

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

TIKTOK INC.,
Petitioner,

v.

NTECH PROPERTIES, INC.,
Patent Owner.

Case No. IPR2024-01339

U.S. Patent No. 9,317,597

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 9,317,597**

TABLE OF CONTENTS

I.	Introduction.....	1
II.	Mandatory Notices.....	2
	A. Real Party-in-Interest	2
	B. Related Matters.....	2
	C. Counsel and Service Information.....	2
	D. Payment of Fees	3
	E. Requirements For Inter Partes Review	3
III.	Grounds.....	3
IV.	Overview of the '597 patent	4
V.	Level of Ordinary Skill in the Art	4
VI.	Claim Construction.....	5
VII.	Summary of the Prior Art	5
	A. Marcus156 (Ex. 1009).....	5
	B. Harbick (Ex. 1010).....	7
	C. Hosken (Ex. 1011).....	7
	D. Byers (Ex. 1013).....	8
	E. Ferman (Ex. 1012).....	8
VIII.	Ground 1: Marcus156 and Harbick Render Claims 1-20 Obvious	8
	A. Proposed Modification	8
	B. Motivation to Combine	9

C.	Reasonable Expectation of Success	11
D.	Claim 1	13
1.	Element 1[pre].....	13
2.	Element 1[a]	13
3.	Element 1[b].....	14
4.	Element 1[c]	15
5.	Element 1[d].....	16
6.	Element 1[e]	19
7.	Element 1[f]	20
8.	Element 1[g].....	21
9.	Element 1[h].....	22
10.	Element 1[i].....	23
11.	Element 1[j].....	23
12.	Element 1[k].....	24
13.	Element 1[l].....	24
14.	Element 1[m].....	24
E.	Claim 2	29
F.	Claim 3	29
G.	Claim 4	29
H.	Claim 5	31
I.	Claim 6	31

J.	Claim 7	32
K.	Claim 8	32
L.	Claim 9	32
M.	Claim 10	33
N.	Claim 11	33
O.	Claim 12	33
P.	Claim 13	33
Q.	Claims 14-20	34
IX.	Ground 2: Marcus156, Harbick, and Hosken Render Claims 1-20 Obvious	35
A.	Proposed Modification	35
B.	Motivation to Combine	35
C.	Reasonable Expectation of Success	36
D.	All Claims and Claim Elements Other Than 1[h] and 1(i)	37
E.	Element 1[h]	37
F.	Element 1[i]	38
X.	Ground 3: Marcus156, Harbick, Hosken, and Byers Render Claims 8 and 16 Obvious.....	39
A.	Proposed Modification	39
B.	Motivation to Combine	39
C.	Reasonable Expectation of Success	40
D.	Byers Discloses and Renders Obvious “One or More Video Clips Comprise Advertising”	41

XI.	Ground 4: Ferman and Harbick Render Claims 1-7, 9-15, and 17-20 Obvious	41
A.	Proposed Modification	41
B.	Motivation to Combine	42
C.	Reasonable Expectation of Success	43
D.	Claim 1	43
1.	Element 1[pre].....	43
2.	Element 1[a].....	43
3.	Element 1[b].....	44
4.	Element 1[c].....	45
5.	Element 1[d].....	47
6.	Element 1[e].....	49
7.	Element 1[f]	51
8.	Element 1[g].....	51
9.	Element 1[h].....	52
10.	Element 1[i].....	54
11.	Element 1[j].....	54
12.	Element 1[k].....	54
13.	Element 1[l].....	55
14.	Element 1[m].....	55
E.	Claim 2	57
F.	Claim 3	57

G.	Claim 4	58
H.	Claim 5	59
I.	Claim 6	59
J.	Claim 7	59
K.	Claim 9	60
L.	Claim 10	61
M.	Claim 11	61
N.	Claim 12	62
O.	Claim 13	62
P.	Claims 14-15 and 17-20	63
XII.	Ground 5: Ferman, Harbick, and Byers Render Claims 8 and 16 Obvious ..	63
A.	Proposed Modification	63
B.	Motivation to Combine	64
C.	Reasonable Expectation of Success	64
D.	Claims 8 and 16: Ferman, Harbick, and Byers Render Obvious “One or More Video Clips Comprise Advertising”	65
XIII.	Instituting This IPR Would Be Equitable.....	65
A.	35 U.S.C. § 325(d) Supports Institution.....	66
B.	<i>Fintiv</i> Supports Institution.....	68
1.	Overlap Between Issues	69
2.	Likelihood of a Stay	69
3.	Proximity of the Court’s Trial Date to a FWD	69

4.	Investment in the Parallel Proceeding	70
5.	Whether the Petitioner and the Defendant in the Parallel Proceeding are the Same Party	70
6.	Other Circumstances, that Impact the Board’s Exercise of Discretion, Including the Merits	70
XIV.	No Secondary Considerations	65
XV.	Conclusion	71

LIST OF EXHIBITS

Exhibit	Description
Ex. 1001	U.S. Patent No. 9,317,597 (“the ’597 patent”)
Ex. 1002	Declaration of Dr. Kevin Almeroth (“Almeroth”)
Ex. 1003	U.S. Provisional Patent Application No. 60/482,597 (“the ’597 provisional”)
Ex. 1004	File History of U.S. Patent Application No. 10/874,819 (now abandoned) (“the ’819 application”)
Ex. 1005	File History of U.S. Patent No. 8,875,185 (U.S. Patent Application No. 13/195,709) (certain prior art omitted) (“the ’709 application”)
Ex. 1006	File History of U.S. Patent No. 9,317,597 (U.S. Patent Application No. 14/494,110) (“the ’110 application”)
Ex. 1007	File History of U.S. Patent No. 11,184,686 (U.S. Patent Application No. 15/131,873) (“the ’873 application”)
Ex. 1008	File History of U.S. Patent Application No. 17/529,374 (currently pending) (“the ’374 application”)
Ex. 1009	U.S. Patent No. 6,032,156 (“Marcus156”)
Ex. 1010	U.S. Patent No. 8,260,656 (“Harbick”)
Ex. 1011	International Patent Application Publication No. WO 01/06398 (“Hosken”)
Ex. 1012	U.S. Patent Publication No. 2002/0059584 (“Ferman”)
Ex. 1013	U.S. Patent No. 7,334,249 (“Byers”)
Ex. 1014	List of Recent Expert Witness Engagements of Dr. Kevin Almeroth
Ex. 1015	Curriculum Vitae of Dr. Kevin Almeroth
Ex. 1016	Complaint for Patent Infringement, <i>NTECH Properties, Inc. v. ByteDance Ltd. et al</i> , Case No. 2:24-cv-130 (E.D. Tex. Feb. 23, 2024), ECF No. 1

Exhibit	Description
Ex. 1017	Amended Complaint for Patent Infringement, <i>NTECH Properties, Inc. v. ByteDance Ltd. et al</i> , Case No. 2:24-cv-130 (E.D. Tex. May 24, 2024), ECF No. 6

LIST OF CHALLENGED CLAIMS

Claim No.	Claim Language
1[pre]	1. A method of providing video programming to a user, comprising the steps of:
1[a]	(a) receiving a first request for video programming at a computing device comprising a video server, by communication with a computing device associated with a user;
1[b]	(b) searching one or more databases of video clips based on the first request, wherein each said database is in communication with said video server computing device;
1[c]	(c) receiving one or more first search results at said video server computing device, wherein said first search results collectively correspond to a plurality of first video clips in said one or more databases, each of which is responsive to the first request, and wherein at least some of said responsive first video clips are each associated with one or more second video clips in said one or more databases;
1[d]	(d) identifying a plurality of first transitional clips in said one or more databases for association with at least some of said responsive first video clips;
1[e]	(e) defining first video programming comprising a selected plurality of said responsive first video clips and a selected plurality of first transitional clips that are concatenated according to a first temporal sequence;
1[f]	(f) providing a presentation of said first video programming to a display associated with said user computing device;
1[g]	(g) receiving an identification of a first video clip of said first video programming at said video server computing device;
1[h]	(h) searching said one or more databases of video clips according to a second request for second video programming

Claim No.	Claim Language
	that is based at least in part on a stored association between the identified first video clip and at least one second video clip;
1[i]	(i) receiving one or more second search results at said video server computing device, wherein said second search results collectively correspond to a plurality of second video clips in said one or more databases that are responsive to the second request, including at least one second video clip that is associated with the identified first video clip;
1[j]	(j) identifying a plurality of second transitional clips in said one or more databases for association with at least some of said responsive second video clips;
1[k]	(k) defining second video programming comprising a selected plurality of second video clips, including the designated second video clip, and a selected plurality of second transitional clips that are concatenated according to a second temporal sequence; and
1[l]	(j) presenting said second video programming to a display associated with said user computing device;
1[m]	wherein at least one searching or defining step is performed according to one or more templates which collectively comprise a plurality of template variables comprising (a) information about a user associated with the second computing device, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c)“style requirements”.
2	A method according to claim 1 wherein at least one of said video clips of said first or second video programming is a transitional clip.
3	A method according to claim 2, wherein at least one of said video clips of each of said first and second video programming is a transitional clip.
4	A method according to claim 3, wherein said transitional clip is a hosting clip consistent with at least one template variable.

Claim No.	Claim Language
5	A method according to claim 3, wherein at least one transitional clip is a hosting clip consistent with at least one template variable.
6	A method according to claim 2, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.
7	A method according to claim 6, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.
8	A method according to claim 6, wherein one or more video clips comprise advertising.
9	A method according to claim 6, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.
10	A method according to claim 6, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.
11	A method according to claim 6, wherein said stored associations between said video clips comprise a plurality of media tags.
12	A method according to claim 6, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.
13	A method according to claim 1, wherein said first and second video programming are each presented as a stream of information to the user.

Claim No.	Claim Language
14	14. A method according to claim 1, wherein said first or second template variables corresponding to viewing scenarios comprise sequencing limitations for incorporating video clips at specific locations in the first or second video programming.
15	15. A method according to claim 14, wherein the sequencing limitations comprise at least one of length requirements, requirements for presenting a set of associated video clips in a predetermined order, and requirement for presenting a set of associated video clips within a predetermined time frame.
16	16. A method according to claim 1, wherein one or more video clips comprise advertising.
17	17. A method according to claim 1, wherein information about the user comprises at least one of demographic information, psychographic information, and a viewing history.
18	18. A method according to claim 1, wherein the step of selecting a second video clip based on said association with at least one first video clip is performed according to at least one of the first or second templates.
19	19. A method according to claim 1, wherein said stored associations between said video clips comprise a plurality of media tags.
20	20. A method according to claim 1, wherein said media tags comprise a plurality of information selected from at least some of context, subject matter, source, region, language, sponsorship, demographics, psychographics, and viewing history.

I. Introduction

The disclosure of the challenged '597 patent is limited and superficial. The '597 patent has three figures and under four columns of specification, of which one column is a description of the prior art. The '597 patent incorporates by reference the contents of the Marcus156 patent, Ex. 1006, which names the same inventor, Dwight Marcus, as the '597 patent. Marcus156 had already issued more than three years before the first provisional application that led to the '597 patent was filed. Despite incorporating it by reference, the '597 patent does not claim priority to Marcus156 or to any patent applications in the same family as Marcus156, rendering Marcus156 prior art to the '597 patent under 35 U.S.C. § 102(b).

The issued claims of the '597 patent, however, are based primarily on material contained in the prior art Marcus156 patent, rather than anything in the specification of the '597 patent itself. For example, claim 1 of the '597 patent requires assembling video streams comprising clips and transitions, using templates containing “(a) information about a user associated with the second computing device, (b) viewing scenarios of two or more selected video clips in a temporal relationship to each other, and (c) ‘style requirements’.” '597 patent at cl. 1. These requirements are discussed at length in Marcus156 using the exact same terminology claimed in the '597 patent. On the other hand, the specification of the '597 patent itself is mostly silent on these topics. Thus, the '597 patent is an improper attempt by the named inventor and

assignee to extend the life of Marcus156 by claiming it again—or at least a minor and obvious modification to it.

Accordingly, Petitioner respectfully requests that the Board institute this IPR and find all claims of the '597 patent invalid.

II. Mandatory Notices

A. Real Party-in-Interest

The real parties-in-interest for Petitioner are TikTok Inc., Bytedance Ltd., Bytedance Pte. Ltd., and TikTok Pte. Ltd.

B. Related Matters

The '597 patent has only been asserted in *NTECH Properties, Inc. v. ByteDance Ltd. et al.*, Case No. 2:24-cv-00130 (E.D. Tex., filed Feb. 23, 2024).

The '597 patent is in the same family as U.S. Patent No. 8,875,185. Petitioner is concurrently filing an IPR petition against that patent.

C. Counsel and Service Information

Electronic service may be made on the email addresses of counsel identified below and in the accompanying Power of Attorney.

Lead Counsel	Backup Counsel
James M. Glass (Reg. No. 46729) jimglass@quinnemanuel.com QUINN EMANUEL URQUHART & SULLIVAN, LLP 51 Madison Avenue, 22 nd Floor New York, NY 10010 Tel: (212) 849-7000	Ognjen Zivojnovic (Reg. No. 69,516) ogizivojnovic@quinnemanuel.com QUINN EMANUEL URQUHART & SULLIVAN, LLP 50 California Street, 22nd Floor San Francisco, California 9411 Tel: (415) 875-6600 Gavin Snyder (<i>pro hac vice</i> to be requested upon grant authorization) gavinsnyder@quinnemanuel.com QUINN EMANUEL URQUHART & SULLIVAN, LLP 1109 First Avenue, Suite 210 Seattle, Washington 98101 Tel: (206) 905-7000

D. Payment of Fees

The Office is authorized to charge the fee required for this Petition (and any additional fees) to Deposit Account No. 50-5708.

E. Requirements For Inter Partes Review

Petitioner certifies that the '597 patent is available for *inter partes* review, that Petitioner is not barred or estopped, and that 35 U.S.C. §§315(a)-(b) is inapplicable.

III. Grounds

Petitioner presents the following grounds of challenge.

Ground	Basis	Prior Art	Claims
1	§103	Marcus156 and Harbick	1-20
2	§103	Marcus156, Harbick, and Hosken	1-20
3	§103	Marcus156, Harbick, Hosken, and Byers	8, 16
4	§103	Ferman and Harbick	1-7, 9-15, 17-20
5	§103	Ferman, Harbick, and Byers	8, 16

IV. Overview of the '597 patent

The '597 patent claims priority to the '597 provisional, filed June 24, 2003. The specification of the '597 patent is exceptionally short. It includes three figures and less than four columns of text. The “Background of the Invention” section contrasts the alleged invention with three different prior art video delivery systems: television, video on demand, and the world wide web. '597 patent at 1:31-2:5. With television, “a user has very little control over what information is presented at a specific time.” *Id.* at 1:50-51. Video on demand used mundane, list-based user interfaces. *Id.* at 1:61-64. The world wide web offered more user control but was “inefficient and non-entertaining.” *Id.* at 2:4-5.

The '597 patent purports to address these deficiencies by automatically customizing video programming based on metadata, which the '597 patent calls “templates” and “headers.” *Id.* at 2:9-21; 3:1-11.

V. Level of Ordinary Skill in the Art

A person of ordinary skill in the art at the time of the '597 patent (“POSITA”) had at least a Bachelor of Science in electrical engineering, computer engineering,

or similar fields and at least two years of practical experience in the field of media processing and/or transmission. More education can supplement for less practical experience, and vice versa.

Petitioner's expert exceeded this level by the priority date.

VI. Claim Construction

The claim terms should be given their plain and ordinary meaning under Phillips, and “construed only to the extent necessary to resolve the controversy.” *Wellman, Inc. v. Eastman Chem. Co.*, 642 F.3d 1355, 1361 (Fed. Cir. 2011). Here, no express constructions are required to find the '597 patent claims unpatentable.

VII. Summary of the Prior Art

A. Marcus156 (Ex. 1009)

Marcus156 issued on February 29, 2000. It is 35 U.S.C § 102(b) prior art to the '597 patent. It names the same inventor, Dwight Marcus, and describes the same technology as the '597 patent. There is no exception to pre-AIA 35 U.S.C. § 102(b) based on commonality of inventorship between the prior art and the challenged patent. Just like the '597 patent, Marcus156 describes a video system where templates are used to select video clips, find associated transition clips, and concatenate them together into a stream of video programming. Marcus156 at Abstract.

During the '709 application, which ultimately issued as the parent '185 patent

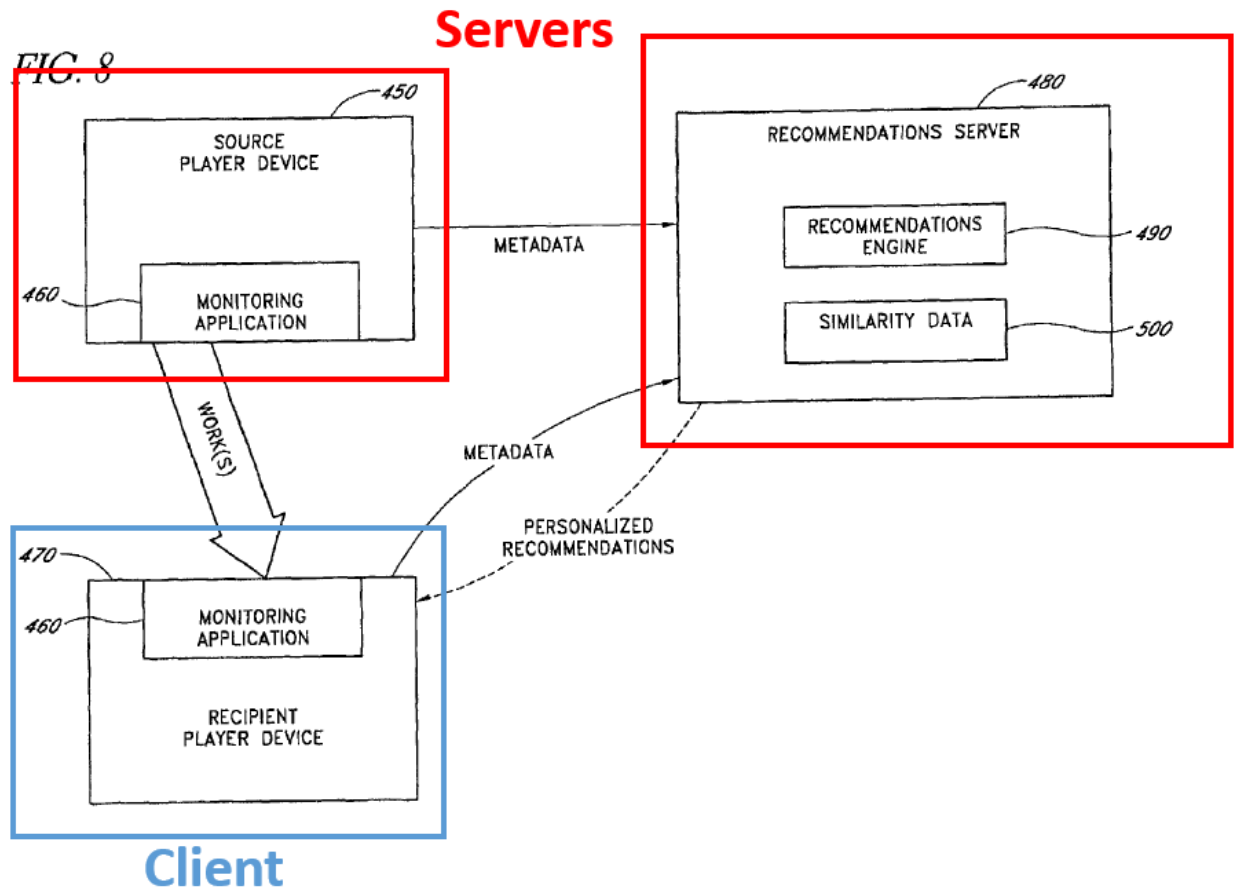
to the challenged '597 patent, the examiner cited Marcus156 as a secondary reference in obviousness rejections (Ex. 1005 at 219-23) but never as a base reference. The applicant acquiesced to these rejections, canceled all pending claims, and added the set of new claims that became issued claims 1-20 of the '185 patent. *Id.* at 232-42. The '185 patent issued without any further office action rejecting these new claims. *Id.* at 253-61, 294-303, 328. In allowing the new claims, the examiner appears to have overlooked several aspects of Marcus156. This petition highlights and explains those overlooked aspects.

Marcus156 was not cited by the examiner as the basis for any claim rejection during the prosecution of the '110 application, which became the '597 patent. However, the examiner did remark: “The prior art made of record and not relied upon is considered pertinent to the applicant’s disclosure: . . . Marcus (Patent No US 6,032,156) – utilizing templates to assemble video clips in a play list, Col. 9 lines 9-21, col. 13 lines 45-51 figures 5 and 6.” Ex. 1006 at 100-01.

Marcus156 was also cited by the examiner during the prosecution of other patents in the same family as the '597 patent, and each time applicant acquiesced to these rejections by rewriting the claims. In the instances where the Examiner considered Marcus156 as a base reference (rather than just a secondary reference), applicant was further forced to add additional, substantive limitations not found in the '597 patent claims.

B. Harbick (Ex. 1010)

Harbick is a U.S. Patent filed on April 19, 2002 and issued on September 4, 2012. It is 35 U.S.C. § 102(e) prior art. Harbick discloses a client-server architecture for a video recommendation service. Harbick at Abstract. Figure 8 depicts an exemplary division of client and server responsibilities in Harbick:



Id. at Fig. 8 (annotated); 18:37-41. Other configurations are possible.

C. Hosken (Ex. 1011)

Hosken is a PCT publication published on January 25, 2001. Hosken is 35 U.S.C. § 102(b) prior art. Hosken discloses content recommendation algorithms.

Hosken at 1:10-14. Hosken describes how items from a first search result set can be selected by the user and used to perform successive searches. *Id.* at 2:9-11; 9:4-8.

D. Byers (Ex. 1013)

Byers is a U.S. Patent that was filed on April 26, 2000 and issued February 19, 2008. Byers is prior art under 35 U.S.C. § 102(e). Byers describes dynamically inserting advertising into a video. Byers at 2:48-52.

E. Ferman (Ex. 1012)

Ferman is a patent publication published on May 16, 2002. Ferman is 35 U.S.C § 102(b) prior art. Ferman uses metadata templates to match user preferences with a database of video clips to create interactive video programming. Ferman at [0001]; [0089]; [0091]; [0093].

VIII. Ground 1: Marcus156 and Harbick Render Claims 1-20 Obvious

A. Proposed Modification

Marcus156's system would be modified, according to the teachings of Harbick, to use Marcus156's computer 20 as a video server as disclosed in Harbick. Harbick at 3:9-12 ("The system includes one or more types of player devices 30 that communicate over a computer network, such as the Internet, with a remote server system 40 ('server')."). As part of this modification, Marcus156's computer 20 would accept and process a request over a network for video programming from a client device (element 1[a]) or an identification of a first video clip of said first video

programming (element 1[g]) from a client device, which could be another instance of Marcus156's computer 20, or a different type of device altogether. The combined system would use existing functionality disclosed in Marcus156 to transmit assembled first and second video programming to the client device, which would present it on an associated display. The resulting system renders all challenged claims obvious. Almeroth ¶144.

B. Motivation to Combine

A POSITA would have been motivated to enhance Marcus156's interactive video delivery system according to the client-server networking teachings of Harbick to arrive at the challenged claims for at least the following reasons:

Scalability: By the time of the alleged invention, the client-server model was well-established for delivering web and network services. Moving video recommendation and video programming compilation to a server would allow the system to scale more easily to serve multiple users simultaneously. Server hardware could be upgraded centrally to handle increased load, rather than requiring upgrades to individual user computers. Almeroth ¶146.

Centralized content management: Storing video clips and metadata on a central server would simplify content updates and management. New clips could be added once to the server rather than needing distribution to many individual computers. The server-based system could have upgraded storage designed to store

more video content than a normal personal computer could. Centralized storage would also allow for more robust backup and disaster recovery solutions. Almeroth ¶147.

Reduced client requirements and increased flexibility: Offloading video processing to a server would lower the hardware requirements for client devices, expanding the user base to weaker computers. Streaming video over a network would allow users to access personalized content from multiple devices, not just a single computer with the full system installed. Streaming video would eliminate the need for local storage of videos on the client. Almeroth ¶148.

Content protection: Streaming video from a server, rather than providing full video files locally, would offer better control over content distribution and help prevent unauthorized copying. Server-side authentication and encryption could provide more robust protection for video content. And it allows regulatory compliance to be tracked better. Almeroth ¶149.

Improved video recommendations: Multiple users communicating with a central server would give the server more data and would allow for more sophisticated collaborative algorithms that could leverage data from all users to improve recommendations. This technique is specifically identified in Harbick as an advantage. Harbick at 1:62-67; 3:63-67; 13:21-24; Almeroth ¶150.

Design choice: Splitting up functionality between a client and server would

have been one of a limited number of design choices available to a software developer. A POSITA would have recognized that software components could optionally be situated on the same or different machines, and that the components could communicate with each other over a network or locally. Almeroth ¶151.

Same field of art: Harbick, Marcus156, and the '597 patent all relate to improved video recommendation systems that tailor video content to the characteristics and interests of the user. And all three contemplate using networking technologies as parts of the system. Almeroth ¶152.

C. Reasonable Expectation of Success

A POSITA would have had a reasonable expectation of success in improving Marcus156's system according to the client-server teachings of Harbick to render the challenged claims obvious, for at least the following reasons.

Marcus156 already contemplates separating logical components on different physical machines: Marcus156 already contemplates that its logical systems do not have to be co-located on the same physical computer. For example, Marcus156 already describes that library 70 may be situated on a different physical machine than system 10. Marcus156 at 2:56-60. Additionally, Marcus156 describes its computer 20 as being "suitably associated" with database 100, rather than necessarily co-located on the same physical machine. *Id.* at 2:66-67. At a minimum, this language conveys that the inventor realized that software components may

optionally be located on the same or different physical machines. Almeroth ¶154.

Marcus156's disclosed hardware is suitable for use as a server or client: The hardware configuration of Marcus156's computer 20 could be duplicated and one instance could be used as a server and another as a client. Marcus156's computer 20 is already described as running all of the system's components, so splitting the roles would increase the available hardware resources available to all components. Additionally, Marcus156's described hardware could be used as either a client or a server. *Id.* at 2:52-57. Marcus156's hardware is commodity ("conventional personal computer"), and those computers were readily available at retail at the time of the invention. Almeroth ¶155.

Emerging broadband adoption: By the time of invention, broadband internet connections were becoming common. Cable and DSL services capable of supporting video streaming were beginning to be rolled out. This made it increasingly feasible to stream video content from remote servers. Networks were fast enough to transmit the MPEG-encoded video described in Marcus156 in real time without appreciable delay, packet loss, or quality degradation. *See* Marcus156 at 9:46-10:9 (output video programming compressed in MPEG format). Almeroth ¶156.

Software modification would be predictable: As discussed above, Marcus156's disclosed hardware would only need to be replicated (at commodity

prices), rather than changed. Only software would change. Given the generally-acknowledged predictability of the software arts, a POSITA would have had a reasonable expectation of success in making the proposed enhancement. The software changes would involve adding more networking functionality to Marcus156's software components. That functionality would use well-understood network protocols, such as HTTP, and associated software libraries. Almeroth ¶157.

D. Claim 1

1. Element 1[pre]

Providing video programming to a user: Marcus156's system "deliver[s] information via customized video programming." Marcus156 at 10:23-27; Abstract; 2:10-11; 14:46-58; Almeroth ¶159.

2. Element 1[a]

Receiving request for video programming at video server: In the combined system of Marcus156 and Harbick, the video server receives a user request (also called a "query"). Marcus156 at 1:67-2:2 ("selecting some of those media elements in response to a request for media programming"); 2:25-28; 4:60-61; 9:9-11; 13:51-52 ("The user can select assets through a suitable query...."); 14:33-34; Almeroth ¶161.

By communication with a computing device associated with a user: In the combined system of Marcus156 and Harbick, the client computer is associated with

a user making the query. The server running database 100 and library 70 receives the query by communicating with the client. Marcus156 at 2:56-58. Harbick provides an example of how a client machine makes a request for video programming to a video server via a web server. Harbick at 7:29-30 (client makes HTTP GET request); 7:52-54 (server responds). This type of communication could have been adapted for use in Marcus156's system. Almeroth ¶162.

3. Element 1[b]

Searching based on the request: A program running on the server computer in the combined system of Marcus156 and Harbick performs the search based on the user request. Marcus156 at 2:6-7 (“selects and orders the media elements according to the data in the request”); 1:67-2:2; 2:25-28; 4:52-5:2; 8:13-32; 9:11-12; 13:51-52; 14:33-34; cl. 1; Harbick at 9:62-67 (search carried out by recommendation algorithm 42 on server); Almeroth ¶¶164-65.

Databases of video clips: Marcus156 discloses a database 100 of video clips to be searched. Marcus156 at 1:63-65 (“A database is provided which contains selected information about each of a large number of media elements.”). The database includes metadata about clips. *Id.* at 2:67-3:2; 2:35-37; 5:40-43; Fig. 2; cl. 1; Harbick at 4:65-5:1 (database of video clips); Almeroth ¶¶166-68.

Marcus156 also discloses a library 70 storing the underlying video data for those video clips. Marcus156 at 2:56-58; 2:61-64. To the extent the claimed

“database of video clips” must include raw video data, library 70, alone or when combined with database 100, constitutes the claimed “one or more databases of video clips.” To the extent PO argues that library 70 is not a “database,” it would have been obvious to integrate the raw video data in library 70 into database 100 (thereby eliminating need to separately access library 70 after selecting video clips based on database 100) or to implement library 70 as a database (thereby speeding up the look-up and retrieval of video clips from library 70). Almeroth ¶¶169-71.

Database in communication with video server: Marcus156’s database 100 and library 70 are in communication with its server. Marcus156 at 2:56-57 (library 70); 1:63-65 (database 100); 2:66-67 (same); Harbick at 4:65-5:1; Almeroth ¶172.

4. Element 1[c]

Receiving search results at video server: The video server receives a set of search results that correspond to the request. Marcus156 at 2:25-28 (“In a first step, from a database containing information concerning the audiovisual media elements, certain audiovisual media elements are selected.”); 2:6-7 (“The system selects and orders the media elements according to the data in the request”); 1:67-2:2 (“The method provides for selecting some of those media elements in response to a request for media programming”); 4:60-61; 6:65-7:2; Harbick at 12:8-11; Almeroth ¶174.

Search results are responsive video clips in databases: As described in the

citations immediately above, the search results are multiple clips in the database, all responsive to the request. Almeroth ¶175.

First video clips associated with second video clips in databases: Metadata, including tags and templates, associates each clip in the database with other clips based on numerous criteria. Any responsive clip in the search results (“first video clip”) is associated with other video clips (“one or more second video clips in said one or more databases”) because they share the same or similar tags, or are identified as similar according to templates. Marcus156 at 3:25-30; 3:13-14; Almeroth ¶176.

Hierarchical categorizations also cause the clips to be associated. Marcus156 at 2:67-3:2; 3:31-33; 3:60-65. Finally, associations between clips will cause them to be concatenated into the same output video stream. Marcus156 at 4:9-14; 8:3-5; Almeroth ¶¶177-78.

5. Element 1[d]

Identifying transitional clips: Once initial clips are selected, Marcus156’s system identifies an accompanying set of transitional clips. Marcus156 at 2:29-31 (“Transitions between the media elements are automatically selected.”); 14:13-17; 14:43-45. The transitions can be standalone clips (a “bridge” or “floater”) inserted between two other clips (for example, inserting a percussion clip between the two clips lacking percussion). *Id.* at 8:51-56; 4:28-29; 4:34-38. Or they can be created by modifying the start or end of other clips. *Id.* at 8:33-35 (“***transition selections*** at

the beginning and end of the audio and video portions of each clip”) (emphasis added);¹ Almeroth ¶¶180-82.

Marcus156’s transitions may be selected either before or after the initial clips have been selected:

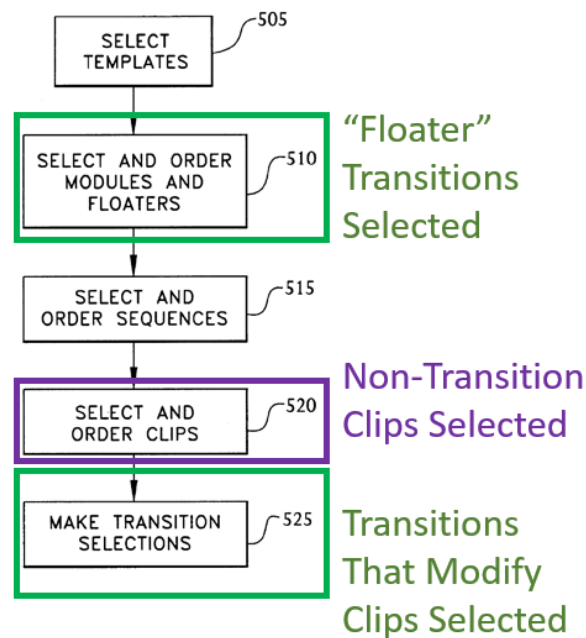


FIG. 5

Id. at Fig. 5 (annotated); 2:42-43; 7:3-7; Almeroth ¶183.

In databases: The bridge/floater clips and the source data for the modified

¹ The appendix to the '597 patent's provisional application confirms that modification of existing clips qualifies as the claimed "transitional clips." Ex. 1003 at 34 ("types of transitions (hard cut, fade-to-black, white-burst, and so-on)").

transition clips are listed in the database. Marcus156 at Abstract (“Transitions between audiovisual clips are determined by the system based on information stored in the database....”); 2:67-3: (“Database 100 contains unique identifying information for each clip and has associated therewith additional information often arranged in a hierarchical manner.”); 13:45-52; Almeroth ¶184.

To the extent that PO argues that the modified transition clips are not contained “in the database,” 1) the relevant claim language is disclosed by Marcus156’s bridge/floater clips alone and does not also require the modified transition clips; 2) a POSITA would have considered the modified transition clips to be “in the database” because their source data is, and the appendix to the ’597 provisional contemplates these types of transitions to be within the scope of the invention (*e.g.*, Ex. 1003 at 34 (“types of transitions (hard cut, fade-to black, white-burst, and so on)”); and 3) it would have been obvious for a POSITA to include more data about the modified transitions in Marcus156’s database 100, including because it would have allowed more granular control over the types of transitions, would have allowed for reuse of the transitions when the underlying clip is reused, and would have entailed a minor software modification. Almeroth ¶185.

For association with responsive first clips: Each transition is associated internally in the system with the other clips that it transitions between. Marcus156 at 1:39-50; 4:26-38; 4:48-51; 5:27-35; 7:47-49. Transitions are also associated with

other clips based on aesthetic requirements derived from the content and format of the clips that are being transitioned between. *Id.* at 8:48-51 (“determine whether a direct transition between any two clips meets aesthetic requirements”); 3:36-39; Almeroth ¶186.

More generally, the transitions are also associated with at least some of the selected first video clips because they meet the system’s transition criteria or share tags. Marcus156 at 4:26-38; 4:48-51; 5:27-33. Transitions that fulfill the criteria for a preferred or required transition of another video clip (and then are slotted next to the other clip in the final video programming) are “associated with” that clip. Almeroth ¶187.

6. Element 1[e]

Defining first video programming comprising responsive first video clips and first transitional clips: An assembly program creates a stream of video from the selected video clips and transitional clips, thus “defin[ing] first video programming.” Marcus156 at 4:13-15; 9:22-39; 14:7-11; Almeroth ¶189.

Concatenated: Marcus156 describes assembling clips and transitions into a sequenced final video as “concatenating.” Marcus156 at 14:13-17 (“create a concatenated stream of audiovisual programming that has transitions”); 3:44-46 (“placement of media elements in the concatenated stream”); 7:5-9; 9:23-32; Almeroth ¶190.

According to a first temporal sequence: Marcus156's system determines a temporal order for the clips and their accompanying transitions. Marcus156 at 1:67-2:3; 2:28-31 ("A temporal sequence for the selected elements is designated."); 4:22-34. Also, control tags for clips may dictate a temporal structure to the concatenated video programming. *Id.* at cl. 8; Almeroth ¶¶191-92.

7. Element 1[f]

Presenting first video programming on user computing device's display: In Marcus156's system, the first video programming is assembled by a "production player" or "assembly program" and then output as a video stream to be displayed. Marcus156 at 7:36-41; 7:59-60; 9:30-32. Users "receiv[e] the programming online through a modem or Internet link" and then "play[...the video sequence." Marcus156 at 11:4-10; 14:35-40. Almeroth ¶194.

In the combined system of Marcus156 and Harbick, the stream of first video programming would be transmitted to the same client machine ("user computing device") that originated the "first request for video programming" specified in 1[a], using Marcus156's disclosed techniques for transmitting MPEG-encoded video over a network to a recipient. The client machine would present the first video programming on its display. Almeroth ¶195.

As one example, the first video programming may be presented as part of the user interface for a "List Editor/Previewer" program running on the user computing

device. Marcus156 at 8:2; 6:3-7; 6:9-12; 7:11-15. Using that program, a video editor sees the assembled first video programming on a display. Marcus156 at 7:51-52. In the combined system of Marcus156 and Harbick, the interface for the “List Editor/Previewer” program could run on a user computing device, while the video editing functionality could run on a video server. Similarly, Marcus156’s “interactive system” could display the programming on a user computing device’s display. Marcus156 at 14:35-67; Almeroth ¶196.

8. Element 1[g]

Receiving identification of video clip at video server: Marcus156 discloses that a user can identify a first video clip in the first video programming by a mouse clip to trigger a second search. Marcus156 at 14:35-43; 14:46-55. The “List Editor/Previewer” has a UI to identify video clips. *Id.* at 7:11-15. In general, Marcus156’s system supports “viewer input” to search for video. *Id.* at 14:59-67; Almeroth ¶¶198-200.

In the combined system of Marcus156 and Harbick, the video server would receive this identification from the user computing device. For example, Harbick discloses reporting an identification of played items (which could have been in the first search result set/first video programming) to the server. Harbick at 6:20-22 (“metadata of...television shows”); Almeroth ¶201.

9. Element 1[h]

Searching databases of video clips according to second request for second video programming: When the user interacts with video programming that was already generated by the system, Marcus156's system performs additional interactive searches and generates a new stream of video programming. Marcus156 at 14:40-45 (new search performed based on displayed information from first search); 14:46-55 (user click on screen of first search results triggers second search); 7:11-15 (List Editor/Previewer 430 can change video programming in response to user interaction). The searches are performed on Marcus156's database 100 (potentially in combination with library 70), just as initial searches are. Almeroth ¶203.

While the example in column 14 mentions "a predetermined program," Marcus156's disclosure is not so limited. Marcus156's "interactive system" is agnostic as to whether video programming is "predetermined" or itself selected based on a first search. At a minimum, it would have been obvious to apply the same teachings to both "predetermined" and dynamically-generated programming. The same benefits of allowing the user to interact with a "predetermined" program to search for "particular information" related to that program equally apply to searching for additional information about a targeted program. There is also a reasonable likelihood of success, since the same selection mechanics can be applied

to either type of video. Almeroth ¶204.

Based on association between first and second clip: The search is based on the stored associations between the prior, selected clip and the new, second clip, including because both videos may have the same or similar tags, including at least at higher levels of the hierarchies described in Marcus156. Marcus156 at 9:16-21; 3:26-30; 14:40-55; Almeroth ¶205.

10. Element 1[i]

Receiving second search results at video server: The results of the second search performed in 1[h] would be received at the video server just as the results of the first search were received, as described in 1[c]. Almeroth ¶207.

Second search results correspond to second video clips in database responsive to second request: As described in 1[h], the search results are a plurality of second video clips in the database 100 that fulfill the request. Almeroth ¶208.

Including a second video clip associated with identified first video clip: All of the second video clips are associated with the identified first video clip because Marcus156's system identified them as similar to the first video clip as part of building the second set of search results. Marcus156 at 2:67-3:2; 3:13-14; 3:25-33; 3:60-65; Almeroth ¶209.

11. Element 1[j]

Identifying second transitional clips in databases for association with

responsive second video clips: As with the first search, Marcus156's system will identify a set of second transitional clips for association with the second video clips. Marcus156 at 14:52-58. Further details are in 1[d] above. Almeroth ¶211.

12. Element 1[k]

Defining second video programming with second video clips and second transitional clips concatenated according to second temporal sequence: The same systems, data, templates, and algorithms that Marcus156's system used to select a plurality of first video clips, determine a temporal sequence for them, and concatenate them are also used in the second search to define second video programming. The material in 1[b], 1[c], 1[d], and 1[e] applies here too. Almeroth ¶213.

13. Element 1[l]

Presenting second video programming on user computing device's display: After performing a second search and defining second video programming, the programming is presented on a display associated with the user computing device, just like the first video programming. Marcus156 at 14:52-55; 14:61-67; Almeroth ¶215. Additional detail is in 1[f].

14. Element 1[m]

Searching or defining step is performed according to templates: A "template layer" is consulted during both the searching and defining steps (which are thus

“performed according to” the templates):

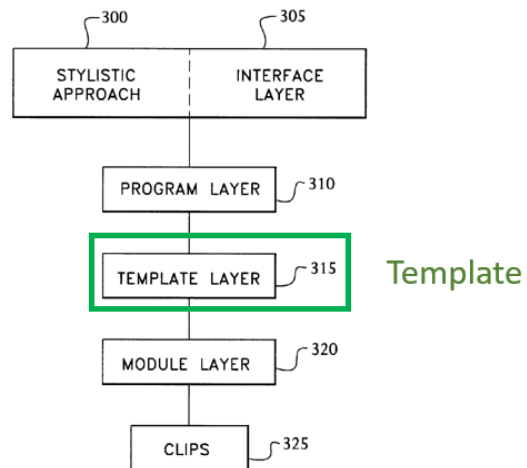


FIG. 3

Marcus156 at Fig. 3 (annotated); 8:13-15 (“one or more suitable templates”); 9:9-21; Abstract. Templates are selected before selecting, ordering, and concatenating video clips and transitional clips:

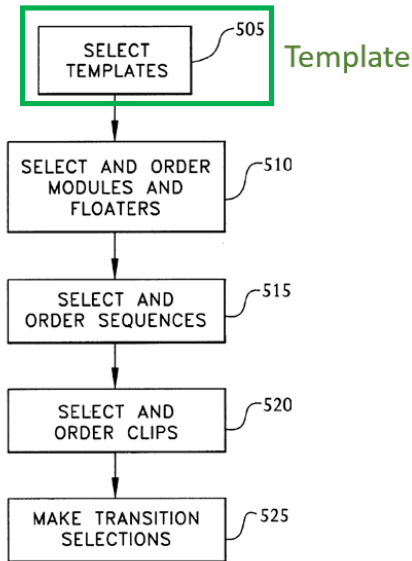


FIG. 5

Id. at Fig. 5 (annotated). The subsequent steps all comply with (“performed according to”) the requirements of applicable templates. *Id.* at 3:67-4:9; 8:19-25; 8:39-43; 9:2-6; 14:22-24; Almeroth ¶¶217-18.

Plurality of template variables: The templates comprise a plurality of template variables, namely the data contained in the templates. Marcus156 at 3:67-4:3 (“a family of defining values”); Almeroth ¶219.

Information about a user associated with the second computing device: Marcus156’s templates include demographic and psychographic parameters. Marcus156 at 3:66-4:5 (template layer can target viewers by “demographic and psychographic terms.”); 6:33-40; 13:47-51 (parameters for “demographic or psychographic characteristics”). This is “information about the user” at least

because it restricts the intended audience for the video programming based on user characteristics. *Id.* at 4:9-14; 14:30-34; Almeroth ¶¶220-21.

Additionally, Marcus156's system tracks individual users in its "database of viewer information" and "patient profile." Marcus156 at 10:12-20; 6:49-52. That information is used to select and sequence clips and transitions. *Id.* at 10:16-20; 4:56-59. While Marcus156 does not explicitly describe the "database of viewer information" or "patient profile" as part of its template layer, a POSITA would not restrict the scope of the "templates" claimed in the '597 patent to the specific data structures in the template layer in Marcus156. The plain and ordinary meaning of "template" does not require any particular data structure. The intrinsic evidence does not define "template." The appendix from the '597 provisional says that a template can be code, in addition to data, and can employ complex programming logic. Ex. 1005 at 34 n.6. The appendix contemplates multiple templates addressed to different types of metadata. *Id.* at 35. PO has accused TikTok's video recommendation algorithms as a whole of being a claimed "template." Ex. 1018 at 68. Applying the same interpretation of "template" that PO uses for infringement purposes, a POSITA would have considered the "database of viewer information" or "patient profile" in Marcus156 to also be a "template." Almeroth ¶¶222-26.

To the extent a specific data format is required for "templates," it would have been obvious to include Marcus156's "database of viewer information" or "patient

profile” in its template layer. A POSITA would be motivated to convert the data into template format to better use the information to select video clips to create video programming, and to simplify the system. A POSITA would have had a reasonable expectation of success in doing so because it would entail a routine conversion of information in one data format to another. Marcus156’s system already uses templates and includes software to import new clip and user data (Marcus156 at 6:15-7:9), which would reduce the time and effort to make the modification. Almeroth ¶¶227-28.

Viewing scenarios of two or more selected video clips in a temporal relationship to each other: Marcus156 includes “temporal organizational templates.” Marcus156 at 4:5-7. These include viewing scenarios of two or more selected clips in a temporal relationship to each other. *Id.* at 4:7-9 (“preferred temporal modes of presentation”); 4:14-21; 8:39-43 (clips sequenced by templates); 14:22-25 (“template that imposes sequential requirements”). They can describe an expected audience and subject matter for a sequence of temporally-ordered clips – a “scenario” for how the video programming is ultimately viewed. Marcus156 at 4:9-14. The ’597 patent specification does not contain the term “viewing scenario,” and it only shows up once in the appendix to the ’597 provisional application, where its meaning is not restricted. Ex. 1003 at 25. The plain and ordinary meaning of “viewing scenario” encompasses this functionality in Marcus156. Almeroth ¶¶229-

31.

Style requirements: Marcus156’s templates constrain the video style. Marcus156 at 9:2-6 (“a specific style of editing”); 14:22-29; 3:41-46 (“aesthetic or programmatic choices”); Almeroth ¶232.

E. Claim 2

As described in 1[d], Marcus156’s system includes transitional clips in the first video programming. As described in 1[j], the same types of transitional clips are also included in the second video programming. Additional details are in 1[b], 1[d], and 1[e]. Almeroth ¶234.

F. Claim 3

Transitional clips can be present in the initial assembled video stream (“first video programming”) and also in the interactive video stream that is assembled in response to further user input (“second video programming”). Therefore, claim 3 is disclosed by Marcus156 for the same reasons claim 2 is. Almeroth ¶236.

G. Claim 4

Marcus156 discloses transitional clips that are hosting clips. Video clips in database 100 can be tagged with their actors. Marcus156 at 3:21-25; 14:30-34. A media tag “TALKING HEAD” can be used to select transitional clips for inclusion in the video. *Id.* at 4:39-48. A talking head clip has a person—a host—talking. The appendix to the ’597 provisional describes such talking head clips as hosting clips.

Ex. 1005 at 102. The term “hosting clip” is also described broadly in the appendix. *Id.* at 34, 71, 74 (“Any form of entertainment inserts/content or hosting segments can be the subject of this model.”). Clips with actors or talking head clips can be “bridge” or “floater” transitional clips that are used to link other clips. Marcus156 at 4:22-38; *see also* claim 2 above. When such a clip is selected as a linking clip, a POSITA would have considered it to be a “hosting clip” that is also a “transitional clip.” Almeroth ¶¶238-39.

Marcus156’s hosting clips are “consistent with at least one template variable.” Like other transitional clips, the hosting clips must comport with the contents of the selected templates in order to be included in the final video programming. If the content of a hosting clip is inconsistent with the templates then it will not be included. Marcus156 at 8:19-25; Almeroth ¶¶240-41; 1[m] above.

To the extent that Marcus156 does not disclose this claim, it renders it obvious.

A POSITA would have been motivated to use hosting clips as transitional clips. Including hosting clips would have allowed the concatenated video stream to take the form of traditional television shows that include a host, such as nature documentaries, late-night talk shows, or sequences of music videos hosted by a video jockey. Anybody familiar with radio DJs, music video ranking shows, America’s Funniest Home Videos, and similar programming compiling multiple clips would

be familiar with the concept of using hosting clips as transitions. Using such a format in the context of Marcus156 would thus breed familiarity for users (the clips would be presented in a format the users are already familiar with) and also expand the potential repertoire of the video programming (allowing for introductions to video programming users may otherwise be unfamiliar with). Including hosting clips would allow for additional types of transitions used for comedic or dramatic effect. Almeroth ¶242.

A POSITA would have had a reasonable expectation of success in making any needed modification to Marcus156. Given Marcus156's existing disclosure of tracking talking head clips, a POSITA would have been able to exploit the existing data tracking those clips. The existing templates could have been easily modified to use the hosting clips as transitional clips, given the extensive sequencing and clip association capabilities detailed in Marcus156. Almeroth ¶243.

H. Claim 5

Marcus156 discloses or renders obvious this claim for the same reasons given above in claim 4. Almeroth ¶245.

I. Claim 6

The temporal organization templates restrict where clips are placed in the programming, as in 1[e]. The template variables define a “preferred sequence” of clips ordered by content type. Marcus156 at 4:14-21; 14:22-25. Because these

“sequential requirements” place clips in temporal relationships to each other, they are “sequencing limitations for incorporating video clips at specific locations” in the programming. Almeroth ¶247.

J. Claim 7

Marcus156’s templated sequencing limitations meet each of the required three categories. The limitations may be length requirements or requirements for presentation within a particular time frame. Marcus156 at 4:65-5:2; 5:27-33. Order of clips can also be limited. *Id.* at 4:14-21; Almeroth ¶249.

K. Claim 8

Video clips in the database (“assets”) may be tagged as advertising or promotion. Marcus156 at 3:52-55. Clips may be classified as “market-driven,” *i.e.* advertising. *Id.* at 3:55-57. Marcus156 says that customized advertising programming is valuable, which is an indication that its system included advertising video clips. *Id.* at 1:19-26. One of the objects of Marcus156’s invention is to facilitate improved customization of video programming, such as the customized advertising example, by automating the process and eliminating the need for human video editors. *Id.* at 1:37-56. Thus, one or more clips in Marcus156’s system contain advertising. Almeroth ¶251.

L. Claim 9

As reviewed in 1[m], Marcus156’s templates contain demographic and

psychographic information. And the viewing history of users is tracked in the “database of viewer information.” Marcus156 at 10:12-16 (“the history of items viewed by the viewer”); Almeroth ¶253.

M. Claim 10

Templates are used to select clips based on associations with other clips, including by using a tag of one clip as a requirement for another clip. Marcus156 at 3:67-4:5; 4:56-64; 4:66-5:2; Almeroth ¶¶255-56.

N. Claim 11

Media tags (“header information”) are attached to clips. Marcus156 at 3:4-8; 5:3-4; 5:27-33; Almeroth ¶258.

O. Claim 12

Tags include context, subject matter, and demographics. Marcus156 at 3:21-26; 4:41-48. Demographics include source or region. *Id.* at 14:30-33 (“geographic origin”). Tags include psychographic information. *Id.* at 13:47-51 (“demographic or psychographic characteristics”); 4:61-63 (“psychographic requirements”); Almeroth ¶¶260-62.

P. Claim 13

The finalized first and second video programming is assembled and presented to the user as a unified video stream. The user sees a sequence of videos where the clips and transitions proceed sequentially, one after the other. Marcus156 at 3:44-

46 (“media elements in the concatenated stream”); 5:35-37 (“data stream before or after the clips are concatenated”); 7:7-10 (“The transitions are then executed and inserted into the stream....”); 14:13-17 (“concatenated stream of audiovisual programming that has transitions”); 9:30-32; cl. 1; Almeroth ¶264.

During prosecution of the related ’185 patent, the examiner stated that “Marcus [is] silent to explicitly disclose that said first and second video programming are each presented as a stream of information to the user.” Ex. 1005 at 222. This was clear error by the examiner. Marcus¹⁵⁶ expressly refers to its video programming as a “stream.” Almeroth ¶265.

Q. Claims 14-20

Marcus¹⁵⁶ and Harbick render claims 14-20 obvious for the same reasons as discussed in corresponding claims. Almeroth ¶¶266-272.

Claim	Corresponding Claim
14	6
15	7
16	8
17	9
18	10
19	11
20	12

IX. Ground 2: Marcus156, Harbick, and Hosken Render Claims 1-20 Obvious

A. Proposed Modification

To the extent that the combined system of Marcus156 and Harbick somehow does not fully render obvious a second video search in 1[h] or 1[i], the combined system of Marcus156 and Harbick would be further enhanced to perform additional searches for video clips (using the existing capabilities of Marcus156 and Harbick, plus the additional teachings of Hosken) based on an identification of a first video clip of said first video programming. The combined system's behavior with respect to elements 1[h] and 1[i] would be enhanced to better use the selection of one item in an initial search result set as the basis for a second, refined search. Almeroth ¶274.

B. Motivation to Combine

A POSITA would have been motivated to modify the combined system of Marcus156 and Harbick according to the teachings of Hosken for at least the following reasons.

Improved video clip recommendations: A POSITA would have been motivated to add Hosken's techniques to the combined system of Marcus156 and Harbick because progressive refinement of a result set through user selections of items from the previous result set would increase the accuracy and pertinence of the video programming. Hosken at 4:17-20 (advantage of personalized recommendations). The use of stored associations between video clips as disclosed

in both Marcus156 and Hosken would additionally improve the relevance of the clips selected for the second video programming. Almeroth ¶276.

Similar field of art: The '597 patent, Marcus156, and Hosken all pertain to automated video recommendation systems. Hosken at 4:9-11; 28:25-27; Almeroth ¶277.

C. Reasonable Expectation of Success

A POSITA would have had a reasonable expectation of success in modifying the combined system of Marcus156 and Harbick, further according to the teachings of Hosken, in the manner discussed above, for at least the following reasons:

Software modification would be predictable: Petitioner's proposed modification can be implemented entirely in software and does not require any hardware modifications. Given the generally-acknowledged predictability of the software arts, a POSITA would have had a reasonable expectation of success in making the proposed modification. Almeroth ¶279.

Marcus156 already teaches some degree of successive searching: Marcus156 already discloses and renders obvious performing successive searches. To the extent it does not, its close similarity with the claim language would have motivated a POSITA to add progressive search refinement to Marcus156 because all of the building blocks are already present. Marcus156 at 14:40-45; Almeroth ¶280.

D. All Claims and Claim Elements Other Than 1[h] and 1(i)

All claims 1-20 and claim elements other than element 1[h] and 1(i) are disclosed or rendered obvious by Marcus156 and Harbick for the reasons given above in ground 1.

E. Element 1[h]

Searching said one or more databases of video clips: Hosken’s system will search a database associated with a server computer for media content items, including videos, in response to a request. Hosken at 3:23-27 (media content database); 7:28-8:3 (database has “logical connections between different media content items”); 28:24-27 (applicable to videos). A database may also store underlying video data (“content samples”). *Id.* at 8:28-9:4; Almeroth ¶282.

Searching . . . according to a second request for second video programming that is based at least in part on a stored association between the identified first video clip and at least one second video clip: An identification of a first video clip in the first video programming (which is a result of an initial search) is used in combination with “associations” between content items, to produce another search results set. Hosken at 9:4-8 (“Based on the review and consideration of the recommendation set, including as applicable any presented content samples, the user may request a further search and production of a new recommendation set, typically identifying a prior recommendation media content item as part of the request”);

10:13-16 (“Input requests from a user 38, such as requests to find media content in some way similar to an identified media content item, are submitted for processing through system processes 42 to produce a responsive recommendation set back to the [user] 38.”); 11:9-21; Almeroth ¶283.

The associations between content items are stored in Hosken’s system, just as they are in Marcus156’s system. Hosken at 7:20-22 (profiling data used to build associations); 7:28-8:3 (database has “logical connections between different media content items”); 9:21-23 (“The implicit information gathered from user actions is preferably processed 22 and stored as an addition and refinement of the profile data 24 previously stored.”); Almeroth ¶284.

F. Element 1[i]

Receiving one or more second search results at said video server computing device: Hosken searches content items at a video server computing device, and similarly receives search results there as well. Hosken at 3:21-23 (“providing media content recommendations through a computer server system connected to a network communications system”); Almeroth ¶285.

Wherein said search results collectively correspond to a plurality of second video clips in said one or more databases that are responsive to the second request: The set of responsive videos come from a database and are responsive to the second request. Hosken at 8:22-25; Almeroth ¶286.

Including at least one second video clip that is associated with the identified first video clip: Each of the content items (which may include videos) is associated with the identified first video clip. Hosken at 8:1-3 (“This database 26 preferably specifies logical connections between different media content items based on the sharing or similarity of characterizing attributes.”); 8:22-25; Almeroth ¶287.

X. Ground 3: Marcus156, Harbick, Hosken, and Byers Render Claims 8 and 16 Obvious

A. Proposed Modification

The combined system of Marcus156, Harbick, and Hosken described in ground 2 would be further modified according to the teachings of Byers to include advertising clips in the video programming, including by dynamically editing other clips. Part of this modification would entail adding advertising material to the clip database and tracking additional advertising-related parameters in Marcus156’s templates. Almeroth ¶289.

B. Motivation to Combine

Commercial benefit to advertising: Because insertion of advertising into video programming would enable a new revenue stream for the video creator, a POSITA would have had an economic motivation to do it. Byers at 2:51-52; Marcus156 at 3:52-57; Almeroth ¶290.

Commercial benefit to user-targeted advertising: The user-targeted recommendation techniques disclosed in Marcus156, Harbick, Hosken, and Byers

would all allow improved advertising that is customized to the demographics, interests, or other facets of the user, thus increasing the pertinence and, ultimately, the click-through or conversion rate of the targeted advertising. Byers at 2:57-67 (“One problem with product placement is that the placement is not tailored for an individual user.”); Almeroth ¶291.

Same field of art: Like Marcus156, Harbick, Hosken, and the ’597 patent, Byers is directed to creation and presentation of dynamic video programming. Byers at 1:40-43 (“Therefore, a need exists for a method and apparatus for dynamically altering digital video images to allow customization of a digital video stream based upon characteristics associated with a user.”); Almeroth ¶292.

C. Reasonable Expectation of Success

Software modification would be predictable: As with the other reference combinations, this change would be accomplished in software and would entail predictable changes to Marcus156’s system. Almeroth ¶293.

Byers provides explicit instructions on how to modify clips to insert advertising: Byers discloses underlying mechanisms to insert the advertising content. Byers at 3:67-4:5; 5:60-66. The hardware used in this process is similar to Marcus156’s disclosed hardware. Byers at 5:45-67; Marcus156 at 2:51-56. Marcus156’s disclosed hardware could perform Byers’ video editing algorithms. Almeroth ¶294.

Byers’ instructions on adding advertising parallel Marcus156’s instructions

on how to modify clips to insert transitions: Marcus156 provides instructions to the user on extracting frames from a video clip and modifying them dynamically according to Marcus156’s templates to add transitions. Marcus156 at 7:47-63. The same technique could be used to add advertising information. Almeroth ¶295.

D. Byers Discloses and Renders Obvious “One or More Video Clips Comprise Advertising”

Byers discloses an “advertiser database” that contains advertising-related information. Byers at 4:19-25. The information in the database is used to dynamically insert ads into a video stream. *Id.* at 2:48-51; 3:1-4; Almeroth ¶¶296-98.

XI. Ground 4: Ferman and Harbick Render Claims 1-7, 9-15, and 17-20 Obvious

This ground relies on Ferman as a base reference, further modified according to the teachings of Harbick to add well-understood and conventional client-server networking functionality. The combined system renders all challenged claims except 8 and 16 obvious.

A. Proposed Modification

Ferman’s system would be modified, according to the teachings of Harbick, to use Ferman’s audiovisual system 16 as a video server as disclosed in Harbick. Harbick at 3:9-12 (“The system includes one or more types of player devices 30 that

communicate over a computer network, such as the Internet, with a remote server system 40 ('server')."). Additionally, Ferman's system 16 would be further modified according to the teachings of Harbick to accept and process a request over a network for video programming from a client device (element 1[a]) or an identification of a first video clip of said first video programming (element 1[g]), from a client device, which could be another instance of Ferman's system 16, or a different type of device altogether. The combined system would use existing functionality disclosed in Ferman to transmit assembled first and second video programming to the client device, which would present it on an associated display. The resulting system would practice and render obvious all challenged claims except 8 and 16. Almeroth ¶300.

B. Motivation to Combine

A POSITA would have been motivated to modify the interactive video generation system described in Ferman according to the client-server teachings of Harbick to achieved the challenged claims for at least the following reasons.

Scalability; centralized content management; reduced client requirements; content protection; improved video recommendations; design choice: Described in detail in Section VII.B above. Each motivation is equally or more applicable to the combination of Ferman and Harbick as it is to the combination of Marcus¹⁵⁶ and Harbick. Almeroth ¶¶302-07.

Same field of art: Ferman, Harbick, and the '597 patent all relate to improved video recommendation systems that tailor video content to the characteristics and interests of the user. And all three contemplate using networking technologies as parts of the system. Almeroth ¶308.

C. Reasonable Expectation of Success

A POSITA would have had a reasonable expectation of success in modifying Ferman's system according to the teachings of Harbick for several reasons.

Ferman already contemplates separating logical components on different physical machines: Ferman already contemplates that its logical systems do not have to be co-located on the same physical computer. Ferman at [0272]; Almeroth ¶310.

Emerging broadband adoption; software modification would be predictable: Described in Section VII.C above and equally or more applicable to the combination of Ferman and Harbick. Almeroth ¶¶311-12.

D. Claim 1

1. Element 1[pre]

Providing video programming to user: Ferman's system provides customized interactive video programming to a user. Ferman at [0092]; [0106]; Almeroth ¶314.

2. Element 1[a]

Receiving request for video programming at video server: In the combined

system of Ferman and Harbick, the video server receives a user request, which may be a specific search, a user preference description, or a set of filtering preferences. Ferman at [0114]; [0264]; [0266]; [0315]; Almeroth ¶316.

By communication with user computing device: In the combined system of Ferman and Harbick, the client computer is associated with a user making the request. The video server computer (Ferman's system 16) communicates with that client computer. Harbick provides an example of how a client machine makes a request for video programming to a video server via a web server. Harbick at 7:29-30 (client makes HTTP GET request); 7:52-54 (server processes HTTP GET request). This communication paradigm could be adapted to Ferman's system. Almeroth ¶317.

3. Element 1[b]

Searching based on request: Ferman's system searches the program description scheme based on a user request. Ferman at [0114] ("enable the user to browse, search, and filter audiovisual programs."); [0264] ("[A] a filtering agent 520 selects the appropriate content for the particular user(s)."). The user request may be triggered by an interaction or may be a user preference description or a set of filtering preferences. Ferman at [0315]; Ferman at [0266] ("[P]referred programs can be determined as a result of filtering program descriptions according to user's filtering preferences."); Almeroth ¶¶319-20.

Databases of video clips: Ferman's system includes a program description scheme, which is a database of video clips. Ferman at [0114]; [0117] ("The search engine uses the information contained in the program description scheme to retrieve programs on the basis of their content."); [0264] ("program descriptions...describing...video content available from which a filtering agent 520 selects the appropriate content for the particular user(s)."); [0266]; [0315]; Almeroth ¶321.

Database in communication with video server: The program description scheme (database) resides inside Ferman's system 16 (video server computing device), so is therefore "in communication with" the video server computing device. Ferman at Figs. 1, 2, 13; [0094]; Almeroth ¶322.

4. Element 1[c]

Receiving search results at video server: Ferman's system 16 selects an initial set of video clips ("preferred programs") that fulfill the user request, and therefore receives those clips as a set of first search results. Ferman at [0117] ("The search engine uses the information contained in the program description scheme to retrieve programs on the basis of their content."); [0264] ("The content is selected based upon the usage preference description for a particular user identification(s) to determine a list of preferred audio and/or video programs."); [0315]; Fig. 28 ("FILTERING AND SEARCH" component leads to "PREFERRED

PROGRAMS”); Almeroth ¶323.

From this initial set, based on the user preferences, Ferman’s system will select (and therefore also “receive”) a subset of video clips corresponding to specific events or highlights. *Id.* at [0103] (generation of 5-minute basketball game highlight); [094] (“a key frame...providing multiple levels of summary ranging from coarse to fine.”); [0104] (“highlight generation functionality for a particular program”); [0107] (describing a user viewing session where some programs (“the Microsoft trial”) are presented un-edited and other programs (“sporting events”) are automatically summarized and culled into sub-segments). Ferman refers to this functionality of breaking down a long-form program into subsequences as “filtering.” *Id.* at [0119]; Almeroth ¶324.

Search results are responsive video clips in databases: The result of these operations detailed above is a set of recommended video clips that are presented to the user. Ferman at [105] (“The SFB module 52 may perform filtering, searching, and browsing of the programs 38, on the basis of the information contained in the description schemes.”). This is the claimed “plurality of first video clips” in Ferman. Each of these clips is “in said one or more database” because their contents are described in the program description scheme. And each is “responsive to the first request” because they are clips that the system identified as the correct search result to respond to the first request. Almeroth ¶325-326.

First video clips associated with second video clips in databases: Each first video clip in is associated with a set of second video clips in Ferman's databases, including the full program, other clips that are part of the program but not included in a filtered set, or other, different filtered subsequences of clips from the same program. Ferman at [0103]; [0104] ("generating the preferred views and storing associated data in the program description scheme."); [0107]. The associations are contained within at least the program view descriptions, program description scheme, and segment description scheme. *Id.* at [0104] (cited above, program description scheme); [0094] (program view descriptions); [0252] (segment description scheme); Almeroth ¶¶327-28.

5. Element 1[d]

Identifying transitional clips in databases: Ferman's system filters and segments a video program "into story units, scenes, and shots." Ferman at [0252]. The set of segments includes multiple major events in the video clip, including events which transition between other clips, and all clips are listed in databases such as the syntactic structure description scheme. *Id.* at [0094]; [0252] (syntactic structure description scheme segments clips via a "segment tree [which] may be used to define[] a temporal structure of a video program"); [255] (semantic description scheme segments clips into events, reflected in an event description scheme). Almeroth ¶¶331.

In particular, the set of sub-segments for given preferred program are related to each other as a sequence of events that pertain to the same overall program. *Id.* at [0094] (“grouping of different events”). Based on the user’s viewing preferences, a set of initial preferred programs can be supplemented by sub-segments from those programs that are automatically selected. A POSITA would consider these video segments, which are thematically and structurally related to other surrounding video clips, to be “transitional clips.” Almeroth ¶331.

To the extent PO contends that these clips do not qualify as “transitional clips,” it would have been obvious for a POSITA to modify Ferman’s system to more explicitly select clips for inclusion in the final video programming based on their ability to transition between other events. A POSITA would have been motivated to do so by at least Ferman’s expressed desire to “include content suited to [the user’s] particular viewing preferences.” Ferman at [0089]. Moreover, doing so would provide more compelling and cohesive video programming. A POSITA would have had a reasonable expectation of success in making any necessary changes to Ferman’s system to more explicitly include transitions, including because Ferman already categorizes sub-sections of a video program by their contextual relationship to other sections. *Id.* at [0094]; Almeroth ¶332.

For association with at least some of said responsive first video clips: The transitional clips are associated with the first video clips via at least the segment

description scheme. Ferman at [0252] (“The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video.”). Additionally, they are associated at least via the program description scheme and program view descriptions, for example by sharing similar media tags. *Id.* at [094]; [0104]; Almeroth ¶333.

6. Element 1[e]

Defining first video programming comprising responsive first video clips and first transitional clips: As above, Ferman’s system receives search results, identifies transitional clips, and concatenates clips to form a final video stream (“first video programming”). Ferman at [0102]; [0105] (filtered video programming output by SFB module); [0107] (summarized highlight video programming output); [0111]; [0114]. To the extent PO contends that Ferman’s system itself fails to define video programming consisting of video data, it would have been obvious for a POSITA to modify Ferman’s system to do so. A POSITA would have been motivated to do so to reduce interdependency among different software components and minimize the chance for service interruptions due to bandwidth congestion between a viewing device and a video source. A POSITA would have a reasonable expectation of success in doing so because techniques for combining and manipulating video data (e.g., via transcoding) were well-known at the time of the invention. Almeroth ¶335-36.

Concatenated: The responsive video clip and transitions are concatenated and presented as a stream of video programming to the user. Ferman at [0607] (“PlayStream Play content from input stream”); [0107]; *id.* (“summarization functionality of the system...prepare[s] a summary”); [0111]; [0114] (“The user at any time can start viewing the program either in parts, or in its entirety.”). The system can generate a highlight video of a given length that consists of multiple subsidiary highlight or event clips concatenated and presented together. *Id.* at [0267] (15-minute highlight); Fig. 10; [0107] (5-minute highlight). The act of generating a single video highlight from multiple previously separate clips qualifies as “concatenating.” Almeroth ¶337.

To the extent that PO may argue that concatenation of the video clips is not explicitly described in Ferman, it would have been obvious for a POSITA to modify Ferman’s system to record and allow viewing of archival copies of the sequence of programming viewed by a user, which would be a concatenated video stream. A POSITA would have been motivated to make the modification because it would allow its system to achieve better VCR-like functionality. Moreover, concatenating clips would lead to a more coherent viewing experience corresponding, for example, to a single highlight video rather than disparate clips. A POSITA would have had a reasonable expectation of success in making the modification, because Ferman’s system already has video recording capability. Ferman at [0107]; Almeroth ¶338.

According to a first temporal sequence: The selection of video clips and transitional clips in turn determines a temporal sequence for the concatenated video programming. Ferman at [0252] (“Such a segment tree may be used to define[] a temporal structure of a video program.”). Additionally, the video programming is sequenced based on information in the program description scheme and/or event description scheme. *Id.* at [0095] (sequencing program views using the program description scheme); [0255] (“The event description scheme 480 may be used to form relationships between different events of the video normally consisting of multiple frames of the video.”); Almeroth ¶339.

7. Element 1[f]

Presenting first video programming on user computing device’s display: In Ferman’s system, the first video programming is presented on a display associated with the user computing device. Ferman at [0607] (“PlayStream Play content from input stream”); [0107]; *id.* (“summarization functionality of the system...prepare[s] a summary”); [0111]; [0114] (“The user at any time can start viewing the program either in parts, or in its entirety.”); Almeroth ¶341.

8. Element 1[g]

Receiving identification of video clip at video server: In Ferman’s user interface, a user can interact with a user interface on a user computing device to identify a video clip, including an identification of a clip that triggers the generation

and subsequent display of a highlight reel or an event reel. Ferman at [0304] (“The frames provide a visual summary. Depending on the browser, they may also provide clickable entry points to the video. The user may click and start playing back the video starting from that frame.”); [0107] (“The user starts interacting with the system with a pointer or voice commands to indicate a desire to view recorded sporting programs....”); [0152]; [0154]; [0112]; [0114]; [0250]; [0105] (“The selections of the desired program(s) to be retrieved, stored, and/or viewed may be programmed, at least in part, through a graphical user interface.”); [0309]; Fig. 10; Almeroth ¶343.

In the combined system of Ferman and Harbick, the video server would receive this identification from the user computing device. For example, Harbick discloses reporting an identification of played items (which could have been in the first search result set/first video programming) to the server. Harbick at 6:20-22 (“[T]he player device 30 may be configured to recognize and to report metadata of individual songs or television shows played....”); Almeroth ¶344.

9. Element 1[h]

Searching databases of video clips according to second request for second video programming: As described above, in Ferman’s user interface, a user can interact with a user interface on a user computing device to identify a video clip, including an identification of a clip that triggers the generation and subsequent display of a highlight reel or an event reel. This causes Ferman’s system to perform

another search similar to that described above in 1[b], but with additional information, namely the identified first video clip. Ferman at [0114]; [0264]; [0266]; [0315]; Almeroth ¶345.

The user identifications are used to populate the user description scheme and therefore to search for and select additional video clips. Ferman at [0097] (user description scheme “includes the user’s personal preferences, and information regarding the user’s viewing history such as for example browsing history, filtering history, searching history, and device setting history”); Almeroth ¶346.

Additionally, the user’s identification of a clip can also directly trigger the generation of new video, such as another highlight reel or event reel. *Id.* at [0304] (“The frames provide a visual summary. Depending on the browser, they may also provide clickable entry points to the video. The user may click and start playing back the video starting from that frame.”); [0107] (“The user starts interacting with the system with a pointer or voice commands to indicate a desire to view recorded sporting programs.”); [0112] (“The user description scheme is generated by direct user input, and by using a software that watches the user to determine his/her usage pattern and usage history.”); Almeroth ¶347

Based on association between first and second clip: The search is based on the stored associations between the prior, selected clip and the new, second clip, including because both videos may have the same or similar tags or entries in the

program description scheme. Almeroth ¶348.

10. Element 1[i]

Receiving second search results at video server: The results of the second search performed in 1[h] would be received at the video server just as the results of the first search were received, as described in 1[c]. Almeroth ¶¶349.

Second search results correspond to second video clips in database responsive to second request: As described in 1[h], the search results are a plurality of second video clips in the program description scheme that fulfill the request. Almeroth ¶¶350.

Including a second video clip associated with identified first video clip: All of the second video clips are associated with the identified first video clip because Ferman's system identified them as similar to the first video clip as part of building the second set of search results. Almeroth ¶351.

11. Element 1[j]

Identifying second transitional clips in databases for association with responsive second video clips: As with the first search, Ferman's system will identify a set of second transitional clips for association with the second video clips. Further details are in 1[d] above. Almeroth ¶353.

12. Element 1[k]

Defining second video programming with second video clips and second

transitional clips concatenated according to second temporal sequence: The same systems, data, templates, and algorithms that Ferman's system used to select a plurality of first video clips, determine a temporal sequence for them, and concatenate them are also used in the second search to define second video programming. In response to a user's identification of a first video clip, Ferman's system can generate a highlight reel or event reel with a plurality of video clips in a temporal order, using Ferman's templates. Ferman at [0107] (generation of highlights). The material in 1[b], 1[c], 1[d], and 1[e] applies here too. Almeroth ¶355-56.

13. Element 1[l]

Presenting second video programming on user computing device's display:

After performing a follow-up search and defining second video programming, the programming is displayed on the user computing device's display, just like the first video programming. Additional detail is in 1[f] above. Almeroth ¶358.

14. Element 1[m]

Searching or defining step is performed according to templates: These searching and defining steps are both performed according to at least a template referred to in Ferman as the "user description scheme." Ferman at [0097] (contents of user description scheme); [0097] ("[C]ompare the information contained in the user description scheme 20 to the available information contained in the program

description scheme 18 to make selections.”); [0097]; Almeroth ¶360. As discussed above in ground 1, element 1[m], the intrinsic evidence does not limit a “template” to a particular type of data structure. The user description scheme is a structured collection of data in XML format, which a POSITA would have understood to be a claimed “template.” Ferman at [0196] (XML format); Almeroth ¶361.

Plurality of template values: The templates comprise a plurality of template variables, namely the data contained in the user description scheme. Ferman at [0196]; [0273]; Almeroth ¶361.

Information about a user associated with the second computing device: The user description scheme template includes data about the user associated with the user computing device that submitted the request. Ferman at [0108] (“user history” and “video viewing habits”); [0333]; [0334] (“user’s preferences”); [0394]; [0395]; [0407]; [0410]; cl. 1 (“user attribute information”); Almeroth ¶362.

Viewing scenarios of two or more selected video clips in a temporal relationship to each other: The user description scheme includes viewing scenarios of multiple clips in a temporal relationship. Ferman at [0267] (“user’s preference for continuous playback of the entire program versus visualizing a short summary of the program”); Almeroth ¶363.

Moreover, other templates in Ferman’s system, including the segment description scheme and the semantic structure description scheme, include temporal

viewing scenarios for clips. Ferman at [0252] (“[A] segment tree may be used to defined a temporal structure of a video program.”); [0255] (“The event description scheme 480 may be used to form relationships between different events of the video normally consisting of multiple frames of the video.”); Almeroth ¶364.

Style requirements: The user description scheme template includes style requirements. Ferman at [0318] (“the desired configuration for presentation, such as volume or any other functionality”). Additionally, other templates that are used in the process of generating video programming also include style requirements. *Id.* at [0266] (“preferences for setting up the device in relation to the type of content being presented”); [0268] (“favorite actors, directors”); *id.* (“favorite program category”); [0269] (“how the user browses and consumes the audio and/or video content.”); [282] (“the preferred theme, or point of view, of a segment...annotated with its theme or emphasis point”); Almeroth ¶365.

E. Claim 2

As described in 1[d] above, Ferman’s system includes transitional clips in the first video programming. And as described in 1[j], the same types of transitional clips are also included in the second video programming. Additional details are in 1[b], 1[d], and 1[e]. Almeroth ¶367.

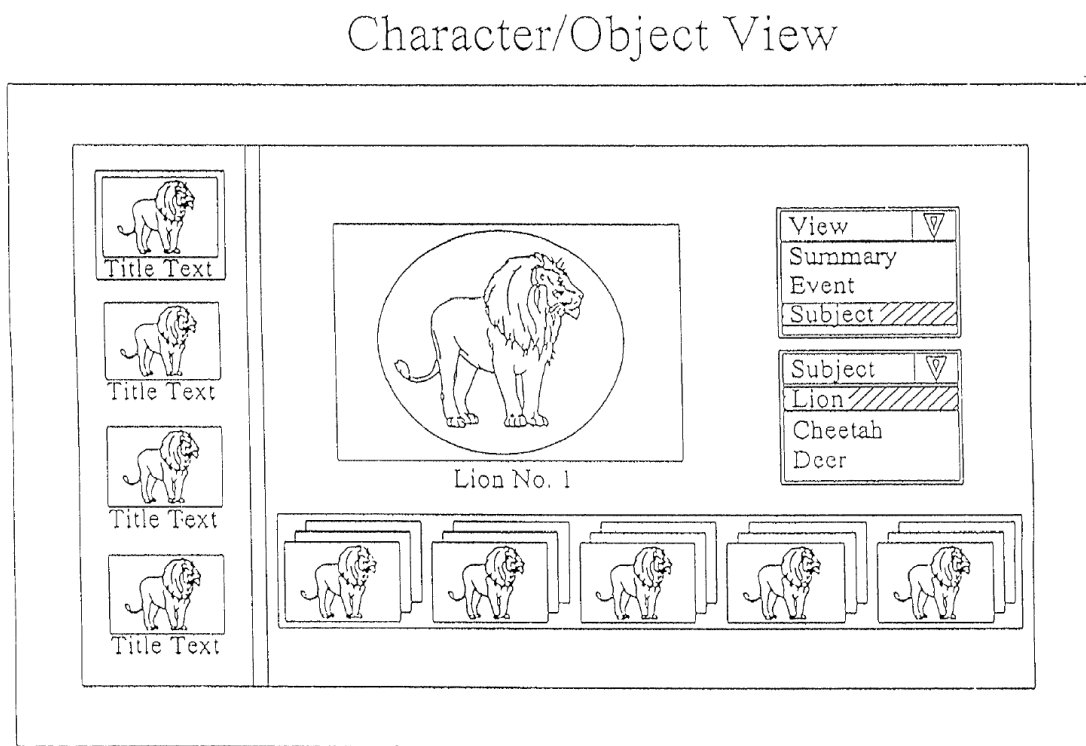
F. Claim 3

Transitional clips will be present in the dynamic video programming produced

by Ferman's system, either for an initial search or a follow-up search. Therefore, for the same reasons as claim 2, claim 3 is also disclosed by Ferman. Almeroth ¶369.

G. Claim 4

Video clips in Ferman's system can focus on a particular person or another host subject, like a lion:



Ferman at Fig. 12; [0124] (character/object view); [0095] (close-up view of Michael Jordan). As with the “talking head” clips in Marcus156 (ground 1), a POSITA would consider these types of clips to be “hosting clips,” as claimed. Almeroth ¶371.

To the extent that Ferman does not disclose this claim, it renders it obvious,

for the same reasons that Marcus156's "talking head" clips would also render this claim obvious, discussed in ground 1, claim 4 above. Almeroth ¶372.

H. Claim 5

Ferman discloses or renders obvious this claim for the same reasons as claim 4. Almeroth ¶374.

I. Claim 6

The template variables corresponding to viewing scenarios discussed in 1[m] also include sequencing limitations for incorporating video clips at specific locations in the first or second video programming. Ferman at [0095]; [0252] ("[A] segment tree may be used to defined a temporal structure of a video program."); [0255] ("The event description scheme 480 may be used to form relationships between different events of the video normally consisting of multiple frames of the video."); [0614] (compilation of sequencing-related variables). Because these "sequential requirements" place clips in temporal relationships to each other, they are "sequencing limitations for incorporating video clips at specific locations" in the programming. Almeroth ¶376.

J. Claim 7

Ferman's templated sequencing limitations meet each of the three categories. The limitations may be length, order, or time frame requirements. Ferman at [0252]; [0255]; [0281] ("The MaxSummaryDuration and MinSummaryDuration descriptors

specify, respectively, the maximum and minimum temporal duration of an audiovisual highlight summary.”); [0614] (compilation of sequencing-related variables, including the temporal duration of clips); Almeroth ¶378.

K. Claim 9

Ferman’s templates include demographic information. Ferman at [0097] (“personal information about the particular user, such as demographic and geographic information, e.g. zip code and age”); *id.* (“demographics”); [0110] (“demographic information such as age, zip code, etc.”); [0118]; [0221] (“The descriptor <Age> specifies the age of a user.”); [0224] (“The descriptor <Gender> specifies the gender of a user.”); Almeroth ¶380.

The templates include psychographic information, namely the user’s “emotional state.” Ferman at [0274] (“different emotional states of the user (e.g., happy mood versus tired or sad)”; [0275] (“different emotional states...may likewise be used as the basis for distinguishing between the user preference descriptions”). A POSITA would have considered an estimated or measured “emotional state” to be “psychographic information,” as claimed. Almeroth ¶381.

The templates also include a viewing history. Ferman at [0097] (“information regarding the user's viewing history such as for example browsing history, filtering history, searching history, and device setting history”); [0212] (“the history of a user’s browsing related activities”); [0218] (“the history of a user’s device related

activities”); Ferman at [0258]; [0259]; [0271]; Almeroth ¶382.

L. Claim 10

Templates are used to select video clips based on associations with other video clips, including by using a tag or attribute of one clip as a necessary requirement for another clip. Ferman at [0105] (“SFB module 52 may perform filtering, searching, and browsing of the programs 38, on the basis of the information contained in the description schemes.”); Almeroth ¶384.

M. Claim 11

Ferman’s system uses media tags associated with clips in the program profile to categorize clips and to associate them with other clips. Ferman at [0094] (program profiles include “characteristics of the content of the program, such as actors, stars, rating, director, release date, time stamps, keyword identification, trigger profile, still profile, event profile, character profile, object profile, color profile, texture profile, shape profile, motion profile, and categories”); [0315] (“Each of the selected programs include a rich set of associated data which may then be applied against user criteria at block 618 to further refine the data by extracting desirable sub-portions of each program.”); [0407]; cl. 1 (program attribute information). These are “media tags” as described in the appendix to the ’597 provisional. Ex. 1003 at 61 (“Described here are tagging considerations.... Media clips will normally bear taxonomic data allowing the clip to be retrieved and sorted by clip content.”);

Almeroth ¶386.

N. Claim 12

Ferman's media tags include context and subject matter data. Ferman at [0278] ("media type, media source, or content"); [0282] ("preferred theme, or point of view, of a segment"); [0303] (same); [0601] ("title 746, location 748, creator 750, date 752, and actor/director"); Almeroth ¶388.

The media tags also include source information. Ferman at [0283] ("source of the content"); [0284] ("source distinction and identification"); [0305] ("source preference description describes the preferred source of multimedia information"); [0602] (history of sources); Almeroth ¶389.

Moreover, the media tags include data about region and language. Ferman at [0306] ("preferred program title, genre, language, actor, creator of the program"); [0328]; [0329] ("Location"); [0584]; [0600] ("genre, language, and country of origin"); Almeroth ¶390.

O. Claim 13

The finalized first and second video programming is assembled and presented to the user as a unified video stream. The user sees a sequence of videos where the clips and transitions proceed sequentially, one after the other and without further user input. Ferman at [0607] ("PlayStream Play content from input stream"). The system generates a sequence of video programming of a given length, composed of

a concatenation of highlight or event clips. [0091]; [0107] (5-minute highlight stream); [0267] (15-minute highlight stream); Fig. 10; Almeroth ¶392.

P. Claims 14-15 and 17-20

Ferman renders claims 14-15 and 17-20 obvious for the same reasons as discussed in corresponding claims. Almeroth ¶¶393-398.

Claim	Corresponding Claim
14	6
15	7
17	9
18	10
29	11
20	12

XII. Ground 5: Ferman, Harbick, and Byers Render Claims 8 and 16 Obvious

A. Proposed Modification

The combined system of Ferman and Harbick described above in ground 4 would be modified according to the teachings of Byers to include advertising clips in the video programming, including by dynamically editing other clips. Part of this modification would entail adding advertising material to the program description scheme and tracking additional advertising-related parameters in both the program description scheme and the user description scheme described in Ferman. Almeroth

¶400.

B. Motivation to Combine

Commercial benefit to advertising: Because insertion of advertising into video programming would enable a new revenue stream for the video creator, a POSITA would have had an economic motivation to do it. *E.g.*, Byers at 2:51-52; Ferman at [0088-92]; Almeroth ¶401.

Commercial benefit to user-targeted advertising: Ferman’s user-targeted recommendation techniques allow improved advertising that is customized to the demographics, interests, or other facets of the user, thus increasing the pertinence and, ultimately, the click-through or conversion rate of the targeted advertising. Byers at 2:57-67 (“One problem with product placement is that the placement is not tailored for an individual user....”); Almeroth ¶402.

Same field of art: Like Ferman and the ’597 patent, Byers is directed to creation and presentation of dynamic video programming. Byers at 1:40-43 (“[A] need exists for a method and apparatus for dynamically altering digital video images to allow customization of a digital video stream based upon characteristics associated with a user.”); Ferman at [0102]; Almeroth ¶403.

C. Reasonable Expectation of Success

Software modification would be predictable: This change would be accomplished in software and would entail predictable changes to Ferman’s system.

Almeroth ¶404.

Byers provides explicit instructions on how to modify clips to insert advertising: Byers discloses underlying mechanisms to insert the advertising content. Byers at 3:67-4:5; 5:60-66. The hardware used in this process is similar to Ferman’s disclosed hardware. Byers at 5:45-67; Ferman at Fig. 2. Ferman’s disclosed hardware could perform Byers’ video editing algorithms. Almeroth ¶405.

**D. Claims 8 and 16: Ferman, Harbick, and Byers Render Obvious
“One or More Video Clips Comprise Advertising”**

Byers discloses an “advertiser database” that contains advertising-related information. Byers at 4:19-25. The information in the database is used to dynamically insert ads into a video stream. *Id.* at 2:48-51; 3:1-4; Almeroth ¶¶406-07.

Moreover, Ferman’s system explicitly tracks advertising clips. For example, Ferman mentions the capability to intentionally remove advertising. Ferman at [0103] (“commercial filter 92 to remove commercials”). A POSITA would have understood that in order to remove commercials, the system must be able to identify commercials in clips, and the clips in the database must include commercials. Almeroth ¶408. If the commercial-removing functionality is not activated, then commercials will remain in the video programming and be displayed to the user.

XIII. No Secondary Considerations

PO has not identified any secondary considerations sufficient to overcome

Petitioner's strong showing of obviousness. Almeroth ¶409.

XIV. Instituting This IPR Would Be Equitable

A. 35 U.S.C. § 325(d) Supports Institution

The Board should not exercise its discretion under § 325(d). *Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6 at 7 (PTAB Feb. 13, 2020). The petition does not present the same or similar grounds previously before the Office, and all six *Becton* factors support institution. *Becton, Dickinson & Co. v. B. Braun Melsungen AG*, IPR2017-01586, Paper 8 at 24 (Dec. 15, 2017).

Becton factors (a) and (b): None of the references relied upon in this petition except for Marcus156 were considered by the examiner. Grounds 4 and 5 do not rely on Marcus156 at all and instead rely on Ferman, Harbick, and Byers. These references were not before the examiner, so any rationale for discretionary denial based on the previous citation of Marcus156 is not applicable to the full set of grounds. And grounds 1-3 combine Marcus156 with additional references (Harbick, Hosken, and Byers) not considered by the examiner. Thus, the PTAB is presented with new, non-cumulative obviousness grounds that were never considered during prosecution of the '597 patent.

Becton factors (c) and (d): Marcus156 was presented differently and for different purposes by the examiner during prosecution than here. For example,

Marcus156 was not cited by the examiner during the prosecution of the '597 patent as the basis for any anticipation or obviousness rejection. Although Marcus156 was cited by the examiner during the prosecution of the related '185 patent, it was only used as a secondary reference for some specific claim limitations. Ex. 1005 at 217-26. In contrast, the Petition relies on Marcus156 as the base reference in grounds 1-3. Additionally, when Marcus156 was cited as a primary reference during the prosecution of other patents in the same family, the wording of the then-pending claims was significantly different from the challenged claims of the '597 patent. Ex. 1007 at 290-96, 320-39; Ex. 1008 at 51-57.

Prosecution of later family members of the '597 patent shows the critical relevance of Marcus156 to the alleged invention. In the subsequent '873 and '374 applications, the examiner rejected claims based on Marcus156 in combination with other prior art, and in both cases PO was unable to successfully distinguish Marcus156 and instead had to further amend the pending claims with limitations not found in the '597 patent. Ex. 1007 at 305-14, 360-67; Ex. 1008 at 78-81. Had the examiner considered Marcus156 as a base reference during prosecution of the '597 patent, the '597 patent claims would not have issued in their current form.

Becton factors (e) and (f): Petitioner raises new arguments and expert analysis concerning the disclosures of Marcus156 and its relationship to the challenged claims. For example, during prosecution of the '185 patent, the examiner

stated that Marcus156 does not “explicitly disclose that said first and second video programming are each presented as a stream of information to the user.” Ex. 1005 at 222. In claim 13 in ground 1 above, Petitioner and its expert show how that is simply not the case, and how the examiner overlooked Marcus156’s disclosure of a stream of video programming. Indeed, Marcus 156 expressly states that its video programming is “streamed.” Petitioner has overcome any presumption of discretionary denial by showing specifically how the examiner erred in evaluating Marcus156 during prosecution, and has submitted additional analysis in the petition and expert declaration that warrants reconsideration of the Marcus156. *Becton* at 24.

Put simply, if there ever could be a case where it would be appropriate to raise references in an IPR that were previously considered during prosecution, this is it. Discretionary denial under § 325(d) is manifestly unwarranted. *Pure Storage, Inc. v. Realtime Data LLC*, Case No. IPR2018-00549, Paper 7 at 11 (PTAB July 23, 2018).

B. *Fintiv* Supports Institution

District court litigation has not even commenced in earnest. Petitioner is filing this IPR petition a month before its answer is due. The six *Fintiv* factors thus strongly favor institution.

1. Overlap Between Issues

While PO has not yet identified the set of asserted claims, this Petition challenges all '597 patent claims. Accordingly, no matter what claims PO asserts in district court, all those claims will also be at issue here. On the other hand, Petitioner has not yet served its invalidity contentions. Thus, there is no basis for concluding that there is substantive overlap in invalidity theories.

2. Likelihood of a Stay

Neither party has requested a stay in the district court in part because the district court case is not yet underway. The likelihood of a stay is increased here, however, by the fact that Petitioner is filing this IPR petition before answering the complaint and is filing concurrent IPR petitions against all asserted patents.

3. Proximity of the Court's Trial Date to a FWD

No trial date has been set. Indeed, no case schedule of any kind has been entered, nor has a Rule 26(f) conference been held. At Judge Gilstrap's most-recent, July 2024 scheduling conferences, the Court assigned trial dates from January to April 2026. *Avant Location Techs., LLC v. Samsung Elec. Co. Ltd.*, No. 2:24-cv-00133-JRC, Dkt. 36 (E.D. Tex. Aug. 5, 2024); *Alto Dynamics, LLC v. Gucci Am., Inc.*, No. 2:24-cv-00375, Dkt. 29 (E.D. Tex. Aug. 2, 2024). This case is scheduled for a case management conference on September 18, 2024. *NTECH Properties, Inc. v. ByteDance Ltd.*, No. 2:24-cv-00130-JRG, Dkt. 13 (E.D. Tex. Aug. 8, 2024). By

then, Judge Gilstrap will presumptively be assigning trial dates starting in May 2026.

If that is the case, then the PTAB will issue its final written decision *before* any district court trial.

4. Investment in the Parallel Proceeding

Although PO filed a district court complaint on February 23, 2024, investment in the parallel proceeding has been nonexistent. PO did not serve the complaint until June 14, 2024. The parties then stipulated to an answer deadline of September 16, 2024. There have been no other case developments and no discovery. The district court has also invested zero time in assessing the parties' claims and defenses.

Moreover, Petitioner is filing this IPR a mere two months after being served with the complaint—almost 10 months before its statutory bar date (June 14, 2025).

5. Whether the Petitioner and the Defendant in the Parallel Proceeding are the Same Party

The parties are the same in this IPR and in the Litigation.

6. Other Circumstances, that Impact the Board's Exercise of Discretion, Including the Merits

This is the Board's opportunity to reassess a facially invalid patent that improperly attempts to extend the lifespan of PO's significantly earlier Marcus156 patent. The merits of the Petition are particularly strong and outweigh any countervailing considerations of efficiency.

XV. Conclusion

For the reasons above, *inter partes* review is requested.

Date: August 23, 2024

Respectfully submitted,

By: /s/ James M. Glass
James M. Glass

Counsel for Petitioner

CERTIFICATION UNDER 37 C.F.R. § 42.24

Under the provisions of 37 C.F.R. § 42.24, the undersigned hereby certifies that the word count for the foregoing Petition for *inter partes* review (excluding the table of contents, table of authorities, mandatory notices, certificate of service or word count, and appendix of exhibits or claim listing) totals 13,706 words, which is within the word limit allowed under 37 C.F.R. § 42.24(a)(i).

Date: August 23, 2024

/s/ James Glass

James Glass (Reg. No. 46729)

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. §§ 42.6(e), 42.105(a), the undersigned hereby certifies service on the PO of a copy of this Petition and its respective exhibits at the official correspondence address for the attorneys of record for the '597 patent as shown in USPTO PAIR via FedEx:

Louis DelJuidice
Troutman Pepper Hamilton Sanders LLP
600 Peachtree St. NE, Suite 3000
Atlanta, GA

Courtesy copies were also sent via electronic mail to Patent Owner's counsel of record in the related district court proceeding:

Elizabeth L. DeRieux
ederieux@capshawlaw.com
Capshaw DeRieux LLP
114 E. Commerce St.
Gladewater, TX 75647

Paul J. Andre
pandrew@kramerlevin.com
Lisa Kobialka
lkobialka@kramerlevin.com
James Hannah
jhannah@kramerlevin.com
Kris Kastens
kkastens@kramerlevin.com
Kramer Levin Naftalis & Frankel LLP
333 Twin Dolphin Drive
Redwood City, CA 94065

Petition for *Inter Partes* Review

U.S. Patent No. 9,317,597

Date: August 23, 2024

/s/ *James Glass*

James Glass (Reg. No. 46729)