Herley682 and Seet.

C. Dependent Claim 2

[2] The method of claim 1 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.

222. Herley682 teaches dependent claim 2. Herley682 discloses that "in one embodiment, when a match is identified, the metadata from the metadata database is associated with the portion of the stream then playing, via an object database, since it can be inferred that the object identified in the database is playing at that point in the stream." EX1005, ¶29. "In one embodiment, this data is provided to a user in real-time as the stream is playing." EX1005, ¶29. "Once an object has been identified in the audio stream, the information describing that object, along with statistical information such as time and date played, and the particular radio station on which it was played is stored to an object or 'station' database, D_s." EX1005, ¶29.

223. Further, as explained above with respect to elements [1.3], [1.5], and [1.7], the data stored Herley682's object database 240 includes at least sampled portions, metadata (content identifiers), and times (log times). Thus, I conclude that Herley682 teaches the sample sequence data as recited in claim 2.

D. Dependent Claim 3

[3] The method of claim 1 wherein: scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content of identification results.

224. The '768 patent describes that channel data may be used in the scrubbing process with the disclosure of "[i]nformation about the sequence of content in a play stream can be used to generate a play list and identify a channel, and the identified channel data can also potentially be used to help further 'scrub' a play stream." EX1001, 9:52-55. I note that the '768 patent continues, with reference to Figure 4: "With respect to further 'scrubbing' the Raw Play Stream 410, in this example, channel data may be used to correct the apparent misidentification of sample 29." EX1001, 10:5-15. "Note that the sample sequence data alone in this example does not provide a clear basis for correcting sample 29." EX1001, 10:5-15. "Sample 28 is identified as 'ad49', sample 29 as 'song81,' and sample 30 as 'song35.'" EX1001, 10:5-15. "However, given the sequence 'song21, ad49, and song35' is associated with a known sequence of the radio broadcast referenced as 'radio 35,' there is a basis in channel data 470 for

believing that the proper content ID for sample 29 is ‘song35’ rather than ‘song81.’” EX1001, 10:5-15. In other words, known channel data may be used to “scrub” the raw play stream to determine the correctness of the identified samples in the “raw play stream” and to correct (e.g., change a result in the sequence) any irregular or incorrect data in the “clean play stream.”

225. The ’768 patent, thus, describes using known information about the identification of songs played on a particular channel to provide information helpful to identify missing or incorrect data in the raw play stream. Although in the embodiment described in the ’768 patent, the method identifies the channel from the play stream information (e.g., EX1001, 8:50-65, FIG. 3 (steps 304, 305, 306)), the claim is not so limited and does not exclude the case where the channel listened to is already known, as in Herley682 (e.g., EX1005, ¶80).

226. I conclude that the Herley682-Seet combination similarly utilizes channel data to facilitate identification of audio content. Thus, the Herley682-Seet combination teaches or at least renders obvious dependent claim 3. Herley682 stores additional statistical information (including channel data) with each “unknown” data entry. Specifically, Herley682 discloses that, “if the fingerprint extracted from the sampled input signal does not match any entries in the metadata/ fingerprint database 235, that fingerprint is stored to the metadata/fingerprint database as an ‘unknown object’ entry, **along with either a**

copy of the sample, or a pointer to a location in the input signal where the sample was taken.” EX1005, ¶88 (emphasis added). “Further, **any statistical information available** for the sample, such as, for example, the broadcast date and time of the sample, and the **station, channel, frequency, etc. where the input signal was monitored.”** EX1005, ¶88 (emphasis added). “If any subsequent occurrences of an object having a matching fingerprint are identified, then both the first instance of the unknown object, and each subsequent instance, will then be added to the object/station database 240 along with any statistical information that is available for that object.” EX1005, ¶88. A POSA would understand these teachings describe that Herley682 uses known information about the content (including unknown content) played on a particular channel to help identify the unknown content.

227. I also note that as in the '768 patent, Herley682 teaches that once a channel is known, using known information about the songs played on a particular channel can be used to facilitate faster identification of audio content. One way in which Herley682 describes identifying unknown objects is by “automatically queuing up a more authoritative and up-to-date fingerprint database.” EX1005, ¶71. Another way is found in Herley682’s disclosure of “when listening to a particular channel, the fingerprints of previously identified objects from that channel are searched first in the database, thereby speeding the

search.” EX1005, ¶80. “For example, when listening to a station that plays pop hits from the 1980's repetitively it makes sense to search the fingerprints of songs previously identified on that station, before searching the larger database.”

EX1005, ¶80.

228. On the other hand, Seet's method for inferring unknown data cannot be used for all data samples using Seet's method. For example, Seet teaches that matching for “the sixth sample ... is not inferred because it is not bounded by matched samples that correspond to the same media item.” EX1006, ¶32; EX1019, 6, FIG. 2 (identity cannot be inferred as insufficient data). In such cases, a POSA would have turned next to querying Herley682's more authoritative database to determine a match in order to quickly compile as complete a data set as possible. A POSA would have desired to facilitate fast identification of “unknown” data in the Herley682-Seet combination system. Therefore, a POSA would have been motivated by Herley682's related teaching of searching previously identified objects from the same channel (e.g., “analyzing channel data”) when querying the additional database to identify the unknown object. This would have been a natural extension of using this channel information when identifying the objects in the first instance.

229. Alternatively, Seet's method for inferring unknown data itself relies on *channel data* because it takes into account previous samples, e.g., from the

channel, when it determines whether it can infer matches. EX1006, ¶¶27 (previous samples may be from a particular station), ¶32 (relying on previous samples to infer matches in current samples database); EX1019, 4, 6. In the example of Seet’s Figure 4, a match is inferred when the sequence in the current samples (database 110) is consistent with the sequence in the previous samples (database 108) from that channel. A POSA would have, thus, understood that when combining Seet’s inference method with Herley682’s teaching of identifying such unmatched samples, that the combined teachings result in *analyzing channel data as part of determining whether to change a result of the sequence of content identification results.*

E. Dependent Claim 5

[5] The method of claim 1 wherein a clean play list is generated from the clean play stream.
--

230. I note that the ’768 patent describes that after scrubbing, “[t]he resulting clean play stream may be converted to a play list listing the content captured and indicating the order in which it was captured.” EX1001, 8:6-8, FIG. 4; EX1001, 9:52-53. The ’768 patent also describes that information about the sequence of content in a play stream can be used to generate a play list. EX1001, 11:27-28 (“[Clean] Play List 560 contains a row for each internally continuous

media consumption event.”). In other words, a POSA would have understood a “clean play list” to encompass a list of the content identified in the audio data.

231. Given these disclosures, I conclude that Herley682 teaches or renders obvious dependent claim 5. Herley682 allows users to query the object/station database 240 to “view information describing monitored input streams.” EX1005, ¶94; EX1005, ¶¶176-186 (describing exemplary queries). In one example, “a predefined database query string” is provided, such as “a dropdown menu listing call signs of each monitored radio station.” EX1005, ¶94. “User selection of a particular radio call sign may then automatically call up a display of statistical information regarding that radio station, and the objects of interest, songs, commercials, etc., identified for that radio station.” EX1005, ¶94. Herley682 also teaches that the “identification information” of individual objects in the signal (the data in object/station database 240) can be “used for any number of purposes” including “generation of play lists or the like for cataloging a media stream type signal.” EX1005, ¶7.

232. Additionally, Herley682 discloses how to translate a series of identified samples into a single identified song via an increment counter. Herley682 discloses “[h]owever, if an object is detected more than once on the same station, it is better to simply increment a counter for that object rather than creating a new entry for the already played object.” EX1005, ¶29. “Note that it is

quite common for a second or subsequent instance of a particular object of interest to be identified in the input signal.” EX1005, ¶87. “For example, it is quite common for a particular song to be played on one or more radio stations throughout the day.” EX1005, ¶87. “Therefore, multiple instances of that song will be identified on monitored radio stations.” EX1005, ¶87. “Therefore, rather than creating a new entry in the object/station database 240, a counter representing a total number of identifications for that object of interest is simply incremented in the existing entry for that object, along with the statistical information documenting the time that the object was identified, and the station, channel, frequency, etc. where the input signal was monitored.” EX1005, ¶87. A POSA would understand that fewer data entries representing the same identified song would increase storage capacities and speed processing, and therefore be motivated to simplify a series of identified song samples into a single identified song. A POSA would also be motivated to generate a list of identified songs (a clean play list), as opposed to a scrubbed list of identified samples, because this data is more useful to companies that analyze media usage. The list of identified songs is more useful for such purposes.

233. Based on these disclosures, a POSA would understand the desirability of creating a clean play list for cataloging a media stream signal. It would have been obvious to generate a clean play list of songs, as claimed, from

the identification information in the updated collection of data in object database 240 (clean play stream, see element [1.8]). Herley682 recognizes that the available “queries are not intended to be limited to the user interface and query examples provided,” and also suggests creating a “play list[]” from the “identification information.” EX1005, ¶¶7, 186. Alternatively, a POSA would have been motivated to generate a clean play list from a clean play stream in order increase storage and processing efficiencies, or to provide more useful information to other companies.

F. Dependent Claim 6

[6] The media measurement method of claim 1 wherein the plurality of audio data representations include a plurality of audio signatures.

234. Herley682 teaches dependent claim 6. Herley682 discloses that “[t]he interactive signal analyzer uses a signal fingerprint extraction algorithm, i.e., a ‘fingerprint engine,’ for deriving traces from segments of one or more signals.” EX1005, ¶67. A POSA would have understood these audio fingerprints are a type of audio signature that allows an audio object to be uniquely identified. In fact, the ’768 patent expressly equates audio signatures and audio fingerprints with “the uploaded audio data may include parameters useable to calculate audio signatures (sometimes referred to as audio fingerprints).” EX1001, 3:46-48. This is consistent with the plain and ordinary meaning of a “signature.” EX1015 (“6:

something ... that serves to set apart or identify”; “7: FINGERPRINT”).

Herley682 describes that “[o]ne common method for automatically identifying such objects involves analyzing an input signal, or a predefined portion or segment of such a signal, to produce a set of parameters or features that are derived from the signal.” EX1005, ¶5. “These parameters or features are then stored as ‘fingerprints’ that uniquely identify such objects.” EX1005, ¶5. “These fingerprints may then be used to identify subsequent occurrences of objects in a similar signal.” EX1005, ¶5.

235. Therefore, in my opinion, a POSA would have understood that Herley682 teaches claim 6.

G. Dependent Claim 8

[8] The media measurement method of claim 1 wherein an audio data representation of the plurality of audio data representations includes a signal usable for generating audio energy corresponding to audio energy captured at the media monitoring device.

236. Initially, I note that the ’768 patent provides no explanation or guidance on the scope or meaning of “a signal usable for generating audio energy corresponding to audio energy captured at the media monitoring device.” Therefore, the ordinary and customary meaning of the terms, as would be understood by one of ordinary skill in the art and the prosecution history will be used herein.

237. Herley682 teaches dependent claim 8. As I explained above for claim element [1.2], the plurality of audio data representations correspond to Herley682's pre-computed fingerprints, as stored in metadata/fingerprint database 235. EX1005, ¶¶88, 121, 161, 162. I note that Herley682 further discloses that unidentified samples are “stored in the metadata/fingerprint database 235 as an ‘unknown object’ entry, along with either a copy of the sample, or a pointer to a location in the input signal where the sample was taken.” EX1005, ¶164. Afterwards, the metadata/fingerprint database 235 would include a copy of the unidentified sample. A POSA would have understood that this copy of the sample is a copy of an audio signal having the ability to play back the captured media because Herley682 describes that “storing the input signals in the archived input signal database 220 allows a user to play back one or more portions of an input signal, or even jump to a point in an archived signal corresponding to an identified object of interest.” EX1005, ¶83. Thus, Herley682 explicitly discloses a “play back” feature of captured signals. Further, Herley682 describes that “audio clips or samples, such as a particular song, commercial, speaker, etc.” are provided to the fingerprint engine with “user entry of signal samples is not to be limited to audio clips” and Herley682 describes that storing a copy of a sample allows the user to “listen to [the] sample.” EX1005, ¶¶26, 121. In other words, the stored sample is and audio clip which is a signal usable for

generating audio energy corresponding to audio energy captured at the media monitoring device. Therefore, in my opinion, a POSA would have understood that Herley682 teaches claim 8.

H. Independent Claim 9

1. Preamble [9.P]

[9.P] A computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers, the one or more computers being coupled to a network to receive, over the network, data for a sequence of audio data samples, the audio data samples comprising representations of audio data captured at a media monitoring device, the plurality of audio data samples being submitted over the network, the instructions comprising instructions for:

238. Initially, I note that the '768 patent explains “[c]omputer program product 760 comprises executable code in an electronically readable medium that may instruct one or more computers such as computer system 700 to perform processing that implements the system 100 of FIG. 1 and/or accomplishes the exemplary method 200 of FIG. 2.” EX1001, 13:57-62. For the remaining aspects of the preamble, see my comments at [1.P] above.

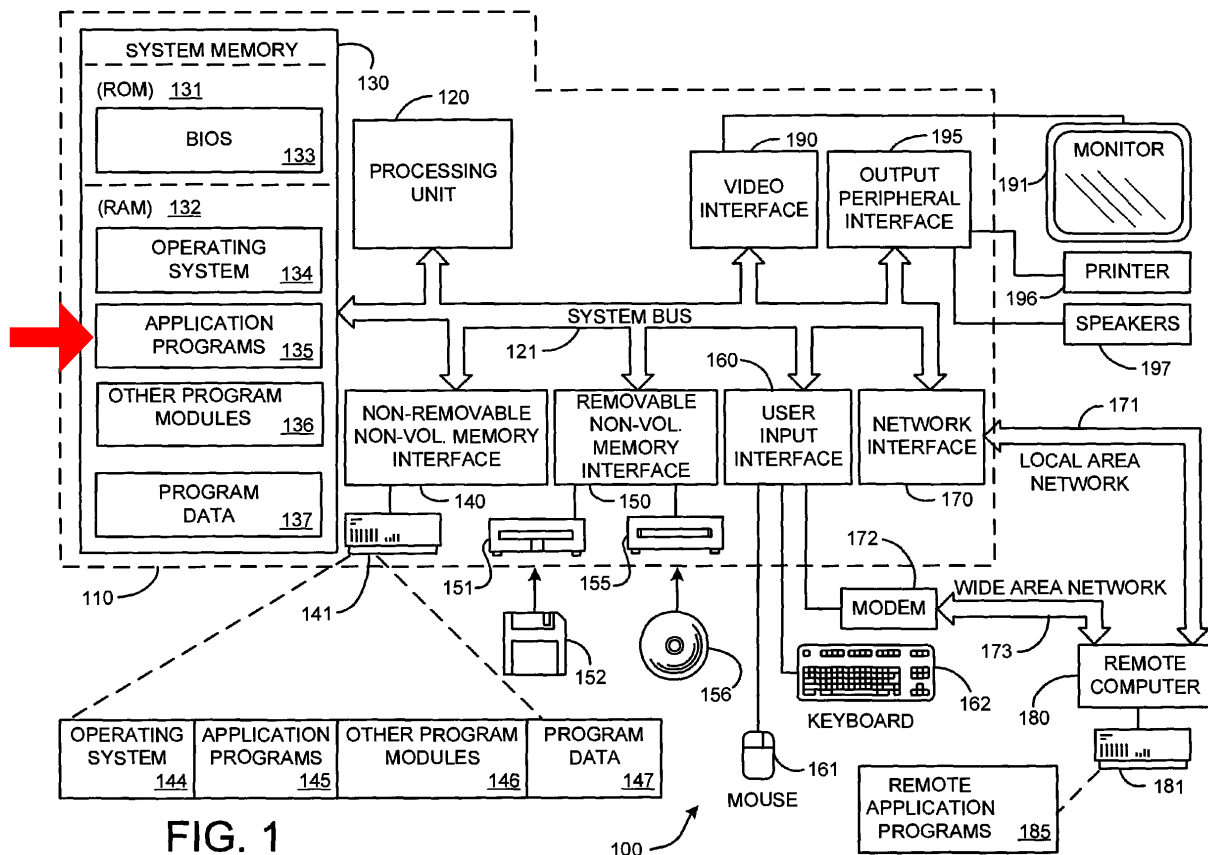
239. To the extent the preamble of claim 9 is limiting, Herley682 teaches the preamble. For example, Herley682, discloses that “[t]he invention may be described in the general context of computer-executable instructions, such as

program modules, being executed by a computer.” EX1005, ¶51. “Generally, program modules include routines, programs, objects, components, data structures, etc., that perform particular tasks or implement particular abstract data types.” EX1005, ¶51. “The invention may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. In a distributed computing environment, program modules may be located in both local and remote computer storage media including memory storage devices.” EX1005, ¶51. “With reference to FIG. 1, an exemplary system for implementing the invention includes a general-purpose computing device in the form of a computer 110.” EX1005, ¶51.

240. Herley682 also discloses that “system memory 130 includes computer storage media in the form of volatile and/or nonvolatile memory such as read only memory (ROM) 131 and random access memory (RAM) 132.” EX1005, ¶55. “A basic input/output system 133 (BIOS), containing the basic routines that help to transfer information between elements within computer 110, such as during start-up, is typically stored in ROM 131. RAM 132 typically contains data and/or program modules that are immediately accessible to and/or presently being operated on by processing unit 120.” EX1005, ¶55. “By way of example, and not limitation, FIG. 1 illustrates operating system 134, application

programs 135, other program modules 136, and program data 137.” EX1005,

¶55. From these teachings, a POSA would have understood that Herley682’s system and method, when implemented on Figure 1’s computing device as program modules such as those shown in Figure 2, teaches a computer program product in a non-transitory computer readable medium comprising instructions executable by one or more processors of one or more computers, as recited in element [9.P].



EX1005, FIG. 1 (annotated).

241. To the extent that the preamble of claim 9 is limiting, and for the reasons I discussed both here, and above with respect to the preamble of claim 1, (element [1.1]) I conclude that Herley682 teaches the preamble of claim 9. The remaining elements of claim 9 recite the specific instructions included in the computer program product. Each of these instructions corresponds to a step in the method recited in claim 1, and is taught by Herley682, or the Herley682-Seet combination, for the same reasons discussed above and noted below.

2. Element [9.1]

[9.1] (a) using the one or more computers to query an electronic database of a plurality of audio data representations and corresponding content identifiers;

242. For the reasons that I discussed above in regard to claim element [1.2] above, I conclude that Herley682 teaches element [9.1].

3. Element [9.2]

[9.2] (b) generating a raw play stream, the raw play stream comprising a sequence of content identification results corresponding to the sequence of audio data samples; wherein:

243. For the reasons that I discussed above in regard to claim element [1.3] above, I conclude that the Herley682-Seet combination teaches element [9.2].

4. Element [9.3]

[9.3] the sequence of content identification results is obtained by querying the electronic database to attempt to determine respective likely matches between respective audio data samples in the sequence of audio data samples and respective audio data representations in the electronic database;

244. For the reasons that I discussed above in regard to claim element [1.4] above, I conclude that Herley682 teaches element [9.3].

5. Element [9.4]

[9.4] a content identification result of the sequence of content identification results comprises either: (i) a content identifier associated in the electronic database with a respective audio data representation that is determined to be a respective likely match with a respective one of the audio data samples; or (ii) an indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database; and

245. For the reasons that I discussed above in regard to claim element [1.5] above, I conclude that Herley682 teaches element [9.4].

6. Element [9.5]

[9.5] the raw play stream includes either: at least two different content identifiers obtained from the electronic database; or at least one content identifier obtained from the electronic database and at least one indication of the absence of a respective likely match between a respective audio data sample and an audio data representation in the electronic database;

246. For the reasons that I discussed above in regard to claim element [1.6] above, I conclude that the Herley682-Seet teaches element [9.5].

7. Element [9.6]

[9.6] (c) scrubbing the raw play stream by analyzing sample sequence data of the raw play stream to determine whether to change a result of the sequence of content of identification results in view of a pattern of the sample sequence data of the raw play stream compared to an expected pattern of sample sequence data; and

247. For the reasons that I discussed above in regard to claim element [1.7] above, I conclude that the Herley682-Seet combination teaches element [9.6].

8. Element [9.7]

[9.7] (d) generating a clean play stream from the raw play stream by making any changes to the raw play stream that are determined to be made by the scrubbing.

248. For the reasons that I discussed above in regard to claim element [1.8] above, I conclude that the Herley682-Seet combination teaches element [9.7].

I. Dependent Claim 10

[10] The computer program product of claim 9 wherein sample sequence data includes at least one of log times, content identifiers, and content offsets.

249. For the reasons that I discussed above in regard to claim [2] above, I conclude that Herley682 teaches claim [10].

J. Dependent Claim 11

[11] The computer program product of claim 9 wherein scrubbing the raw play stream further comprises analyzing channel data as part of determining whether to change a result of the sequence of content identification results.

250. For the reasons that I discussed above in regard to claim [3] above, I conclude that the Herley682-Seet combination teaches claim [11].

K. Dependent Claim 13

[13] The computer program product of claim 9 wherein the computer program product further comprises instructions for generating a clean play list is generated from the clean play stream.

251. For the reasons that I discussed above in regard to claim [5] above, I conclude that Herley682 teaches claim [13]. A POSA would have understood a predefined database query that results in a play list to teach a computer program product having instructions for generating a clean play list.

L. Dependent Claim 14

[14] The computer program product of claim 9 wherein the plurality of audio data representations include a plurality of audio signatures.

252. For the reasons that I discussed above in regard to claim [6] above, I conclude that Herley682 teaches claim [14].

IX. GROUND 2: CLAIMS 1-6 AND 8-14 OF THE '768 PATENT ARE UNPATENTABLE UNDER 35 U.S.C. § 103 UNDER HERLEY682 IN VIEW OF SEET AND HERLEY864

A. Motivation to Combine Herley864 with Herley682 and Seet

253. A POSA would have found it obvious to combine Herley682 and Seet for the reasons discussed above in Section VIII.A. A POSA would have found it obvious to further look to Herley864 for multiple reasons. Both Herley682 and Herley864 share many similarities, structures, goals, and methodologies. For example, both Herley682 and Herley864 relate to methods

and systems of identifying objects of interest (e.g., songs) within signals (e.g., a radio or television (media) broadcast). They both do so by receiving signal samples, extracting sample “fingerprints,” and comparing these to fingerprints of known objects in order to determine if there is a match and thereby identify the received signal. Both Herley682 and Herley864 also describe how to handle the situation when there is not a match.

254. For example, Herley682 discloses a “system and method for sampling one or more channels of a broadcast spectrum, such as a radio frequency spectrum, identifying and storing content information for each sampled channel, and providing a user interface for allowing interactive user queries and display of the stored content information.” EX1005, ¶2. Herley682’s “interactive signal analyzer” among other things, “monitors one or more signals, identifies objects of interest within such signals, [and] stores statistical and metadata information describing such objects to an object database.” EX1005, ¶10. And Herley682 teaches that these monitored signals may be, for example, audio signals, video signals, or combined audio/video signals. EX1005, ¶11.

255. Herley864 discloses a “system and method for jointly segmenting and identifying repeating objects from one or more broadcast audio streams.” EX1007, 1:16-19. Herley864’s “repeating object extractor” “automatically segments and identifies audio objects such as songs, station identifiers, jingles,

advertisements, emergency broadcast signals, etc., in an audio stream.” EX1007, Abstract. And Herley864 also describes performing analysis of the content of at least radio or television signals. *See, e.g.*, EX1007, 2:13-26.

256. I also note that Herley682 and Herley864 are in the same field of endeavor, and even share an inventor. *Compare* EX1005, ¶2, *with* EX1007, 1:16-19; *compare* EX1005, (75), *with* EX1007, (75). Because of these similarities and more, a POSA would have been motivated to combine these references, and would have had a reasonable expectation of success in doing so, because they are closely related and both disclose methods of identifying objects of interest in a broadcast signal. To apply the teachings of Herley864 to the system and method taught in the Herley682-Seet combination would have simply been the combination of prior-art elements according to known methods to yield predictable results, a concept that I understand as a motivation in the legal construct of obviousness. Similarly, the combination would have merely been the application of known techniques (Herley864’s system/method) to a known system/method ready for improvement (Herley682’s system/method) to yield the predictable result of identifying objects of interest in a broadcast signal, which I understand as a motivation in the legal construct of obviousness.

257. I also conclude that a POSA further would have had a reasonable expectation of success in combining these references because they both teach

implementing their inventions in similar operating environments. In particular, both teach performing the methods thereof in the same “general-purpose computing device constituting an exemplary system for automatically identifying” either “sampled signals” (in Herley682) or “audio objects in an audio stream” (in Herley864). EX1005, ¶37; EX1007, 4:11-14; *compare* EX1005, FIG. 1, *with* EX1007, FIG. 1. I believe that it would have been well within the level of ordinary skill to implement Herley864’s teachings in the system of the Herley682-Seet combination because these exemplary suitable operating environments are described in the same way. See, for example, EX1005, ¶¶48-61; EX1007, 4:34-7:15.

258. For the reasons outlined below, A POSA would have found claims 1-6 and 8-14 obvious over Herley682, Seet, and Herley864.

B. Independent Claims 1 and 9

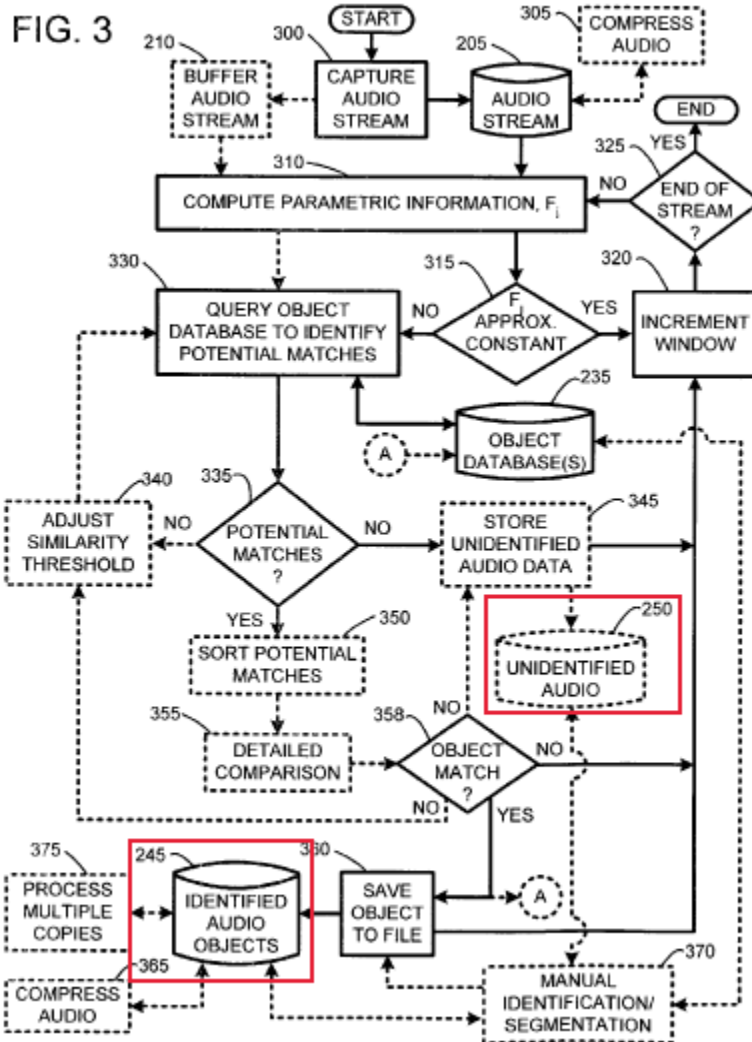
259. Claims 1 and 9 would have been obvious over the Herley682-Seet combination for all the reasons discussed above. *See* Sections VIII.B and VIII.H. Additionally, elements [1.8] and [9.7] further would have been obvious in view of Herley682, Seet, and Herley864 for the reasons discussed below.

260. The Herley682-Seet combination teaches element [1.8], [9.7] as discussed above. *See* Sections VIII.B.9 and VIII.H.8. I have considered Patent Owner’s possible response to my conclusions, and to the extent Patent Owner

argues that generating a clean play stream as recited in element [1.8], [9.7] requires generation of a new data set, rather than simply updating data in the Herley682-Seet combination's object database. I conclude that this element would have nonetheless been obvious in view of Herley682, Seet, and Herley864.

261. For example, In Herley682, once an “unknown object” is identified (either identified by a user change to the data in the object database 240, or alternatively in the metadata/fingerprint database 235) the “corresponding entry in the object/station database 240 is automatically updated to reflect the changes in the metadata.” EX1005, ¶¶ 17, 18, 71, 89, 165. Therefore, in Herley682, the raw play stream (data in object/station database 240 before updating) and the clean play stream (data in object/station database 240 after updating) are stored in the same place.

262. Herley864 teaches an alternative way of handling the unknown objects. As seen in Herley864's Figure 3, reproduced below, “unidentified audio portions 250 are saved to a computer readable medium for later manual ... identification via a user interface module” and the “identified audio objects 245 are stored to a [different] computer readable medium.” EX1007, 16:19-45, FIG. 3.



EX1007, FIG. 3 (annotated).

263. As shown in Herley864's Figure 3, once an unidentified audio object is identified, it is stored to the identified audio objects database. EX1007, FIG. 3. Separate storage, according to Herley864, allows for later manual segmenting, identification, or processing via a user interface. EX1007, 18:31-43. Herley864 continues with a disclosure of automated vector processing in attempts to identify the unknown object. EX1007, 18:44-57. When a match is determined,

information about the sample is then stored in object database 235 and identified audio object database 245. EX1007, 18:58–19:10. A POSA would understand that storage in two databases results in generates a new data set which represents the clean play stream. The net result of Herley864’s data separation arrangement is that separate processing allows for faster unknown sample searching and faster unknown sample identification resolution. A POSA would have found it obvious to implement Herley864’s separation of identified and unidentified data objects in the Herley682-Seet combination’s system, to provide faster unknown sample searching and faster unknown sample identification resolution.

264. A POSA would understand that this would have been a matter of combining prior art elements according to known methods to yield predictable results and/or a simple substitution of one known element for another to obtain predictable results. In this combination, my interpretation of the “unidentified audio portions 250” correspond to the “raw play stream,” and “identified audio objects 245” correspond to the “clean play stream.”

C. Dependent Claims 4 and 12

[4], [12]: [preamble] wherein scrubbing the raw play stream further comprises using audience data as part of determining whether to change a result of the sequence of content of identification results.

265. I conclude that the Herley682-Seet combination in view of Herley864 teaches these claims. Turning to the ’768 patent for an understanding

of “audience data”, the patent describes that “[a]udience data may also be used in the ‘scrubbing’ process.” EX1001, 10:28-29. The ’768 patent continues, with reference to Figure 4: “for example, if ‘song81’ is radically different from the user’s known consumption habits and ‘song35’ is within the user’s known consumption habits, another basis might be provided for suspecting that ‘song35’ rather than ‘song81’ is the correct content ID for sample 29 of Play Stream 410.” EX1001, 10:29-33. Genre is commonly used to identify and/or classify user consumption habits in signal content determination art. Thus, a POSA would have understood this passage as teaching that one example of audience data is data related to a user’s “music genre” preferences. I also note that the ’768 patent provides additional examples of what type of data “audience data” comprises, all of which are known in the art. Specifically, the ’768 patent discloses that audience data comprises “age, gender, and other demographic information about the audience member,” “media consumption patterns of the audience member,” or “location of the audience member.” EX1001, 5:11-18. I review the Herley682 and Herley864 references with these disclosures in mind.

266. As I explained above in the sections for claims [3] and [11], a POSA would have been motivated to use Herley682’s teaching of searching previously identified objects from the channel (analyzing channel data) when querying the

additional database to identify the unknown object in order to facilitate faster identification as taught as desirable by Herley682. *See* EX1005, ¶80.

267. In a similar manner, Herley864 teaches limiting searches for unknown objects to a particular music genre as a way to reduce search and comparison times. Specifically, Herley864 discloses “separate databases for different music genres are maintained in one embodiment, thereby reducing search and comparison times when interested in particular genres.” EX1007, 9:18-20. Herley864 also discloses additional methodologies to speed matching. EX1007, 12:44-52. Thus, reducing the search time for matches is important to both Herley682 and Herley864.

268. With these teachings in mind, a POSA would have been motivated to search for unknown object matches based on genre data (e.g., audience data or known media consumption pattern) when identifying unknown objects, rather than channel data used in Herley682’s system, in order to facilitate faster identification searches. A POSA would understand that this modification would have been the use of known technique to improve similar devices (methods, or products) in the same way. Further, a POSA would have understood that the genre of music a user typically listens to is one example of “media consumption patterns of the audience member” because genre was commonly used to identify and/or classify user consumption habits. EX1001, 5:17-20. Thus, using genre as a

way to facilitate faster identification of unknown objects teaches the claim limitation of *using audience data*.

D. Dependent Claims 2, 3, 5, 6, 8, 10, 11, 13, and 14

269. I have previously concluded that the Herley682-Seet combination teaches the limitations of dependent claims 2, 3, 5, 6, 8, 10, 11, 13, and 14 for the reasons discussed above. *See* Sections VIII.C-VIII.G and VIII.I-VIII.L. Therefore, because independent claims 1 and 9 would have been obvious in view of Herley682, Seet, and Herley864, I conclude that dependent claims 2, 3, 5, 6, 8, 10, 11, 13, and 14 likewise would have been obvious over Herley682, Seet, and Herley864.

X. GROUND 3: CLAIMS 7 AND 15 ARE OBVIOUS IN VIEW OF HERLEY682 AND SEET (WITH OR WITHOUT HERLEY864) AND WANG060

A. Claims 7 and 15

270. Dependent claims 7 and 15 each recite “wherein an audio data representation of the plurality of audio data representations includes a plurality of signal landmarks.” As I have previously noted, the ’768 patent explains that signal landmarks are a part of the stored versions of the media content. In particular, the ’768 patent explains “audio signatures (sometimes referred to as audio fingerprints) and landmarks or other values useful in the content recognition process,” EX1001, 3:7-49, and “content may be stored signatures and

landmarks derivable from a song's audio, but not necessarily an audio file useable to play the song itself.” EX1001, 4:23-24. I have reviewed the prior art documents with this description of “signal landmark” in mind.

271. I conclude that Wang060 teaches this element, and would have rendered this element obvious when viewed in combination with the Herley682 reference, with or without Herley864 and the data imputation references.

272. Wang060, like both of the Herley references, teaches a method for recognizing a media sample, such as audio, by matching the media sample to one in a database index of known media files. EX1008, ¶18. And like the Herley references, Wang060 utilizes audio fingerprints. Wang060's database contains “fingerprints representing features at particular locations of the indexed media files.” EX1008, ¶18. The “particular locations” are preferably “reproducibly computable locations ... called ‘landmarks.’” EX1008, ¶20. Wang060 explicitly recognizes that its “method is particularly useful for recognizing audio samples.” EX1008, ¶21.

273. A POSA would have understood that signal landmarks were a known method by which audio data could be analyzed, as Wang060 discloses that “the specific technical processes used to implement the landmarking schemes of the invention are known in the art and will not be discussed in detail.” EX1008, ¶51. This is consistent with the '768 patent's description that “the uploaded audio data

may include parameters useable to calculate audio signatures (sometimes referred to as audio fingerprints) and landmarks or other values useful in the content recognition process,” without further elaboration. EX1001, 3:46-49.

274. Landmarks are expressly taught in Wang060. In particular, Wang060 teaches: “Landmarks occur at specific timepoints of the sound and have values in time units offset from the beginning of the file, while fingerprints characterize the sound at or near a particular landmark.” EX1008, ¶49. Wang060 uses its landmarks as a starting point for fingerprint computation. EX1008, ¶55. In addition, Wang060 teaches a method for constructing a searchable sound database index, in which “landmarks are first computed, and then fingerprints are computed at or near the landmarks.” EX1008, ¶48, FIG. 3; *see also* EX1008, ¶49 (“To index the sound database, each recording in the collection is subjected to a landmarking and fingerprinting analysis that generates an index set for each audio file.”), ¶74 (“An index set for a given sound recording is a list of pairs of values (fingerprint, landmark).”). In addition, Wang060 also teaches using the index to perform sound recognition. EX1008, ¶79.

275. Based on Wang060’s disclosure, a POSA would have been motivated to implement Herley682’s fingerprint-based content identification method using landmarks, as taught in Wang060, as including data to identify the starting point of a fingerprint computation, given that Herley682’s system calculates “trace

fingerprints” several times per second on an incoming signal (Herley682, ¶63).

The use of landmarks would allow identification and tracking of the samples used to generate the trace fingerprints. A POSA would have had a reasonable expectation of success in making this modification to Herley682 because both Herley682 and Wang060 are fingerprint-based methods of identifying content in a media, e.g., audio, stream. Wang060 further describes that “specific technical processes used to implement the landmarking schemes of the invention are known in the art.” EX1008, ¶51. It would have been within the skill of a POSA to implement these known landmarking schemes into Herley682’s content identification method. Further, a POSA, looking to improve the content-identification method of Herley682 would be motivated to find techniques to improve identification and a POSA would be motivated to seek out and adopt the teaching of Wang060, since it is in the same field of art and would improve identification in a known, easily implemented way.

276. Accordingly, based on claim 7’s dependence from claim 1, and based on claim 15’s dependence from claim 9, and the explanations regarding these independent claims above, claims 7 and 15 would have been obvious in view of either; (1) Herley682, Seet, and Wang060; or (2) and Herley682, Seet, Herley864, and Wang060.

XI. SECONDARY CONSIDERATIONS OF NON-OBVIOUSNESS

277. I understand that the obviousness analysis involves evaluating secondary considerations (also known as objective indicia) of nonobviousness (if available). I understand that any such considerations must have a nexus, or a close connection, to the novel aspects of the claimed invention. I understand that secondary considerations of nonobviousness include the following: commercial success, long-felt but unmet need, failure of others, copying, praise in the industry, and unexpectedly superior results as compared with the closest prior art.

278. Considering the information I have reviewed, as well as my general knowledge of the field of media usage monitoring and measurement, I am not aware of any objective evidence that might support the patentability of the claims of the '768 patent. Further, from my review of the file history (EX1002) of the '768 patent, I did not see that the applicant raised any secondary considerations of nonobviousness during prosecution. I am not otherwise aware of any publicly available evidence of secondary considerations of nonobviousness. Should any alleged considerations be presented, I reserve the right to opine on them.

XII. CONCLUSION

279. In signing this declaration, I recognize that the declaration will be filed as evidence in an *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 cross-examination in the case and that cross-examination will take place within the United States. If cross-examination is required, I will appear for cross-examination within the United States during the time allotted.

280. I hereby declare that all statements made herein of 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.

Executed on this 12th day of July, 2024, in Mobile, AL.



Samuel H. Russ, Ph.D.