

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC

Petitioner

v.

ART RESEARCH AND TECHNOLOGY LLC

(record) Patent Owner

Patent No. 10,609,442

DECLARATION OF DR. HENRY HOUH

TABLE OF CONTENTS

TABLE OF CONTENTS	2
I. ENGAGEMENT AND COMPENSATION	5
II. SUMMARY OF OPINIONS	7
III. QUALIFICATIONS	7
IV. MY UNDERSTANDING OF THE RELEVANT LAW	10
A. Claim Construction	10
B. Anticipation	10
C. Obviousness	11
V. TECHNICAL INTRODUCTION	14
VI. OVERVIEW OF THE '442 PATENT CLAIMS	24
VII. RELEVANT TIMEFRAME FOR DETERMINING OBVIOUSNESS	31
VIII. LEVEL OF ORDINARY SKILL IN the ART	32
IX. DETAILED EXPLANATION OF THE REASONS FOR UNPATENTABILITY	32
Ground 1. Claims 1 and 7 are obvious over Junee in view of De Jong and further in view of Houh	32
A. Overview of the Ground	33
1. Overview of Junee	34
2. Overview of De Jong	40
3. Overview of Houh	44
B. Rationale (Motivation) Supporting Obviousness	45
C. Reasonable Expectation of Success	55
D. Graham Factors	55

E.	Analogous Art.....	56
F.	Claim Mapping	56
Ground 2.	Claims 5 and 6 are obvious over June, De Jong, Houh, and further in view of Lee	98
A.	Overview of the Ground	98
1.	Overview of Lee	99
B.	Rationale (Motivation) Supporting Obviousness	102
C.	Reasonable Expectation of Success.....	105
D.	Graham Factors.....	106
E.	Analogous Art.....	106
F.	Claim Mapping	107
Ground 3.	Claim 8 is obvious over June, De Jong, Houh and further in view of Syed, Giunio-Zorkin, and Kwon	114
A.	Overview of the Ground	115
1.	Overview of Syed	116
2.	Overview of Giunio-Zorkin.....	120
3.	Overview of Kwon	123
B.	Rationale (Motivation) Supporting Obviousness	126
C.	Reasonable Expectation of Success.....	136
D.	Graham Factors.....	136
E.	Analogous Art.....	137
F.	Claim Mapping	137
Ground 4.	Claim 9 is obvious over June, De Jong, Houh, Syed, Giunio-Zorkin, Kwon, and further in view of Dayan	148

A.	Overview of the Ground	149
1.	Overview of Dayan.....	150
B.	Rationale (Motivation) Supporting Obviousness	152
C.	Reasonable Expectation of Success.....	157
D.	Graham Factors.....	158
E.	Analogous Art.....	158
F.	Claim Mapping	158
Ground 5.	Claims 10 and 11 are obvious over Syed in view of Giunio-Zorkin and Kwon.....	167
A.	Overview of the Ground	167
1.	Overview of Syed.....	167
2.	Overview of Giunio-Zorkin.....	169
3.	Overview of Kwon	170
B.	Rationale (Motivation) Supporting Obviousness	171
C.	Reasonable Expectation of Success.....	178
D.	Graham Factors.....	179
E.	Analogous Art.....	179
F.	Claim Mapping	179
X.	OATH.....	197

I, Henry Houh, hereby declare as follows:

I. ENGAGEMENT AND COMPENSATION

1. I have been retained by Google LLC (“Petitioner”) as an expert in the *inter partes* review proceeding described above. I have been asked to provide my technical review, analysis, insights, and opinions regarding the references that form the basis for the grounds of unpatentability set forth in the Petition for *inter partes* review of U.S. Pat. No. 10,609,442 (“the ’442 patent”). For this service, my billing rate is my standard hourly consulting rate, and I am reimbursed for actual expenses. My compensation in no way depends on the outcome of this matter.

2. This declaration sets forth my analyses and opinions based on the materials I have considered thus far and the bases for my opinions. I understand that this declaration will be used in the above mentioned *inter partes* review. In forming my opinions, I have considered the following Exhibits:

Exhibit No.	Description
1001	U.S. Patent No. 10,609,442 B2 (“ the ’442 patent ”)
1002	Declaration of Dr. Henry Houh
1003	C.V. of Dr. Henry Houh
1004	Prosecution History of U.S. Application No. 15/479,774 (which Issued as the ’442 patent)
1005	U.S. Patent No. 8,826,117 B1 (“ June ”)
1006	U.S. Patent Pub. No. 2018/0190323 A1 (“ De Jong ”)

1007	Intentionally left blank
1008	Intentionally left blank
1009	U.S. Patent Pub. No. 2015/0139610 A1 (“ Syed ”)
1010	Intentionally left blank
1011	U.S. Patent Pub. No. 2015/0180820 A1 (“ Lee ”)
1012	U.S. Patent Pub. No. 2016/0350430 A1 (“ Kwon ”)
1013	U.S. Patent No. 10,031,949 B2 (“ Giunio-Zorkin ”)
1014	U.S. Patent No. 10,180,989 B2 (“ Dayan ”)
1015	U.S. Patent No. 8,341,123 B2
1016	Marcus, “ <i>Graphical user interfaces.</i> ” Handbook of human-computer interaction. North-Holland, 1997. 423-440.
1017	Jansen, “ <i>The graphical user interface.</i> ” ACM SIGCHI Bulletin 30.2 (1998): 22–26.
1018	Galitz, “ <i>The essential guide to user interface design: an introduction to GUI design principles and techniques.</i> ” John Wiley & Sons, 2007.
1019	Microsoft Press, “ <i>Microsoft computer dictionary.</i> ” Microsoft Press, 2002.
1020	U.S. Patent No. 9,697,230 B2 (“ Houh ”)
1021	Intl. Pub. No. WO2016/174524 A2
1022	Redline of Ex-1006 and Ex-1021
1023	U.S. Patent No. 7,801,910 B2

II. SUMMARY OF OPINIONS

3. It is my opinion that:

- **Ground 1:** Claims 1 and 7 are obvious over Junea in view of De Jong and further in view of Houh.
- **Ground 2:** Claims 5 and 6 are obvious over Junea, De Jong, Houh, and further in view of Lee.
- **Ground 3:** Claim 8 is obvious over Junea, De Jong, Houh and further in view of Syed, Giunio-Zorkin, and Kwon.
- **Ground 4:** Claim 9 is obvious over Junea, De Jong, Houh, Syed, Giunio-Zorkin, Kwon, and further in view of Dayan.
- **Ground 5:** Claims 10 and 11 are obvious over Syed in view of Giunio-Zorkin and Kwon.

III. QUALIFICATIONS

4. I hold a Ph.D. in electrical engineering and computer science from the Massachusetts Institute of Technology (“MIT”), awarded in 1998. I also hold a Master’s Degree in electrical engineering and computer science from MIT, awarded in 1991, and two Bachelor’s degrees (the first in electrical engineering and computer science and the second in physics) also from MIT, awarded in 1989 and 1990, respectively.

5. I have approximately 25 years of work experience related to topics in networking, Internet communications, Web development, social networking, Internet advertising, and Web-based digital media. I am currently a technical consultant specializing in Social Networking, Web 2.0, Web Site Development, Data Networking, Optical Networking, Telecommunications, Media Streaming and Voice over IP.

6. I have extensive experience with Web programming, Web-based digital media, and Web-based advertisement technologies. I am familiar with social networking, annotation of media, including static and playable media.

7. From 1994-1996, I was the founder and CEO of Agora Technology Group, where I conceived and oversaw development of targeted advertising-supported web sites. From 1997-1999, I was a Software Engineer and Senior Scientist and Engineer for NBX (acquired by 3Com), where I worked on IP telephony (telephony using Internet protocol). From 1999-2004, I worked for Teradyne and later its spinout (Empirix). During this period, I held the position of Chief Technologist of the Web Application Test Group. In that position, I was the Chief Architect for all Web testing products, re-architected the core testing product, and helped write a JavaScript interpreter.

8. In 2006, I was the Vice President of Operations and Technology for PodZinger Inc. There, I helped create and launch a consumer-facing search site for

audio and video and redeployed the web site from our company's machine room to a co-location facility. I was also responsible for evaluating advertisement models of revenue. The PodZinger search site indexed audio and video located all over the web, and used speech recognition and natural language processing to create a spoken-word transcripts of the media files, and augmented the audio/video search results with snippets of the transcripts. When a search result was selected from the results, the Podzinger web site provided a media page with a player integrated with a full transcript of the spoken words, where each word in the transcript was selectable by the user and would jump the player to the exact location of the selected spoken word within the audio or video file. Podzinger was named by Forbes Magazine as their favorite Video & Audio Search Engine in 2006, beating Google Video Search, Yahoo Video Search and others.

9. From 2007-2008, I was Vice President of Technology for the Delta Division of BBN Technologies, where I oversaw the development and deployment of a new fully-featured social networking web site. This social networking site allowed for user to share, discuss, and annotate media. From 2008-2009, I was the Chief Technology Officer for Eons. With Eons, I integrated a product I created at BBN that Eons acquired. I integrated the product into the Eons social networking platform. I also helped evaluate advertisement platform offerings and brought the "Boom Network" advertising network to market.

10. My qualifications, publications, and patents are set forth more fully in my curriculum vitae, attached as Ex-1003.

IV. MY UNDERSTANDING OF THE RELEVANT LAW

A. Claim Construction

11. I understand that a claim term in an *inter partes* review is to be interpreted according to standards applied by district courts. First, the language of the claims themselves is of primary importance in the effort to determine precisely what it is that is patented. The terms used in a claim are generally given the ordinary and customary meaning that the terms would have to a person of ordinary skill in the art in question at the time of the alleged invention, unless the term is expressly defined in the patent. The person of ordinary skill in the art reads the claim in the context of the entire patent, including the specification. Next to the language of the claims, the specification is the single best source for interpreting the claim terms. The claims may also be interpreted using the record of correspondence between the patent applicant and the Patent Office, and also with reference to other sources of evidence that can help to define the meaning of terms to a person of ordinary skill in the relevant time frame. I define the relevant time frame in ¶45, below.

B. Anticipation

12. I understand that a claim in an issued patent can be invalid if it is anticipated. In this case, “anticipation” means that there is a single prior art

reference that discloses every element of the claim, arranged in the way required by the claim.

13. I understand that an anticipating prior art reference must disclose each of the claim elements expressly or inherently. I understand that “inherent” disclosure means that the claim element, although not expressly described by the prior art reference, must necessarily be present based on the disclosure. I understand that a mere probability that the element is present is not sufficient to qualify as “inherent disclosure”.

C. Obviousness

14. I understand that a claim in an issued patent can be invalid if it is obvious. Unlike anticipation, obviousness does not require that every element of the claim be in a single prior art reference. Instead, it is possible for claim elements to be described in different prior art references, so long as there is motivation or sufficient reasoning to combine the references.

15. I understand that a claim is invalid for obviousness if the differences between the claimed subject matter and the prior art are such that the subject matter as a whole would have been obvious at the time the alleged invention was made to a person having ordinary skill in the art to which said subject matter pertains.

16. I understand, therefore, that when evaluating obviousness, one must consider obviousness of the claim “as a whole”. This consideration must be from the perspective of the person of ordinary skill in the relevant art, and that such perspective must be considered as of the “time the invention was made”.

17. The level of ordinary skill in the art is discussed in ¶46 below.

18. The relevant time frame for obviousness, the “time the invention was made”, is discussed in ¶45, below.

19. I understand that in considering the obviousness of a claim, one must consider four things. These include the scope and content of the prior art, the level of ordinary skill in the art at the relevant time, the differences between the prior art and the claim, and any “secondary considerations”. I understand that these four things were specified by the U.S. Supreme Court in a case called *Graham v. John Deere*, and are therefore called the “*Graham* factors”.

20. I understand that “secondary considerations” include real-world evidence that can tend to make a conclusion of obviousness either more probable or less probable. For example, the commercial success of a product embodying a claim of the patent could provide evidence tending to show that the claimed invention is not obvious. In order to understand the strength of the evidence, one would want to know whether the commercial success is traceable to a certain aspect of the claim not disclosed in a single prior art reference (i.e., whether there

is a causal “nexus” to the claim language). One would also want to know how the market reacted to disclosure of the invention, and whether commercial success might be traceable to things other than innovation, for example the market power of the seller, an advertising campaign, or the existence of a complex system having many features beyond the claims that might be desirable to a consumer. One would also want to know how the product compared to similar products not embodying the claim. I understand that commercial success evidence should be reasonably commensurate with the scope of the claim, but that it is not necessary for a commercial product to embody the full scope of the claim.

21. Other kinds of secondary considerations are possible. For example, evidence that the relevant field had a long-established, unsolved problem or need that was later provided by the claimed invention could be indicative of non-obviousness. Evidence that others had tried, but failed to make an aspect of the claim might indicate that the art lacked the requisite skill to do so. Evidence of copying of the patent owner’s products before the patent was published might also indicate that its approach to solving a particular problem was not obvious. Evidence that the art recognized the value of products embodying a claim, for example, by praising the named inventors’ work, might tend to show that the claim was non-obvious.

22. I further understand that prior art references can be combined where there is an express or implied rationale to do so. Such a rationale might include an expected advantage to be obtained, or might be implied under the circumstances. For example, a claim is likely obvious if design needs or market pressures existing in the prior art make it natural for one or more known components to be combined, where each component continues to function in the expected manner when combined (i.e., when there are no unpredictable results). A claim is also likely invalid where it is the combination of a known base system with a known technique that can be applied to the base system without an unpredictable result. In these cases, the combination must be within the capabilities of a person of ordinary skill in the art.

23. I understand that when considering obviousness, one must not refer to teachings in the specification of the patent itself. One can, however, refer to portions of the specification admitted to be prior art, including the “Background” section. Furthermore, a lack of discussion in the patent specification concerning how to implement a disclosed technique can support an inference that the ability to implement the technique was within the ordinary skill in the prior art.

V. TECHNICAL INTRODUCTION

24. I have reviewed the '442 patent. The '442 patent is titled “Method and apparatus for generating and annotating virtual clips associated with a playable

media file.” Ex-1001. According to the ’442 patent, its subject matter relates “to assemblies, methods, devices, and systems for sharing and annotating content between members of a social network.” Ex-1001, 1:20-22. The ’442 patent states, “Applicants’ system and method can be used by multiple members of a social network to associate annotations with a Playable Media File including a composite digital clip, and/or to initiate discussion threads associated with that Playable Media File including a composite digital clip.” Ex-1001, 3:65-4:3.

25. The social network, according to the ’442 patent, is implemented using an “[a]lgorithm 136,” which “comprises Applicants’ source code to operate a public or private social network, and when implemented by computing device 110 causes a graphic user interface (‘GUI’) to be displayed.” Ex-1001, 5:4-7. The ’442 patent further states, “[a] member and/or user can further utilize the GUI to access Applicants’ social network as described herein.” Ex-1001, 5:14-15. “Applicants’ algorithm 136, and its functions, can be accessed by users of Applicants’ network 100 to create, share, edit, associate one or more annotations with, and/or associate one or more discussion threads with, a Playable Media File.” Ex-1001, 7:7-11.

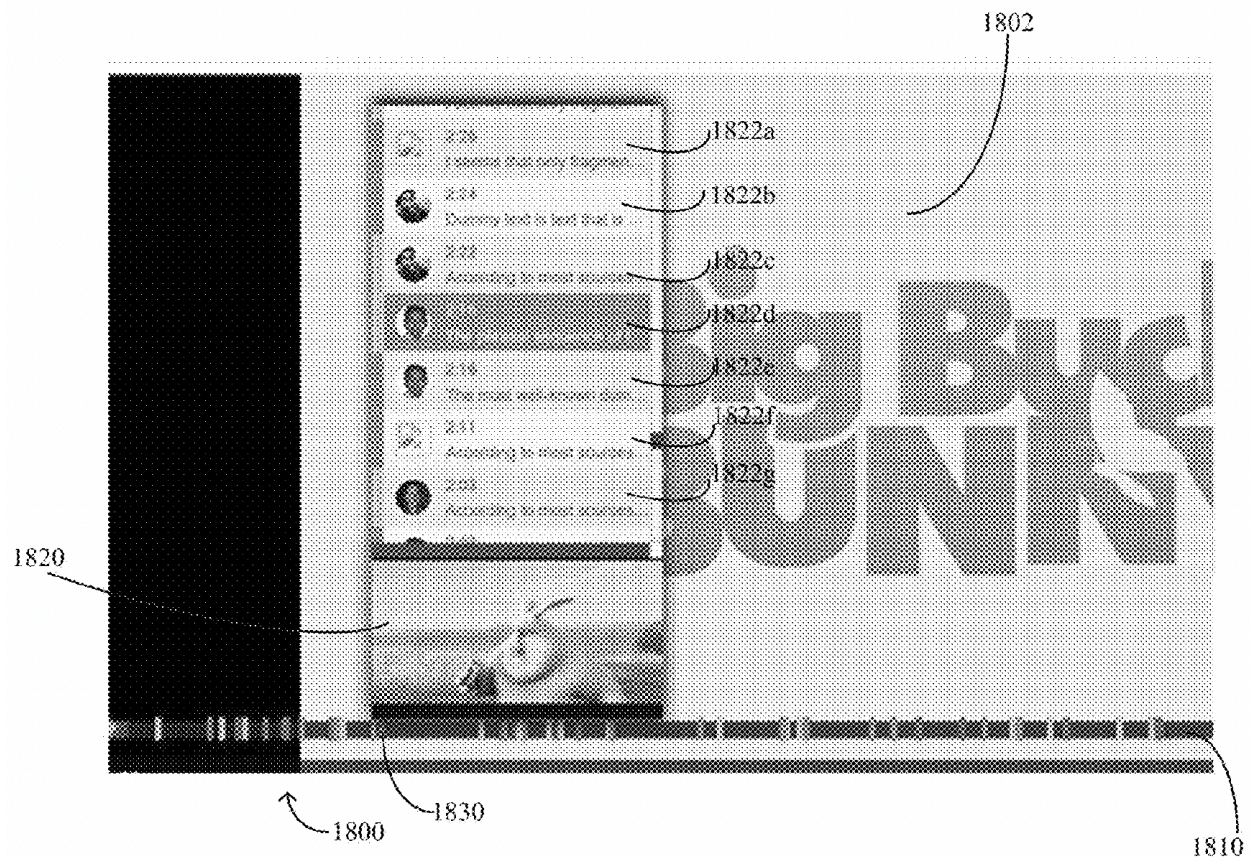
26. According to the ’442 patent, “a plurality of virtual clips [is] each associated with the playable media file.” Ex-1001, 1:30-31. The ’442 patent describes a “virtual clip” as “a set of data points that together delineate a particular

subset of the content encoded in the media file. Thus, the virtual clip is comprised of references that identify specific content in the corresponding media file, but the virtual clip is not necessarily itself a stored media file containing the content data.”

Ex-1001, 9:60-66.

27. The '442 patent further states that a “graphical user interface 1800 may include a display window 1802 for displaying content encoded by the playable media file.” Ex-1001, 17:66-18:1. According to the '442 patent, “the GUI 1800 may include a first interactable graphical object 1810, which displays a timeline representing a duration of play of a playable media file...the first interactable graphical object 1810 may display a plurality of visible annotation indicators. For example, a visible annotation indicator may be a clip indicator 1830 associated with a corresponding virtual clip that is associated with the playable media file.” Ex-1001, 18:13-23. The GUI 1800 is shown in Figure 18A, reproduced below:

FIG. 18A



Ex-1001, FIG. 18A.

28. According to the '442 patent, "[t]he GUI 1800 may further include a second interactable graphical object 1820 ... configured to dynamically display ... graphic elements each associated with a corresponding virtual clip of the plurality of virtual clip." Ex-1001, 18:40-46. "[W]hen the system receives a user input indicating that the user interacted with ... the timeline at a first display position ..., the system may identify one, some, or all of the virtual clips as displayable virtual clips: the virtual clip having its start time closest to the time at the first display position may be selected as a first clip; then, one or more virtual clips preceding

(e.g., sequentially preceding) the first clip and/or one or more virtual clips subsequent (e.g., sequentially subsequent) may be selected.” Ex-1001, 18:48-62.

“Then, in order of their start times, the displayable virtual clips are each associated with one of the graphic elements.” Ex-1001, 18:64-66.

29. The '442 patent further describes a “a data profile 300 [which] is created for the Playable Media File” when “a user having permission... associates an annotation with the Playable Media File.” Ex-1001, 7:25-28. According to the '442 patent, annotations are entered by a user via a displayed “Annotation Panel.” Ex-1001, 13:13-14. The '442 patent further describes “user input associated with the virtual clip, such as an annotation,” which is evaluated to determine whether the user input “conforms to a predetermined format.” Ex-1001, 13:39-41.

30. According to the '442 patent, “the format comprises a tag identifier indicating the user input includes taxonomy tags, a first tag following the tag identifier and identifying a first category of the composite virtual clip, and (P) subtag(s) sequentially following the first tag and each including a delimiter indicating a previous tag is complete and an additional tag follows, identifying an additional category of the virtual clip.” Ex-1001, 13:42-48. “The virtual clip may be associated with the taxonomy tags to produce a searchable virtual clip that is delivered to a requesting device in response to a query from the requesting device for any of the plurality of virtual clips that are associated with the taxonomy tags.”

Ex-1001, 14:19-24. The '442 patent states, “[t]he tag identifier may be a character (the ‘#’ or hash symbol in the examples herein)... In certain embodiments, the taxonomy tag comprises a form ‘#content:TITLE.’ in other embodiments, the taxonomy tag comprises a form ‘#first tag:subtag1:subtag2: . . . :subtagP,’ where the first tag and each subtag(1 . . . P) are character strings separated by the delimiter character ‘:.’” Ex-1001, 13:48-63.

31. In my opinion, the claims of the '442 patent are directed to different methods related to annotated playable media files. As set forth below, in my opinion, the challenged claims represent nothing more than the use of prior art techniques for annotating, storing, and displaying playable media files. Ex-1001.

32. I note specifically that I have considerable experience with annotating playable media and have obtained patents relating to the same. For example, U.S. Patent No. 9,697,230 B2 (“**Houh**”), on which I am the first named inventor, concerns annotating playable media. Ex-1020.

33. The Houh patent issued from an application filed on March 31, 2006 (and claims priority to a provisional filed on November 9, 2005). The Houh patent is broadly directed to “dynamic presentation of advertising, factual, [and] informational content” “according to the playback of corresponding segments identified within a media file or stream.” Ex-1020, Abstract, 1:65-2:4; *see generally*, Ex-1020, 1:48-3:33. For example, Figure 7A of the Houh patent,

reproduced below, illustrates an example “graphical user interface 610 [which] includes a media player 630 capable of audio/video playback and a user actuated display element 640 that serves as a navigational control.” Ex-1020, 14:40-53.

“Advertising content 900a, factual content 900b and informational content 900c (advertising content 900, generally) are shown positioned about the media player 630. The advertising content can also overlay video content, for example, at fixed positions 900d or scrolling across a portion of the video 900e.” *Id.*

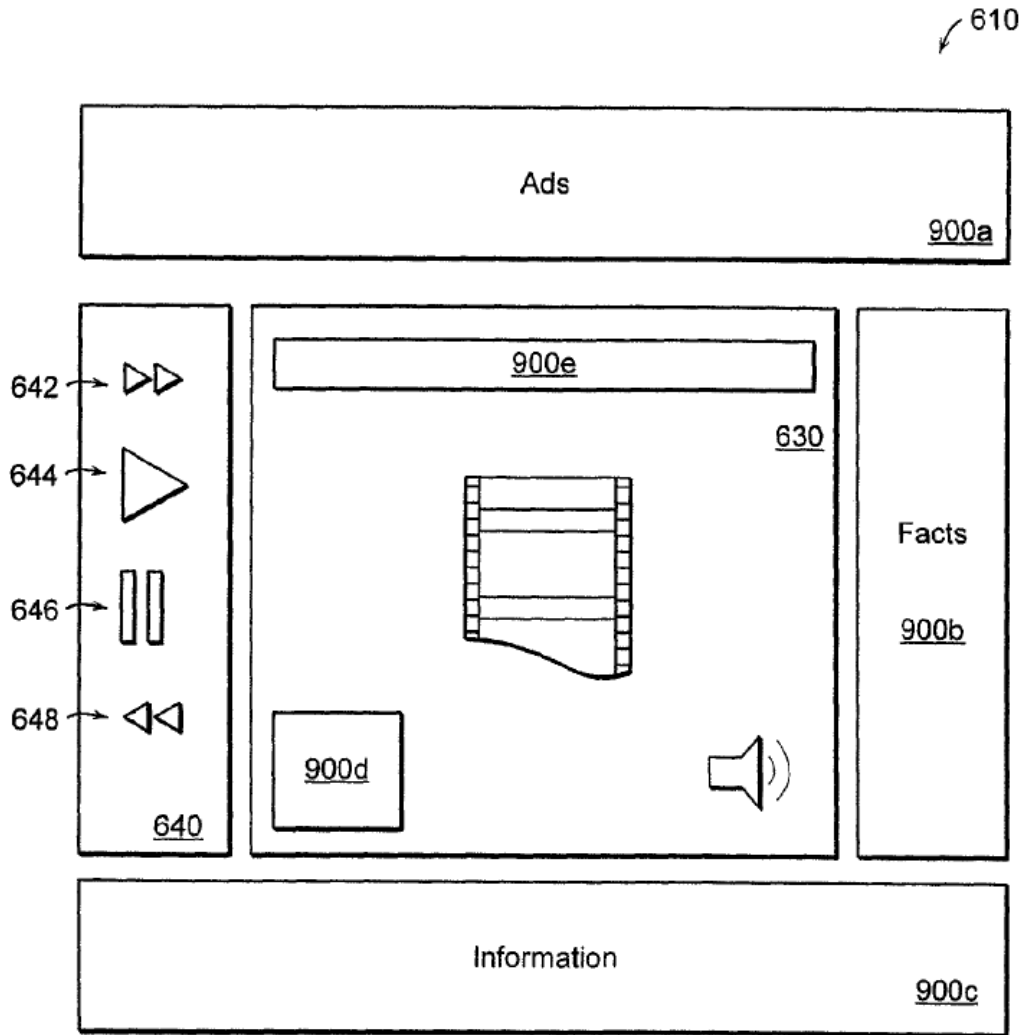


FIG. 7A

Ex-1020, Fig. 7A. Content 900 generally are annotations which relate to audio or video content displayed via the media player 630. The Houh patent explains that, for example, “with each transition from one content segment to another, an appropriate set of advertising, factual and information content can be presented that is related to the topics, stories, scenes or other segment characteristics of the underlying media content.” Ex-1020, 15:60-64.

34. As another example, U.S. Patent No. 7,801,910 B2 (“the ’910 patent”), on which I am the first named inventor, also concerns annotating playable media. Ex-1023.

35. The ’910 patent issued from an application filed on June 1, 2006 (and claims priority to a provisional filed on November 9, 2005). The ’910 patent is broadly directed to “generating metadata enhanced for audio, video or both (“audio/video”) search driven applications” by “appl[ying] one or more automated media processing techniques to the media file/stream, combin[ing] the results of the media processing into metadata enhanced for audio/video search, and stor[ing] the enhanced metadata in a searchable index or other data repository” Ex-1023, 1:45-61. The ’910 patent is also directed to “generating a timed segment index of discrete media content, the timed segment index identifying content segments of the discrete media content and corresponding timing boundaries of the content segments” where “the timed segment index identif[ies] text of audible words from content segments of the discrete media content and corresponding timing boundaries of the content segments” Ex-1023, 1:62-2:9. Furthermore, “[o]ne or more of the content segments identified in the timed segment index can include word segments, audio speech segments, video segments, or marker segments. For example, one or more of the content segments identified in the timed segment index can include audio corresponding to an individual word, audio corresponding

to a phrase, audio corresponding to a sentence, audio corresponding to a paragraph, audio corresponding to a story, audio corresponding to a topic, audio within a range of volume levels, audio of an identified speaker, audio during a speaker turn, audio associated with a speaker emotion, audio separated by sound gaps, audio separated by markers embedded within the media content or audio corresponding to a named entity. One or more of the content segments identified in the timed segment index can also include video of individual scenes, watermarks, recognized objects, recognized faces, overlay text or video separated by markers embedded within the media content.” Ex-1023, 2:10-27.

36. For example, Figure 6A of the '910 patent, reproduced below, illustrates an example graphical user interface that “includes a user actuated display element 640 that serves as a navigational control. The navigational control 640 enables a user to control playback of the underlying media content. The text area 620 is optional for displaying the text 625 of the words spoken during one or more segments of the underlying media content.” Ex-1023, 11:41-44.

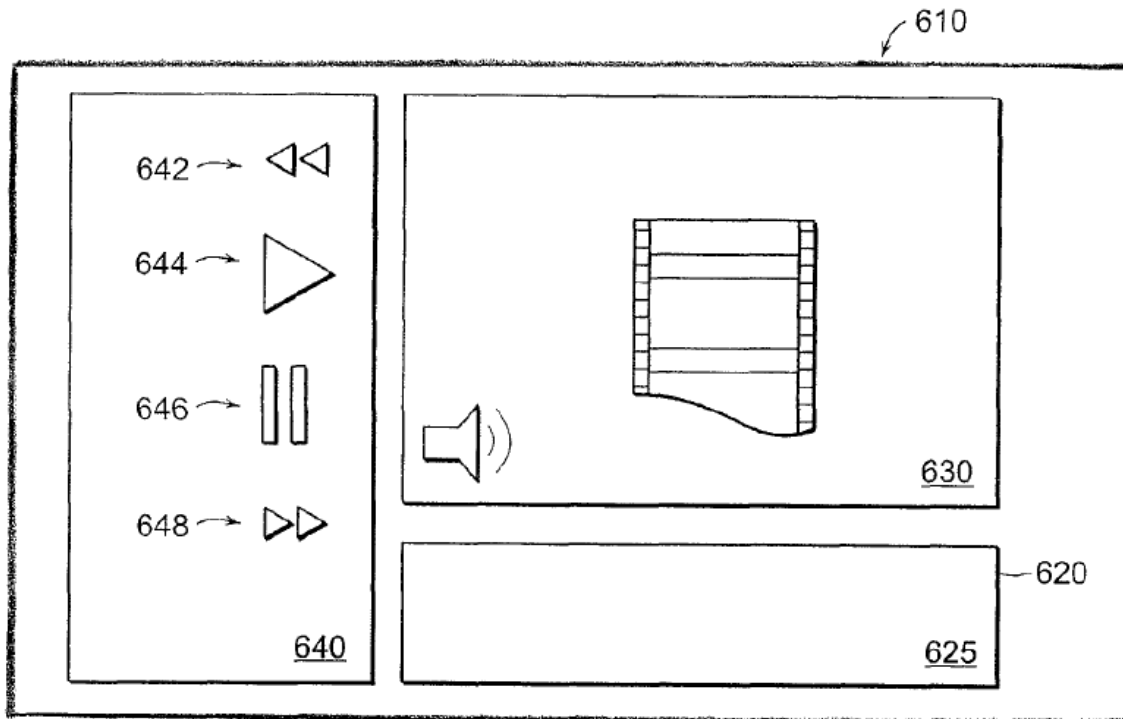


FIG. 6A

Ex-1023, Fig. 6A. “The text area 620 is optional for displaying the text 625 of the words spoken during one or more segments of the underlying media content.” Ex-1023, 11:44-47.

VI. OVERVIEW OF THE '442 PATENT CLAIMS

37. The '442 patent has two independent claims and thirteen claims in total.

38. Independent claim 1 is directed to a “method for displaying information associated with a Playable Media File.” Independent claim 1 is reproduced below, with bracketed numbering added to each claim limitation:

1[a]. A method for displaying information associated with a Playable Media File, comprising:

[1b] obtaining stored data describing the information,

[1c] the stored data comprising a storage location of the playable media file and

[1d] a plurality of virtual clips each associated with the playable media file and including

[1e] a first data element identifying a first time within the playable media file at which a corresponding virtual clip begins, and

[1f] a second data element identifying a first user profile associated with creating the corresponding virtual clip;

[1g] accessing the playable media file at the storage location;

[1h] causing a graphical user interface (GUI) to be displayed on a computing device of a user, wherein said GUI enables the user to generate user inputs by interacting with the GUI, and

[1i] the GUI comprises: a display window for displaying content encoded by the playable media file;

[1j] a first interactable graphical object, wherein the first

interactable graphical object overlays a first portion of the display window and displays a timeline representing a duration of the playable media file and a plurality of clip indicators each associated with a corresponding virtual clip of the plurality of virtual clips, each clip indicator appearing on the timeline at a display position corresponding to the first time identified by the first data element of the corresponding virtual clip; and

[1k] a second interactable graphical object, wherein the second interactable graphical object overlays a second portion of the display window, is configured to display up to a first number of graphic elements each associated with a corresponding virtual clip of the plurality of virtual clips, and is initially not displayed in the GUI;

[1l] receiving a first user input indicating a first interaction of the user with a first display position on the timeline;

[1m] determining a selected time within the playable media file that corresponds to the first display position;

[1n] identifying, as a plurality of displayable virtual clips:

[1o] a first virtual clip of the plurality of virtual clips, the corresponding first time of the first virtual clip being the closest, of the plurality of virtual clips, to the selected time; and

[1p] one or more of the virtual clips wherein the corresponding first time precedes and is approximate to the first time of

the first virtual clip, and one or more of the virtual clips wherein the corresponding first time is subsequent and approximate to the first time of the first virtual clip, such that at most the first number of the plurality of virtual clips are selected as the plurality of displayable virtual clips; and

[1q] updating the user interface on the computing device to display a list of the plurality of displayable virtual clips in the second interactable graphical object.

Ex-1001, claim 1.

39. Independent claim 10 is directed to a “method of annotating a playable media file.” Independent claim 10 is reproduced below, with bracketed numbering added to each claim limitation:

10[a]. A method of annotating a playable media file, the method comprising:

[10b] obtaining a virtual clip comprising a first location within the playable media file and a second location within the playable media file, the first and second locations together defining a clip of the playable media file occurring between the first and second locations;

[10c] causing, using the virtual clip, the clip to be displayed on a computing device of a user;

[10d] receiving a first user input associated with the virtual clip;

[10e] determining that the first user input conforms to a

predetermined format defining taxonomy tags;

[10f] identifying one or more taxonomy tags from the user input, wherein identifying the one or more taxonomy tags comprises determining the one or more taxonomy tags using the predetermined format comprising:

[10g] a tag identifier indicating that the user input includes taxonomy tags; and

[10h] one or more tags following the tag identifier and separated from each other by a delimiter, the one or more tags including a first tag and zero or more subtags arranged in sequence; and

[10i] associating, in an account of the user, the virtual clip with each of the one or more taxonomy tags identified from the user input.

Ex-1001, claim 10.

40. As seen above, the independent claims are directed to different methods related to annotated playable media files. Claim 1 is directed to a method for displaying information associated with a Playable Media File, the method displaying a GUI that includes (i) a first interactable graphical object to display a timeline and (ii) a second interactable graphical object to display a plurality of displayable virtual clips based on user interactions within the timeline of the first interactable graphical object. By contrast, independent claim 10 is directed to a method of annotating a playable media file, the method including (i) the

identification of one or more taxonomy tags from user input using a predetermined format, which includes a tag identifier and one or more tags separated by a delimiter, and (ii) associating the one or more taxonomy tags with a virtual clip.

41. I provide the following review of the prosecution history of the application leading to the '442 patent. I understand that the '442 patent issued on March 31, 2020, from an application filed on April 5, 2017. Ex-1001. I also understand that the '442 patent claims priority to a provisional application filed on July 20, 2016. Ex-1001.

42. I understand that during prosecution, a Restriction Requirement was issued on February 27, 2019. Ex-1004, 105-109. The Restriction Requirement determined originally filed claims 1-9 and 16-20, which cover the same subject matter as the issued claims, are “drawn to displaying information associated with playable media files.” Ex-1004, 107. I understand that in Applicant’s response filed on April 23, 2019, Applicant elected, without traverse, originally filed claims 1-9 and 16-20 for further examination. Ex-1004, 102. I understand that originally filed claims 1-9 and claims 16-20 led to issued claims 1-9 and claims 10-13, respectively, of the '442 patent. Ex-1004, 53-59.

43. I understand that during prosecution, an Office Action was mailed on August 1, 2019, which rejected claim 16 as anticipated by Petrov,¹ but allowed claims 1-9 and indicated claims 17-20 would be allowable if rewritten in independent form. Ex-1004, 86-91. I understand that in Applicant's response filed on October 30, 2019, Applicant addressed the rejection of claim 16 by amending claim 16 to incorporate the subject matter of claim 17. Ex-1004, 57-58, 60-61. I understand that Applicant did not dispute the limitations of originally filed claim 16 were disclosed in the prior art and, hence, were not novel. I understand that Applicant further amended claims 1, 3, and 5 to change the recitation of "one or more displayable virtual clips" to a "plurality of displayable virtual clips" and amended the abstract and the title. Ex-1004, 53-56.

44. I understand that a Notice of Allowance was mailed shortly thereafter on November 20, 2019. Ex-1004, 17-19. The Examiner concluded that the prior art disclosed generating annotation for video content, among other limitations. Ex-1004, 22. In the Examiner's reasons for allowance, I understand that the Examiner stated the prior art of record "does not teach or suggest or render obvious" the limitations of claim 1 directed to "identifying...a plurality of displayable virtual clips" and "updating the user interface...to display a list of the plurality of

¹ U.S. Patent No. 9,767,087.

displayable virtual clips.” Ex-1004, 22-23. The Examiner further stated the prior art of record “does not teach or suggest or render obvious” the limitations of claim 16 directed to “determining the one or more taxonomy tags using the predetermined format” where the predetermined format comprises “a tag identifier indicating that the user input includes taxonomy tags” and “one or more tags following the tag identifier and separated from each other by a delimiter” as well as “associating, in an account of the user, the virtual clip with each of the one or more taxonomy tags identified from the user input.” Ex-1004, 23.

VII. RELEVANT TIMEFRAME FOR DETERMINING OBVIOUSNESS

45. I understand that obviousness must be evaluated at the time the alleged invention was made. From the second page of the '442 patent (left column, field “[60] Related U.S. Application Data”), I understand that the first patent application (a provisional application) was filed on July 20, 2016. I will assume for the purpose of this declaration that July 20, 2016, is the date the purported invention was made, with the understanding that my testimony will also be applicable to some time prior to July 20, 2016, approximately in the 2015 to early 2016 timeframe. I may refer to this time period as the “relevant time frame”, and my testimony concerning obviousness is directed to this timeframe, even if I occasionally do not explicitly use a past tense.

VIII. LEVEL OF ORDINARY SKILL IN THE ART

46. I understand that the legal question of obviousness must be assessed from the perspective of a person having ordinary skill in the art (“PHOSITA”) at the time of the invention. In my opinion, the level of ordinary skill in the art was equal to an undergraduate degree in electrical engineering, computer engineering, or computer science and two years of computer programming experience and experience with web-based digital media. Relevant practical experience could offset less or different education and vice versa.

IX. DETAILED EXPLANATION OF THE REASONS FOR UNPATENTABILITY

Ground 1. Claims 1 and 7 are obvious over Junee in view of De Jong and further in view of Houh

47. In my opinion, independent claim 1 and dependent claim 7 would have been obvious over U.S. Pat. 8,826,117 B1 (“**Junee**”) (Ex-1005) in view of U.S. Pat. Pub. 2018/0190323 (“**De Jong**”) (Ex-1006) and further in view of U.S. Pat. 9,697,230 B2 (“**Houh**”) (Ex-1020).

48. Junee is a U.S. patent that issued on September 2, 2014. Ex-1005.

49. De Jong is a U.S. patent publication that published on July 5, 2018. Ex-1006. The application that led to the De Jong publication was filed on October 30, 2017, and is a national phase application under 35 U.S.C. § 371 claiming priority to International Application No. PCT/IB2016/000715 (Ex-1021) filed on

April 29, 2016. (Ex-1006). The disclosures in the PCT application (Ex-1021) and De Jong (Ex-1006) are substantially the same. (Ex-1022)

50. Houh is a U.S. patent that issued on July 4, 2017. Ex-1020. The application that led to the Houh patent was filed on March 31, 2006. Ex-2020. I am the first named inventor on the Houh patent.

51. I understand that Junea, De Jong, and Houh are each prior art.

52. I further understand that neither Junea nor De Jong nor Houh was cited during prosecution of the application leading to the '442 patent. Ex-1004.

A. Overview of the Ground

53. In my opinion, the challenged claims would have been obvious over the combination of Junea and De Jong. Junea discloses “web-based systems and methods for editing digital videos” including “[a] graphical editing interface allows designating one or more videos to assemble into a video compilation.” Ex-1005, Abstract. Junea discloses or renders obvious most of the elements of the challenged claims. Furthermore, to the extent Junea does not disclose or render obvious limitations directed to the overlay of interactable graphical objects onto a display window, user interactions with a timeline to select a time in a playable media file, the identification of a plurality of displayable virtual clips that includes virtual clips preceding and subsequent to a selected virtual clip, and the display of the

plurality of displayable virtual clips in a GUI, in my opinion, these limitations would have been obvious over Junee in view of De Jong and Houh.

54. De Jong teaches the overlay of a semi-transparent timeline and playback controls onto a display window, a user interactable timeline, identification of multiple highlight clips based on a selected time from the timeline, and the display of the highlight clips in a GUI. Ex-1006, ¶¶ 0232, 0379-0380, FIG. 13.

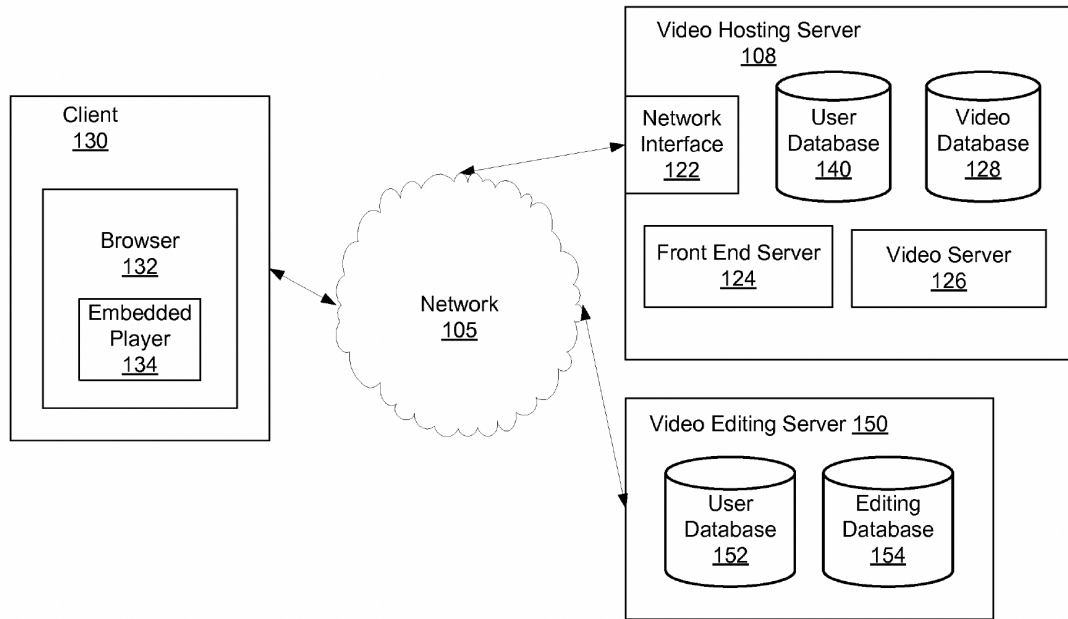
55. Houh further teaches the overlay of semi-transparent icons overlaid onto video content in a GUI. Ex-1020, 14:28-39.

56. In my opinion, dependent claim 7 would have been obvious over Junee, De Jong, and Houh, because De Jong teaches capturing a recording of live content using a camera, and the creation of highlight clips while the live content is captured.

1. Overview of Junee

57. Junee is directed to “web-based editing, composition, and **annotation of digital videos.**” Ex-1005, 1:5-6.² Junee’s system is illustrated in FIG. 1, reproduced below, which shows “a block diagram of [the] system architecture.” Ex-1005, 1:51.

² Emphases added throughout unless indicated otherwise.

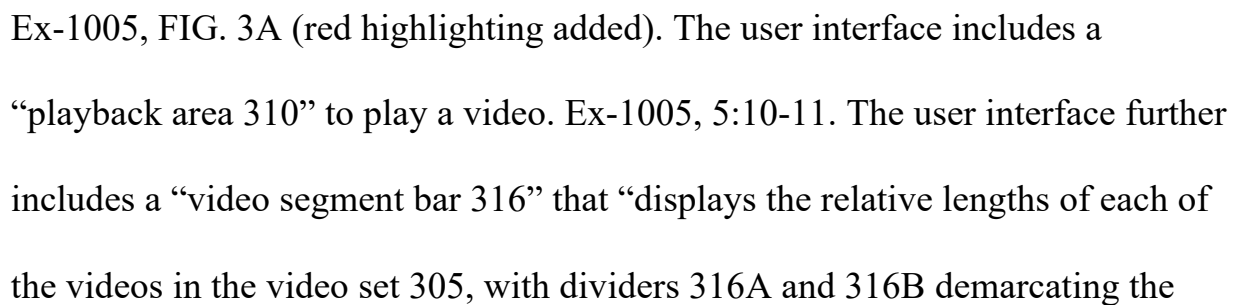


Ex-1005, FIG. 1. As shown in FIG. 1, Junee’s system includes a “client 130,” a “video hosting server 108,” and a “video editing server 150,” all connected together via a “network 105.” *Id.* The video hosting server 108 includes a “[v]ideo server 126 [that] receives uploaded media content from content providers and allows content to be viewed by [a] client 130” and a “video database 128” that stores the uploaded media content. Ex-1005, 2:55-57, 3:46. Examples of the video hosting server 108 include “the YouTube™ and Google Video™ websites.” Ex-1005, 2:15-16. “Users of clients 130 can also search for videos based on keywords, tags or other metadata” by submitting queries “to the video server 126, which is responsible for searching the video database 128 for videos that satisfy the user queries.” Ex-1005, 3:7-11.

58. The “video editing server 150 provides the ability to create compilations from, and add annotation to, videos in the video database 128.” Ex-1005, 3:51-53. “The video editing server 150 keeps a record of various user video editing actions, such as aggregating videos into a compilation, clipping videos to more restricted portions, annotating the videos with information about slides, people or events, adding popup visuals such as text boxes, and the like. It then stores these records within an editing database 154 in association with the user ID from the user database 152.” Ex-1005, 4:4-10.

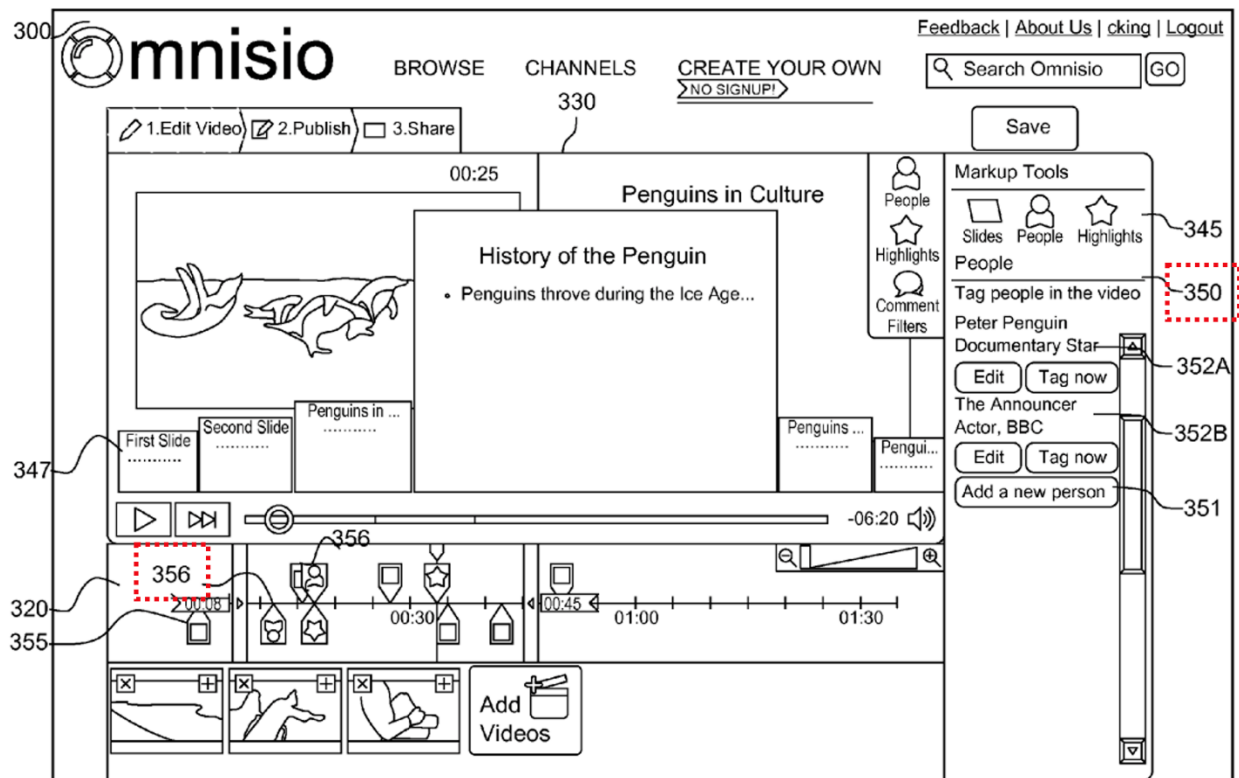
59. “The editing database 154 stores information on video compilations, which are ordered collections of videos treated as a single composite video, and may include a single video. For example, editing database 154 stores ... an identifier of the user who created of the compilation ... an identifier of each video to be played as part of the compilation, as well as an indicator (such as a time range) of the portion of each video to be played.” Ex-1005, 4:21-29. “The editing database also maintains an association between each annotation and the appropriate portion of the annotated video or video compilation.” Ex-1005, 4:31-34.

60. User interactions with Junee’s system are facilitated by “[a] graphical editing interface” that provides various functions, such as “designating one or more videos to assemble into a video compilation,” “specifying the portion of a



boundaries of the videos.” Ex-1005, 5:19-22. “Timeline bar 320 represents the video currently being played, and thus corresponds to one of the segments on the video segment bar 316.” Ex-1005, 5:31-34.

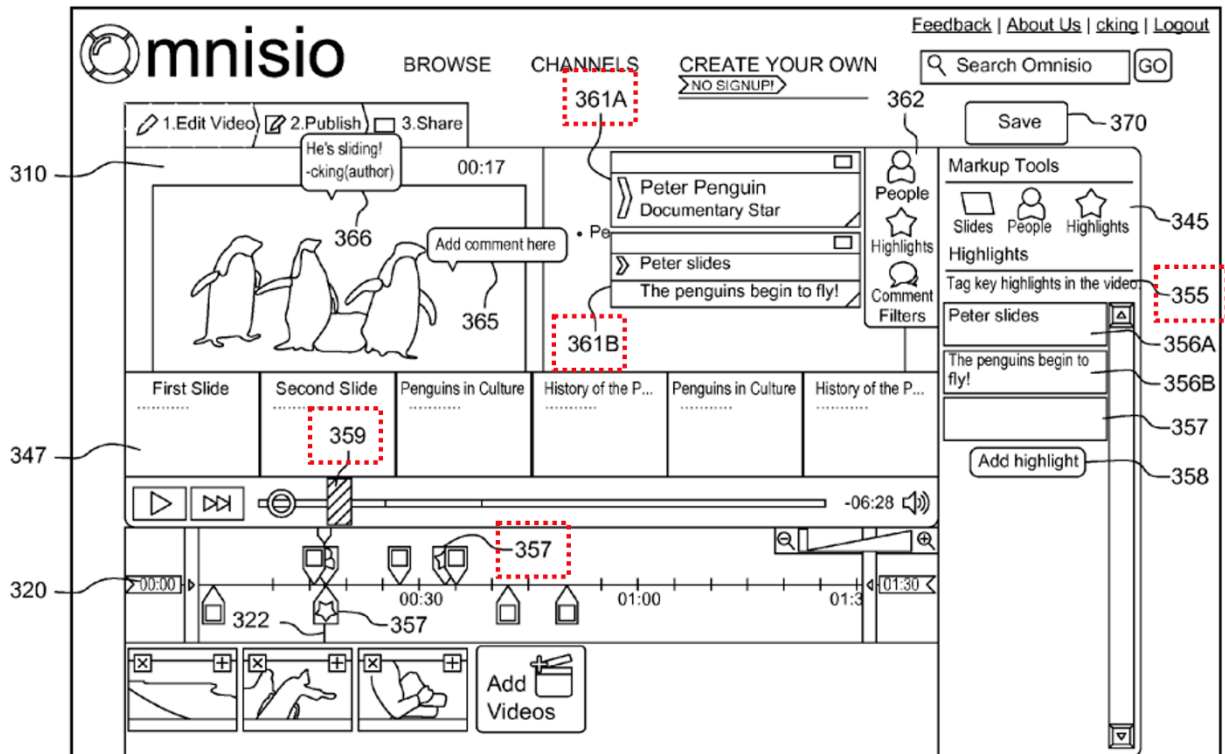
61. FIG. 3D of Junee, reproduced below, further shows a “user interface region 350 for defining annotation tools and adding annotation relating to people in the video.” Ex-1005, 6:54-56.



Ex-1005, FIG. 3D (red highlighting added). Junee’s user interface further allows the creation of a person tag 356, which is “an annotation for the person corresponding to that entry and associates it with the current time of the current

video, including displaying a person tag at the corresponding locations on the timeline 320.” Ex-1005, 6:65-7:1.

62. Additionally, FIG. 3E of Junee, reproduced below, shows the user interface includes “a user interface region 355 for adding annotation related to highlights in a video,” which include “a visual indication in the form of highlight annotation tag 357 added on the timeline 320 at a location corresponding to the current time of the video.” Ex-1005, 7:6-8, 9:43-45.



Ex-1005, FIG. 3E (red highlighting added). Junee explains that “[a]ny of the annotations displayed on the timeline 320—e.g., slide, people, and highlight tags—can be associated with different portions of the video compilation by dragging their

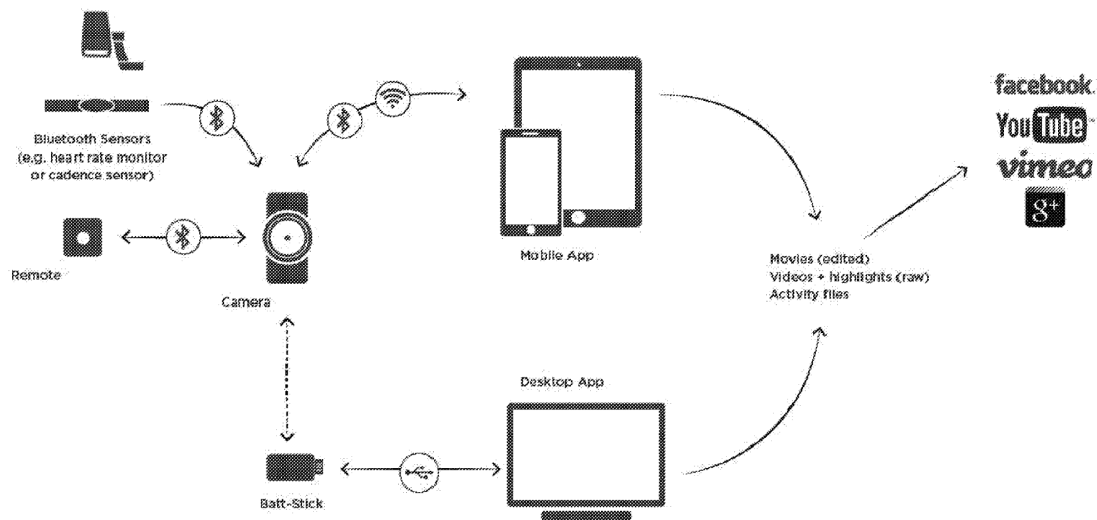
tag icons 355-357 to a different location on the timeline. Similarly, the annotations may be removed entirely by dragging their icons off the timeline, selecting them and pressing the ‘Delete’ key, or the like.” Ex-1005, 7:61-67. Additionally, Junea discloses, as an example, a “menu 361A contains an item for ‘Peter Penguin’, the only people annotation that has been added to the compilation. Each selection of the ‘Peter Penguin’ item causes video playback to begin at the next video portion (following the position indicated by the current time marker 322) previously associated with the corresponding people tag. Similarly, menu 361B contains a list of each highlight with which the video compilation has been tagged, the selection of any of which will cause playback of the video compilation at the corresponding time.” Ex-1005, 7:27-36.

63. The user interface of Junea further includes an “indicator 359... located on the video segment bar 316.” Ex-1005, 7:40-41. “The indicator 359 can appear in response to a variety of events, such as hovering the mouse pointer over a visual representation of the annotation...Moving the mouse pointer away from the visual representation of the annotation then causes the indicator 359 to disappear.” Ex-1005, 7:43-50.

2. Overview of De Jong

64. De Jong is directed to “the processing of video and sensor data recorded by a video camera.” Ex-1006, ¶0001. In particular, De Jong discloses “the

creation of highlight data based on sensor data,” which includes, in part, “one or more time periods in the video image data” also referred to as a highlight clip. Ex-1006, Abstract, ¶0113. According to De Jong, “[t]he highlight clips can then be manipulated by a user to create a ‘movie’, e.g. comprising several highlight clips.” Ex-1006, ¶0379. De Jong’s system is illustrated in FIG. 4, reproduced below, which shows “a schematic depiction of a video camera system.” Ex-1006, ¶0335.

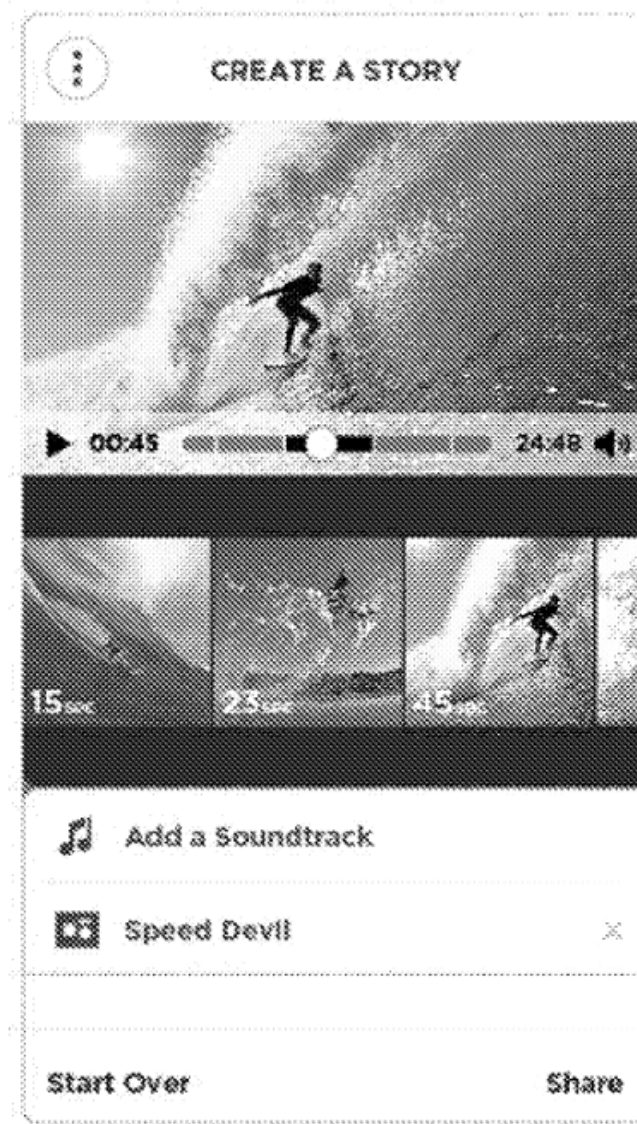


Ex-1006, FIG. 4. As shown, De Jong’s video camera system includes “a video camera having a number of inbuilt sensors including a GPS device.” Ex-1006, ¶0355. “The video camera is also arranged to be wirelessly connectable to an application (‘app’) running on a mobile computing device; the app can be used to control functions of the camera, play (or preview) videos, etc.” *Id.* De Jong explains that “highlights can be used to automatically create a media file with the most interesting parts of recorded video image data, and...can then be quickly

shared with other users via a video sharing platform, such as YouTube®.” Ex-1006, ¶0280.

65. De Jong teaches: “A highlight, and preferably each highlight, comprises one or more of the following information...(iv) a start time identifying the start of the highlight, e.g. as an offset from the beginning of the video image data; (v) an end time identifying the end of the highlight, e.g. as an offset from the beginning of the video image data; and (vi) additional information, which can be based on the type of tag or highlight, e.g. name of a user.” Ex-1006, ¶0115.

66. FIG. 19 of De Jong, reproduced below, further “shows an exemplary screen where the movie formed from the various highlights can be viewed by a user.” Ex-1006, ¶0347.



Ex-1006, FIG. 19. As shown in FIG. 19, “the movie formed from the various highlights can be viewed by a user, at the same time as the user being able to manipulate the highlights, e.g. by adding highlights, deleting highlights and reordering highlights,[sic] The timeline of the movie is shown divided into a plurality of segments, each segment relating to a different highlight. An icon is shown on the timeline indicating the position within the movie currently being

displayed. The relevant segment of the timeline is also shown differently from the other segments, such that the user can easily see which highlight is currently being viewed, and the relative length of the highlight in comparison to the other highlights forming the movie and the full movie itself.” Ex-1006, ¶0380.

67. De Jong teaches: “the representation of the highlight data displayed to the user preferably comprises a representation of one or more or all highlights in the one or more media files. This representation can be selected by the user, e.g. using a computer pointing device... to allow the highlight associated with the representation to be played, deleted and/or selected... for combining with other highlights to create a highlights video.” Ex-1006, ¶0178. “[T]he computing device can display a representation of the selected highlights in the ordered sequence. The representation can include a thumbnail image, which is preferably a frame from the video image data contained in the respective media file for the time window defined by the respective highlight.” Ex-1006, ¶0299.

3. Overview of Houh

68. Houh is directed to “methods and apparatus for generating and using enhanced metadata in search-driven applications.” Ex-1020, 1:17-19. Houh teaches: “a computerized method and apparatus for dynamic presentation of advertising, factual, informational content and combinations thereof (hereinafter referred to as “advertising content” generally) using shifting content in a

multimedia file to drive advertising placement. In particular embodiments, the advertising content is dynamically presented according to the playback of corresponding segments identified within a media file or stream.” Ex-2020, 13:58-66.

69. Houh further teaches “[t]he advertising content can be placed around the frame of the video, if viewing on a computer screen. The advertisements can be placed on the small screen of a portable device such as an iPod.” Ex-2020, 13:58-14:2. Specifically, “[t]he screen of the portable device can be used to display such advertising content. Semi-transparent icons can overlay a portion of video content to indicate the availability of additional content, such as advertisements, information or facts. User action while the icon is present, such as pressing a button on the portable device or another remote control may interrupt the current playback of the underlying media file/stream to switch to an advertisement...Advertising content in the form of text can overlay on top of video or scrolled at the bottom of the video.” Ex-2020, 14:26-39.

B. Rationale (Motivation) Supporting Obviousness

70. In my opinion, the combination of Junea, De Jong, and Houh renders obvious the elements of independent claim 1. The combination of Junea, De Jong, and Houh further renders obvious dependent claim 7 of the '442 patent, which

depends from claim 1 and is directed to capturing a recording of live content and the creation of highlight clips while the live content is captured.

71. Junee explains: “[c]onventional web-based systems permitting the storage and display of digital videos typically allow video playback and have some rudimentary tools for supplementing or altering the original video. These tools typically oblige a user to manually specify aspects such as time ranges during which a particular condition applies. Conventional systems also lack mechanisms for selecting video clips from a larger video and compositing the videos or video clips into a single compilation.” Ex-1005, 1:10-18.

72. Accordingly, Junee teaches a system that allows users to upload, view, and share media content, create video clips and compilations from the media content, and add annotations to the compilation. Ex-1005, 1:5-6, 2:49-57, 3:51-53, 4:4-10, 4:21-34, 8:9-10. Junee further teaches a graphical user interface (GUI) to facilitate the creation and editing of compilations, which includes various graphical interface objects, such as a playback area to view the media content, a video segment bar that “displays the relative lengths of each of the videos” and includes dividers between video clips, and thumbnail representations of videos included in the video compilation. Ex-1005, Abstract, 4:63-11, 5:19-22.

73. To the extent Junee alone does not disclose or render obvious an interactable graphical object that overlays a portion of a display window (claim

elements [1j] and [1k]), determining a selected time within a playable media file based on user input indicating user interaction with a display position on a timeline (claim elements [1l] and [1m]), identifying a plurality of displayable virtual clips (claim elements [1m], [1n], and [1o]), and updating a user interface to display a list of the plurality of displayable virtual clips in a second interactable graphical object (claim element [1q]), in my opinion, it would have been obvious for a PHOSITA to incorporate such features into a system such as Junee's in view of De Jong's and Houh's teachings.

Overlay of an interactable graphical object onto a portion of a display window [1j], [1k]

74. De Jong discloses in FIG. 13, reproduced below, a GUI with a **display window** that displays video content, and **an interactable graphical object** (the playback controls and the timeline bar indicated by the red arrow) overlaid onto a portion of the display window. Ex-1006, FIG. 13.



Id. (red highlighting added). As shown above, the playback controls and the timeline bar of De Jong are semitransparent such that a portion of the video content is visible behind the playback controls and the timeline bar.

75. Houh discloses in its “apparatus for presentation of...‘advertising content’,” “[s]emi-transparent icons can overlay a portion of video content to indicate the availability of additional content, such as advertisements, information or facts... Advertising content in the form of text can overlay on top of video or scrolled at the bottom of the video.” Ex-1020, 13:59-52, 14:28-39. Thus, Houh explicitly discloses overlaying an interactable graphical object (semi-transparent icons, text) onto a display window (GUI element displaying video content) of a GUI.

76. In my opinion, it would have been obvious for a PHOSITA to modify June’s GUI to overlay the playback control area 315 and the portion of June’s

GUI that displays thumbnail representations of the videos in the video set 305 onto the playback area 310 according to the teachings of De Jong and Houh because this would allow for a larger-sized playback area 310 to view video content and any annotations associated with the video content without sacrificing the playback controls provided by the playback control area 315 and the management of videos in the video set 305.

77. These modifications represent nothing more than the predictable combination of graphical interface objects taught by Junea, De Jong, and Houh. In my opinion, a PHOSITA would have been able to combine the elements taught by Junea and the overlay of interactable graphical objects onto a portion of a display window in elements [11] and [1m] taught by De Jong and Houh using known GUI design principles and standard computer programming, particularly since Junea already teaches user interactable graphical objects overlaid onto a display window. Ex-1005, 8:28-32 (“comments (such as the comment area 366 of FIG. 3E), the name of a people tag entry (e.g. ‘Peter Penguin’), and the title of an event (e.g. ‘Peter slides’) are displayed along with the video, such as **over top of the playing video**, or adjacent to it, for some given time period, e.g. 3 seconds.”). Furthermore, in my opinion, the respective functions of the elements taught by Junea, De Jong, and Houh would not change when combined together.

78. Further, in my opinion, these modifications represent the use of a known technique to improve Junee's system based on a similar system taught by De Jong and Houh in the same way. As explained above, Junee discloses a first interactable graphical object (the video segment bar 316), a second interactable graphical object (the portion of Junee's GUI that displays thumbnail representations of the videos 305A-305C), and a display window (the playback area 310). De Jong and Houh each teach a comparable system that includes semi-transparent graphical interface objects overlaid onto a display window. In my opinion, a PHOSITA could have modified Junee's computing system to incorporate the improvements in De Jong's and Houh's systems using nothing more than known GUI design principles and standard computer programming.

User interaction with a display position on a timeline [1l] and determining a selected time within a playable media corresponding to the display position [1m]

79. De Jong describes a video camera system that includes a computing device (e.g., a desktop computer, laptop, tablet, mobile phone) with a GUI displaying a timeline. Ex-1006, ¶¶ 0178, 0179, 0186. De Jong teaches "[t]he timeline preferably further includes an icon (or slider) that moves along the timeline as the video is played, so as to show the location along the timeline of the video image data currently being displayed. The icon (or slider) can **preferably be**

manipulated by the user, i.e. by moving the icon along the timeline, so as to allow the user to select the video image data being displayed.” Ex-1006, ¶0232.

80. To the extent Junee does not disclose or render obvious user interactions with a timeline and determination of a selected time based on said interaction, in my opinion, a PHOSITA would have been motivated to modify the GUI of Junee to (1) allow user interactions with a display position on a timeline and (2) determine a selected time in a media file based on the display position according to the teachings of De Jong. In my opinion, a PHOSITA would have recognized these features can help a user more easily navigate to different points of interest within a video.

81. These modifications represent nothing more than the predictable combination of graphical interface objects taught by Junee and De Jong. In my opinion, a PHOSITA would have been able to combine the elements taught by Junee and elements [11] and [1m] taught by De Jong using known GUI design principles and standard computer programming, particularly since Junee already teaches user interactable annotation tags displayed on a timeline and menus to facilitate playback of a video compilation at times corresponding to the annotation tags. Furthermore, in my opinion, the respective functions of the elements taught by Junee and De Jong would not change when combined together: Junee’s timeline

would still be able to display annotation tags, and users would be able to interact with the timeline according to De Jong's teachings.

82. Further, in my opinion, these modifications represent the use of a known technique to improve June's system based on a similar system taught by De Jong in the same way. As explained above, June discloses a computing system with a user interface that includes a timeline for a video compilation. De Jong teaches a comparable system that allows user interactions with a display position on a timeline and determines a selected time in a media file based on the display position. In my opinion, a PHOSITA could have modified June's computing system to incorporate the improvements in De Jong's system using nothing more than known GUI design principles and standard computer programming.

Identifying [1n] and displaying [1q] a list containing a selected video clip [1o] and video clips preceding and subsequent to the selected video clip [1p]

83. De Jong further discloses that the GUI can display multiple highlight clips corresponding to different moments in time including clips preceding and subsequent to a selected time. De Jong explains that with the GUI, "the movie formed from the various highlights can be viewed by a user, at the same time as the user being able to manipulate the highlights, e.g. by adding highlights, deleting highlights and reordering highlights,[sic] The timeline of the movie is shown

divided into a plurality of segments, each segment relating to a different highlight.”

Ex-1006, ¶0380.

84. To the extent Junee does not disclose or render obvious the identification and display of a list of virtual clips, in my opinion, a PHOSITA would have also been motivated to modify the GUI of Junee to identify a plurality of displayable virtual clips and update Junee’s GUI to display a list of the plurality of displayable virtual clips in the portion of Junee’s GUI where thumbnail representations are displayed according to the teachings of De Jong. In my opinion, a PHOSITA would have recognized that these features provide a mechanism to more easily select video clips to aid the creation of a video compilation pursuant to Junee’s teachings. As explained by De Jong, “[h]ighlight clips can...be used to give users a **quick and effortless** overview of the most interesting moments in the recordings they made.” Ex-1006, ¶0369. De Jong further explains “it is likely that a user will be interested in times of ‘extreme’ events.” Accordingly, the highlight clips allows users to “**quickly find and play these moments and/or share the moment with others.**” Ex-1006, ¶0120.

85. These modifications represent nothing more than the predictable combination and use of graphical interface objects taught by Junee and De Jong. In my opinion, a PHOSITA would have been able to combine the elements taught by Junee and elements [1n]-[1p] as taught by De Jong using known GUI design

principles and standard computer programming. Furthermore, the respective functions of the elements taught by Junee and De Jong would not change when combined together. In particular, the process of identifying a selected video clip and video clips preceding and subsequent to the selected video clip may be readily executed by Junee's computing system. Further, the display of the video clips may be readily integrated into Junee's user interface, particularly since Junee's user interface already displays thumbnail representations of videos used in a video compilation as "ordered representations." Ex-1005, 8:59-65.

86. Further, in my opinion, these modifications represent the use of a known technique to improve Junee's system based on a similar system taught by De Jong in the same way. As explained above, Junee discloses a computing system with a user interface that displays thumbnail representations of videos included in a video compilation. De Jong teaches a comparable system that identifies and displays a selected video clip and video clips preceding and subsequent to the selected video clip. In my opinion, a PHOSITA would have and could have modified Junee's computing system to incorporate the improvements in De Jong's system using nothing more than known GUI design principles and standard computer programming.

C. Reasonable Expectation of Success

87. In my opinion, a PHOSITA would have had a reasonable expectation of success in modifying Junee to identify and display a virtual clip at a selected time as well as virtual clips preceding and subsequent to the selected time according to De Jong's teachings. I am informed that "reasonable expectation of success" is required only for the limitations actually recited in the claims.

88. The art in the relevant timeframe was predictable. Specifically, in the relevant timeframe (no earlier than 2016), video compilations and GUIs like those disclosed in Junee and De Jong were well-known in the art.

89. In my opinion, the '442 patent's general and high-level disclosures support this understanding of the relevant art. *See, e.g.*, Ex-1001, 3:65-7:6, 7:38-8:5, 9:25-9:59.

D. Graham Factors

90. The **level of ordinary skill** in the art is addressed above under the "Level of Ordinary Skill in the Art" Section.

91. The **scope and content of the prior art** are discussed throughout my declaration.

92. The differences between the prior art and the claims are discussed throughout my declaration.

93. I am not aware of any **secondary considerations** that would make an inference of non-obviousness more likely.

94. This discussion of Graham factors applies to all Grounds.

E. Analogous Art

95. In my view, Junee, De Jong, and Houh are analogous art because they are in the same field as the '442 patent (relating to “assemblies, methods, devices, and systems for managing information” generally, which includes “sharing and annotating content,” Ex-1001, 1:18-22), and their methods would have been reasonably pertinent to the problems facing the named inventors, specifically “displaying information associated with a playable media file.” Ex-1005, Title, Abstract, 1:5-6; Ex-1006, Title, Abstract, ¶0001; Ex-1001, 1:26-28.

F. Claim Mapping

96. This section maps the challenged claims to the relevant disclosures of Junee, De Jong, and Houh, where the claim text appears in bold-italics, and the relevant mapping follows.

CLAIM 1

“1[a]. A method for displaying information associated with a Playable Media File, comprising:”

97. In my opinion, to the extent the preamble is limiting, Junee discloses or, at least renders obvious, **a method for displaying information associated with a Playable Media File**. Junee’s system is directed to “web-based editing,

composition, and **annotation of digital videos.**” Ex-1005, 1:5-6. Junee teaches “[a] graphical editing interface allows designating one or more videos to assemble into a **video compilation**...The graphical editing interface additionally allows the **association of annotations--specifying, for example, slides, people, and highlights--with portions of the video. The associated annotations alter the appearance of the video compilation when it is played**, such as by **displaying slides, or text associated with the annotations**, along with the video at times associated with the annotations.” Ex-1005, Abstract.

98. Junee teaches that video compilations are “ordered collections of videos treated as a single composite video, and may include a single video.” Ex-1005, 4:22-23. Junee further teaches that “[u]sers...can upload content to the video hosting server ...The uploaded content can include, for example, video, audio or a combination of video and audio.” Ex-1005, 3:15-18. I thus reference a video compilation as an example of a **playable media file** throughout, but note that Junee’s teachings apply broadly to different types of playable media files. Ex-1001, 3:65-4:1 (“Applicants’ system and method can be used by multiple members of a social network to associate annotations with a Playable Media File including a composite digital clip”).

“[1b] obtaining stored data describing the information,”

99. In my opinion, Junee discloses or, at least renders obvious, **obtaining stored data describing the information**, which comprises, *e.g.*, annotations. Junee teaches that a “video editing server 150 **keeps a record of various user video editing actions**, such as aggregating videos into a compilation, clipping videos to more restricted portions, **annotating the videos with information about slides, people or events, adding popup visuals such as text boxes**, and the like. It then **stores these records within an editing database 154** in association with the user ID from the user database 152. **The video editing server 150 also provides to entities such as the client 130** or the video hosting server 108, for a given video, **records of editing actions stored within the editing database 154 for that video.**” Ex-1005, 4:4-14. Junee further teaches that “[t]he editing database 154 **stores information on video compilations.**” Ex-1005, 4:21-22.

“[1c] the stored data comprising a storage location of the playable media file and”

100. In my opinion, Junee discloses, or at least renders obvious, the stored data comprises **a storage location of the playable media file**. Junee describes “[t]he video database 128 **stores video content** and associated metadata ... such as a video ID, artist, video title, label, genre, and time length.” Ex-1005, 3:46-50. Junee further discloses “[t]he video editing server 150 **keeps a record** of various

user video editing actions, such as **aggregating videos into a compilation.**” Ex-1005, 4:4-6. According to June, “[t]he editing database 154 **stores information on video compilations.**... For example, editing database 154 stores in association with, e.g., **a unique compilation identifier.**” Ex-1005, 4:21-25. Because June’s system stores and knows where to retrieve the video content and compilations, June’s system thus provides a storage location (the servers 128 and 150 and the database 154) to store video content and compilations along with respective, unique identifiers for the stored video content.

101. To the extent this limitation is not explicitly disclosed in June, in my opinion, a PHOSITA would have recognized the need for a file system to store video content and compilations at storage locations in the servers 128 and 150 and the database 154 for later retrieval. Conventional file systems typically included lookup tables containing identifiers for files on a storage volume. (*see* Ex-1015, 1:17-21 (“Computer file systems have traditionally involved some mechanism of data storage at a physical location (for instance, a disk sector) and a lookup table or index of some sort identifying the data and its physical address (or a logical address from which that may be inferred.”)). Thus, it would have been well within the knowledge and skill of a PHOSITA to implement a conventional file system in June to store a storage location of the playable media file.

“[1d] a plurality of virtual clips each associated with the playable

media file and including”

102. In my opinion, Junee discloses or, at least renders obvious, stored data comprising **a plurality of virtual clips each associated with the playable media file**. The '442 patent describes a “virtual clip” as “a set of data points that together delineate a particular subset of the content encoded in the media file”; “the virtual clip is comprised of **references** that identify specific content **in the corresponding media file**, but the virtual clip is not necessarily itself a stored media file containing the content data.” Ex-1001, 9:60-66.

103. As explained under claim element 1[a], Junee’s system allows users to “designat[e] **one or more videos** to assemble into **a video compilation**.” Ex-1005, Abstract. As explained under claim element [1b], a “video editing server 150 keeps a record of various user video editing actions, such as **aggregating videos into a compilation, clipping videos to more restricted portions**...It then stores these records within an editing database 154.” Ex-1005, 4:4-9. Junee further explains the “editing database 154 stores in association with, e.g., a unique compilation identifier...**an identifier of each video to be played as part of the compilation**, as well as **an indicator (such as a time range) of the portion of each video to be played**.” Ex-1005, 4:21-29. Thus, Junee discloses a plurality of virtual clips, i.e., **references** (an identifier, an indicator) that identify the content in a video, **associated with a playable media file** (a video compilation). *see* Ex-1019 at 410

(defining a pointer as a “variable that contains the memory location (address) of some data rather than the data itself”).

“[1e] a first data element identifying a first time within the playable media file at which a corresponding virtual clip begins, and”

104. In my opinion, Junea discloses or, at least renders obvious, this claim element. As shown in FIG. 3A, the GUI of Junea includes a “video segment bar 316 [that] displays **the relative lengths of each of the videos in the video set 305, with dividers 316A and 316B demarcating the boundaries of the videos.**” Ex-1005, 5:19-22. Thus, Junea discloses **a first data element** (the dividers 316A and 316B), which identifies **a first time** (playback location along the video segment bar 316) within **the playable media file** (video compilation) where **the virtual clip** (the videos 305A-305C in the video set 305) **begins**.

105. As explained under claim element [1b], the editing database 154 of Junea “**stores information on video compilations.**” Ex-1005, 4:21-22. This information includes “a total number of videos within the compilation, an identifier of each video to be played as part of the compilation, as well as **an indicator (such as a time range) of the portion of each video to be played and an indicator of the order of the video within the compilation (such as ‘1’ to indicate the first video in the compilation ordering).**” Ex-1005, 4:26-31.

106. As shown in FIG. 3A, the GUI of Junee includes a “video segment bar 316 [that] displays **the relative lengths of each of the videos in the video set 305, with dividers 316A and 316B demarcating the boundaries of the videos.**” Ex-1005, 5:19-22. The GUI further includes a timeline bar 320 for a particular video in the video compilation with “[d]rag bars 321A and 321B [that] indicate **start and end times**, respectively, for playback of the video. Thus, a user can drag the start drag bar 321A to some time after the beginning of the current video, and/or the end drag bar 321B to some time before the actual end of the video, and **subsequent playback of the video will then begin and end at corresponding times.**” Ex-1005, 5:41-47.

107. To the extent the identification of a start time within the playable media file for each video or video clip in the video compilation is not express or inherent in Junee, in my opinion, it would have been obvious for a PHOSITA to identify the start time based on Junee’s teachings. As explained under claim element [1b], the editing database 154 of Junee “**stores information on video compilations.**” Ex-1005, 4:21-22. This information includes “a total number of videos within the compilation, an identifier of each video to be played as part of the compilation, as well as **an indicator (such as a time range) of the portion of each video to be played and an indicator of the order of the video within the compilation (such as ‘1’ to indicate the first video in the compilation**

ordering).” Ex-1005, 4:26-31. In my opinion, a PHOSITA would have been readily able to determine the start time based on the start and end times of each video included in the video compilation and the order of the videos to be played in the video compilation as disclosed in Junee. The identification of a start time and the indicator of the order of the videos in the video compilation represent a finite number of predictable ways to ensure that the videos in the video compilation are played in the correct sequence.

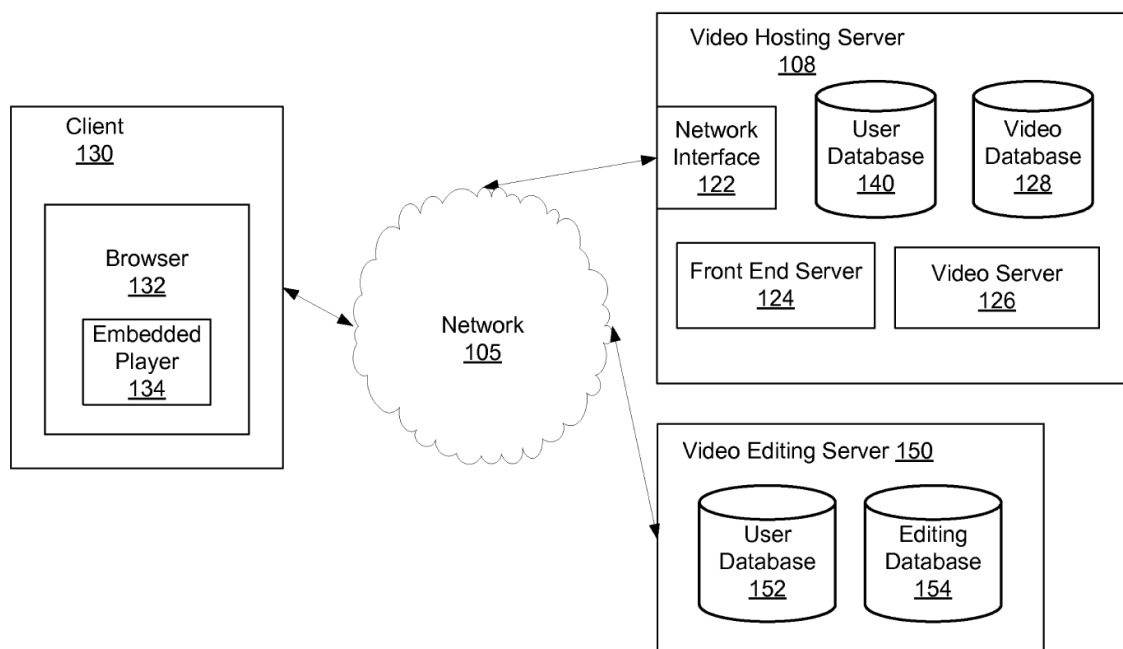
“[1f] a second data element identifying a first user profile associated with creating the corresponding virtual clip;”

108. In my opinion, Junee discloses **a second data element identifying a first user profile associated with creating the corresponding virtual clip.** Junee teaches that “[t]he video editing server 150 has a user database 152 that maintains a record of **all users** using the video editing system.” Ex-1005, 3:53-55. Junee further explains that “[t]he video editing server 150 keeps **a record of various user video editing actions**, such as **aggregating videos into a compilation, clipping videos to more restricted portions...** It then stores these records within an editing database 154 **in association with the user ID** from the user database 152.” Ex-1005, 4:4-10. As an example, Junee discloses the “editing database 154 stores in association with, e.g., a unique compilation identifier, **an identifier of the user who created of the compilation.**” Ex-1005, 4:24-26. Thus, Junee discloses a

second data element (the record stored in the editing database 154 of the video editing server 150) identifying a **first user profile** (user ID) associated with creating **the corresponding virtual clip** (the videos 305A-305C in the video set 305).

“[1g] accessing the playable media file at the storage location;”

109. In my opinion, June discloses **accessing the playable media file at the storage location**. In June’s system, users interface with the system through a client 130, which “executes a browser 132, and connects to the front end server 124 [of a video hosting server 108] via a network 105.” Ex-1005, 2:9-11, 2:30-31. The client 130 connects to the video editing server 150 via the network 105, as depicted in FIG. 1 below:



Ex-1005, FIG. 1.

110. As explained under claim element [1c], the video database 128, which is in the video hosting server 108, “stores video content and associated metadata.” Ex-1005, 3:47-48. Junee explains “[a] user can **access a video from the video hosting server 108.**” Ex-1005, 2:49-50. As an example, Junee discloses “in one embodiment media content is provided as a file download to a client 130; in an alternative embodiment, media content is streamed [to] client 130.” Ex-1005, 2:62-64.

111. As explained under claim element [1b], “[t]he editing database 154 stores information on video compilations.” Ex-1005, 4:21-22. Junee discloses “[t]he video editing server 150 also provides to entities such as **the client 130...**for a given video, **records of editing actions stored within the editing database 154 for that video.**” Ex-1005, 4:10-14. In Junee, video compilations “are ordered collections of videos treated as a single composite video, and may include a single video.” Ex-1005, 4:22-23. Thus, in Junee’s system, the client 130 **accesses a playable media file** (a video compilation) **at a storage location** (video content stored in the video database 128 and editing actions related to the creation of the video compilation stored in the editing database 154).

“[1h] causing a graphical user interface (GUI) to be displayed on a computing device of a user, wherein said GUI enables the user to generate user inputs by interacting with the GUI, and”

112. In my opinion, Junee discloses this claim element. Junee discloses in FIGS. 3A-3E “an example **user interface** [300] used for video editing,” which is referred to by Junee as a GUI. Ex-1005, 1:56-57, 4:62, 9:20-24 (“A **graphical user interface** is then displayed 520, the graphical user interface including the video, a timeline associated with the video, and annotation tools. For example, the **user interface 300** of FIG. 3D displays people annotation tools 352A and 352B.”); *see also* Ex-1017 at 22 (“A GUI is a type of computer human interface on a computer”). Junee’s user interface is displayed on a client device 130, which “may include a variety of different **computing devices**.” Ex-1005, 2:38-39. Thus, Junee discloses a **GUI (the user interface 300) displayed on a computing device of a user** (client device 130).

113. Junee’s description of FIGS. 3A-3E makes clear that the user interface is designed for users to **generate user inputs** through **interactions** with the GUI. Junee teaches a variety of user inputs facilitated by the GUI. Junee discloses the GUI includes a video set 305, which includes videos selected for inclusion in a video compilation. Ex-1005, 4:62-65. Each video in the video set 305 includes “a remove button 306A and a duplicate button 306B; **selection** of the former removes the video from the video set 305, and **selection** of the latter creates a new, separate instance of the video and adds it to the video set 305.” Ex-1005, 5:1-4. Junee also discloses an “Add Videos button 325” “to **select** additional videos” for inclusion in

a video compilation. Ex-1005, 5:5-7. Additional examples of user interactable graphical objects are discussed further below in connection with claim elements [1j] and [1k]. Junee's description of FIGS. 3A-3E makes clear that the user interface is designed for users to **generate user inputs** (at least addition and removal of videos from the video set 305) through **interactions** with the GUI (selection of the buttons 306A, 306B, or 325). Ex-1005, 41:56-57 ("FIGS. 3A-3E illustrate an example user interface used for video editing").

"[1i] the GUI comprises: a display window for displaying content encoded by the playable media file;"

114. In my opinion, Junee discloses **a display window for displaying content encoded by the playable media file**. In particular, Junee shows in FIG. 3A, reproduced below, a user interface 300 with a playback area 310 to display the video content of videos in a video compilation. Ex-1005, 5:7-11, FIG. 3A.

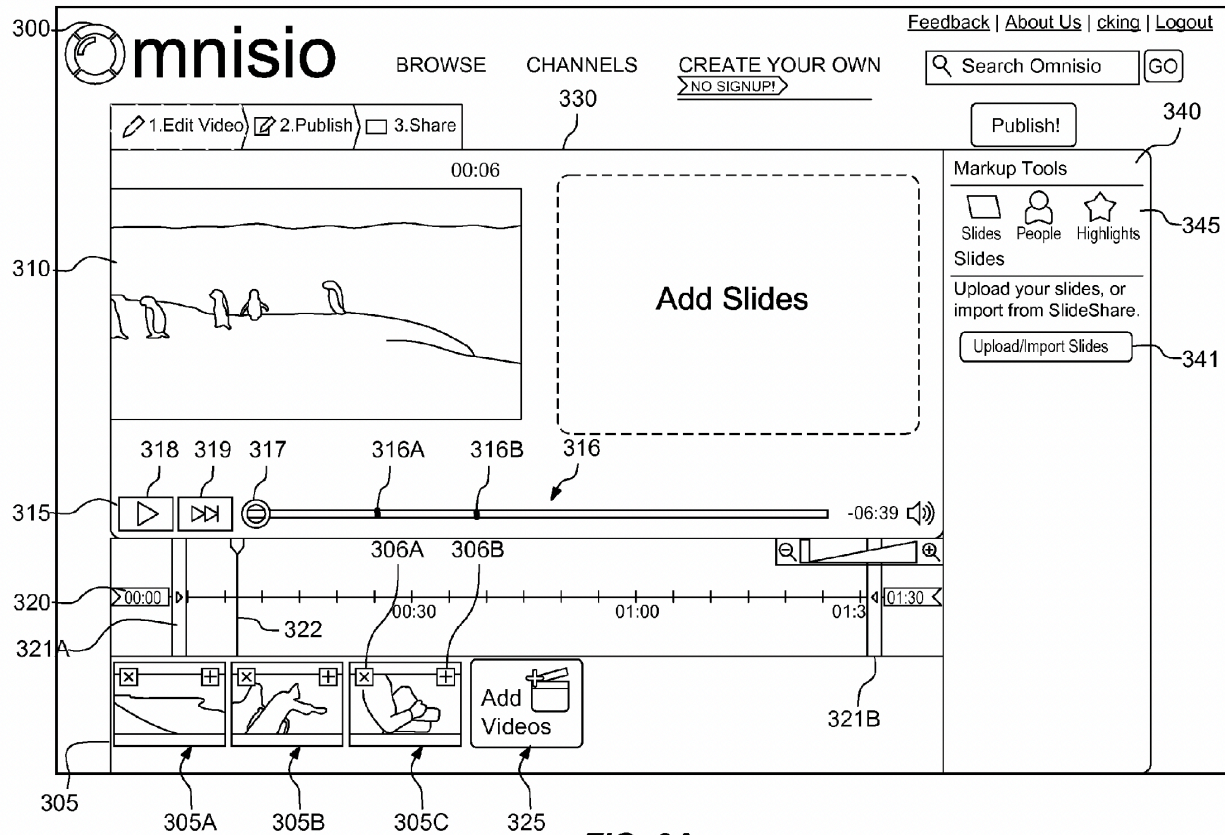


FIG. 3A

Ex-1005, FIG. 3A.

115. As seen above, **“the compilation includes three videos 305A-C.”** Ex-1005, 4:62–66. “One of the videos—video 305A in the example of FIG. 3A—... is shown in **playback area 310**. The controls in playback control area 315 both **control playback** and display playback position **within the set of selected videos.**” Ex-1005, 5:7-11. Junee further discloses “[t]he video segment bar 316 displays **the relative lengths of each of the videos in the video set 305**, with dividers 316A and 316B demarcating the boundaries of the videos... the position of the thumb marker 317 in FIG. 3A indicates that **the current playback location**

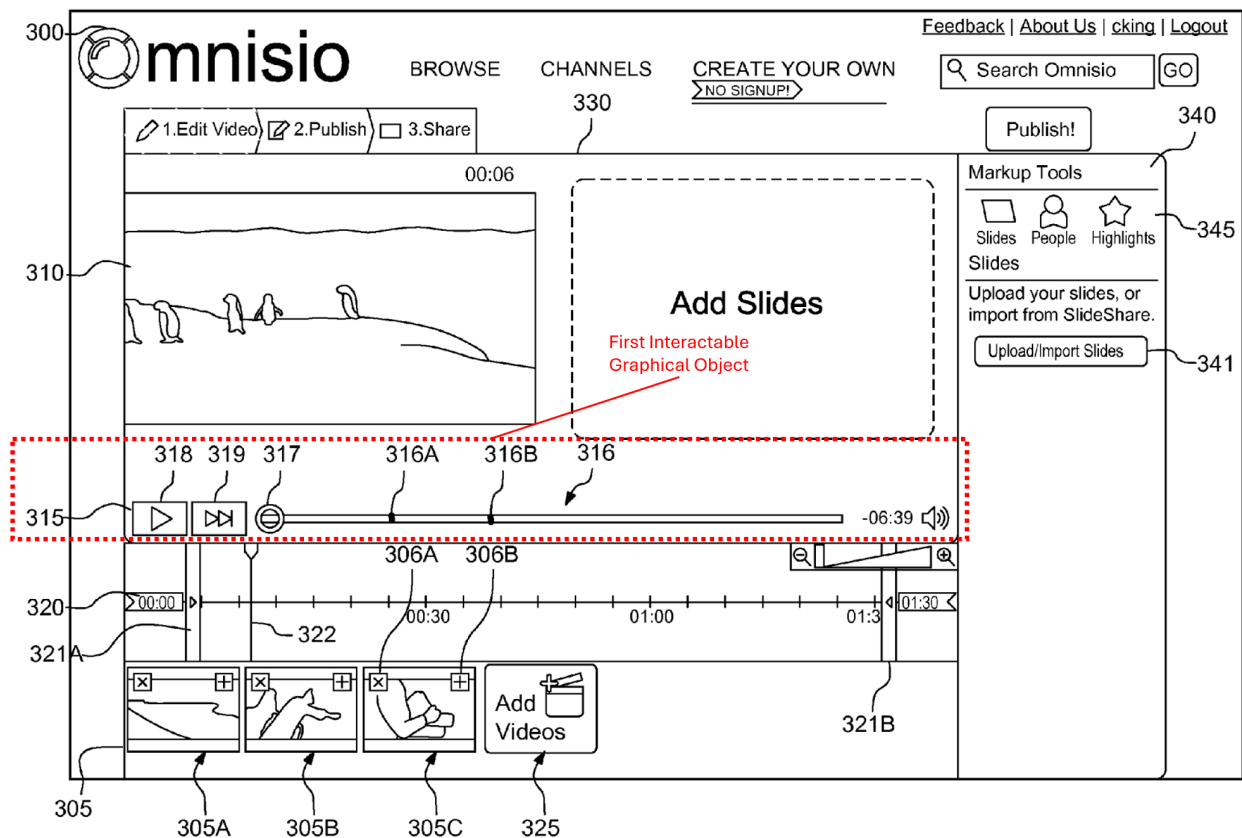
is near the beginning of the first of the three videos 305.” Ex-1005, 5:19-30.

June’s user interface 300 is designed to display content in the video compilation via the playback area 310 with controls (playback control area 315, video segment bar 316) affecting which video is displayed in the playback area 310. Thus, June’s user interface 300 comprises **a display window** (the playback area 310) **for displaying content encoded by the playable media file** (the video content in the video compilation).

*“[1j] a first interactable graphical object, wherein
the first interactable graphical object overlays a first portion of
the display window and
displays a timeline representing a duration of the playable media
file and
a plurality of clip indicators each associated with a corresponding
virtual clip of the plurality of virtual clips, each clip indicator
appearing on the timeline at a display position corresponding to the
first time identified by the first data element of the corresponding
virtual clip; and”*

116. In my opinion, June in view of De Jong and Houh renders obvious this claim element. June explains in connection with FIGS. 3A-3E that “controls in playback control area 315 both control playback and **display playback position within the set of selected videos.”** Ex-1005, 5:11-13. In particular, June describes the playback control area 315 includes a “**video segment bar 316** [that] displays the relative lengths of each of the videos in the video set 305.” Ex-1005, 5:19-21. The video segment bar 316 further includes a “thumb marker 317 [that]

displays the current playback location, including both the video that is being played and the location within that video.” Ex-1005, 5:25-27. Thus, Junee discloses a **first interactable graphical object** (the playback control area 315), which **displays a timeline** (the video segment bar 316) **representing a duration of the playable media file** (the relative lengths of each video in the video set 305). For reference, FIG. 3A of Junee is reproduced below with annotations to indicate the first interactable graphical object.



Ex-1005, FIG. 3A (red highlighting added).

117. In Junee, the GUI shown in FIGS. 3A-3E is for a “compilation [that] includes three videos 305A-C.” Ex-1005, 4:65-66. Junee explains the video

segment bar 316 includes “**dividers 316A and 316B demarcating the boundaries of the videos**. That is, dividers 316A and 316B divide the video segment bar 316 into three distinct portions, the length of each reflecting the relative lengths of the played portions of the videos 305.” Ex-1005, 5:22-25. The dividers 316A and 316B represent the start times for the videos 305B and 305C in Junee’s example video compilation. Ex-1005, FIG. 3A. Thus, Junee discloses **a plurality of clip indicators** (dividers 316A and 316B) **each associated with a corresponding virtual clip** (videos 305B and 305C) **of the plurality of virtual clips**, where each clip indicator **appears on the timeline** (the video segment bar 316) at a **display position corresponding to the first time of the corresponding virtual clip** (the start times for the videos 305B and 305C). A PHOSITA would readily appreciate a clip indicator for the first video in Junee’s example video compilation is not needed because the left end of the video segment bar 310 represents the start time for the first video.

118. De Jong discloses in FIG. 13, reproduced below, a GUI with **a display window** that displays video content, and **an interactable graphical object** (the playback controls and the timeline bar indicated by the red arrow) overlaid onto a portion of the display window. Ex-1006, FIG. 13.



Id. (red highlighting added). As shown above, the playback controls and the timeline bar of De Jong are semitransparent such that a portion of the video content is visible behind the playback controls and the timeline bar.

119. Houh discloses in its “apparatus for presentation of... ‘advertising content’,” “[s]emi-transparent icons can overlay a portion of video content to indicate the availability of additional content, such as advertisements, information or facts... Advertising content in the form of **text can overlay on top of video** or scrolled at the bottom of the video.” Ex-1020, 13:59-52, 14:28-39. Thus, Houh explicitly discloses overlaying **an interactable graphical object** (semi-transparent icons, text) onto **a display window** (GUI element displaying video content) of a GUI.

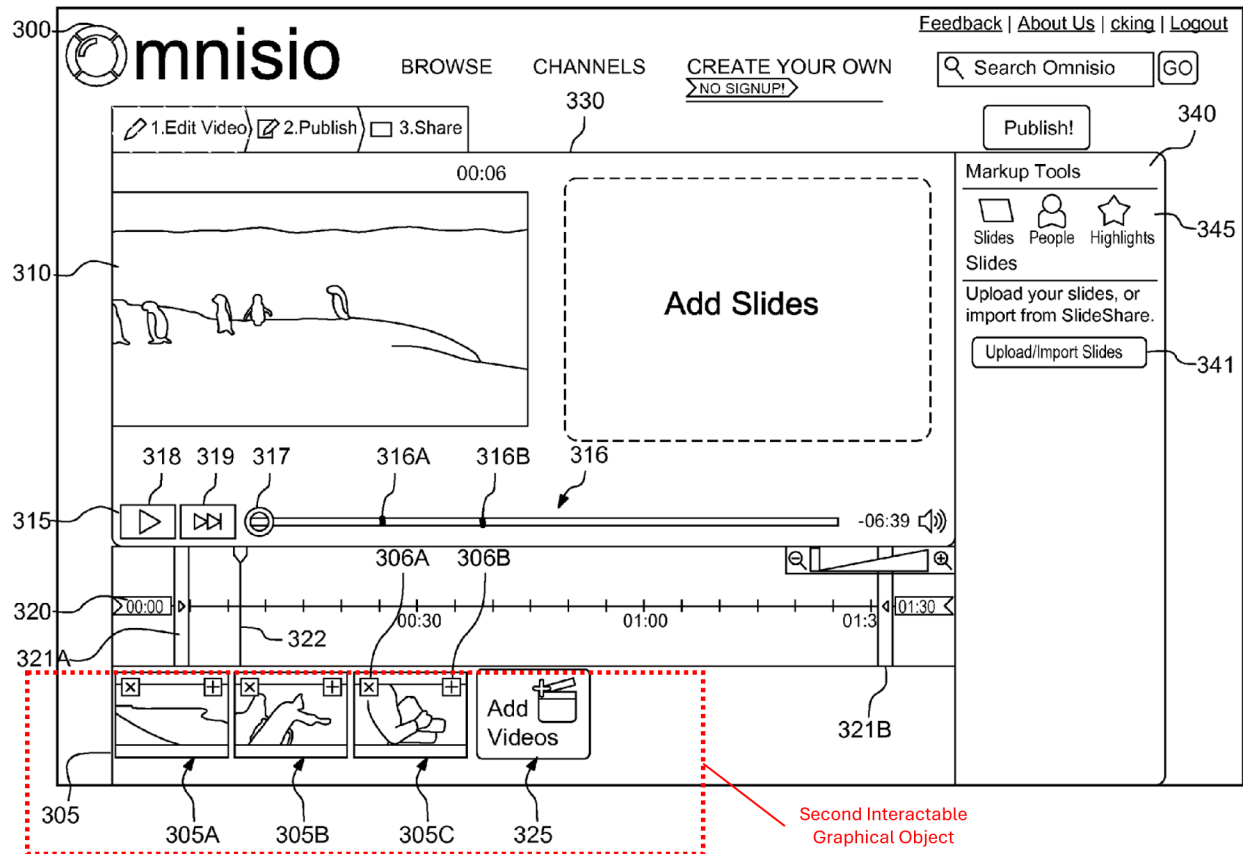
120. In my opinion, it would have been obvious for a PHOSITA to modify Junee’s GUI to overlay the playback control area 315 onto the playback area 310

according to the teachings of De Jong and Houh because this would allow for a larger-sized playback area 310 to view video content and any annotations associated with the video content without sacrificing the playback controls provided by the playback control area 315. *See* Ex-1016 at 430-431 (explaining that overlapping windows can be used to “take advantage of existing spatial management skills to push one or more windows to the rear, or to bring one or more forward”); *see also* Ex-1017 at 23-24 (noting that the amount of information a user can absorb from a GUI is limited and that the GUI “should display only what the user needs to perform the current operation”). Moreover, it would have been well within the knowledge and skill of a PHOSITA to overlay graphical interface objects in a GUI configured to display video content because this was common practice at the relevant time. *See* Ex-1018 at 590 (suggesting the use of an overlay for error messages), 401 (noting that secondary windows may use the same action controls as the primary window).

“[1k] a second interactable graphical object, wherein the second interactable graphical object overlays a second portion of the display window, is configured to display up to a first number of graphic elements each associated with a corresponding virtual clip of the plurality of virtual clips, and is initially not displayed in the GUI;”

121. In my opinion, Junea in view of De Jong and Houh renders obvious this claim element. Junea teaches in FIGS. 3A-3E a GUI that displays “videos selected via the interface of FIG. 2” for inclusion in a video compilation. Ex-1005,

4:63-65. Junee discloses “[i]n the example of FIG. 3A, the compilation includes three videos 305A-C, which are depicted using thumbnails representations of their content. Each video in the video set 305 has a remove button 306A and a duplicate button 306B; selection of the former removes the video from the video set 305, and selection of the latter creates a new, separate instance of the video and adds it to the video set 305...One of the videos—video 305A in the example of FIG. 3A—is visually distinguished, for example by use of a highlighted box or other characteristic to designate that it is the currently selected video, and is shown in playback area 310.” Ex-1005, 4:65-5:11. Thus, Junee teaches an example of **a second interactable graphical object** (the portion of Junee’s GUI that displays thumbnail representations of the videos 305A-305C) that is configured to display **a number of graphic elements** (the thumbnail representations) each associated with **a virtual clip of the plurality of virtual clips** (the videos 305A-305C). For reference, FIG. 3A of Junee is reproduced below with annotations to indicate the second interactable graphical object.



Ex-1005, FIG. 3A (red highlighting added).

122. As explained under claim element [1j], De Jong discloses a GUI with an interactable graphical object overlaid onto a portion of a display window. Ex-1006, FIG. 13. Houh further explicitly discloses the use of semi-transparent icons for overlay onto video content. Ex-1020, 14:28-39. In my opinion, it would have been obvious for a PHOSITA to modify Junee's GUI to further overlay the portion of Junee's GUI that displays thumbnail representations of the videos 305A-305C together with the playback control area 315 onto the playback area 310 according to the teachings of De Jong and Houh because this would allow for an even larger-

sized playback area 310 to view video content and any annotations associated with the video content without sacrificing the user's ability to interact with and manage the videos in the video compilation.

123. By overlaying both the playback control area 315 and the portion with the thumbnail representations of the video, a PHOSITA would have immediately recognized the need to control the visibility of these graphical interface objects in order to reduce potential obstructions to the playback area 310. Accordingly, in my opinion, it would have been obvious for a PHOSITA to modify Junee's GUI such that the portion of Junee's GUI that displays the thumbnail representations is **initially not displayed in the GUI** so that users can view video content without this portion obstructing the playback area 310.

124. Although not required in claim 1 of the '442 patent, in my opinion, a PHOSITA would have been motivated to further provide controls that dynamically alter the visibility of the portion with the thumbnail representations so that users could toggle off this portion to provide a less obstructed view of the playback area 310 and toggle on this portion to allow users to change the videos selected for the video compilation. Dynamic alterations to the visibility of graphical interface objects in a GUI would have been a routine and predictable practice. For instance, Junee discloses an "indicator 359... located on the video segment bar 316," which "can appear in response to a variety of events, such as hovering the mouse pointer

over a visual representation of the annotation, such as an entry for annotations in menus 361, a slide displayed in associated slides region 347, or one of the annotation tags on timeline 320. Moving the mouse pointer away from the visual representation of the annotation then causes the indicator 359 to disappear.” Ex-1005, 7:40-41, 7:43-50. Thus, Junea provides an example of a GUI where the visibility of certain graphical interface objects (the indicator 359) can be dynamically changed based on user interactions with the GUI.

125. In my opinion, it would also have been obvious for a PHOSITA to limit the number of graphic elements displayed in the second interactable graphical object (the portion of Junea’s GUI that displays the thumbnail representations) so that the graphic elements (the thumbnail representations of the videos 305A-305C) remain intelligible and easily interactable for a user. A PHOSITA would have readily appreciated the number of graphic elements that can be displayed in a GUI is limited by a variety of factors, such as the resolution and the size of a display screen.

“[11] receiving a first user input indicating a first interaction of the user with a first display position on the timeline;”

126. In my opinion, Junea alone or in view of De Jong renders obvious this claim element. Junea discloses the video segment bar 316 includes a “thumb marker 317,” which “displays the current playback location, including both the

video that is being played and the location within that video.” Ex-1005, 5:25-27.

Thus, Junee provides a graphical interface object (the thumb marker 317) that tracks **a display position** along **a timeline** (the video segment bar 316) as the video compilation is played.

127. To the extent user interactions with the video segment bar 316, *e.g.*, to select a particular playback location in a video, is not express or inherent in Junee, in my opinion, it would have been obvious for a PHOSITA to configure the thumb marker 317 to be user interactable according to Junee’s teachings because this provides users a convenient way to select a playback location within a video of the video compilation. Moreover, Junee teaches that “[a]ny of the annotations displayed on the timeline 320—*e.g.*, slide, people, and highlight tags—can be associated with different portions of the video compilation by dragging their tag icons 355-357 to a different location on the timeline.” Ex-1005, 7:61-64. Junee’s GUI further provides “[m]enus 361 [to] display the people and highlights annotations, respectively...For example, menu 361A contains an item for ‘Peter Penguin’...Each selection of the ‘Peter Penguin’ item causes video playback to begin at the next video portion (following the position indicated by the current time marker 322) previously associated with the corresponding people tag. Similarly, menu 361B contains a list of each highlight...the selection of any of which will cause playback of the video compilation at the corresponding time.” Ex-1005,

7:24-36. Thus, Junee already discloses graphical interface objects displayed on a timeline that are **user interactable** and control playback of a video. Accordingly, Junee's GUI supports user interactable graphical interface objects that indicate **a display position on a timeline** (the thumb marker 317) and **user interactions with graphical interface objects** (the tag icons 355-357, the menu items 361A and 361B) that facilitate selection of a time in a video to control playback of the video.

128. To the extent Junee does not disclose or render obvious this claim element, De Jong discloses in FIG. 19 “an exemplary screen where the movie formed from the various highlights can be viewed by a user,” which includes a “timeline of the movie...shown divided into a plurality of segments, each segment relating to a different highlight. An icon is shown on the timeline indicating the position within the movie currently being displayed.” Ex-1006, ¶0380. De Jong further teaches that “[t]he icon (or slider) [on a timeline] can **preferably be manipulated by the user**, i.e. by **moving the icon along the timeline**, so as to **allow the user to select the video image data being displayed**.” Ex-1006, ¶0232. Thus, De Jong teaches **receiving a first user input** (interactions with a user interactable icon displayed on a timeline) **indicating a first interaction of the user with a first display position on the timeline** (location of icon along the timeline).

129. As set forth above under the “Rationale (Motivation) to Combine” Section, a PHOSITA would have been motivated to modify the GUI of Junee to allow user interactions with a display position on a timeline according to De Jong’s teachings because a PHOSITA would have recognized these features can help a user more easily navigate to different points of interest within a video.

130. In my opinion, a PHOSITA would have had a reasonable expectation of success in modifying Junee’s GUI to allow user interactions with the display position on a timeline because Junee, as explained above, already contemplates graphical interface objects to facilitate user interactions with a timeline and to control playback of a video.

“[1m] determining a selected time within the playable media file that corresponds to the first display position;”

131. In my opinion, Junee alone or in view of De Jong renders obvious this claim element. As explained under claim element [11], Junee discloses the video segment bar 316 includes a “thumb marker 317,” which “displays the current playback location, including both the video that is being played and the location within that video.” Ex-1005, 5:25-27. As explained under claim element [1b], the editing database 154 of Junee **“stores information on video compilations”** including **“an indicator (such as a time range) of the portion of each video to be played and an indicator of the order of the video within the compilation**

(such as ‘1’ to indicate the first video in the compilation ordering).” Ex-1005, 4:21-22, 4:29-31. In my opinion, it would have been obvious to determine a selected time within the playable media file using **the selected display position on a timeline** (the current playback location), which graphically relates position with time, and information on the time range and order of the videos in the video compilation.

132. To the extent determination of a selected time based on a display position is not obvious in view of June, De Jong teaches “[t]he icon (or slider) [on a timeline] can **preferably be manipulated by the user, i.e. by moving the icon along the timeline**, so as to **allow the user to select the video image data being displayed.**” Ex-1006, ¶0232. De Jong teaches “[a] timeline (or playback bar) is preferably also displayed simultaneously with the video image data. One end of the timeline indicates the start of the video currently being played, with the other end of the timeline indicating the end of the video currently being played. Thus, the timeline can be said to form **a schematic representation of the video image data over the duration of the video.**” Ex-1006, ¶0232. In my opinion, it would have been obvious to determine a selected time based on a selected display position on a timeline, which graphically relates position with time, to facilitate selection of video image data corresponding to the selected time as taught by De Jong.

“[1n] identifying, as a plurality of displayable virtual clips:”

133. In my opinion, Junee discloses this claim element. As explained under claim element [1e], Junee discloses a “video segment bar 316 [that] displays **the relative lengths of each of the videos in the video set 305, with dividers 316A and 316B demarcating the boundaries of the videos.**” Ex-1005, 5:19-22.

Because displaying the relative lengths of each video on the video segment bar 316 and an appropriate number of dividers at respective boundaries of the videos requires knowing which videos are in the video compilation, Junee’s system provides identifying the plurality of displayable virtual clips.

134. Junee further teaches a GUI that displays a video set 305, which includes videos selected for inclusion in a video compilation. Ex-1005, 4:62-65. Junee discloses the videos in the video set 305 are “**selected** via the interface of FIG. 2,” which allows the user to “**specify[] the videos to be included in the compilation.**” Ex-1005, 44:46-47, 4:64-65. Junee further discloses that “[o]**rd**ered **representations of the videos** are then displayed.” Ex-1005, 8:59. Junee explains the ordered representations may be based on “an order specified by a user, such as the order in which the videos were added using the user interface of FIG. 2, or a new order resulting from the user dragging the thumbnail representations into a new order.” Ex-1005, 8:62-65. Thus, Junee discloses an example mechanism for **identifying a plurality of displayable virtual clips** (the videos 305A-305C) for display in an ordered manner in a GUI.

135. To the extent identifying a plurality of displayable virtual clips is not disclosed in Junee, in my opinion, Junee in view of De Jong renders obvious this claim element. De Jong teaches a video camera system where “highlight clips are generated by the video camera” or, “[a]lternatively, the computing device can generate the highlight clips itself after receiving the ‘raw’ data from the video camera.” Ex-1006, ¶0371. De Jong discloses “[t]he highlight clips can then be manipulated by a user to create a ‘movie’, e.g. comprising several highlight clips.” Ex-1006, ¶0379. De Jong discloses in FIG. 19 an exemplary screen where “**the movie formed from the various highlights can be viewed by a user**, at the same time as the user being able to manipulate the highlights, e.g. by adding highlights, deleting highlights and reordering highlights.” Ex-1006, ¶0380. In FIG. 19 of De Jong, the highlight clips are arranged in chronological order based on a selected time in a timeline for the movie and displayed beneath the timeline. Ex-1006, FIG. 19. Accordingly, De Jong discloses **identifying virtual clips** (the highlight clips) that span different portions of a timeline based on a user selected time.

136. As set forth above under the “Rationale (Motivation) to Combine” Section, in my opinion, a PHOSITA would have also been motivated to modify the GUI of Junee to identify video clips for display in the GUI according to the teachings of De Jong because a PHOSITA would have recognized the identification and display of highlight clips in chronological order provide a

mechanism to more easily select video clips to aid the creation of a video compilation pursuant to June's teachings. De Jong explains "[h]ighlight clips can...be used to give users a **quick and effortless** overview of the most interesting moments in the recordings they made." Ex-1006, ¶0369. For example, De Jong explains the highlight clips allows users to "**quickly find and play these moments [‘extreme’ events] and/or share the moment with others.**" Ex-1006, ¶0120. A PHOSITA would have had a reasonable expectation of success in modifying June's GUI because June already teaches displaying the videos 305A-305C according to an ordered representation.

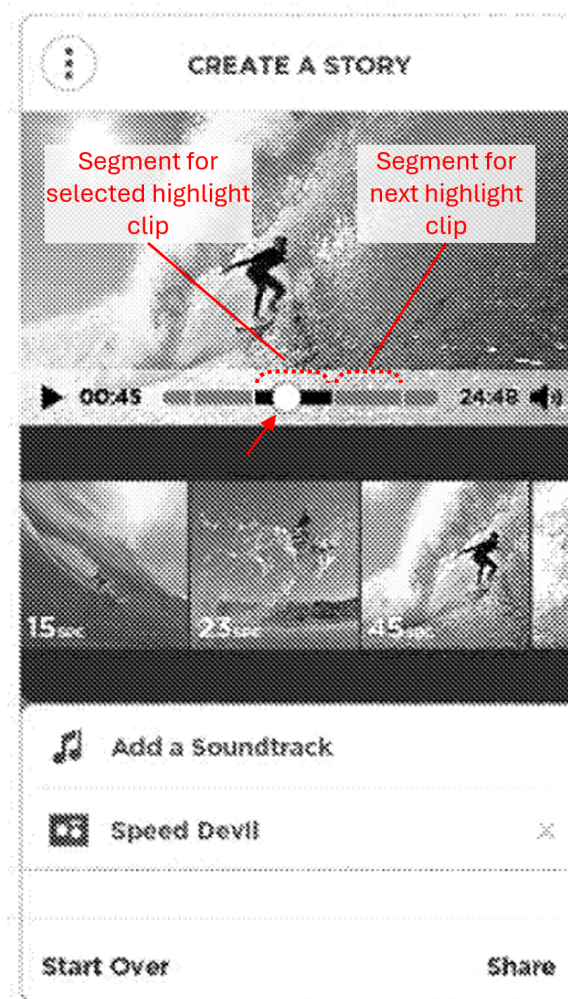
“[1o] a first virtual clip of the plurality of virtual clips, the corresponding first time of the first virtual clip being the closest, of the plurality of virtual clips, to the selected time; and”

137. In my opinion, June alone or in view of De Jong renders obvious this claim element. As explained under claim element [11], June discloses the video segment bar 316 includes a “thumb marker 317,” which “displays the current playback location, including both the video that is being played and the location within that video.” Ex-1005, 5:25-27. As explained under claim element [1o], June discloses “the video segment bar 316 displays **the relative lengths of each of the videos in the video set 305, with dividers 316A and 316B demarcating the boundaries of the videos.**” Ex-1005, 5:19-22.

138. In Junee, a video is a first virtual clip in accordance with this claim element when the following conditions are satisfied: (a) the display position of the thumb marker 317 overlaps the portion of the video segment bar 316 corresponding to the video; and (b) the display position is closer to the start time of the video than any other video. As shown in FIG. 3A of Junee, the video set 305 includes multiple videos (videos 305A-305C) such that for any display position on the video segment bar 316, there may be videos preceding and subsequent to the selected video. Ex-1005, FIG. 3A. Thus, even if the video that is being played does not qualify as a first virtual clip, the next video on the video segment bar 316 would qualify as a first virtual clip. Accordingly, Junee teaches **a first virtual clip** (the video currently being played, the next video) **of a plurality of virtual clips** (the videos) where **a first time of the virtual clip** (the start time of each video) is closest to the selected time.

139. To the extent June alone does not render this claim element obvious, De Jong teaches a GUI that includes a timeline and an icon where “[t]he icon (or slider) [on a timeline] can preferably be manipulated by the user, i.e. by moving the icon along the timeline, so as to allow the user **to select the video image data being displayed.**” Ex-1006, ¶0232. Thus, the display position of the icon in De Jong represents a selected time in De Jong’s movie.

140. De Jong further teaches “[t]he timeline of the movie [in FIG. 19] is shown divided into a plurality of segments, **each segment relating to a different highlight**. An icon is shown on the timeline indicating the position within the movie currently being displayed. The **relevant segment of the timeline is also shown differently** from the other segments, such that **the user can easily see which highlight is currently being viewed.**” Ex-1006, ¶0380. FIG. 19 of De Jong is reproduced below.



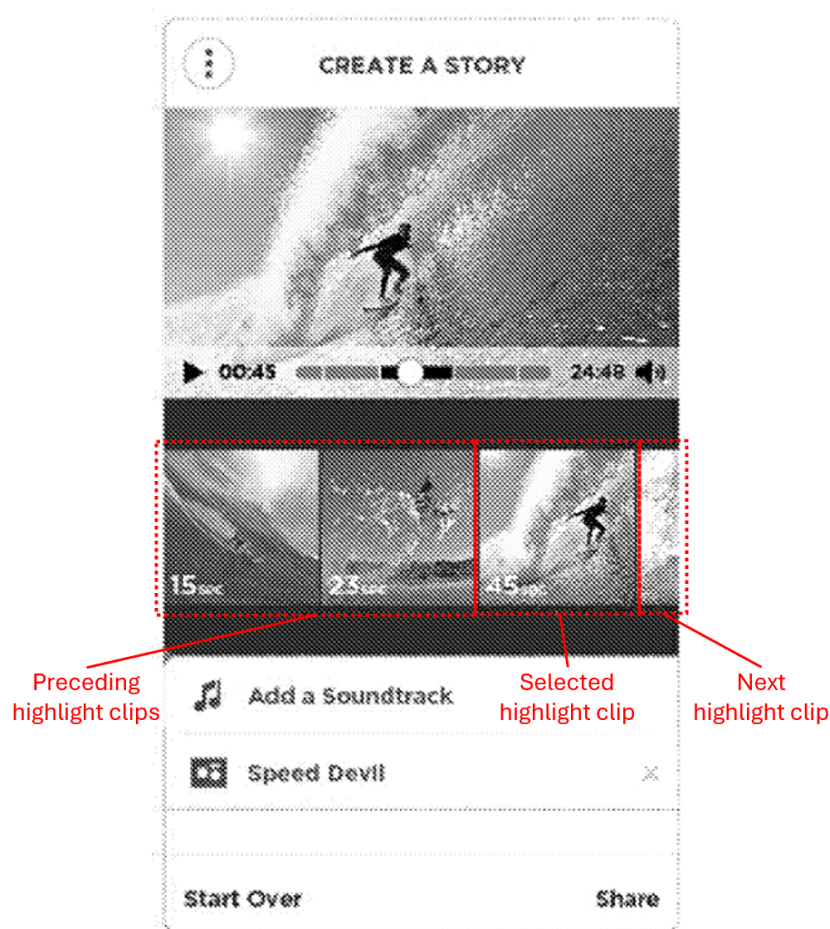
Ex-1006, FIG. 19 (red highlighting added). Thus, in De Jong, the display position of the icon on the timeline coincides with a segment corresponding to the selected highlight clip being viewed.

141. The selected highlight clip is a first virtual clip in accordance with this claim element when the following conditions are satisfied: (a) the display position of the icon overlaps the segment corresponding to the selected highlight clip; and (b) the display position is closer to the start time of the selected highlight clip than any other highlight clip. Figure 19 of De Jong also shows multiple highlight clips are identified and displayed preceding and subsequent to the selected time. Ex-1006, FIG. 19 (see selected time of 45 seconds and the display of multiple segments and thumbnails for highlight clips before and after 45 seconds); Thus, even if the selected highlight clip does not qualify as a first virtual clip, the next subsequent highlight clip, which is also identified in De Jong, would qualify as a first virtual clip. Accordingly, De Jong teaches **a first virtual clip** (the selected highlight clip, the highlight clip subsequent to the selected highlight clip) **of a plurality of virtual clips** (the multiple highlight clips) where **a first time of the virtual clip** (the beginning of each segment) is closest to the selected time.

“[1p] one or more of the virtual clips wherein the corresponding first time precedes and is approximate to the first time of the first virtual clip, and one or more of the virtual clips wherein the corresponding first time is subsequent and approximate to the first time of the first virtual clip, such that at most the first number of the

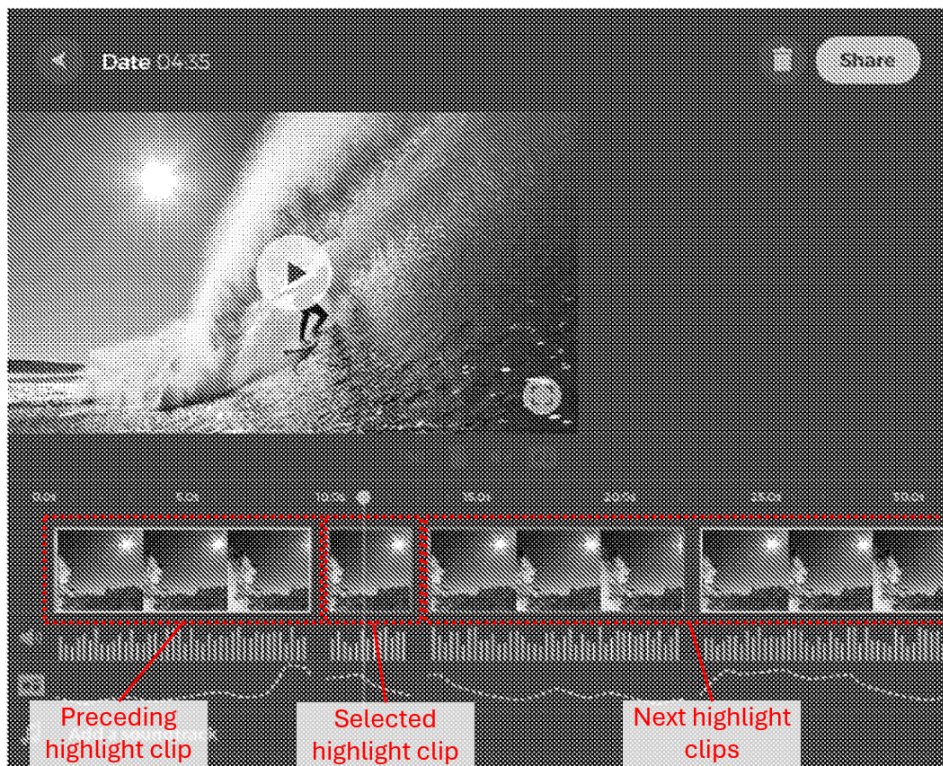
plurality of virtual clips are selected as the plurality of displayable virtual clips; and”

142. See claim elements [1n] and [1o]. As explained under claim element [1o], Figure 19 of De Jong, reproduced below, shows multiple highlight clips are identified and displayed preceding and subsequent to the selected time in the format of a highlight reel. Ex-1006, FIG. 19. As seen below, there are two highlight clips preceding and one highlight clip subsequent to the selected highlight clip at 45 seconds.



Ex-1006, FIG. 19 (red highlighting added). As seen above, there are two highlight clips preceding and one highlight clip subsequent to the selected highlight clip at 45 seconds.

143. De Jong discloses another example user interface in FIG. 21, reproduced below, which “includes a vertical line indicating the current position in the movie relative to a horizontal timeline, together with a series of thumbnails showing the individual highlights that make up the movie, each highlight being shown relative to the timeline, such that the user can see the start and end times of each highlight.” Ex-1006, ¶0381.



Ex-1006, FIG. 21 (red highlighting added). In the example shown in FIG. 21, there is one highlight clip preceding and two highlight clips subsequent to the selected highlight clip indicated by the vertical line. Ex-1006, FIG. 21.

144. Thus, De Jong discloses **identifying (i) one or more of the virtual clips wherein the corresponding first time precedes and is approximate to the first time of the first virtual clip** (the highlight clips preceding the selected highlight clip), and **(ii) one or more of the virtual clips wherein the corresponding first time is subsequent and approximate to the first time of the first virtual clip** (the highlight clips following the selected highlight clip). De Jong further discloses limiting the selection of highlight clips displayed in a user interface.

145. In my opinion, in view of the above, it would have been obvious for a PHOSITA to modify June's system to identify and display a limited the number of highlight clips for selection for the same reasons set forth under claim element [1k], i.e., to display highlight clips that are intelligible and easily interactable, and claim element [1n], i.e., to more easily select video clips to aid the creation of a video compilation pursuant to June's teachings.

“[1q] updating the user interface on the computing device to display a list of the plurality of displayable virtual clips in the second

interactable graphical object.”

146. In my opinion, Junee in view of De Jong renders obvious this claim element. As explained under claim element [1k], Junee discloses an example of a **second interactable graphical object** (the portion of Junee’s GUI that displays thumbnail representations of the videos 305A-305C), which displays a number of graphic elements (the thumbnail representations) each associated with a virtual clip of the plurality of virtual clips (the videos 305A-305C). Thus, a PHOSITA would have appreciated Junee’s user interface readily provides support for multiple interactable graphical objects to display graphical elements associated with corresponding virtual clips.

147. Accordingly, it would have been obvious for a PHOSITA to update Junee’s GUI to display the list of highlight clips in the user interface of De Jong because, as explained under claim element [1n], the display of multiple video clips in Junee would aid the creation of a video compilation pursuant to Junee’s teachings. In particular, Junee teaches that “[c]onventional systems also lack mechanisms for selecting video clips from a larger video and compositing the videos or video clips into a single compilation.” Ex-1005, 1:15-18. De Jong, for its part, teaches that “[h]ighlight clips can...be used to give users a **quick and effortless** overview of the most interesting moments in the recordings they made.” Ex-1006, ¶0369. For example, De Jong explains the highlight clips allows users to

“quickly find and play these moments [‘extreme’ events] and/or share the moment with others.” Ex-1006, ¶0120. Moreover, De Jong shows GUIs with multiple graphical objects overlaid on a display window, including a timeline overlay (FIG. 13) and other functional icons (tags, highlights, “my story,” back button) in FIGS. 12–13. Indeed, Junee already discloses all essential claimed GUI features—timeline (video segment bar 316), virtual clips (thumbnails 305A–C), and annotation menus (361A, 361B). Adding a discussion thread menu is consistent with Junee’s contemplated elements, and overlaying them on playback area 310 merely rearranges existing elements to enlarge the playback area.

148. A PHOSITA would have had a reasonable expectation of success because Junee already discloses the portion of Junee’s GUI that displays thumbnail representations of the videos 305A-305C can display **“[o]rdered representations of the videos.”** Ex-1005, 8:59. Junee explains the ordered representations may be based on “an order specified by a user, such as the order in which the videos were added using the user interface of FIG. 2, or a new order resulting from the user dragging the thumbnail representations into a new order.” Ex-1005, 8:62-65. Junee also discloses one of the thumbnail representations can be “visually distinguished, for example by use of a highlighted box or other characteristic to designate that it is the currently selected video, and is shown in playback area 310.” Ex-1005, 5:8-5:11. Thus, the modification would amount to displaying a list of videos as

thumbnail representations where the thumbnail representations are associated with the currently selected video and videos preceding and subsequent to the currently selected video. In some instances, the list of videos displayed in Juneé's GUI may be limited to a subset of the videos in the video compilation to ensure the thumbnail representations remain clear and intelligible. I also note that the claims themselves do not limit how big or small the graphical objects are when overlaid onto the display window. Moreover, the claims permit the visibility of the second and third interactable graphical objects to be changed.

CLAIM 7

“7[a]. The method of claim 1, wherein obtaining the stored data comprises: causing a recording device to begin capturing, as the playable media file, a recording of live content;”

149. See claim 1. Juneé in view of De Jong renders obvious this claim element. De Jong discloses a method of storing data identifying one or more times of interest in video image data collected by a digital video camera where the method begins with “receiving an **first input to cause the camera to start recording.**” Ex-1006, ¶0097. De Jong's method further includes “**storing highlight data** identifying one or more times of interest in the video image data” and thereafter receiving “**a second input to cause the camera to stop recording.**” Ex-1006, ¶0100. Thus, De Jong discloses **causing a recording device** (digital

video camera) to begin capturing **a recording of live content** (recording from the digital video camera) in response to a first input.

150. In my opinion, it would have been obvious for a PHOSITA to modify Junee's system to give users the option to capture and upload a video recording of live content according to De Jong's teachings because this would enable users to create and edit a video compilation that includes live content as the live content is broadcast in real time. Accordingly, it would have also been obvious for a PHOSITA to modify Junee's system to include a recording device, such as the digital video camera of De Jong, to facilitate the capture of the video recording, which can then be uploaded to the video hosting server 108 of Junee. Ex-1005, 3:15-18 ("[u]sers of the clients 130 and browser 132 can upload content to the video hosting server 108 via network 105. The uploaded content can include, for example, video, audio or a combination of video and audio."), 2:55-57 ("[v]ideo server 126 [of the video hosting server 108] receives uploaded media content from content providers and allows content to be viewed by client 130."). For example, users could use Junee's modified system to generate instant replays with annotations during a live broadcast, which can then be shared with other users of a social networking site. In another example, Junee's modified system could be used by users to facilitate interactions with other users during a livestream through the

creation and publication of video compilations based on content from the livestream.

151. Moreover, in my opinion, a PHOSITA would have had a reasonable expectation of success in modifying Junee’s system because Junee already teaches a video compilation may be created from a single video, which in this case would be a recording of the live content. Ex-1005, 4:22-23 (“ordered collections of videos treated as a single composite video, and may include a single video.”).

“[7b] while the recording device is capturing the live content, receiving a second user input at a first time, the recording device continuing to capture the live content subsequent to the second user input; while the recording device is capturing the live content, receiving on the user interface a third user input at a second time; and creating a first virtual clip of the plurality of virtual clips, the first virtual clip having the first time as the corresponding first position and the second time as the corresponding second position.”

152. In my opinion, Junee in view of De Jong renders obvious these claim elements. De Jong discloses “[t]he highlight data identifies **one or more times of interest** in the video image data, e.g. for use by the user when editing the video image data. The highlight data can comprise...one or more time periods in the video image data...**each time period is referred to herein as a ‘highlight’.**” Ex-1006, ¶0113. De Jong further describes “[a] highlight...comprises one or more of the following information:...**(iv) a start time identifying the start of the highlight**, e.g. as an offset from the beginning of the video image data; **(v) an end**

time identifying the end of the highlight, e.g. as an offset from the beginning of the video image data.” Ex-1006, ¶0115.

153. De Jong further discloses that “[t]he one or more times of interest in the video image data can be determined based on received data **indicative of a manual input, i.e. a user input.**” Ex-1006, ¶0117. For instance, De Jong describes a scenario where “**a first user, who may be performing an activity and is carrying or is associated with the camera, can actuate a user input on the camera or a remote device**, such as one worn on the user's wrist, while **one or more second users, who may be watching the first user performing the activity, can actuate a user input on their own remote control devices** (and that are in wireless communication with the camera). De Jong further discloses “data indicative of such manual inputs will be received **during the recording of the video image data.**” Ex-1006, ¶0119.

154. Thus, De Jong teaches while **the recording device** (digital video camera) is **capturing a live event** (recording from the digital video camera), a user can provide user inputs manually to **create a virtual clip** (a highlight) where the virtual clip has **a first time** (a start time of a highlight) and **a second time** (an end time of the highlight). The manually created highlights in De Jong are disclosed as having a start time and an end time, confirming that De Jong’s system discloses

receipt of separate manual user inputs defining a start time and an end time for the highlight.

155. I note the recitations of a “first position” corresponding to the first time and a “second position” corresponding to the second time lack antecedent basis. I further note the only other recitation of the term “position” is in claim 1, which recites a “display position” on a timeline that corresponds to a user selected time. Ex-1001, cl. 1. Thus, I herein assume, while not conceding that the claim language is not indefinite, the first position and the second position in this claim element refer to the start and the end, respectively, of the first virtual clip.

156. As explained under claim element [1o], De Jong teaches “[t]he timeline of the movie [in FIG. 19] is shown divided into a plurality of segments, **each segment relating to a different highlight.**” Ex-1006, ¶0380. Thus, De Jong discloses highlights having a first time corresponding to **a first display position** (the beginning of the segment) and a second time corresponding to **a second display position** (the end of the segment). Ex-1006, ¶0380, FIG. 19.

157. In my opinion, it would have been obvious for a PHOSITA to modify Juneo’s system to support the creation of video content for a video compilation via manual user inputs while capturing a recording of a live content as taught by De Jong because this would allow users to rapidly create videos, such as instant

replays, in real time without sacrificing control over the portions of the live content included in the video.

Ground 2. Claims 5 and 6 are obvious over June, De Jong, Houh, and further in view of Lee

158. Claim 5 depends from claim 1. Claim 6 depends from claim 5.

159. In my opinion, claims 5 and 6 would have been obvious over **June, De Jong, and Houh** as discussed under Ground 1 and incorporated herein, further in view of US Pat. Pub. 2015/0180820 (“**Lee**”) (Ex-1011).

160. Lee is a U.S. patent publication that published on June 25, 2015. Ex-1011. The application that led to the Lee publication was filed on March 4, 2015. Ex-1011.

161. I understand that Lee is prior art.

162. I further understand that Lee was not cited during prosecution of any application leading to the ’442 patent.

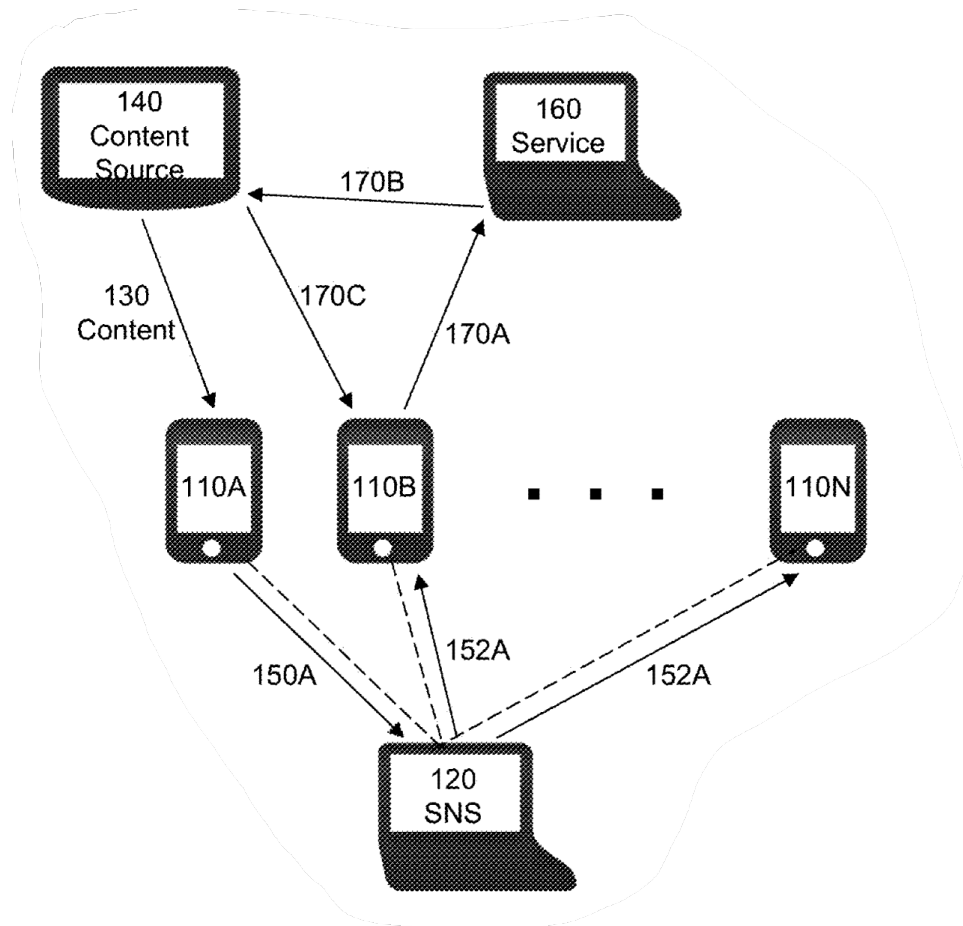
A. Overview of the Ground

163. In my opinion, claims 5 and 6 would have been obvious over the combination of June, De Jong, Houh, and further in view of Lee. To the extent the combination of June, De Jong, and Houh does not disclose or render obvious the limitations of claim 5 directed to the display of a discussion thread, these limitations would have been obvious over Lee. Lee teaches a user interface that

displays a discussion thread associated with video content. In my opinion, dependent claim 6 would have been obvious over Junee, De Jong, Houh, and further in view of Lee because Junee teaches user interactable thumbnail representations of videos and GUI elements (icons in the icon set 362) that can toggle the display of interactable graphical objects.

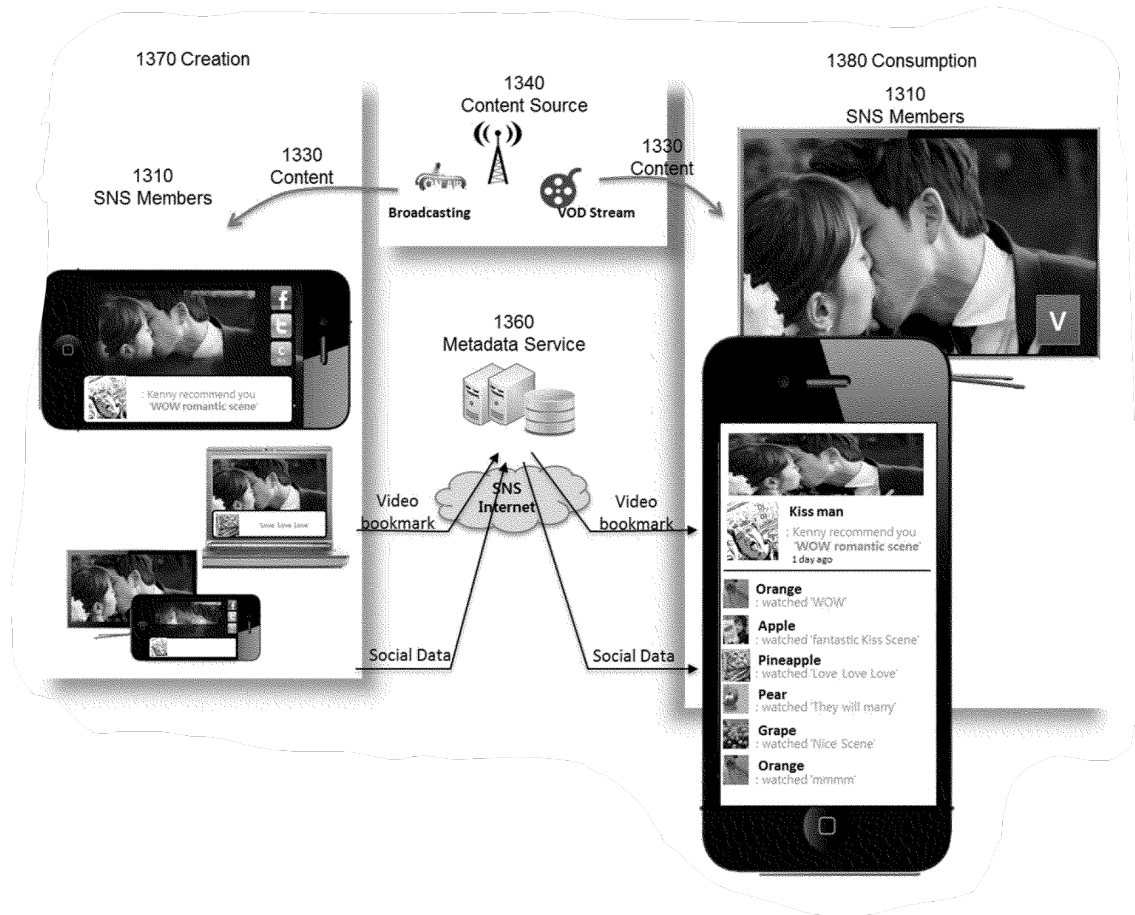
1. Overview of Lee

164. Lee is directed to “sharing personal information about multimedia content, including for example using other people's comments about videos as a basis for a personalized program guide to videos.” Ex-1011, ¶0003. In Lee’s system, “[m]embers of a social network service may share real-time comments about multimedia content they are viewing. The comments are attached to bookmarks that allow other members to view the video and add to the discussion thread.” Ex-1011, Abstract. Lee’s system is illustrated in FIG. 1, reproduced below, which shows “a block diagram of [the] system.” Ex-1011, ¶0011.



Ex-1011, FIG. 1. As shown, Lee discloses that “a member 110A of a social network service (SNS 120) is viewing multimedia content 130 from source 140...While the member 110A is viewing the content 130, he makes comments 150A about the content 130...These comments 150A are shared with other members 110B-N.” Ex-1011, ¶0027. Lee explains that “the comments 150A are attached to bookmarks that mark corresponding time points in the content 130...The comments and bookmarks (collectively, referred to as a **discussion thread 152A**) are forwarded to other members 110B-N.” Ex-1011, ¶0028.

165. FIG. 13 of Lee, reproduced below, shows an example discussion thread. Ex-1011, ¶0057.



Ex-1011, FIG. 13. “In the example of FIG. 13, the ‘V’ in the screen indicates there is a discussion thread on this scene, and **the discussion thread appears in the SNS member's separate device**. Other implementations will be apparent. For example, **the discussion thread could appear within the playback screen**. Or it could be **toggled on and off**, or retrieved only upon request.” Ex-1011, ¶0057.

B. Rationale (Motivation) Supporting Obviousness

166. In my opinion, the combination of Junee, De Jong, and Houh renders obvious claim element 5[a]. As explained under Ground 1, claim elements [1j] and [1k], Junee discloses a first interactable graphical object (the playback control area 315) and a second interactable graphical object (the portion of Junee's GUI that displays thumbnail representations of the videos 305A-305C). Junee's GUI further includes interactable graphical objects that each have a visible state and are configured to display annotations associated with a particular video in the video compilation. Ex-1005, 7:21-27 (menus 361A and 361B), 8:27-32 (comment area 366). Thus, Junee discloses examples of **a third interactable graphical object** (e.g., the menus 361A and 361B) having **a visible state controlling whether the third interactable graphical object is visible or hidden** (toggling the visibility of the menus 361A and 361B using the icons in the icon set 362). In my opinion, it would have been obvious for a PHOSITA to add, for example, an additional menu to the menus 361A and 361B to display other information associated with a video in the video compilation, such as a discussion thread as discussed under claim element [5b].

167. For claim element [5b], Junee further discloses that, when publishing a video compilation, "[r]elated information, such as...**whether users other than the author can provide comments**, can also be specified as part of publishing,

e.g. as depicted in the sample user interface of FIG. 3F.” Ex-1005, 8:2-9. Junee also discloses “[d]uring video playback, either while the video is being edited, or **during post-publication viewing**, any previously-specified annotations are displayed and can be used to alter the flow of control of the video compilation” via the GUI of at least FIGS. 3D and 3E. Ex-1005, 8:11-32. Thus, Junee teaches users can add comments to a published video compilation and the author of a video compilation can use Junee’s GUI to view annotations associated with the videos in the video compilation after publication.

168. Lee discloses a system where “[m]embers of a social network service may share real-time comments about multimedia content they are viewing.” Ex-1011, Abstract. Lee explains the comments “are attached to bookmarks that mark corresponding time points in the content 130...**The comments and bookmarks (collectively, referred to as a discussion thread 152A)** are forwarded to other members 110B-N.” Ex-1011, ¶0028. Thus, Lee discloses displaying **a discussion thread** (the discussion thread 152A) comprising **annotations** (comments and bookmarks) associated with **a selected virtual clip of the plurality of displayable virtual clips** (multimedia content).

169. In my opinion, it would have been obvious for a PHOSITA to modify Junee’s GUI to include a third interactable graphical object to display a discussion thread comprising annotations (comments and bookmarks) associated with a

selected virtual clip of the plurality of virtual clips according to Lee's teachings. For example, the discussion thread can be implemented as an additional menu to the menus 361A and 361B. This modification would facilitate the display of comments by different users in a format that can be easily viewed by the author when, for example, the author views the video compilation post-publication. Additionally, by displaying different discussion threads for each video in a video compilation, authors can more quickly assess user reactions and, in turn, gauge user engagement to different portions of the compilation. The selective display of a discussion thread associated with a particular video would also reduce the risk of cognitive overload and obstructions to the playback area 310.

170. These modifications represent nothing more than the predictable combination of elements taught by June, De Jong, Houh, and Lee. In my opinion, a PHOSITA would have been able to combine the elements taught by the combination of June, De Jong, and Houh with element [5b] taught by Lee using known GUI design principles and standard computer programming, particularly since June already teaches a third interactable graphical object (e.g., the menus 361A and 361B) and a system that allows users to add comments to a video compilation after publication. Furthermore, in my opinion, the respective functions of the elements taught by June and Lee would not change when combined together: users would be able view comments and bookmarks for a video

compilation according to Lee's teachings, and the addition of another menu to display a discussion thread would not affect the use of the menus 361A and 361B to view and select annotations in the video compilation according to June's teachings.

171. Further, in my opinion, these modifications represent the use of a known technique to improve June's system based on a similar system taught by Lee in the same way. As explained above, June discloses a system that provides a third interactable graphical object (e.g., the menus 361A and 361B) and allowing user comments on a video compilation after publication. Lee teaches a comparable system that displays a discussion thread in a user interface where the discussion thread is associated with multimedia content. In my opinion, a PHOSITA could have modified June's computing system to incorporate the improvements in Lee's system using nothing more than known GUI design principles and standard computer programming.

C. Reasonable Expectation of Success

172. A PHOSITA would have had a reasonable expectation of success in incorporating the use of a discussion thread, as taught by Lee, into the system of June as modified by De Jong and Houh. I understand that a "reasonable expectation of success" is required only for the limitations actually recited in the claims.

173. The art in the relevant timeframe was predictable. Specifically, in the relevant timeframe (no earlier than 2016), basic features to create and display comments associated with media content, like those in Junee and Lee, were well-known in the art.

174. In my opinion, the '442 patent's general and high-level disclosures support this understanding of the relevant art. *See, e.g.*, Ex-1001, 3:65-7:6, 7:38-8:5, 9:25-9:59.

D. Graham Factors

175. The Graham factors were discussed under Ground 1 and apply similarly to all grounds.

E. Analogous Art

176. Lee is analogous art because it is in the same field as the '442 patent (relating to “assemblies, methods, devices, and systems for managing information” generally, which includes “sharing and annotating content,” Ex-1001, 1:18-22), and its methods would have been reasonably pertinent to the problems facing the named inventors, specifically “initiat[ing] discussion threads associated with [a] Playable Media File including a composite digital clip.” Ex-1011, Title, Abstract, ¶0003.

F. Claim Mapping

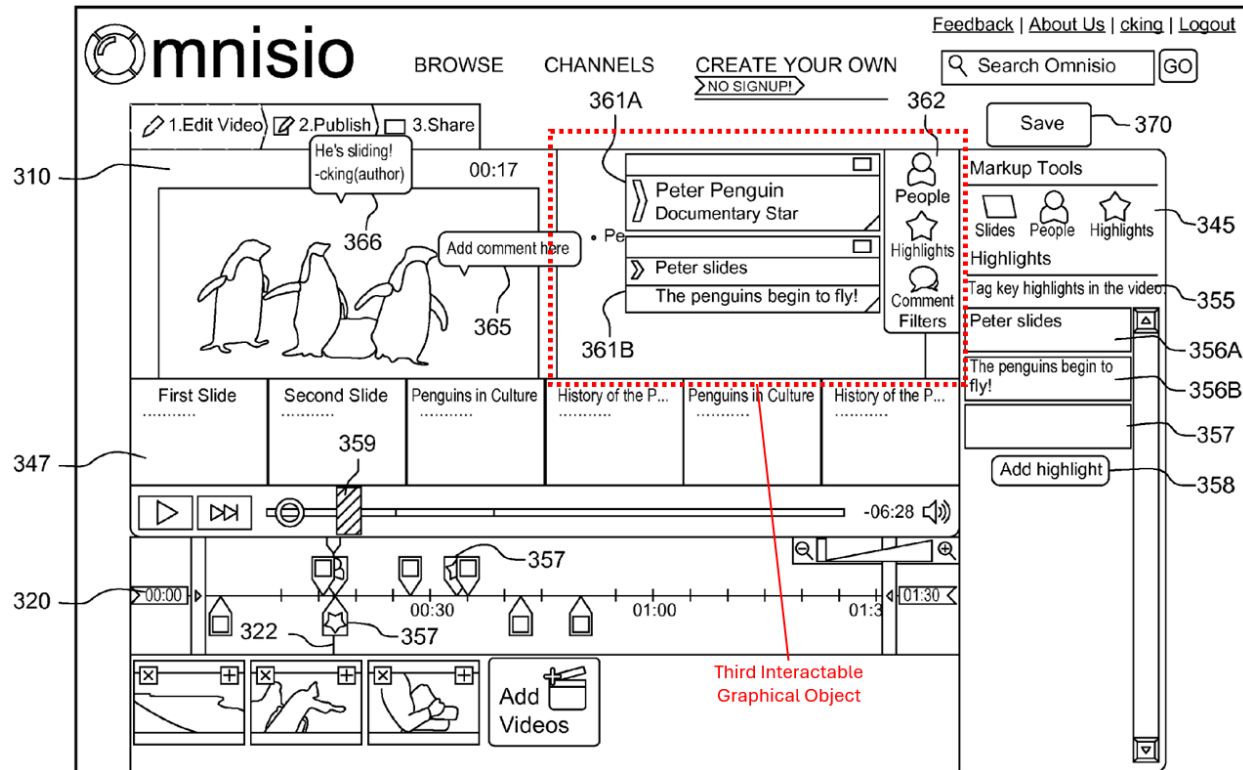
177. This section maps the challenged claims to the relevant disclosures of Junee, De Jong, Houh, and Lee, where the claim text appears in bold-italics, and the relevant mapping follows.

CLAIM 5

“5[a]. The method of claim 1, wherein the graphical user interface further comprises a third interactable graphical object overlaying a third portion of the display window, having a visible state controlling whether the third interactable graphical object is visible or hidden, and”

178. In my opinion, Junee in view of De Jong and Houh renders obvious this claim element. As explained under Ground 1, claim elements [1j] and [1k], Junee discloses a first interactable graphical object (the playback control area 315) and a second interactable graphical object (the portion of Junee’s GUI that displays thumbnail representations of the videos 305A-305C). Junee’s GUI further includes interactable graphical objects that each have a visible state and are configured to display annotations associated with a particular video in the video compilation. Ex-1005, 7:21-27 (menus 361A and 361B), 8:27-32 (comment area 366). In one example, Junee discloses in FIG. 3E, reproduced below, “[s]electing the icons 362 **toggles display of their associated menus**. For example, **selecting the ‘People’ or ‘Highlights’ icons within icon set 362 causes menus 361A and 361B to appear or disappear, respectively**. Menus 361 display the **people and highlights**

annotations, respectively, with which videos of the video compilation have been annotated.” Ex-1005, 7:21-27.



Ex-1005, FIG. 3E (red highlighting added). Thus, Junee discloses examples of a **third interactable graphical object** (e.g., the menus 361A and 361B) having a **visible state controlling whether the third interactable graphical object is visible or hidden** (toggling the visibility of the menus 361A and 361B using the icons in the icon set 362). In my opinion, it would have been obvious for a PHOSITA to add, for example, an additional menu to the menus 361A and 361B to display other information associated with a video in the video compilation, such as a discussion thread as discussed under claim element [5b].

179. As explained under Ground 1, claim elements [1j] and [1k], De Jong discloses a GUI with an interactable graphical object overlaid onto a portion of a display window. Ex-1006, FIG. 13. Houh further explicitly discloses the use of semi-transparent icons for overlay onto video content. Ex-1020, 14:28-39. In my opinion, it would have been obvious for a PHOSITA to modify Junee's GUI to further overlay the menus 361A and 361B together with the playback control area 315 and the portion of Junee's GUI that displays thumbnail representations of the videos 305A-305C onto the playback area 310 according to the teachings of De Jong and Houh. This modification would allow for an even larger-sized playback area 310 to view video content without sacrificing the user's ability to display and interact with the annotations for a particular video. As explained under Ground 1, claim element [1k], the visibility state of the menus 361A and 361B provides a way to reduce potential obstructions to the playback area 310 by allowing users to turn off display of the menus 361A and 361B when they are not used.

“[5b] when visible displaying a discussion thread comprising annotations associated with a selected virtual clip of the plurality of displayable virtual clips.”

180. As a preliminary matter, this claim element is not described in the specification of the '442 patent. Rather, the '442 patent describes annotations and discussion threads as separate two kinds of information that can be associated with a playable media file. Ex-1001, FIG. 3, 3:65-4:3, 7:7-11, 7:35-37. Notably, the

'442 patent, when discussing a third interactable graphical object, does not refer to annotations, but rather “entries” and “information.” Ex-1001, 19:39-44 (“the system may obtain **entries** in a discussion thread associated with the first clip, and may render **information** from the discussion thread into the third interactable graphical object 1840 and update the GUI 1800 to display the third interactable graphical object 1840.”). Accordingly, herein I interpret “annotations” in this claim element as including entries and/or information in a discussion thread, such as user comments posted in a discussion thread.

181. Under this interpretation, the combination of Junee, De Jong, and Houh further in view of Lee renders obvious this claim element. Junee discloses that, when publishing a video compilation, “[r]elated information, such as...**whether users other than the author can provide comments**, can also be specified as part of publishing, e.g. as depicted in the sample user interface of FIG. 3F.” Ex-1005, 8:2-9. Junee further discloses “[d]uring video playback, either while the video is being edited, or **during post-publication viewing**, any previously-specified annotations are displayed and can be used to alter the flow of control of the video compilation” via the GUI of at least FIGS. 3D and 3E. Ex-1005, 8:11-32. Thus, Junee teaches users can add comments to a published video compilation and the author of a video compilation can use Junee’s GUI to view annotations associated with the videos in the video compilation after publication.

182. Lee discloses a system where “[m]embers of a social network service may share real-time comments about multimedia content they are viewing.” Ex-1011, Abstract. Lee explains the comments “are attached to bookmarks that mark corresponding time points in the content 130...**The comments and bookmarks (collectively, referred to as a discussion thread 152A)** are forwarded to other members 110B-N.” Ex-1011, ¶0028. Lee states, “[i]n the example of FIG. 13, the ‘V’ in the screen indicates there is a discussion thread on this scene, and **the discussion thread appears in the SNS member's separate device**. Other implementations will be apparent. For example, **the discussion thread could appear within the playback screen**. Or it could be **toggled on and off**, or retrieved only upon request.” Ex-1011, ¶0057. Thus, Lee discloses displaying a **discussion thread** (the discussion thread 152A) comprising **annotations** (comments and bookmarks) associated with **a selected virtual clip of the plurality of displayable virtual clips** (multimedia content).

183. As set forth above under the “Rationale (Motivation) to Combine” Section, it would have been obvious for a PHOSITA to modify June’s GUI to include a third interactable graphical object to display a discussion thread comprising annotations (comments and bookmarks) associated with a selected virtual clip of the plurality of virtual clips according to Lee’s teachings. For example, the discussion thread can be implemented as an additional menu to the

menus 361A and 361B. This modification would facilitate the display of comments by different users in a format that can be easily viewed by the author when, for example, the author views the video compilation post-publication. Additionally, by displaying different discussion threads for each video in a video compilation, authors can more quickly assess user reactions and, in turn, gauge user engagement to different portions of the compilation. The selective display of a discussion thread associated with a particular video would also reduce the risk of cognitive overload and obstructions to the playback area 310.

184. As set forth above, a PHOSITA would have had a reasonable expectation of success in modifying Junee's GUI to display a discussion thread because Junee, as explained above, already discloses user comments can be added to a video compilation after publication. Thus, the modifications to Junee amount to presenting information already available in Junee in a format that can be easily digested by users.

CLAIM 6

“6. The method of claim 5, further comprising: receiving a second user input describing an interaction with the selected virtual clip within the second interactable graphical object; and responsive to the second user input, displaying the discussion thread within the third interactable graphical object.”

185. In my opinion, the combination of Junee, De Jong, Houh, and Lee renders obvious this claim. As explained under claim element 5[b], the

combination of Junee, De Jong, Houh, and Lee renders obvious a **third interactable graphical object** (e.g., an additional menu to the menus 361A and 361B of Junee) that displays a **discussion thread** (discussion thread of Lee). Junee further discloses “**selecting** the ‘People’ or ‘Highlights’ icons within **icon set 362 causes menus 361A and 361B to appear or disappear**, respectively.” Ex-1005, 7:21-24. Thus, Junee discloses a GUI with GUI elements (the icon set 362) to receive user input that toggles the display of the menus 361A and 361B.

186. In my opinion, it would have been obvious for a PHOSITA, when modifying Junee’s GUI to include a third interactable graphical object (e.g., an additional menu to the menus 361A and 361B of Junee), to also include a GUI element that provides the same functionality as the icons in the icon set 362 according to Junee’s teachings, i.e., to receive user input that toggles the display of the third interactable graphical object. A PHOSITA would have been motivated to make this modification so that users could toggle off this portion to provide a less obstructed view of the playback area 310 and toggle on this portion to allow users to view a discussion thread.

187. In my opinion, it would have also been obvious for a PHOSITA to integrate the GUI element into the thumbnail representations of the videos in the second interactable graphical object such that user interactions with a particular thumbnail representation of a video triggers the display of a discussion thread for

that video. Combining together the GUI element responsible for toggling the display of the third interactable graphical object and the user interactable thumbnail representations taught by Junee would reduce the number of GUI elements in Junee's GUI, thus minimizing obstructions to the playback area 310 when these GUI elements are visible.

188. As explained above, a PHOSITA would have had a reasonable expectation of success in modifying Junee's GUI to display a discussion thread because Junee, as explained above, already discloses user interactable thumbnail representations for each video 305A-305C. Ex-1005, 5:55-5:57 (“[a] video is made active either explicitly, such as by **the user clicking on the corresponding thumbnail representation** of the video 305.”). Accordingly, it would have been well within the knowledge and skill of a PHOSITA to trigger the display of a discussion thread within the third interactable graphical object based on user interactions with the thumbnail representations of a video.

Ground 3. Claim 8 is obvious over Junee, De Jong, Houh and further in view of Syed, Giunio-Zorkin, and Kwon

189. In my opinion, claim 8, which depends from claim 1, would have been obvious over Junee, De Jong, and Houh as discussed under Ground 1 and incorporated herein, further in view of U.S. Pat. Pub. 2015/0139610 (“**Syed**”)(Ex-1009), U.S. Pat. 10,031,949 (“**Giunio-Zorkin**”)(Ex-1013), and U.S. Pat. Pub. 2016/0350430 (“**Kwon**”)(Ex-1012).

190. Syed is a U.S. patent publication that published on May 21, 2015. Ex-1009. The application that led to the Syed publication was filed on November 14, 2011. Ex-1009.

191. Giunio-Zorkin is a U.S. patent that issued on July 24, 2018. Ex-1013. The application that led to the Giunio-Zorkin patent was filed on March 3, 2016. Ex-1013.

192. Kwon is a U.S. patent publication that published on December 1, 2016. Ex-1012. The application that led to the Kwon publication was filed on May 25, 2016. Ex-1012.

193. I understand that Syed, Giunio-Zorkin, and Kwon are each prior art.

194. I further understand that neither Syed nor Giunio-Zorkin nor Kwon was cited during prosecution of any application leading to the '442 patent.

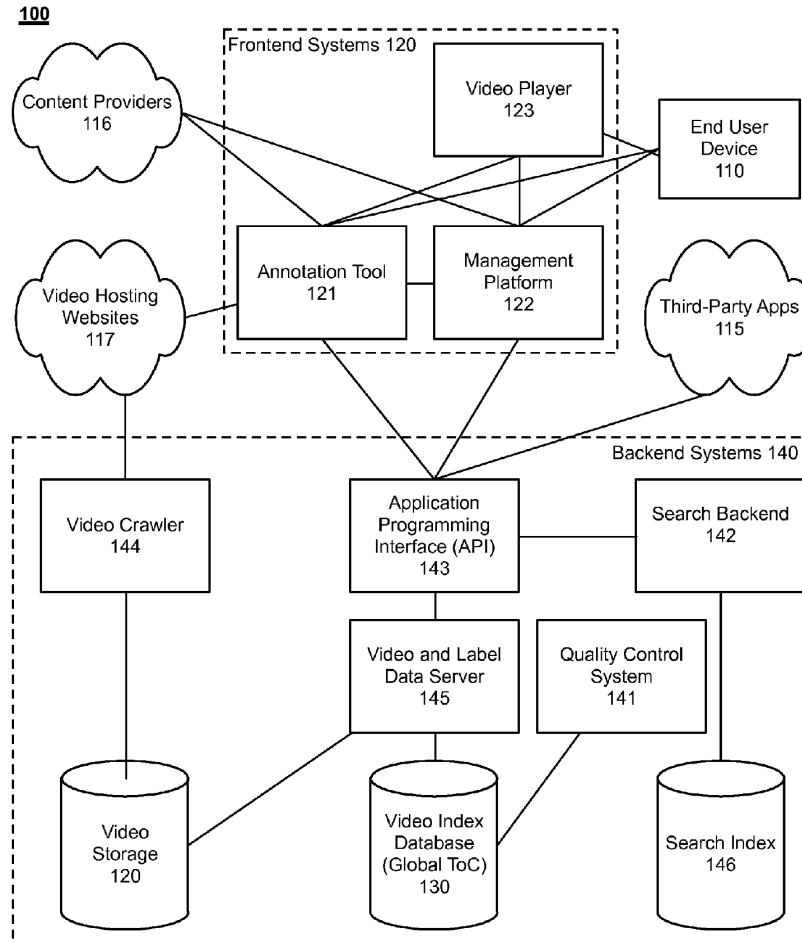
A. Overview of the Ground

195. Claim 8 would have been obvious over the combination of June, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon. To the extent the combination of June, De Jong, and Houh does not disclose or render obvious the limitations of claim 8 directed to determining a user input comprises a character string in a predetermined format, identifying one or more taxonomy tags from the user input, and associating a virtual clip with the one or more taxonomy tags to produce a searchable virtual clip, in my opinion, these limitations would have been obvious

over Syed, Giunio-Zorkin, and Kwon. Syed teaches adding a tag or a label to a video using a tag format comprising a tag identifier and a first tag, identifying a tag from user input, and associating tags with a virtual clip to create a searchable virtual clip. Giunio-Zorkin teaches a taxonomy tag format with a first tag and multiple subtags separated by a delimiter. Kwon teaches identifying multiple tags from user input.

1. Overview of Syed

196. Syed is directed to “systems and methods for discovering and consuming video content, including annotating, organizing, searching, indexing and sharing information about such content. Syed’s system is illustrated in FIG. 1, reproduced below, which shows “a block diagram illustrating frontend systems 120 and backend systems of a video consumption and discovery service 100 and interactions with external systems 140.” Ex-1009, ¶0049.

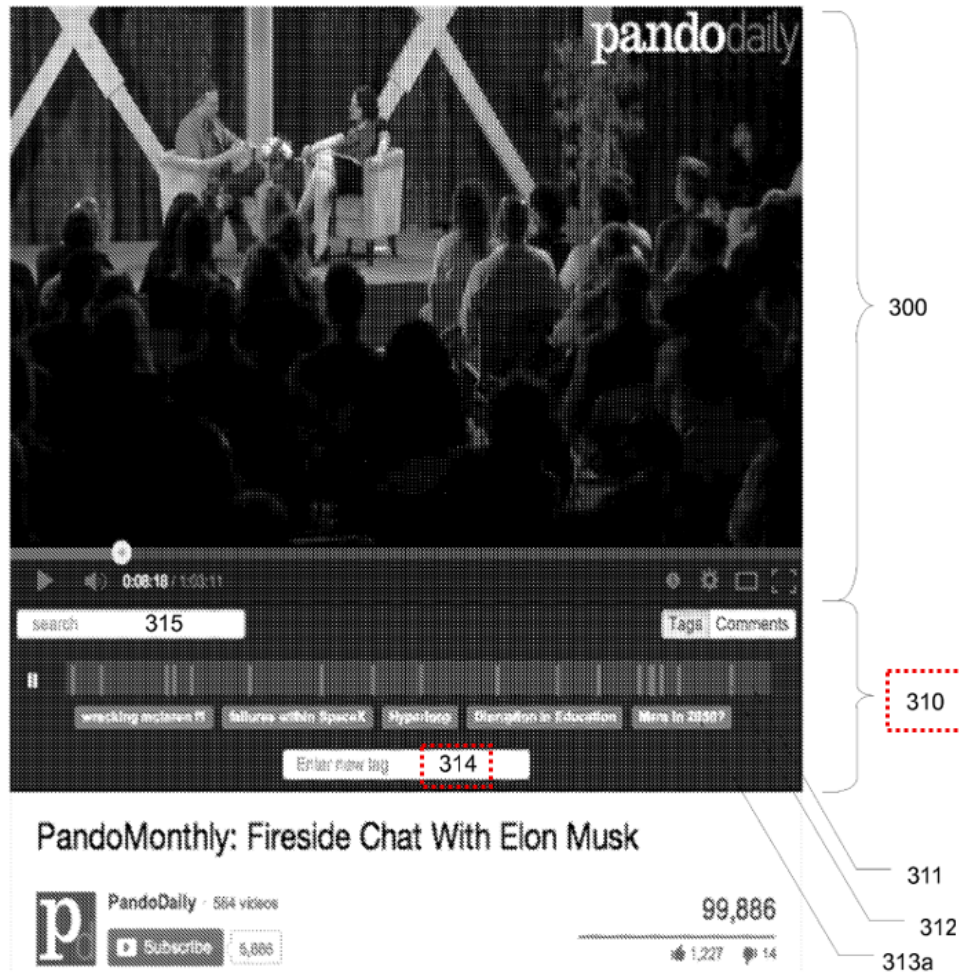


Ex-1009, FIG. 1. As shown, Syed discloses an “annotation tool 121 is presented to users of client devices (e.g., end user device 110) while they are watching a video and allows users to communicate facts, opinions and/or emotions, collectively referred to herein as annotations, about specific moments, portions or segments of a video.” Ex-1009, ¶0057.

197. As used by Syed, “[t]he term ‘annotation’ broadly refers to user-supplied or automatically generated content that is associated with **a portion of video content or a video ‘clip’** (e.g., one or more frames of video content corresponding to a particular time or period of time within a video)” and may

include “**one or more of labels, tags**, comments and additional content.” Ex-1009, ¶0037. Syed explains “a tag or a label may be in the form of **a hash tag that is similar to hash tags on Twitter.**” *Id.* Moreover, “the label may include **one or more words and/or be in the form of one or more hash tags** (e.g., hash tags similar to those used in the context of Twitter).” Ex-1009, ¶0042.

198. To receive annotations from users, Syed discloses an “annotation tool (either integrated within the video player or separate) allows users to **specify content** and placement of **a label or tag** (e.g., a short description of a video segment) **within a video while watching the video.**” Ex-1009, ¶0031. Syed explains the “annotation tool is presented to a first subscriber of a network-based video discovery and consumption service as video content is being consumed by the first subscriber. A first proposed label to be associated with a first portion of the video content is received by the annotation tool from the first subscriber. The first proposed label is incorporated within a global table of contents (ToC) for the video content...and associating with the video content the first proposed label and a start time of the first portion of the video content.” Ex-1009, ¶0013. Syed illustrates an example of the annotation tool in FIG. 3A, reproduced below, which shows “a video player 300, an annotation tool 310 and a table of contents 313 a-n for a video.” Ex-1009, ¶0083.



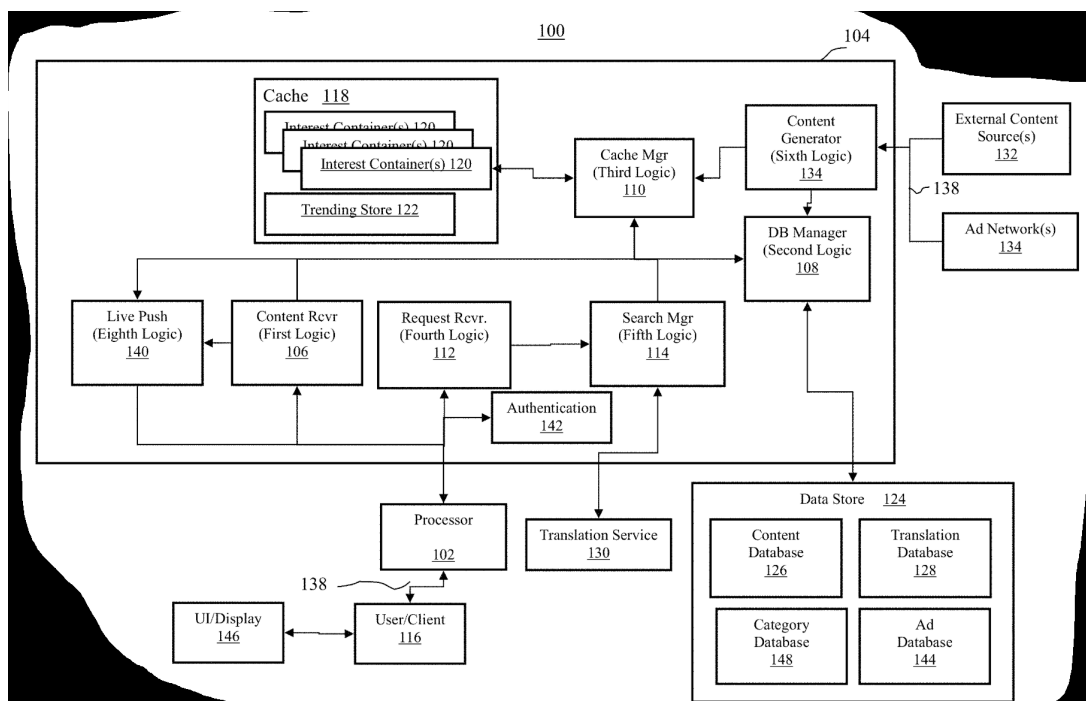
Ex-1009, FIG. 3A (red highlighting added). As shown, the annotation tool 310 includes a “[l]abel text entry field 314 [that] allows the user to propose a new label for a portion of the video being played or edit an existing label.” Ex-1009, ¶0085.

199. Syed further explains that “[a]ny label added or vote cast by the viewer” is ultimately “sent to video label integration and indexing module 247,” which provides “automated label integration/filtering processing.” Ex-1009, ¶0075. As an example, Syed discloses in FIG. 9, “a flow chart illustrating automated label integration/filtering processing,” which “assumes automated label

integration/filtering processing has been triggered as a result of the addition of a new label (e.g., manually by a user or automatically).” Ex-1009, ¶0119. Syed explains that the “automated label integration/filtering processing [in FIG. 9] may involve, among other things, a weighted integration of labels via clustering and spam filtering.” Ex-1009, ¶0093.

2. Overview of Giunio-Zorkin

200. Giunio-Zorkin is directed to “systems and methods for efficiently and automatically distributing content among a plurality of users which may be used to implement a social networking service.” Ex-1013, 4:46-49. Giunio-Zorkin’s system is illustrated in FIG. 1, reproduced below, which shows “a block diagram of a system 100...for automated electronic distribution of a content item.” Ex-1013, 6:5-7.



Ex-1013, FIG. 1. Giunio-Zorkin explains “a ‘content item’ may refer to a data item or data structure which includes content or links/references thereto, e.g. data such as one or more text strings, images, etc., generated by the system, e.g. via the content generator 134, and/or by a user of the system, such as a ‘post’, comment, statement, picture, URL/web link, hashtag, etc.” Ex-1013, 6:66-7:4. The content item is “provided by an external or internal source, e.g. a client 116 or content generator 134, to the content distribution system 100, to a subset of a plurality of users, e.g. clients 116, such as via a network 138.” Ex-1013, 6:7-11.

201. Giunio-Zorkin discloses “**a taxonomy of tags**, e.g. keyword metadata or keywords, is provided...the available keyword metadata of the system 100 is organized into a plurality of keyword sets/hierarchies, where each keyword set includes **a hierarchical sequence of keywords**, wherein for each keyword set a first keyword of the sequence of keywords defines semantics or a context...and each subsequent keyword of the sequence of the keyword set defines a narrowing of the semantics/context over the previous keyword of the sequence of the keyword set. For example, one keyword set may include the sequence of keywords: **Drinks; Alcohol; Wine; White; Pino Grigio; California; Berringer**. Another keyword set may include the sequence of keywords: **World: North America: United States: Illinois: Cook County: Chicago: Lincoln Park**.” Ex-1013, 11:59-12:11. A PHOSITA would have readily understood that the semi-

colon (“;”) and the colon (“:”) in the foregoing examples are delimiters (as were common in the prior art). *See* Ex-1019 at 152 (defining delimiter and providing examples of commonly used delimiters, including semi-colon and colon).

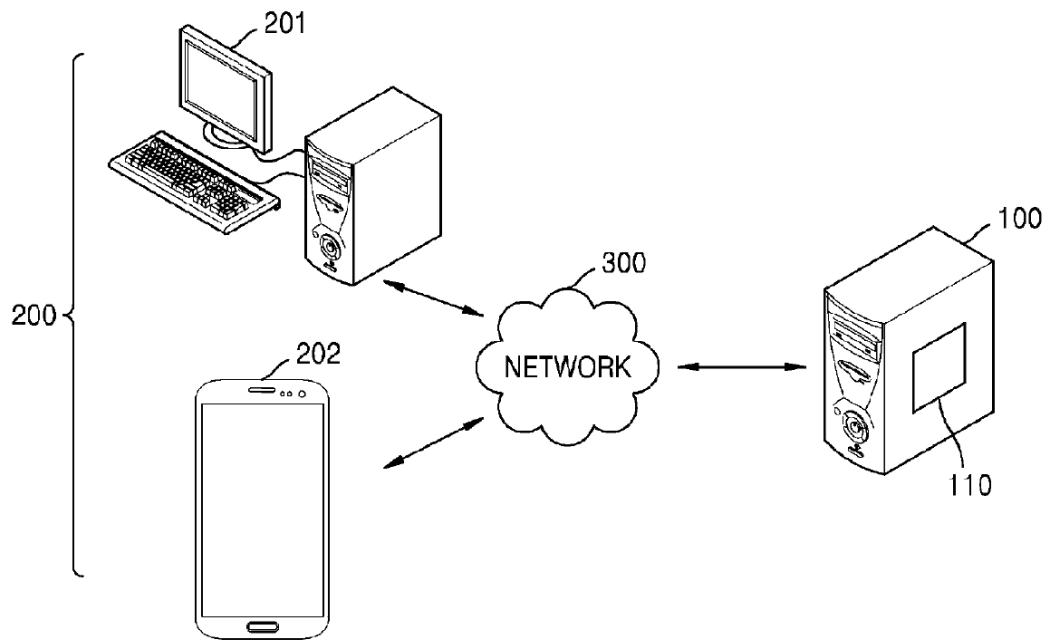
202. Giunio-Zorkin teaches that “[c]ontent items are associated with data indicative of a subset of one or more of the keywords selected from the plurality of keyword sets.” Ex-1013, 12:30-32. Giunio-Zorkin further explains that “[g]enerally, the association of keywords **facilitates the discovery of content** that is of interest to the user by enabling the system 100 to **correlate the interests of the user against the context/relevance of the available content items and content items related thereto.**” Ex-1013, 13:63-67.

203. As an example, Giunio-Zorkin discloses in FIG. 3, “a method of distributing a content item, provided by a source to a content distribution system 100, to a subset of a plurality of users via a network 138 coupled therewith.” Ex-1013, 30:5-8. In this method, Giunio-Zorkin discloses the steps of: “**receiving...a request for content from a requestor** via the network [Block 308]; **identifying...based on the received request, one or more keywords selected from the plurality of keyword sets** [Block 310]; determining...based on at least a subset of the identified keywords, whether there are any previously received content items stored...in association with the subset of the identified keywords and, if there are..., causing...the transmission of at least a subset of those

determined previously received content items...to the requestor via the network [Block 314].” Ex-1013, 30:38-51.

3. Overview of Kwon

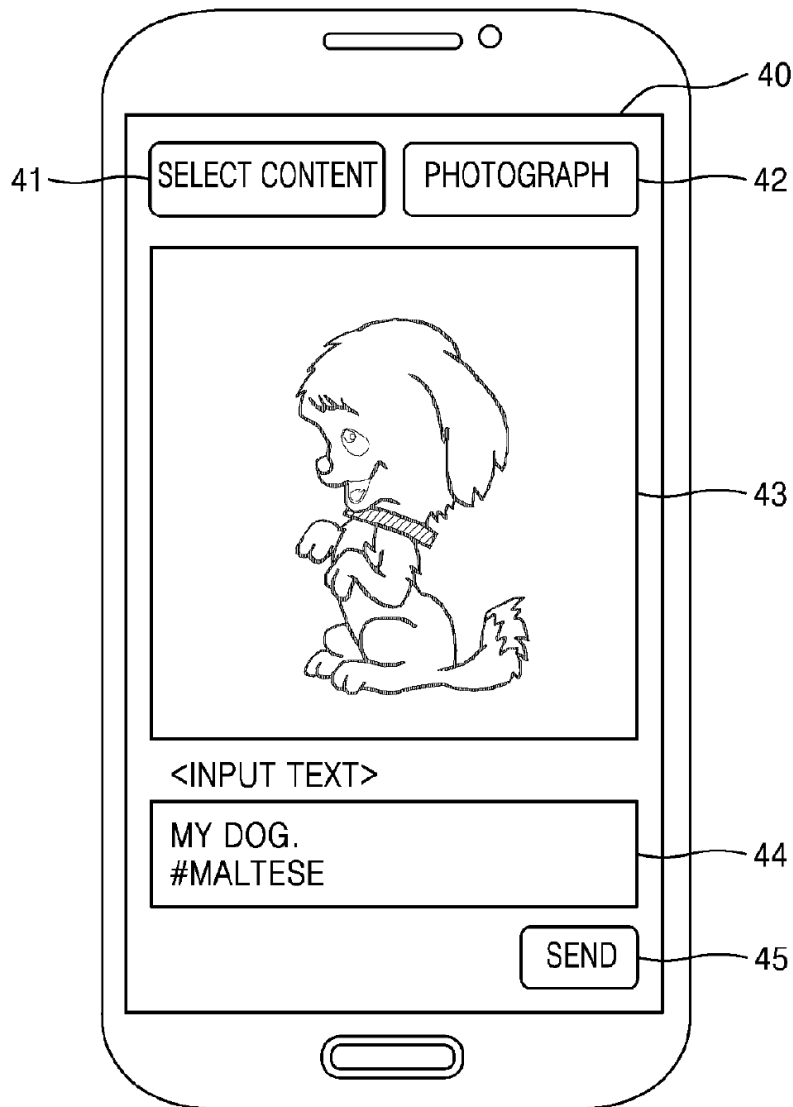
204. Kwon is directed to “methods, apparatuses, and/or computer programs for managing content.” Ex-1012, ¶0003. Kwon’s system is illustrated in FIG. 1, reproduced below, which shows “a diagram of a structure of a content sharing system.” Ex-1012, ¶0013.



Ex-1012, FIG. 1. As shown, “the content sharing system...includes a server 100, a user terminal 200, and a network 300 connecting the server 100 and the user terminal 200.” Ex-1012, ¶0025. Kwon discloses the server 100 includes “an apparatus 110 for managing contents.” Ex-1012, ¶0033. “[T]he apparatus 110 according to an example embodiment includes a receiver 111, a content registerer

112.” Ex-1012, ¶0036. “The receiver 111...may receive content and a user tag from, for example, the user terminal 200. The receiver 111 may receive a content and one user tag or a plurality of user tags to be assigned to the content from the user terminal 200. A type of the content may be an image, a video, or a sound source, but is not limited thereto.” Ex-1012, ¶0037. “The content registerer 112...may register the content and the user tag received from the user terminal 200 in the server 100.” Ex-1012, ¶0038.

205. Kwon discloses in FIG. 4, reproduced below, “an example of a screen 40 provided by the receiver 111.” Ex-1012, ¶0059.



Ex-1012, FIG. 4. As shown, “[t]he user may input text to be uploaded together with the content 43 to a text input region 44. **The user may input a tag to be assigned to the content 43 by using a certain identifier.** In FIG. 4, **the certain identifier may be a hash tag ‘#’.** The user inputs a word together with the identifier ‘#’ to input the word as the tag of the content 43. In FIG. 4, the user input ‘Maltese’ as the tag. **In FIG. 4, the user input one tag, but the user may input a**

plurality of tags. For example, **when the user inputs ‘#Maltese #Dog’, both a tag ‘Maltese’ and a tag ‘Dog’ are input.**” Ex-1012, ¶0061. Kwon further discloses “[t]he content registerer 112 may **use the identifier** included in the text to **distinguish the user tag**, and assign the tag as a general tag of the content 43.” Ex-1012, ¶0063.

B. Rationale (Motivation) Supporting Obviousness

206. Junee discloses element 8[a]. In particular, Junee discloses “a user interface region 355 for adding annotation related to highlights in a video...Selecting the button 358 creates a new highlight for the text within the current text field 357, associating it with the current time of the current video of the compilation and creating a marker tag at the corresponding time on the video.” Ex-1005, 7:6-16. Thus, Junee discloses **receiving a second user input** (annotations related to highlights via text field 357 and button 358) **associated with a first virtual clip** (a video in the video compilation).

207. To the extent the combination of Junee, De Jong, and Houh does not disclose or render obvious the remaining elements of dependent claim 8—directed to determining a user input comprises a character string in a predetermined format (claim elements [8b] and [8c]), identifying one or more taxonomy tags from the user input (claim element [8d]), and associating a virtual clip with the one or more taxonomy tags to produce a searchable virtual clip that is delivered in response to a

query for virtual clips associated with a user profile and the one or more taxonomy tags (claim element [8e])—in my opinion, it would have been obvious for a PHOSITA to incorporate such features into a system such as Junee’s in view of the teachings of Syed, Giunio-Zorkin, and Kwon.

Determining a user input comprises a character string in a predetermined format comprising a tag identifier, a first tag [8b], and one or more subtags sequentially following the first tag and delimited by a delimiter character [8c]

208. Syed describes an “annotation tool (either integrated within the video player or separate) which allows users to **specify content** and placement of **a label or tag** (e.g., a short description of a video segment) **within a video while watching the video.**” Ex-1009, ¶0031. Syed states, “a tag or a label may be in the form of **a hash tag that is similar to hash tags on Twitter.**” Ex-1009, ¶0037. Thus, Syed discloses **determining a user input comprises a character string in a predetermined format** (a tag or a label in the form of a hash tag).

209. Syed further explains that “a label or a tag” can include “a short description of a video segment.” Ex-1009, ¶0031. Thus, Syed discloses **a tag identifier** (the “#” symbol) and **a first tag** (the short description of a video segment) in sequence when the user input is in the form of a hash tag.

210. Giunio-Zorkin discloses “keyword metadata,” which are “**keywords or tags associated with the content items.**” Ex-1013, 11:15-16. Giunio-Zorkin

teaches the keyword metadata may be in the form of “**a taxonomy of tags**” where “the available keyword metadata...is organized into a plurality of keyword sets/hierarchies, where **each keyword set includes a hierarchical sequence of keywords**, wherein for each keyword set **a first keyword of the sequence of keywords defines semantics or a context...**and **each subsequent keyword of the sequence of the keyword set defines a narrowing of the semantics/context** over the previous keyword of the sequence of the keyword set.” Ex-1013, 11:59-12:6.

Giunio-Zorkin further discloses examples of a taxonomy tag: “one keyword set may include the sequence of keywords: Drinks; Alcohol; Wine; White; Pino Grigio; California; Berringer. Another keyword set may include the sequence of keywords: World: North America: United States: Illinois: Cook County: Chicago: Lincoln Park.” Ex-1013, 12:6-12:11.

211. In my opinion, a PHOSITA would have recognized each of the foregoing examples of keyword sets disclosed in Giunio-Zorkin includes commonly-used delimiters, such as a semi-colon (“;”) for the Drinks keyword set and a colon (“:”) for the World keyword set. *See* Ex-1019 at 152 (defining delimiter and providing examples of commonly used delimiters, including semi-colon and colon). Thus, Giunio-Zorkin teaches a predetermined format that includes **one or more subtags** (the keywords in a keyword set that follow the first keyword) **sequentially following a first tag** (the first keyword in a keyword set)

and **delimited by a delimiter character** (the delimiters between the keywords in the “Drinks” and “World” examples).

212. In my opinion, it would have been obvious for a PHOSITA to modify Junee’s system to receive annotations in the form of a character string that conforms to a predetermined format that indicates an annotation, such as the hash tag format according to Syed’s teachings, and supports taxonomy tags according to Giunio-Zorkin’s teachings. These modifications would provide a way for users to include annotations that are not necessarily displayed in the video compilation of Junee, but rather as metadata associated with the video compilation to facilitate retrieval of the video compilation in response to a query by a user of Junee’s system. The addition of taxonomy tags as metadata to a video compilation in Junee’s system would further facilitate retrieval of video compilations more aligned to the interests of the user submitting a query. Ex-1013, 11:6-7 (“for the purpose of distributing and/or facilitating discovery of that content item”).

213. These modifications represent nothing more than the predictable combination of elements taught by Junee, De Jong, Houh, Syed, and Giunio-Zorkin. A PHOSITA would have been able to combine the elements taught by the combination of Junee, De Jong, and Houh with elements [8b] and [8c] taught by Syed and Giunio-Zorkin using standard computer programming. Furthermore, the respective functions of the elements taught by Junee, Syed, and Giunio-Zorkin

would not change when combined together: users would be able generate annotations for different portions of a video using a hashtag format according to Syed's teachings, users would be able to input annotations using the taxonomy tag format according to Giunio-Zorkin's teachings, and users would still be able to generate annotations for display with a video in the video compilation using the user interface region 355 according to June's teachings.

214. Further, in my opinion, these modifications represent the use of a known technique to improve June's system based on a similar system taught by Syed and Giunio-Zorkin in the same way. As explained above, June discloses a system that receives user input to add annotations to a video compilation. Syed teaches a comparable system that allows users to create annotations for a video if the user's input conforms to a particular format (a hashtag format). Likewise, Giunio-Zorkin teaches a comparable system that allows users to associate keyword sets with content items. In my opinion, a PHOSITA could have modified June's computing system to incorporate the improvements in Syed's system and Giunio-Zorkin's system using nothing more than standard computer programming.

Identifying one or more taxonomy tags from the user input [8d]

215. Syed discloses in FIG. 9, "a flow chart illustrating automated label integration/filtering processing," which "assumes automated label integration/filtering processing has been triggered as a result of the addition of a

new label (e.g., manually by a user or automatically).” Ex-1009, ¶0119. Because the execution of the automated label integration/filtering processing disclosed in Syed requires the “addition of a new label,” this shows the automated label integration/filtering processing is preceded by a process to identify a tag or a label from user input to ensure a new label was added. Ex-1009, ¶¶0037, 0119, FIG. 9. Thus, Syed discloses **identifying one taxonomy tag from the user input** (identifying tags from user input to facilitate label integration/filtering processing).

216. In my opinion, it would have been obvious for a PHOSITA to implement a process in Junee to identify a tag from user input according to the teachings of Syed to facilitate the use of taxonomy tags for association with a video compilation, such as by differentiating user input that includes a tag or a label from other user input that would otherwise not be eligible for use as a tag or a label.

217. To the extent this claim element is construed to require the capability to identify multiple taxonomy tags in the user input, Kwon discloses in connection with FIG. 4 that “[t]he user may input text to be uploaded together with the content 43 to a text input region 44.” Ex-1014, ¶0061. Kwon further discloses “[i]n FIG. 4, the user input one tag, but **the user may input a plurality of tags**. For example, **when the user inputs ‘#Maltese #Dog’, both a tag ‘Maltese’ and a tag ‘Dog’ are input.**” Ex-1014, ¶0061. Once user input is submitted, Kwon discloses “[u]pon

receiving the content 43 and the text, the content registerer 112 may register the content 43 and the text in the DB 114 of the apparatus 110. The content registerer 112 may **use the identifier included in the text to distinguish the user tag**, and assign the tag as a general tag of the content 43.” Ex-1014, ¶0063. Because Kwon’s system supports user inputs with multiple tags and includes a content register that identifies tags for registration in a database based on the presence of tag identifiers in the user input, it would have been obvious for Kwon’s system to **identify *multiple* taxonomy tags from user input** (multiple tags that each have a tag identifier, such as “#Maltese #Dog”).

218. In my opinion, it would have been obvious for a PHOSITA to further modify June’s system by implementing a process to identify multiple tags or labels from a single user input where each tag is in hash tag form according to the teachings of Kwon. The support and identification of multiple tags in a single user input would further expedite the process of associating multiple tags with a video or a video compilation in June.

219. These modifications represent nothing more than the predictable combination of elements taught by June, De Jong, Houh, Syed and, to the extent necessary, Kwon. In my opinion, a PHOSITA would have been able to combine the elements taught by the combination of June, De Jong, and Houh with element [8d] taught by Syed and, to the extent necessary, Kwon using standard computer

programming. Furthermore, the respective functions of the elements taught by Junea, Syed, and Kwon would not change when combined together: user input would be processed to identify whether one or multiple tags are present according to Syed's teachings and Kwon's teachings, respectively, and users would be able to add tags for association with a video compilation according to Junea's teachings.

220. Further, these modifications represent the use of a known technique to improve Junea's system based on a similar system taught by Syed and, to the extent necessary, Kwon in the same way. As explained above, Junea discloses a system that receives user input to add annotations to a video compilation. Syed teaches a comparable system that identifies annotations for a video if the user's input conforms to a particular format (a hashtag format). Kwon teaches a comparable system that identifies tags for association with a video when a user input includes text in a hashtag format. In my opinion, a PHOSITA could have modified Junea's computing system to incorporate the improvements in Syed's system and, to the extent necessary, Kwon's system using nothing more than standard computer programming.

Associating a virtual clip with the one or more taxonomy tags to produce a searchable virtual clip that is delivered in response to a query for virtual clips associated with a user profile and the one or more taxonomy tags [8e]

221. Syed discloses that "[t]ags proposed by the subscribers to be associated with various portions of the video content are received, processed

and integrated within a global ToC for the video.” Ex-1009, Abstract. Syed further discloses the video player 300 shown in FIGS. 3A and 3B includes a “[s]earch text entry field 315 [that] allows the user to initiate a **keyword search**. If the user has chosen to perform a search, the user is then presented with results in the form of labels and/or **clips** by management platform.” Ex-1009, ¶0084. Syed further teaches “[s]ome embodiments allow the user to filter **annotations by the type of user** that provided it. For example, **users can filter by themselves**, everyone, friends, those automatically generated by the service provider and so forth.” Ex-1009, ¶0064. In Syed, a video clip may be “defined by a user by **proposing a label to be associated with a particular time offset within the video**.” Ex-1009, ¶0086.

222. Thus, in Syed, filtering annotations, such as a label, by user type is equivalent to filtering clips by user type when the clips are defined by the labels. Thus, Syed teaches **associating the first virtual clip** (portion of video content) with **one or more taxonomy tags** (tags/labels) to produce a **searchable virtual clip** (clip) that is delivered to a requesting device in response to a **query** (keyword search, filter) from the requesting device for **any of the plurality of virtual clips** (clips) that are associated with the **user profile** (users that filter themselves) and the **one or more taxonomy tags** (tags/labels).

223. In my opinion, it would have been obvious for a PHOSITA to modify Junee's system by associating the taxonomy tags with a virtual clip (a video in the video compilation) to facilitate retrieval of the video compilation in response to a query by a user of Junee's system. It would have also been obvious for a PHOSITA to implement features that allow users to search for video content based on a particular user and the taxonomy tags to further improve the quality and relevance of search results in response to a user's search query.

224. These modifications represent nothing more than the predictable combination of elements taught by Junee, De Jong, Houh, and Syed. In my opinion, a PHOSITA would have been able to combine the elements taught by the combination of Junee, De Jong, and Houh with element [8e] taught by Syed using standard computer programming. Furthermore, the respective functions of the elements taught by Junee and Syed would not change when combined together: users would be able to associate annotations with different portions of a video for purposes of facilitating retrieval of the video according to Syed's teachings, and users would still be able to associate annotations as editing actions in the editing database 154 according to Junee's teachings.

225. Further, these modifications represent the use of a known technique to improve Junee's system based on a similar system taught by Syed in the same way. As explained above, Junee discloses a system that associates annotations to a video

compilation in a database (editing database 154). Syed teaches a comparable system that allows users to associate annotations to a video to facilitate retrieval of labels and/or clips. In my opinion, a PHOSITA could have modified June's computing system to incorporate the improvements in Syed's system using nothing more than standard computer programming.

C. Reasonable Expectation of Success

226. In my opinion, a PHOSITA would have had a reasonable expectation of success in modifying June to incorporate annotations that categorizes content according to a particular tag format based on the teachings of Syed, Giunio-Zorkin, and Kwon.

227. The art in the relevant timeframe was predictable. Specifically, in the relevant timeframe (no earlier than 2016), basic formats for annotations like those disclosed in Syed, Giunio-Zorkin, and Kwon were well-known in the art.

228. In particular, in my opinion, the '442 patent's general and high-level disclosures support this understanding of the relevant art. *See, e.g.*, Ex-1001, 3:65-7:6, 7:38-8:5, 9:25-9:59.

D. Graham Factors

229. The Graham factors were discussed under Ground 1 and apply equally to all grounds.

E. Analogous Art

230. Syed, Giunio-Zorkin, and Kwon are analogous art because they are in the same field as the '442 patent (relating to “assemblies, methods, devices, and systems for managing information” generally, which includes “sharing and annotating content,” Ex-1001, 1:18-22), and the methods would have been reasonably pertinent to the problems facing the named inventors, specifically “annotating a playable media file.” Ex-1009, Title, Abstract, ¶0004; Ex-1013, Title, Abstract, 4:46-49; Ex-1012, Title, Abstract, ¶0003; Ex-1001, 1:63.

F. Claim Mapping

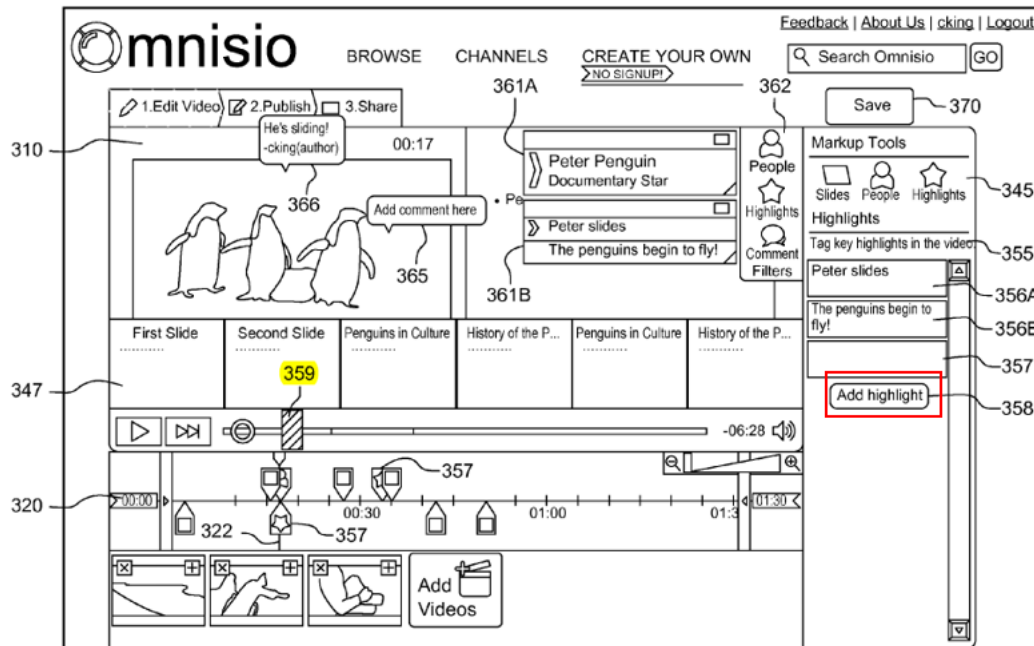
231. This section maps the challenged claims to the relevant disclosures of June, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon where the claim text appears in bold-italics, and the relevant mapping follows.

CLAIM 8

“8[a]. The method of claim 1, further comprising: receiving a second user input associated with a first virtual clip of the plurality of virtual clips;”

232. See claim 1. June discloses this claim element. FIG. 3E of June, reproduced below, “depicts a user interface region 355 for **adding annotation related to highlights in a video...Selecting the button 358 creates a new highlight for the text within the current text field 357, associating it with the**

current time of the current video of the compilation and creating a marker tag at the corresponding time on the video.” Ex-1005, 7:6-16.



Ex-1005, FIG. 3E (highlighting and red box added). Thus, Junee discloses **receiving a second user input** (annotations related to highlights via text field 357 and button 358) **associated with a first virtual clip** (a video in the video compilation).

“[8b] determining that the second user input comprises a character string in a predetermined format comprising: a tag identifier indicating that the user input includes one or more taxonomy tags; a first tag following the tag identifier; and”

233. In my opinion, the combination of Junee, De Jong, Houh, and Syed renders obvious this claim element. Syed describes an “annotation tool (either integrated within the video player or separate) which allows users to **specify**

content and placement of **a label or tag** (e.g., a short description of a video segment) **within a video while watching the video.**” Ex-1009, ¶0031. Syed states, “a tag or a label may be in the form of **a hash tag that is similar to hash tags on Twitter.**” Ex-1009, ¶0037. Syed further states, “[d]epending upon the particular implementation, the label may include one or more words and/or **be in the form of one or more hash tags** (e.g., hash tags similar to those used in the context of Twitter).” Ex-1009, ¶0042. Thus, Syed discloses **determining a user input comprises a character string in a predetermined format** (a tag or a label in the form of a hash tag).

234. Syed further explains that “a label or a tag” can include “a short description of a video segment.” Ex-1009, ¶0031. The terms “label” and “tag” are used interchangeably in Syed and thus are equivalent in the context of Syed’s teachings. Thus, Syed discloses **a tag identifier** (the “#” symbol) and a **first tag** (the short description of a video segment) in sequence when the user input is in the form of a hash tag.

235. As set forth above under the “Rationale (Motivation) to Combine” Section, it would have been obvious for a PHOSITA to modify June’s system to receive annotations in the form of a character string that conforms to a predetermined format indicative of an annotation, such as the hash tag format according to Syed’s teachings. This modification would provide a way for users to

include annotations that are not necessarily displayed in the video compilation of Junee, but rather as metadata associated with the video compilation to facilitate retrieval of the video compilation in response to a query by a user of Junee's system. Moreover, this modification would not be limited to compilations with multiple videos, but can also apply to single videos as well. As explained under claim element 7[a] in Ground 1, Junee discloses video compilations "may include a single video." Ex-1005, 4:23. Thus, adding metadata to a video compilation containing a single video to facilitate retrieval of the video compilation in response to a query is functionally equivalent to adding metadata to a single video to facilitate retrieval of the video in response to a query.

236. In my opinion, a PHOSITA would have had a reasonable expectation of success in modifying Junee's system because Junee already teaches the addition of metadata tagging to video content uploaded to the video hosting server 108 of Junee. Ex-1005, 3:18-21 ("The uploaded content is processed and stored in the video database 128. This processing can include format conversion (transcoding), compression, **metadata tagging**, and other data processing.").

237. To the extent this claim element is construed to require the tag identifier to indicate the user input includes one or more taxonomy tags, the combination of Junee, De Jong, Houh, and Syed further in view of Giunio-Zorkin nevertheless renders obvious this claim element. Giunio-Zorkin discloses

“keyword metadata,” which are **“keywords or tags associated with the content items.”** Ex-1013, 11:15-16. In Giunio-Zorkin, a content item “includes content or links/references thereto, e.g. data such as one or more text strings, images, etc., generated by the system, e.g. via the content generator 134, and/or by a user of the system, such as a ‘post’, comment, statement, picture, URL/web link, hashtag, etc.” Ex-1013, 6:67-7:4.

238. Giunio-Zorkin teaches the keyword metadata may be in the form of **“a taxonomy of tags”** where **“the available keyword metadata...is organized into a plurality of keyword sets/hierarchies, where each keyword set includes a hierarchical sequence of keywords, wherein for each keyword set a first keyword of the sequence of keywords defines semantics or a context...and each subsequent keyword of the sequence of the keyword set defines a narrowing of the semantics/context over the previous keyword of the sequence of the keyword set.”** Ex-1013, 11:59-12:6.

239. In my opinion, it would have been obvious for a PHOSITA to further modify Junee’s system to use a predetermined format that supports taxonomy tags according to Giunio-Zorkin’s teachings **“for the purpose of distributing and/or facilitating discovery of that content item.”** Ex-1013, 11:6-7. In particular, Giunio-Zorkin teaches **“content items are associated with one or more interest classifications”** and its system **“maps or otherwise correlates the interests of each**

user against the interest classification(s) of content items to determine which content items to transmit to a user.” Ex-1013, 11:3-4, 11:10-13. The addition of taxonomy tags as metadata to a video compilation in Junee’s system would further facilitate retrieval of video compilations more aligned to the interests of the user submitting a query.

“[8c] one or more subtags sequentially following the first tag and delimited by a delimiter character;”

240. In my opinion, the combination of Junee, De Jong, Houh, Syed, and Giunio-Zorkin renders obvious this claim element. As explained under claim element [8b], Giunio-Zorkin teaches keyword metadata may be in the form of “**a taxonomy of tags.**” Ex-1013, 11:59.

241. Giunio-Zorkin further discloses examples of a taxonomy tag: “one keyword set may include the sequence of keywords: Drinks; Alcohol; Wine; White; Pino Grigio; California; Berringer. Another keyword set may include the sequence of keywords: World: North America: United States: Illinois: Cook County: Chicago: Lincoln Park.” Ex-1013, 12:6-12:11. A PHOSITA would have recognized each of the foregoing examples of keyword sets disclosed in Giunio-Zorkin includes commonly-used delimiters, such as a semi-colon (“;”) for the Drinks keyword set and a colon (“:”) for the World keyword set. *See* Ex-1019 at 152 (defining delimiter and providing examples of commonly used delimiters,

including semi-colon and colon); Thus, Giunio-Zorkin teaches a predetermined format that includes **one or more subtags** (the keywords in a keyword set that follow the first keyword) **sequentially following a first tag** (the first keyword in a keyword set) and **delimited by a delimiter character** (the delimiters between the keywords in the “Drinks” and “World” examples).

242. In my opinion, it would have been obvious for a PHOSITA to modify Junee’s system to use the taxonomy tag format according to Giunio-Zorkin’s teachings for the same reasons set forth under claim element [8b], i.e., to facilitate retrieval of video compilations more aligned to the interests of the user submitting a query.

“[8d] identifying one or more taxonomy tags from the second user input; and”

243. In my opinion, the combination of Junee, De Jong, Houh, and Syed renders obvious this claim element. Syed discloses “[a]ny label added or vote cast by the viewer” is ultimately “sent to video label integration and indexing module 247,” which provides “automated label integration/filtering processing.” Ex-1009, ¶0075. As an example, Syed discloses in FIG. 9, “a flow chart illustrating automated label integration/filtering processing,” which “assumes automated label integration/filtering processing has been triggered as a result of the addition of a new label (e.g., manually by a user or automatically).” Ex-1009, ¶0119. Syed

explains the “automated label integration/filtering processing [in FIG. 9] may involve, among other things, a weighted integration of labels via clustering and spam filtering.” Ex-1009, ¶¶0093. Because the execution of the automated label integration/filtering processing disclosed in Syed requires the “addition of a new label,” this shows the automated label integration/filtering processing is preceded by a process to identify a tag or a label from user input to ensure a new label was actually added or not. Ex-1009, ¶¶0037, 0119, FIG. 9. Thus, Syed discloses **identifying *one* taxonomy tag from the user input** (identifying tags from user input to facilitate label integration/filtering processing).

244. In my opinion, it would have been obvious for a PHOSITA to implement a process in Junee to identify a tag from user input according to the teachings of Syed to facilitate the use of taxonomy tags for association with a video compilation, such as differentiating user input that includes a tag from other user input that would otherwise not be eligible for use as a tag. A PHOSITA would have had a reasonable expectation of success because the predetermined formats disclosed by Syed are standard formats commonly used to store tabular data, such as data associated with a video.

245. To the extent this claim element is construed to require the capability to identify *multiple* taxonomy tags in the user input, in my opinion, this claim element is nevertheless obvious over the combination of Junee, De Jong, Houh,

and Syed further in view of Kwon. Kwon discloses in connection with FIG. 4 that “[t]he user may input text to be uploaded together with the content 43 to a text input region 44.” Ex-1014, ¶0061. Kwon further discloses “[i]n FIG. 4, the user input one tag, but **the user may input a plurality of tags**. For example, **when the user inputs ‘#Maltese #Dog’, both a tag ‘Maltese’ and a tag ‘Dog’ are input.**” *Id.* Once user input is submitted, Kwon discloses “[u]pon receiving the content 43 and the text, the content registerer 112 may register the content 43 and the text in the DB 114 of the apparatus 110. The content registerer 112 may **use the identifier included in the text to distinguish the user tag**, and assign the tag as a general tag of the content 43.” Ex-1014, ¶0063. Because Kwon’s system supports user inputs with multiple tags and includes a content register that identifies tags for registration in a database based on the presence of tag identifiers in the user input, it would have been obvious for Kwon’s system to **identify *multiple taxonomy tags from user input*** (multiple tags that each have a tag identifier, such as “#Maltese #Dog”).

246. As set forth above under the “Rationale (Motivation) to Combine” Section, in my opinion, it would have been obvious for a PHOSITA to further modify Junee’s system by implementing a process to identify multiple tags from a single user input where each tag is in hash tag form according to the teachings of Kwon. As explained above, a process to identify tags provides a way to

differentiate user inputs that includes tags from other user inputs that do not include tags. The support and identification of multiple tags in a single user input would further expedite the process of associating multiple tags with a video or a video compilation in Junee. By allowing multiple tags to be associated with the same video in Junee, the video and/or the video compilation would be more readily discoverable in response to different search queries, thus improving the quality and relevance of search results in response to a user's search query. In my opinion, a PHOSITA would have had a reasonable expectation of success because the modifications of Junee represent nothing more than the predictable and routine intake of user input using standard programming.

“[8e] associating, in the stored data, the first virtual clip with the one or more taxonomy tags to produce a searchable virtual clip that is delivered to a requesting device in response to a query from the requesting device for any of the plurality of virtual clips that are associated with the user profile and the one or more taxonomy tags.”

247. In my opinion, the combination of Junee, De Jong, Houh, and Syed renders obvious this claim element. In Syed's system, “[t]ags **proposed by the subscribers to be associated with various portions of the video content** are received, processed and integrated within a global ToC for the video.” Ex-1009, Abstract. As an example, Syed discloses that “[a] **first proposed label to be associated with a first portion of the video content is received by the**

annotation tool from the first subscriber. The first proposed label is incorporated within a global table of contents (ToC) for the video content by... **associating with the video content the first proposed label and a start time of the first portion of the video content.”** Ex-1009, ¶0013. Thus, Syed discloses **associating a first virtual clip** (portion of video content) with **one or more taxonomy tags** (tags/labels).

248. Syed further discloses the video player 300 shown in FIGS. 3A and 3B includes a “[s]earch text entry field 315 [that] allows the user to initiate a **keyword search**. If the user has chosen to perform a search, the user is then presented with results in the form of labels and/or **clips** by management platform.” Ex-1009, ¶0084. Syed further teaches “[s]ome embodiments allow the user to filter **annotations by the type of user** that provided it. For example, **users can filter by themselves**, everyone, friends, those automatically generated by the service provider and so forth.” Ex-1009, ¶0064. In Syed, a video clip may be “defined by a user by **proposing a label to be associated with a particular time offset within the video.**” Ex-1009, ¶0086. Thus, in Syed, filtering annotations, such as a label, by user type is equivalent to filtering clips by user type when the clips are defined by the labels. Thus, Syed teaches a **searchable virtual clip** (clip) that is delivered to a requesting device in response to a **query** (keyword search, filter) from the requesting device for **any of the plurality of virtual clips** (clips) that are

associated with the **user profile** (users that filter themselves) and the **one or more taxonomy tags** (tags/labels).

249. As set forth above under the “Rationale (Motivation) to Combine” Section, in my opinion, it would have been obvious for a PHOSITA to further modify Junee’s system by associating the taxonomy tags with a virtual clip (a video in the video compilation) to facilitate retrieval of the virtual clip in response to a query by a user of Junee’s system. In my opinion, it would have also been obvious for a PHOSITA to implement features that allow users to search for video content based on a particular user and the taxonomy tags to further improve the quality and relevance of search results in response to a user’s query. A PHOSITA would have had a reasonable expectation of success in modifying Junee’s system because Junee already teaches “[u]sers of clients 130 can also search for videos based on keywords, tags or other metadata” and “supports searching on any fielded data for a video, including its title, description, tags, author, category and so forth.” Ex-1005, 3:7-8, 3:12-14.

Ground 4. Claim 9 is obvious over Junee, De Jong, Houh, Syed, Giunio-Zorkin, Kwon, and further in view of Dayan

250. In my opinion, claim 9, which depends from claim 8, would have been obvious over Junee, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon as

discussed under Ground 3 and incorporated herein, further in view of U.S. Pat. 10,180,989 (“**Dayan**”) (Ex-1014).

251. Dayan is a U.S. patent that issued on January 15, 2019. Ex-1014. The application that led to the Dayan patent was filed on July 24, 2015. Ex-1014.

252. I understand that Dayan is prior art.

253. I further understand that Dayan was not cited during prosecution of any application leading to the ’442 patent.

A. Overview of the Ground

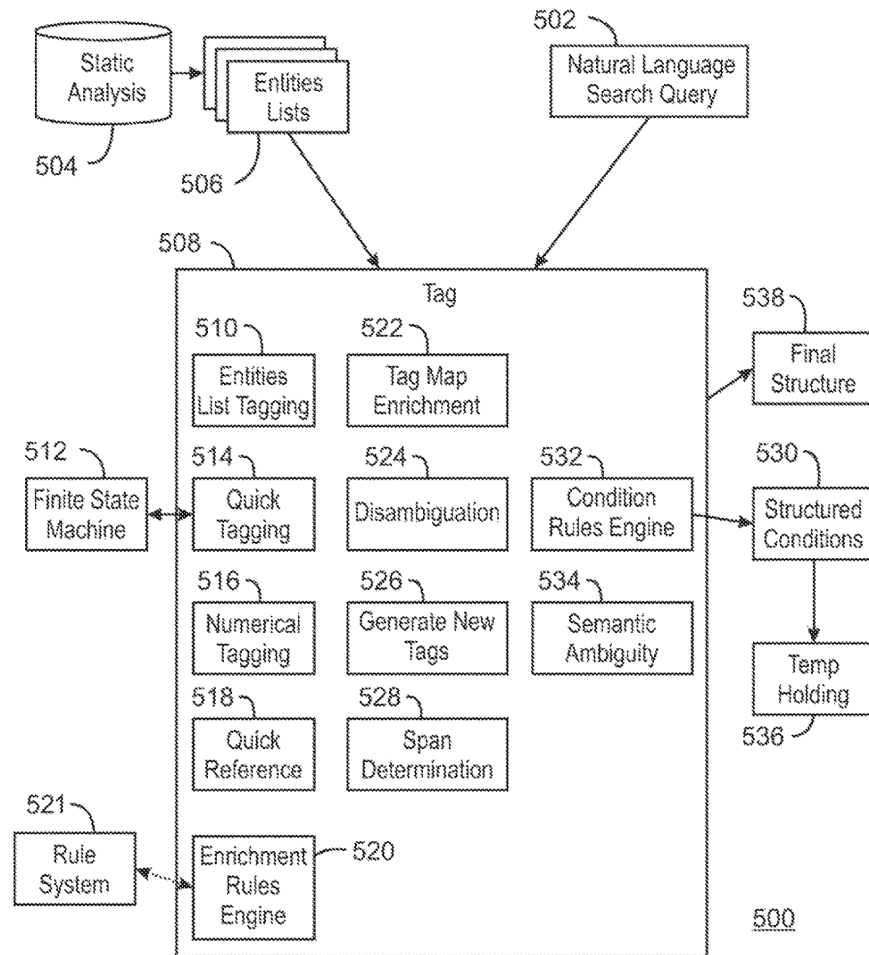
254. In my opinion, under a first interpretation that the term “directed relationship” recited in claim element [9a] does not imply a hierarchical relationship between the “first taxonomy tag” and the “second taxonomy tag,” but instead refers only to the sequence in which the first taxonomy tag and the second taxonomy tag are presented in the “second user input,” claim 9 would have been obvious over the combination of Junee, De Jong, Houh, Syed, Giunio-Zorkin, Kwon, and Dayan. Ex-1001, cl. 9. Dayan discloses and renders obvious creating a directed relationship between tags associated with a virtual clip and using an ordered combination of tags based on the directed relationship as a query to retrieve a virtual clip.

255. In my opinion, under a second interpretation that the term “directed relationship” implies a hierarchical relationship between the first taxonomy tag and

the second taxonomy tag, claim 9 would have been obvious over the combination of June, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon.

1. Overview of Dayan

256. Dayan is directed to “[t]echniques for generating query language statements for a document repository.” Ex-1014, Abstract. In particular, Dayan discloses “convert[ing] a natural language search query (also referred to as a search query) into formal constraints based on repository content and structure (as determined by the repository static analysis), domain knowledge, personal information, and rules.” Ex-1014, 2:60-64. Dayan’s system is illustrated in FIG. 5, reproduced below, which shows “a block diagram illustration of an example system for generating a query language statement.” Ex-1014, 2:46-47.



Ex-1014, FIG. 5. As shown, Dayan’s system includes a “tag module 508 [that] can generate a modified search query by **adding structured data (also referred to herein as tags) to the search query**. Tags can include any suitable terms or logical expressions that can be added to the natural language query. The tags can **improve the accuracy of the results returned by performing a search.**” Ex-1014, 11:5-10.

257. Dayan explains “the tag module 508 can use a rule engine 520 to generate a rule system 521 that **searches for sequences of tags** and/or original

words from the natural language search query 502. The rule system 521 can include any number of rules that can **detect adjacent tags that are at a given proximity from each other**. For example, a word list from among the word lists 506 may be ‘next_last.txt; containing the phrase ‘the last,’ or similar. A second word list from the word lists 506 may be ‘periods.txt’ as defined above. A rule (<Action>, ... last, next, Number_Cardinal, periods) may be added to the rule engine 520 that can find **word sequences such as ‘the next five weeks’ or ‘the last thirty days.’**” Ex-1014, 12:55-67. Dayan notes “a text word or phrase may correspond to several tags.” Ex-1014, 11:31-32.

B. Rationale (Motivation) Supporting Obviousness

258. In my opinion, under the first interpretation that the term “directed relationship” recited in claim element [9a] does not imply a hierarchical relationship between the “first taxonomy tag” and the “second taxonomy tag,” the combination of Junee, De Jong, Houh, Syed, Giunio-Zorkin, Kwon, and Dayan renders obvious claim 9.

259. Dayan’s system is directed to “convert[ing] a natural language search query (also referred to as a search query) into formal constraints” and using the “formal constraints...to generate a query language statement to retrieve and display documents from a document repository.” Ex-1014, 2:60-62, 2:64-67. Dayan teaches “the tag module 508 can use a rule engine 520 to generate a rule

system 521 that searches for **sequences of tags**...from the natural language search query 502. The rule system 521 can include any number of rules that can detect **adjacent tags that are at a given proximity from each other.**” Ex-1014, 12:55-60. Thus, Dayan discloses **a directed relationship between a first taxonomy tag and a second taxonomy tag sequentially following the first taxonomy tag** (sequence of tags) and **retrieving a first virtual clip** (documents from a document repository) using **an ordered combination of first and second taxonomy tags** (a sequence of tags) as a **query** (search query).

260. In my opinion, it would have been obvious for a PHOSITA to create a directed relationship between multiple tags according to Dayan’s teachings for association with a video in the video compilation of Junea to promote retrieval of the video in response to a query. It would have further been obvious for a PHOSITA to enable users to retrieve a video in the video compilation of Junea using a sequence of tags associated with the video in a query according to the teachings of Dayan. For instance, Dayan teaches: “Tags can include any suitable terms or **logical expressions** that can be added to the natural language query. The tags can **improve the accuracy of the results returned by performing a search.**” Ex-1014, 11:5-11. Moreover, directed relationships between tags, examples of which are disclosed in Dayan, provide semantic context, which can improve the quality and relevance of search results in response to a user’s search query.

261. Alternatively, the combination of Juneo, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon also renders obvious this claim element. Kwon further teaches “receiving, from a user terminal, content and a user tag” (wherein “the receiving comprises **receiving a plurality of user tags**”) and “registering the content and the received user tag” (wherein “registering comprises, when a number of the received plurality of user tags is higher than a threshold number, **assigning the threshold number of user tags**, from among the plurality of user tags, as general tags of the content **in an order in which the received plurality of tags are received**”). Ex-1012, cl. 1, 5. Thus, Kwon discloses **creating a directed relationship between a first taxonomy tag and a second taxonomy tag** (assignment of the sequence of user tags) based on **an order in the second user input of the one or more taxonomy tags** (order in which the tags are received).

262. In my opinion, it would have also been obvious for a PHOSITA to create a directed relationship between multiple tags according to Kwon’s teachings for association with a video in the video compilation of Juneo to promote retrieval of the video in response to a query because, as explained above, this would allow users to produce down search results that are more relevant to their query.

263. In my opinion, under the second interpretation that the term “directed relationship” recited in claim element [9a] does imply a hierarchical relationship between the “first taxonomy tag” and the “second taxonomy tag,” the combination

of Junee, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon renders obvious claim 9.

264. Giunio-Zorkin teaches “[c]ontent items are associated with data indicative of a subset of one or more of the keywords selected from the plurality of keyword sets” where “each keyword set includes a hierarchical sequence of keywords.” Ex-1013, 12:30-32, 11:66-67. As explained under claim element [8b], Giunio-Zorkin discloses keyword metadata may be in the form of “**a taxonomy of tags.**” Ex-1013, 11:59. Giunio-Zorkin further discloses “keyword sets **may intersect, i.e. have one or more of the same keywords in common,** but the appearance of those keywords in each keyword set sequence may vary.” Ex-1013, 12:24-27. Giunio-Zorkin discloses “[s]earch requests, which may be referred to as explicit search requests, specify a query...to cause the system 100 to return specific content items. Search requests may be **manually generated by the user.**” Ex-1013, 22:20-23. Thus, Giunio-Zorkin discloses **a directed relationship between a first taxonomy tag and a second taxonomy tag sequentially following the first taxonomy tag** (multiple, different keyword sets for association with a content item where the keyword sets are in the form of a taxonomy tag and intersect one another) and **retrieving a first virtual clip** (content item) using **an ordered combination of first and second taxonomy tags** (multiple keywords from a plurality of keyword sets) as a **query** (manual query).

265. In my opinion, it would have been obvious for a PHOSITA to create a directed relationship between multiple tags according to Giunio-Zorkin's teachings for association with a video in the video compilation of Junee to promote retrieval of the video in response to a query. It would have been further obvious for a PHOSITA to enable users to retrieve a video in the video compilation of Junee using multiple keyword sets associated with the video in a query where the keyword sets intersect one another according to the teachings of Giunio-Zorkin.

266. In my opinion, under both interpretations, these modifications represent nothing more than the predictable combination of elements taught by Junee, De Jong, Houh, Syed, Giunio-Zorkin, Kwon and, to the extent necessary, Dayan. A PHOSITA would have been able to combine the elements taught by the combination of Junee, De Jong, Houh, Syed, and Kwon with elements 9[a] and [9b] taught by Giunio-Zorkin and, to the extent necessary, Dayan using standard computer programming. Furthermore, the respective functions of the elements taught by Junee, Giunio-Zorkin, and Dayan would not change when combined together: users would be able to specify and use the directed relationships according to the teachings of Giunio-Zorkin and, to the extent necessary, Dayan, without altering operation of Junee's system.

267. Further, these modifications represent the use of a known technique to improve Junee's system based on a similar system taught by Giunio-Zorkin and, to

the extent necessary, Dayan in the same way. As explained above, Junea discloses a system that allows users to associate multiple tags with a video in a video compilation. Giunio-Zorkin teaches a comparable system that allows users to associate multiple keyword sets with a content item and queries containing all the keywords from multiple keyword sets. Dayan teaches a comparable system that considers sequences of tags in a query when searching for video content. In my opinion, a PHOSITA could have modified Junea's computing system to incorporate the improvements in Giunio-Zorkin's system and, to the extent necessary, Dayan's system using nothing more than standard computer programming.

C. Reasonable Expectation of Success

268. In my opinion, a PHOSITA would have had a reasonable expectation of success in modifying Junea to implement search queries that included ordered sequences of tags based on Dayan's teachings, using nothing more than standard computing programming (including Dayan's programming executing on sequences of tags in a query when searching for video content), with each reference in the combination performing the same respective function when combined as they perform separately in the prior art, with no unpredictable results.

269. The art in the relevant timeframe was predictable. Specifically, in the relevant timeframe (no earlier than 2016), search queries with ordered sequences of words or phrases like those disclosed in Dayan were well-known in the art.

270. In particular, in my opinion, the '442 patent's general and high-level disclosures support this understanding of the relevant art. *See, e.g.*, Ex-1001, 3:65-7:6, 7:38-8:5, 9:25-9:59.

D. Graham Factors

271. The Graham factors were discussed under Ground 1.

E. Analogous Art

272. Dayan is analogous art because it is in the same field as the '442 patent (relating to “assemblies, methods, devices, and systems for managing information” generally, which includes “sharing and annotating content,” Ex-1001, 1:18-22), and its methods would have been reasonably pertinent to the problems facing the named inventors, specifically retrieval of content in response to a query. Ex-1014, Title, Abstract; Ex-1001, 14:30-36.

F. Claim Mapping

273. This section maps the challenged claims to the relevant disclosures of Junee, De Jong, Houh, Syed, Giunio-Zorkin, Kwon, and Dayan, where the claim text appears in bold-italics, and the relevant mapping follows.

CLAIM 9

“9[a]. The method of claim 8, wherein associating the first virtual clip with the one or more taxonomy tags comprises creating, based on an order in the second user input of the one or more taxonomy tags, a directed relationship between a first taxonomy tag and a second taxonomy tag sequentially following the first taxonomy tag in the character string,”

274. As a preliminary matter, the meaning of the term “directed relationship” in this claim element is unclear; hence, this claim element lacks clarity. I note the ’442 patent does not provide any insight into the meaning of “directed relationship” with the term only appearing in the specification where the language of claim 9 is presented as prose without any further elaboration. Ex-1001, 14:27-36.

275. A first interpretation is that the directed relationship does not imply a hierarchical relationship exists between the first taxonomy tag and the second taxonomy tag. Instead, the directed relationship refers only to the sequence in which the first taxonomy tag and the second taxonomy tag are presented in the second user input regardless of their content, as suggested by the language “based on an **order** in the second user input of the one or more taxonomy tags. Ex-1001, cl. 9. Associating a virtual clip with the sequence of the first and second taxonomy tags implies the virtual clip can be retrieved if the user’s query contains both taxonomy tags in the exact same sequence, as suggested by the language of claim element [9b], “using an **ordered combination** of the first and second taxonomy

tags as the query.” This is analogous to a conventional phrase search where a sequence of keywords enclosed by quotation marks causes a search engine to search for results containing only that exact sequence of keywords. Under this interpretation, the first taxonomy tag and the second taxonomy would each have a tag identifier, a first tag, and one or more subtags.

276. A second interpretation is that the directed relationship implies a hierarchical relationship between the first taxonomy tag and the second taxonomy tag. In taxonomy, directed relationships typically include, for example, a genus and a species, or a whole and a part. Under this interpretation, however, the first taxonomy tag and the second taxonomy tag would serve a similar function as the first tag and the subtags in a single taxonomy tag. In its discussion of the taxonomy tag format, the '442 patent discloses “[i]n one embodiment, each tag immediately following a tag identifier corresponds to a main category, and each subtag corresponds to a sub-category of the (sub)category corresponding to the immediately preceding tag (i.e., the tag to the left of the delimiter). Thus, one or more categories are arranged into a hierarchy determined from a sequence of the corresponding tags identified in the user input.” Ex-1001, 13:67-14:7. To reconcile this overlap, the directed relationship is construed to mean the first taxonomy tag and the second taxonomy tag share common first tags and/or subtags and/or the first tag and subtags of the first taxonomy tag have a hierarchical relationship with

one or more of the first tag and subtags of the second taxonomy tag and vice versa. For example, a first taxonomy tag of “Drinks” and a second taxonomy tag of “Drinks; Alcohol; Wine” would have a directed relationship because they share the first tag “Drinks” under this interpretation. For example, “Drinks” is a grandparent of “Drinks; Alcohol; Wine” whereas “Drinks; Alcohol; Wine” is a grandchild of “Drinks.” For purposes of querying a virtual clip, the first taxonomy tag and the second taxonomy tag are considered separate search terms connected by a Boolean OR operator, i.e., the query retrieves results based on whether a virtual clip is associated with the first taxonomy tag or the second taxonomy tag.

277. Because the term “directed relationship” is subject to multiple interpretations that significantly affect the scope of this claim, and the specification does not provide any clarity to a PHOSITA about where the boundaries of “directed relationship” are, I submit claim 9 is indefinite under 35 U.S.C. § 112. Notwithstanding the lack of clarity, in my opinion, this claim nevertheless would have been obvious under either the first interpretation or the second interpretation. The first and second interpretations are discussed in turn below.

278. In my opinion, under the first interpretation, the combination of June, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon further in view of Dayan renders obvious this claim element. Dayan discloses a “tag module 508 can generate a modified search query by adding structured data (also referred to herein as tags) to

the search query.” Ex-1014, 11:5-7. Dayan further teaches: “the tag module 508 can use a rule engine 520 to generate a rule system 521 that searches for **sequences of tags**...from the natural language search query 502. The rule system 521 can include any number of rules that can detect **adjacent tags that are at a given proximity from each other**. For example, a word list from among the word lists 506 may be ‘next_last.txt’ containing the phrase ‘the last,’ or similar. A second word list from the word lists 506 may be ‘periods.txt’ as defined above. A rule (<Action>, . . . last, next, Number_Cardinal, periods) may be added to the rule engine 520 that can **find word sequences such as ‘the next five weeks’ or ‘the last thirty days.’**” Ex-1014, 12:55-67. Thus, Dayan discloses **a directed relationship between a first taxonomy tag and a second taxonomy tag sequentially following the first taxonomy tag** (sequence of tags).

279. In my opinion, it would have been obvious for a PHOSITA to create a directed relationship between multiple tags according to Dayan’s teachings for association with a video in the video compilation of Junea to promote retrieval of the video in response to a query by giving users the ability to produce search results that are more relevant to their query. For instance, Dayan teaches: “Tags can include any suitable terms or **logical expressions** that can be added to the natural language query. The tags can **improve the accuracy of the results returned by performing a search.**” Ex-1014, 11:5-11. I further note the storage

of relationships between tags was common at the relevant time. For example, Giunio-Zorkin also teaches “[c]ontent items are associated with data indicative of a subset of one or more of the keywords selected from the plurality of keyword sets” where “each keyword set includes a hierarchical sequence of keywords.” Ex-1013, 12:30-32, 11:66-67.

280. Alternatively, the combination of Junea, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon also renders obvious this claim element. Kwon further teaches “receiving, from a user terminal, content and a user tag” (wherein “the receiving comprises **receiving a plurality of user tags**”) and “registering the content and the received user tag” (wherein “registering comprises, when a number of the received plurality of user tags is higher than a threshold number, **assigning the threshold number of user tags**, from among the plurality of user tags, as general tags of the content **in an order in which the received plurality of tags are received**”). Ex-1012, cl. 1, 5. Thus, Kwon discloses **creating a directed relationship between a first taxonomy tag and a second taxonomy tag** (assignment of the sequence of user tags) based on **an order in the second user input of the one or more taxonomy tags** (order in which the tags are received).

281. In my opinion, it would have also been obvious for a PHOSITA to create a directed relationship between multiple tags according to Kwon’s teachings for association with a video in the video compilation of Junea to promote retrieval

of the video in response to a query because, as explained above, this would allow users to produce search results that are more relevant to their query.

282. In my opinion, under the second interpretation, the combination of Junee, De Jong, Houh, Syed, Giunio-Zorkin, and Kwon renders obvious this claim element. As explained above, Giunio-Zorkin teaches associating content items with at least a subset of one or more of the keywords selected from a plurality of keyword sets. Ex-1013, 12:30-32. As explained under claim element [8b], Giunio-Zorkin discloses keyword metadata may be in the form of “**a taxonomy of tags.**” Ex-1013, 11:59. Giunio-Zorkin further discloses “keyword sets **may intersect, i.e. have one or more of the same keywords in common**, but the appearance of those keywords in each keyword set sequence may vary.” Ex-1013, 12:24-27.

283. The intersection between keyword sets in Giunio-Zorkin implies a hierarchical relation can exist between the keyword sets. Referring to the examples disclosed in Giunio-Zorkin, if a first keyword set includes the keywords “Drinks” and a second keyword set includes the keywords “Drinks; Alcohol; Wine,” the first and second keyword sets intersect because they have the common keyword of “Drinks.” A hierarchical relationship also exists between the first and second keyword sets since the first keyword set includes a genus (“Drinks”) and the second key word set includes a species (“Wine”). Thus, Giunio-Zorkin discloses **a directed relationship between a first taxonomy tag and a second taxonomy tag**

sequentially following the first taxonomy tag (multiple, different keyword sets for association with a content item where the keyword sets are in the form of a taxonomy tag and intersect one another).

284. In my opinion, it would have been obvious for a PHOSITA to create a directed relationship between multiple tags according to Giunio-Zorkin's teachings for association with a video in the video compilation of Junee to promote retrieval of the video in response to a query.

“[9b] the directed relationship enabling a user to retrieve the first virtual clip from the stored data using an ordered combination of the first and second taxonomy tags as the query.”

285. See claim element 9[a]. Under the first interpretation of the term “directed relationship” discussed under claim element 9[a], the combination of Junee, De Jong, Houh, Syed, Kwon, Giunio-Zorkin, and Dayan renders obvious this claim element. Dayan's system is directed to “convert[ing] a natural language search query (also referred to as a search query) into formal constraints” and using the “formal constraints...to generate a query language statement to retrieve and display documents from a document repository.” Ex-1014, 2:60-62, 2:64-67. Thus, Dayan discloses **retrieving a first virtual clip** (documents from a document repository) using **an ordered combination of first and second taxonomy tags** (a sequence of tags) as a **query** (search query). In my opinion, it would have been obvious for a PHOSITA to enable users to retrieve a video in the video

compilation of Junee using a sequence of tags associated with the video in a query according to the teachings of Dayan.

286. In my opinion, under the second interpretation of the term “directed relationship” discussed under claim element 9[a], the combination of Junee, De Jong, Houh, Syed, Kwon, and Giunio-Zorkin renders obvious this claim element. Giunio-Zorkin discloses “[s]earch requests, which may be referred to as explicit search requests, specify a query...to cause the system 100 to return specific content items. Search requests may be **manually generated by the user.**” Ex-1013, 22:20-23. Giunio-Zorkin further discloses an example method involving the steps of: “receiving...a request for content from a requestor via the network [Block 308]; **identifying...based on the received request, one or more keywords selected from the plurality of keyword sets** [Block 310].” Ex-1013, 30:38-42. Thus, Giunio-Zorkin discloses **retrieving a first virtual clip** (content item) using **an ordered combination of first and second taxonomy tags** (multiple keywords from a plurality of keyword sets) as a **query** (manual query). In my opinion, it would have been obvious for a PHOSITA to enable users to retrieve a video in the video compilation of Junee using multiple keyword sets associated with the video in a query where the keyword sets intersect one another according to the teachings of Giunio-Zorkin.

Ground 5. Claims 10 and 11 are obvious over Syed in view of Giunio-Zorkin and Kwon

287. In my opinion, independent claim 10 and dependent claim 11 would have been obvious over Syed (Ex-1009) in view of Giunio-Zorkin (Ex-1013) and Kwon (Ex-1012).

A. Overview of the Ground

288. In my opinion, the challenged claims would have been obvious over the combination of Syed, Giunio-Zorkin, and Kwon. Syed discloses or renders obvious most of the elements of the challenged claims. To the extent Syed does not disclose or render obvious limitations directed to determining a user input conforms to a predetermined format defining taxonomy tags, identifying multiple taxonomy tags from user input, and associating taxonomy tags to a virtual clip in an account of the user, these limitations would have been obvious over Syed in view of Kwon and Giunio-Zorkin. Giunio-Zorkin teaches a taxonomy tag format with a first tag and multiple subtags separated by a delimiter, and associating the taxonomy tags with user profiles to facilitate retrieval of virtual clips associated with the user profile and the taxonomy tags. Kwon teaches identifying multiple tags from user input.

1. Overview of Syed

289. *See* discussion of Syed under Ground 3. Syed discloses “[m]ethods and systems...for **generation of an index for video content based on user input**

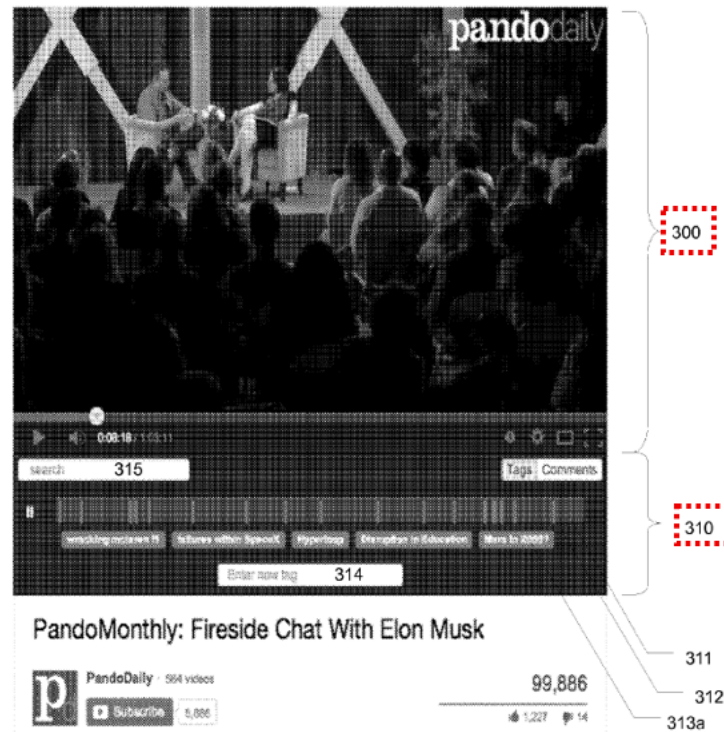
(e.g., crowd-sourced input).” Ex-1009, ¶0031. In Syed’s system, “**a video is displayed to a user** (e.g., a subscriber) of a video consumption and discovery service via a video player running within a web browser.” *Id.* Syed further teaches that “[c]oncurrently with the presentation of the video, **an associated table of contents (ToC), if available, may also be displayed to the user.** The ToC may be displayed within **a separate annotation tool or may be integrated within the video player.** The annotation tool (either integrated within the video player or separate) **allows users to specify content and placement of a label or tag (e.g., a short description of a video segment) within a video** while watching the video.” *Id.*

290. Syed states “[t]he term ‘annotation’ broadly refers to **user-supplied or automatically generated content that is associated with a portion of video content or a video ‘clip’** (e.g., one or more frames of video content corresponding to a particular time or period of time within a video). Depending upon the particular implementation, annotations may include, but are not limited to, one or more of labels, tags, comments and additional content.” Ex-1009, ¶0037.

291. Syed further states “[t]he term ‘clip’ generally refers to a continuous portion or segment of a video having a start time and an end time. A clip may have one or more annotations associated therewith. In some embodiments, users may

modify the start and/or end time of the clip. In one embodiment, clips may be shared with other users and a clip can be part of a playlist.” Ex-1009, ¶0040.

292. As an example, FIG. 3A of Syed, reproduced below, “illustrate a video player 300, an annotation tool 310 and a table of contents 313 a-n for a video.” Ex-1009, ¶0083.



Ex-1009, FIG. 3A (red highlighting added). As shown, “video consumption and discovery service 100 may present the user with a video selected for viewing by the user within video player 300 concurrently with annotation tool 310.” *Id.*

2. Overview of Giunio-Zorkin

293. See discussion of Giunio-Zorkin under Ground 3. Giunio-Zorkin is directed to “systems and methods for efficiently and automatically distributing

content among a plurality of users which may be used to implement a social networking service.” Ex-1013, 4:46-49. Giunio-Zorkin discloses “**a taxonomy of tags**, e.g. keyword metadata or keywords, is provided...the available keyword metadata of the system 100 is organized into a plurality of keyword sets/hierarchies, where each keyword set includes **a hierarchical sequence of keywords**, wherein for each keyword set a first keyword of the sequence of keywords defines semantics or a context...and each subsequent keyword of the sequence of the keyword set defines a narrowing of the semantics/context over the previous keyword of the sequence of the keyword set. Ex-1013, 11:59-12:6. Giunio-Zorkin further discloses “**[u]sers are also associated with a subset of one or more of the keywords** selected from [sic] the plurality of keyword sets wherein the keywords are selected, explicitly or implicitly by user and/or system 100, based on their interests.” Ex-1013, 12:47-50.

3. Overview of Kwon

294. *See* discussion of Kwon under Ground 3. Kwon is directed to “methods, apparatuses, and/or computer programs for managing content.” Ex-1012, ¶0003. Kwon discloses “[t]he user may input a tag to be assigned to the content 43 by using a certain identifier. In FIG. 4, the certain identifier may be a hash tag ‘#’. The user inputs a word together with the identifier ‘#’ to input the word as the tag of the content 43. In FIG. 4, the user input ‘Maltese’ as the tag. **In**

FIG. 4, the user input one tag, but the user may input a plurality of tags. For example, when the user inputs ‘#Maltese #Dog’, both a tag ‘Maltese’ and a tag ‘Dog’ are input.” Ex-1012, ¶0061.

B. Rationale (Motivation) Supporting Obviousness

295. In my opinion, the combination of Syed, Giunio-Zorkin, and Kwon renders obvious the elements of independent claim 10 of the ’442 patent. The combination of Syed, Giunio-Zorkin, and Kwon further renders obvious dependent claim 11, which depends from claim 10 and is directed to arranging the one or more taxonomy tags according to a sequence in the user input of the one or more taxonomy tags.

296. To the extent Syed does not disclose or render obvious determining a user input conforms to a predetermined format defining taxonomy tags (claim elements [10e] and [10h]), identifying multiple taxonomy tags from user input (claim element [10f]), and associating taxonomy tags to a virtual clip in an account of the user (claim element [10i]), it would have been obvious for a PHOSITA to incorporate such features into a system such as Syed’s in view of Giunio-Zorkin’s and Kwon’s teachings.

Determining user input conforms to a predetermined format defining taxonomy tags [10e], [10h]

297. To the extent the term “taxonomy tag” is construed to require support of a first tag and one or more subtags identifying various categories separated by a

delimiter, Syed in view of Giunio-Zorkin renders obvious these claim elements.

Giunio-Zorkin discloses “keyword metadata,” which are “**keywords or tags associated with the content items.**” Ex-1013, 11:15-16. Giunio-Zorkin teaches the keyword metadata may be in the form of “**a taxonomy of tags**” where “the available keyword metadata...is organized into a plurality of keyword sets/hierarchies, where **each keyword set includes a hierarchical sequence of keywords**, wherein for each keyword set **a first keyword of the sequence of keywords defines semantics or a context...**and **each subsequent keyword of the sequence of the keyword set defines a narrowing of the semantics/context** over the previous keyword of the sequence of the keyword set. For example, one keyword set may include the sequence of keywords: Drinks; Alcohol; Wine; White; Pino Grigio; California; Berringer.” Ex-1013, 11:59-12:8.

298. A PHOSITA would have recognized each of the examples of keyword sets disclosed in Giunio-Zorkin includes commonly-used delimiters. *See* Ex-1019 at 152 (defining delimiter and providing examples of commonly used delimiters, including semi-colon and colon). Thus, Giunio-Zorkin teaches a predetermined format defining taxonomy tags that includes **one or more subtags** (the keywords in a keyword set that follow the first keyword) **sequentially following a first tag** (the first keyword in a keyword set) and **delimited by a delimiter character** (the delimiters between the keywords in the “Drinks” example).

299. In my opinion, it would have been obvious for a PHOSITA to modify Syed's system to include the taxonomy tags according to Giunio-Zorkin's teachings "for the purpose of distributing and/or facilitating discovery of that content item." Ex-1013, 11:6-7. Giunio-Zorkin teaches "content items are associated with one or more interest classifications" and its system "maps or otherwise correlates the interests of each user against the interest classification(s) of content items to determine which content items to transmit to a user." Ex-1013, 11:3-4, 11:10-13. Thus, the addition of taxonomy tags that contain multiple keywords in a keyword set as taught by Giunio-Zorkin would facilitate retrieval of videos in Syed's system that are more aligned to a requesting user's interests.

300. These modifications represent nothing more than the predictable combination of elements taught by Syed and Giunio-Zorkin. In my opinion, a PHOSITA would have been able to combine the elements taught by Syed with elements [10e] and [10h] taught by Giunio-Zorkin using standard computer programming. Furthermore, the respective functions of the elements taught by Syed and Giunio-Zorkin would not change when combined together: users would still be able to propose a new label for a portion of a video via a label text entry field according to Syed's teachings, and users would be able propose labels using keyword sets in the form of a taxonomy of tags according to Giunio-Zorkin's teachings.

301. Further, these modifications represent the use of a known technique to improve Syed's system based on a similar system taught by Giunio-Zorkin in the same way. As explained above, Syed discloses a system that supports taxonomy tags with a single tag. Giunio-Zorkin teaches a comparable system that supports a taxonomy tag (keyword sets with a hierarchical sequence of keywords separated by delimiters). In my opinion, a PHOSITA could have modified Syed's computing system to incorporate the improvements in Giunio-Zorkin's system using nothing more than standard computer programming.

Identifying one or more taxonomy tags from the user input [10f]

302. To the extent this claim element is construed to require the capability to identify *multiple* taxonomy tags in the user input, then this claim element is obvious over the combination of Syed and Kwon. Kwon discloses "[i]n FIG. 4, the user input one tag, but **the user may input a plurality of tags**. For example, **when the user inputs '#Maltese #Dog', both a tag 'Maltese' and a tag 'Dog' are input.**" *Id.* Once user input is submitted, Kwon discloses "[u]pon receiving the content 43 and the text, the content registerer 112 may register the content 43 and the text in the DB 114 of the apparatus 110. The content registerer 112 may **use the identifier included in the text to distinguish the user tag**, and assign the tag as a general tag of the content 43." Ex-1014, ¶0063. Because Kwon's system supports user inputs with multiple tags and includes a content register that

identifies tags for registration in a database based on the presence of tag identifiers in the user input, it would have been obvious for Kwon's system to **identify multiple taxonomy tags from user input** (multiple tags that each have a tag identifier, such as "#Maltese #Dog").

303. In my opinion, it would have been obvious for a PHOSITA to further modify Syed's system by implementing a process to identify multiple tags or labels from a single user input where each tag or label is in hash tag form according to the teachings of Kwon. A process to identify tags provides a way to differentiate user inputs that includes tags from other user inputs that do not include tags. The support and identification of multiple tags in a single user input would further expedite the process of associating multiple tags with a video or a video compilation in June.

304. These modifications represent nothing more than the predictable combination of elements taught by Syed and Kwon. In my opinion, a PHOSITA would have been able to combine the elements taught by Syed with element [10f] taught by Kwon using standard computer programming. Furthermore, the respective functions of the elements taught by Syed and Kwon would not change when combined together: user inputs would be processed to assess whether one or multiple tags are present according to Syed's teachings and Kwon's teachings, respectively.

305. Further, these modifications represent the use of a known technique to improve Syed's system based on a similar system taught by Kwon in the same way. Syed discloses a system that identifies one tag from user input. Giunio-Zorkin teaches a comparable system that identifies multiple tags from user input. In my opinion, a PHOSITA could have modified Syed's computing system to incorporate the improvements in Kwon's system using nothing more than standard computer programming.

Associating the virtual clip with each of the one or more taxonomy tags in an account of the user [10i]

306. In Syed's system, "[t]ags **proposed by the subscribers to be associated with various portions of the video content** are received, processed and integrated within a global ToC for the video." Ex-1009, Abstract. Thus, Syed discloses **associating the virtual clip** (video clips) with **one or more taxonomy tags** (tags/labels).

307. Giunio-Zorkin discloses "[u]sers **are also associated with a subset of one or more of the keywords** selected from [sic] the plurality of keyword sets wherein the keywords are selected, explicitly or implicitly by user and/or system 100, based on their interests." Ex-1013, 12:47-50. Giunio-Zorkin further discloses "the keyword sets are stored in an interest database or data structure stored in the data store 124 wherein each keyword is associated with a unique identifier.

Associations between users or content items and keywords are implemented

by storing the associated identifiers of the selected keywords in association therewith.” Ex-1013, 13:30-35. Thus, Giunio-Zorkin discloses associating, in **an account of the user** (user), **the virtual clip** (content item) with **each of the one or more taxonomy tags** (keywords from keyword sets).

308. In my opinion, it would have been obvious for a PHOSITA to modify Syed based on the teachings of Giunio-Zorkin to store tags associated with video clips in an account of the user in order to provide media content to the user that is more aligned with their interests. For instance, Giunio-Zorkin teaches “content items are associated with one or more interest classifications” and its system “maps or otherwise correlates **the interests of each user against the interest classification(s) of content items to determine which content items to transmit to a user.**” Ex-1013, 11:3-13.

309. These modifications represent nothing more than the predictable combination of elements taught by Syed and Giunio-Zorkin. In my opinion, a PHOSITA would have been able to combine the elements taught by Syed with element [10i] taught by Giunio-Zorkin using standard computer programming. Furthermore, the respective functions of the elements taught by Syed and Giunio-Zorkin would not change when combined together: users would be able to associate annotations for different portions of a video according to Syed’s

teachings, and users would be able to associate taxonomy tags with content and user accounts according to Giunio-Zorkin's teachings.

310. Further, these modifications represent the use of a known technique to improve Syed's system based on a similar system taught by Giunio-Zorkin in the same way. As explained above, Syed discloses a system that associates annotations with video content. Giunio-Zorkin teaches a comparable system that allows multiple tags (keywords in a keyword set) to be associated with a virtual clip (a content item) and a user. In my opinion, a PHOSITA could have modified Syed's computing system to incorporate the improvements in Giunio-Zorkin's system using nothing more than standard computer programming.

C. Reasonable Expectation of Success

311. In my opinion, a PHOSITA would have had a reasonable expectation of success in modifying Syed to incorporate annotations that categorize content according to a particular tag format based on the teachings of Giunio-Zorkin and Kwon.

312. The art in the relevant timeframe was predictable. Specifically, in the relevant timeframe (no earlier than 2016), basic formats for annotations like those disclosed in Syed, Giunio-Zorkin, and Kwon were well-known in the art.

313. In particular, in my opinion, the '442 patent's general and high-level disclosures support this understanding of the relevant art. *See, e.g.*, Ex-1001, 3:65-7:6, 7:38-8:5, 9:25-9:59.

D. Graham Factors

314. The Graham factors were discussed under Ground 1 and apply equally to all grounds.

E. Analogous Art

315. Syed, Giunio-Zorkin, and Kwon are analogous art for the same reasons discussed under Ground 3 and incorporated herein.

F. Claim Mapping

316. This section maps the challenged claims to the relevant disclosures of Syed, Giunio-Zorkin, and Kwon, where the claim text appears in bold-italics, and the relevant mapping follows.

CLAIM 10

“10[a]. A method of annotating a playable media file, the method comprising:”

317. To the extent the preamble is limiting, Syed discloses and renders obvious **a method of annotating a playable media file**. Syed is directed to “[m]ethods and systems...for **generation of an index for video content based on user input** (e.g., crowd-sourced input).” Ex-1009, ¶0031. In Syed's system, **“a video is displayed to a user** (e.g., a subscriber) of a video consumption and

discovery service via a video player running within a web browser.” *Id.* Syed further teaches that “[c]oncurrently with the presentation of the video, **an associated table of contents (ToC), if available, may also be displayed to the user.** The ToC may be displayed within **a separate annotation tool or may be integrated within the video player.** The annotation tool (either integrated within the video player or separate) **allows users to specify content and placement of a label or tag (e.g., a short description of a video segment) within a video** while watching the video.” *Id.*

318. Syed further teaches that “[t]he term ‘annotation’ broadly refers to **user-supplied or automatically generated content that is associated with a portion of video content or a video ‘clip’** (e.g., one or more frames of video content corresponding to a particular time or period of time within a video). Depending upon the particular implementation, **annotations** may include, but are not limited to, one or more of **labels, tags, comments** and additional content.” Ex-1009, ¶0037. Thus, Syed discloses **a method of annotating a playable media file** (video content).

“[10b] obtaining a virtual clip comprising a first location within the playable media file and a second location within the playable media file, the first and second locations together defining a clip of the playable media file occurring between the first and second locations;”

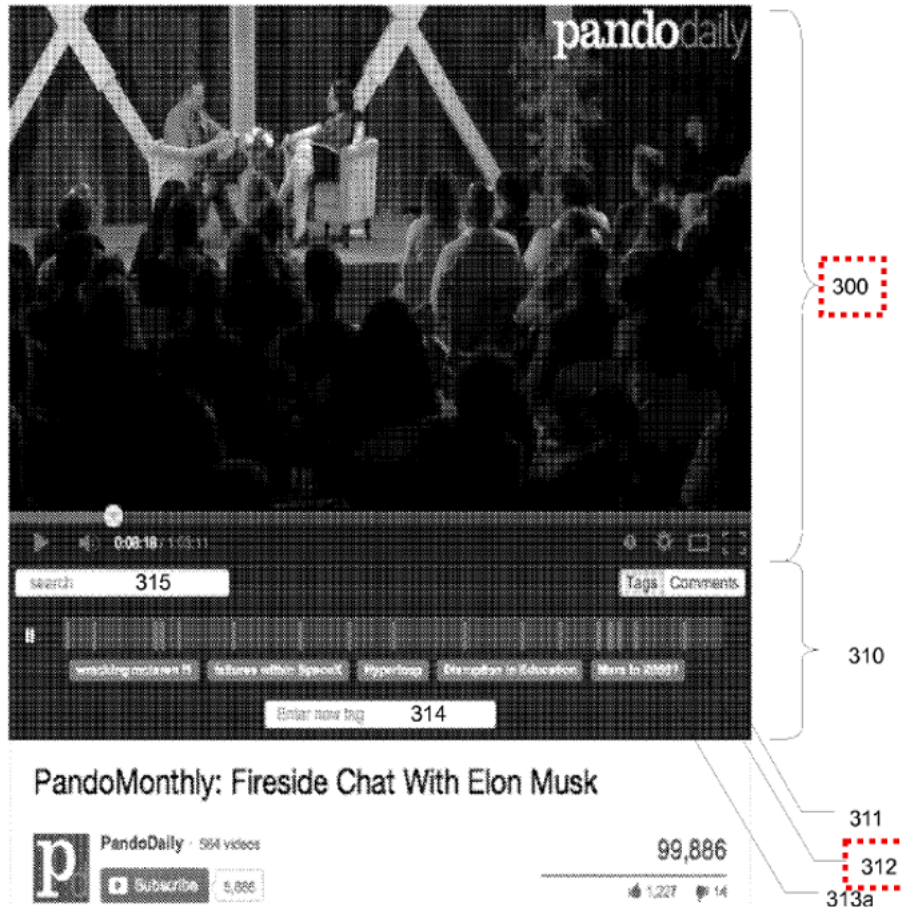
319. Syed discloses this claim element. The '442 patent describes a “virtual clip” as “a set of data points that together delineate a particular subset of the content encoded in the media file”; “the virtual clip is comprised of references that identify specific content in the corresponding media file, but the virtual clip **is not necessarily itself a stored media file containing the content data.**” Ex-1001, 9:60-66. Thus, a virtual clip may optionally include content from a media file according to the '442 patent's teachings. *See* Ex-1001, 11:25-27 (“the user may elect to create a plurality of virtual clips comprising content from one or more Media File”); In Syed, “[t]he term ‘clip’ generally refers to **a continuous portion or segment of a video having a start time and an end time.**” Ex-1009, ¶0040. Thus, Syed discloses a virtual clip, i.e., **references that identify the content in a video** (a start time, an end time, a video from which a clip is created), comprising **a first location** (a start time) and **a second location** (an end time) **within a playable media file** (a video).

320. Syed further teaches that users of the frontend systems 120 **obtain video clips** from the backend systems 140. In Syed, “[f]rontend systems 120 generally include functional modules with which users (e.g., subscribers of video discovery and consumption service) interact and through which they view videos.” Ex-1009, ¶0049. The “frontend systems 120 include three subsystems, an annotation tool 121, a management platform 122 and a video player 123.” Ex-

1009, ¶0056. Syed explains the “[m]anagement platform 122 may allow users to **view their video libraries, clips they have created.**” Ex-1009, ¶0067. The “backend systems 140 comprise several sub components including an application programming interface (API) 143.” Ex-1009, ¶0050. Syed discloses the “API 143 is an interface to backend systems 140 provided to video player 123, management platform 122 and annotation tool 121 that allows access to storage, retrieval and/or search functionality. For example, API 143 may provide **methods to retrieve information**, such as information regarding users, videos, annotations/labels, **video clips and playlists.**” Ex-1009, ¶0051.

“[10c] causing, using the virtual clip, the clip to be displayed on a computing device of a user;”

321. Syed discloses this claim element. As explained under claim element [10b], “[f]rontend systems 120 generally include functional modules with which users (e.g., subscribers of video discovery and consumption service) interact and through which they **view videos.**” Ex-1009, ¶0049. As an example, FIG. 3A of Syed, reproduced below with annotations, shows “**a video player 300**, an annotation tool 310 and a table of contents 313 a-n for a video.” Ex-1009, ¶0083.



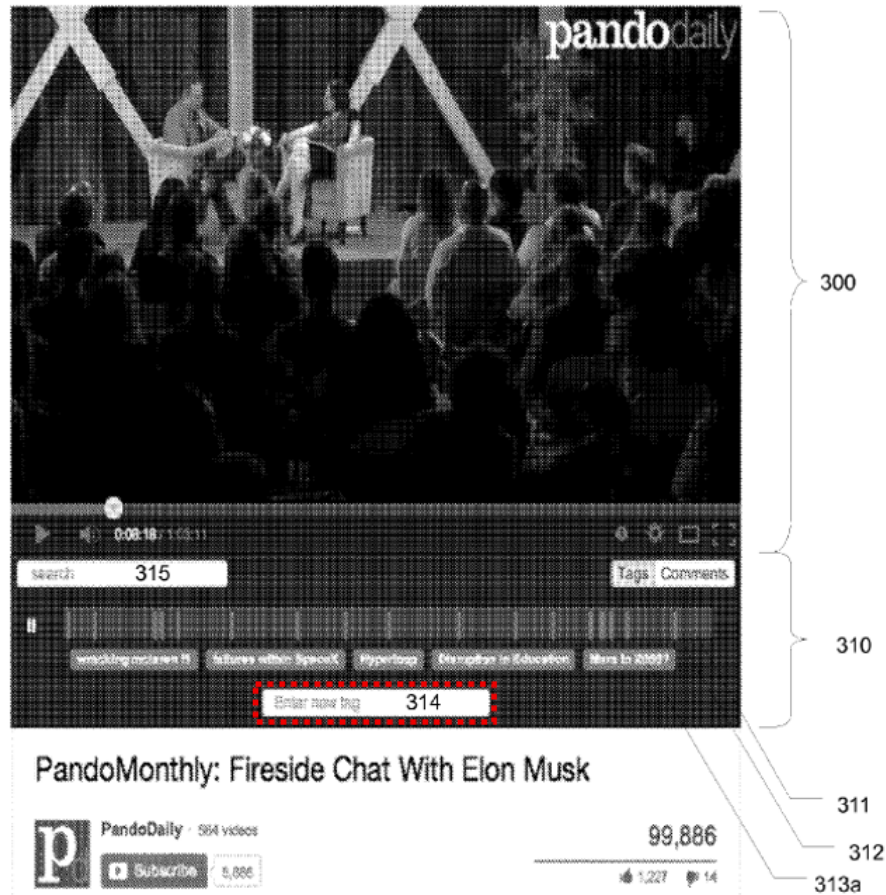
Ex-1009, FIG. 3A (red highlighting added). As shown, “video consumption and discovery service 100 may **present the user with a video selected for viewing by the user within video player 300** concurrently with annotation tool 310.

Annotation tool includes...**clip starting point indicators 312**” Ex-1009, ¶0083.

Thus, Syed discloses **causing, using the virtual clip (video), the clip to be displayed on a computing device of a user (GUI of FIG. 3A).**

“[10d] receiving a first user input associated with the virtual clip;”

322. Syed discloses this claim element. As explained under claim element [10c], FIG. 3A of Syed, reproduced below with annotations, includes an annotation tool 310. Ex-1009, ¶0083.



Ex-1009, FIG. 3A (red highlighting added). As shown, the “[a]nnotation tool 310 also includes a search text entry field 315 and a **label text entry field 314.**” Ex-1009, ¶0083. Syed explains the “[l]abel text entry field 314 **allows the user to propose a new label for a portion of the video being played or edit an existing**

label.” Ex-1009, ¶0085. Thus, the label text entry field 314 of Syed facilitates receipt of **user input associated with a virtual clip.**

“[10e] determining that the first user input conforms to a predetermined format defining taxonomy tags;”

323. Syed discloses this claim element. As explained under claim element [10d], Syed discloses receiving user input from a “[l]abel text entry field 314 [that] **allows the user to propose a new label for a portion of the video being played or edit an existing label.**” Ex-1009, ¶0085. Syed further discloses **“a tag or a label may be in the form of a hash tag** that is similar to hash tags on Twitter.” Ex-1009, ¶0037. Additionally, **“the label may include one or more words and/or be in the form of one or more hash tags** (e.g., hash tags similar to those used in the context of Twitter).” Ex-1009, ¶0042. Thus, in instances where the taxonomy tag includes only a first tag, Syed discloses **determining user input** (user input in the label text entry field 314) **conforms to a predetermined format defining a taxonomy tag** (a tag or a label in the form of a hash tag).

324. To the extent the term “taxonomy tag” is construed to require support of a first tag and one or more subtags identifying various categories (see discussion of claim element [10h]), Syed in view of Giunio-Zorkin also renders obvious this claim element. Giunio-Zorkin discloses “keyword metadata,” which are **“keywords or tags associated with the content items.”** Ex-1013, 11:15-16. In

Giunio-Zorkin, a content item “includes content or links/references thereto, e.g. data such as one or more text strings, images, etc., generated by the system, e.g. via the content generator 134, and/or by a user of the system, such as a ‘post’, comment, statement, picture, URL/web link, hashtag, etc.” Ex-1013, 6:67-7:4.

325. Giunio-Zorkin teaches the keyword metadata may be in the form of “**a taxonomy of tags**” where “the available keyword metadata...is organized into a plurality of keyword sets/hierarchies, where **each keyword set includes a hierarchical sequence of keywords**, wherein for each keyword set **a first keyword of the sequence of keywords defines semantics or a context...**and **each subsequent keyword of the sequence of the keyword set defines a narrowing of the semantics/context** over the previous keyword of the sequence of the keyword set. For example, one keyword set may include the sequence of keywords: Drinks; Alcohol; Wine; White; Pino Grigio; California; Berringer.” Ex-1013, 11:59-12:8.

326. As set forth above under the “Rationale (Motivation) to Combine” Section, in my opinion, it would have been obvious for a PHOSITA to modify Syed’s system to include the taxonomy tags according to Giunio-Zorkin’s teachings “for the purpose of distributing and/or facilitating discovery of that content item.” Ex-1013, 11:6-7. Giunio-Zorkin teaches “content items are associated with one or more interest classifications” and its system “maps or

otherwise correlates the interests of each user against the interest classification(s) of content items to determine which content items to transmit to a user.” Ex-1013, 11:3-4, 11:10-13. Thus, the addition of taxonomy tags that contain multiple keywords in a keyword set as taught by Giunio-Zorkin would facilitate retrieval of videos in Syed’s system that are more aligned to a requesting user’s interests. In my opinion, a PHOSITA would have had a reasonable expectation of success because the modifications of Syed represent nothing more than the use of predictable, commonly used formats for data.

“[10f] identifying one or more taxonomy tags from the user input, wherein identifying the one or more taxonomy tags comprises determining the one or more taxonomy tags using the predetermined format comprising:”

327. In my opinion, Syed alone or in view of Kwon renders obvious this claim element. Syed discloses “[a]ny label added or vote cast by the viewer” is ultimately “sent to video label integration and indexing module 247,” which provides “automated label integration/filtering processing.” Ex-1009, ¶0075. As an example, Syed discloses in FIG. 9, “a flow chart illustrating automated label integration/filtering processing,” which “assumes automated label integration/filtering processing has been triggered as a result of the addition of a new label (e.g., manually by a user or automatically).” Ex-1009, ¶0119. Syed explains the “automated label integration/filtering processing [in FIG. 9] may

involve, among other things, a weighted integration of labels via clustering and spam filtering.” Ex-1009, ¶0093. Because the execution of the automated label integration/filtering processing disclosed in Syed requires the “addition of a new label,” this shows the automated label integration/filtering processing is preceded by **a process to identify a tag or a label from user input** to ensure a new label was added. Ex-1009, ¶¶0037, 0119, FIG. 9. Thus, Syed discloses **identifying *one* taxonomy tag from the user input** (identifying tags from user input to facilitate label integration/filtering processing).

328. In my opinion, it would have been obvious for a PHOSITA to implement a process in Syed to identify a tag or a label from user input to facilitate the use of taxonomy tags by differentiating user input that includes a tag or a label from other user input that would otherwise not be eligible for use as a tag or a label. In my opinion, a PHOSITA would have had a reasonable expectation of success because the predetermined formats disclosed by Syed are standard formats commonly used to store tabular data, such as data associated with a video.

329. To the extent this claim element is construed to require the capability to identify *multiple* taxonomy tags in the user input, then this claim element is nevertheless obvious over the combination of Syed and Kwon. Kwon discloses in connection with FIG. 4 that “[t]he user may input text to be uploaded together with the content 43 to a text input region 44.” Ex-1014, ¶0061. Kwon further discloses

“[i]n FIG. 4, the user input one tag, but **the user may input a plurality of tags.**

For example, **when the user inputs ‘#Maltese #Dog’, both a tag ‘Maltese’ and a tag ‘Dog’ are input.**” *Id.* Once user input is submitted, Kwon discloses “[u]pon receiving the content 43 and the text, the content registerer 112 may register the content 43 and the text in the DB 114 of the apparatus 110. The content registerer 112 may **use the identifier included in the text to distinguish the user tag**, and assign the tag as a general tag of the content 43.” Ex-1014, ¶0063. Because Kwon’s system supports user inputs with multiple tags and includes a content register that identifies tags for registration in a database based on the presence of tag identifiers in the user input, it would have been obvious for Kwon’s system to **identify *multiple* taxonomy tags from user input** (multiple tags that each have a tag identifier, such as “#Maltese #Dog”).

330. As set forth above under the “Rationale (Motivation) to Combine” Section, in my opinion, it would have been obvious for a PHOSITA to further modify Syed’s system by implementing a process to identify multiple tags or labels from a single user input where each tag or label is in hash tag form according to the teachings of Kwon. As explained above, a process to identify tags provides a way to differentiate user inputs that includes tags from other user inputs that do not include tags. The support and identification of multiple tags in a single user input would further expedite the process of associating multiple tags with a video or a

video compilation in Junee. By allowing multiple tags to be associated with the video clip in Syed, the video clip and/or the video compilation would be more readily discoverable in response to different search queries, thus improving the quality and relevance of search results in response to a user's search query. In my opinion, a PHOSITA would have had a reasonable expectation of success because the modifications of Syed represent nothing more than predictable and routine intake of user input using standard programming.

“[10g] a tag identifier indicating that the user input includes taxonomy tags; and”

331. Syed discloses this claim element. As explained under claim element 10[a], Syed discloses “a label or a tag” can include “a short description of a video segment.” Ex-1009, ¶0031. The terms “label” and “tag” are used interchangeably in Syed and thus are equivalent in the context of Syed's teachings. As explained under claim element [10e], Syed further discloses user input that includes “**a tag or a label may be in the form of a hash tag.**” Ex-1009, ¶¶0037, 0042. Thus, Syed discloses **a tag identifier** (the “#” symbol) when the user input is in the form of a hash tag.

“[10h] one or more tags following the tag identifier and separated from each other by a delimiter, the one or more tags including a first tag and zero or more subtags arranged in sequence; and”

332. Syed discloses this claim element. I note the “subtags” recited in this claim element are optional; hence, this claim element only requires the “one or more tags” to include the “first tag” without any delimiters.

333. As explained under claim element 10[a], Syed discloses “a label or a tag” can include “a short description of a video segment.” Ex-1009, ¶0031. As explained under claim element [10e], Syed further discloses user input that includes “**a tag or a label may be in the form of a hash tag.**” Ex-1009, ¶¶0037, 0042. Thus, Syed discloses a **first tag** (the short description of a video segment) following the tag identifier when the user input is in the form of a hash tag.

334. To the extent this claim element requires the capability to include one or more subtags in the “one or more tags,” Syed in view of Giunio-Zorkin renders obvious this claim element. As explained under claim element [10e], Giunio-Zorkin discloses “**a taxonomy of tags**” where “the available keyword metadata...is organized into a plurality of keyword sets/hierarchies.” Ex-1013, 11:59, 11:64-65. Giunio-Zorkin further discloses examples of a taxonomy tag: “one keyword set may include the sequence of keywords: Drinks; Alcohol; Wine; White; Pino Grigio; California; Berringer. Another keyword set may include the sequence of keywords: World: North America: United States: Illinois: Cook County: Chicago: Lincoln Park.” Ex-1013, 12:6-11. In my opinion, a PHOSITA would have recognized each of the foregoing examples of keyword sets disclosed

in Giunio-Zorkin includes commonly-used delimiters, such as a semi-colon (“;”) for the Drinks keyword set and a colon (“:”) for the World keyword set. *See* Ex-1019 at 152 (defining delimiter and providing examples of commonly used delimiters, including semi-colon and colon). Thus, Giunio-Zorkin teaches a predetermined format that includes **one or more subtags** (the keywords in a keyword set that follow the first keyword) **sequentially following a first tag** (the first keyword in a keyword set) and **delimited by a delimiter character** (the delimiters between the keywords in the “Drinks” and “World” examples).

335. In my opinion, it would have been obvious for a PHOSITA to modify Syed to support user input with multiple tags separated by a delimiter as taught by Giunio-Zorkin for the same reasons discussed under claim element [10e], i.e., to facilitate retrieval of video compilations more aligned to the interests of the user submitting a query.

“[10i] associating, in an account of the user, the virtual clip with each of the one or more taxonomy tags identified from the user input.”

336. In my opinion, Syed in view of Giunio-Zorkin renders obvious this claim element. In Syed’s system, “[t]ags **proposed by the subscribers to be associated with various portions of the video content** are received, processed and integrated within a global ToC for the video.” Ex-1009, Abstract. Syed further discloses “an annotation tool is presented to a first subscriber...as video content is

being consumed by the first subscriber. **A first proposed label to be associated with a first portion of the video content is received by the annotation tool from the first subscriber.** The first proposed label is incorporated within a global table of contents (ToC) for the video content by storing within a database maintained by the network-based video discovery and consumption service and **associating with the video content the first proposed label and a start time of the first portion of the video content.**” Ex-1009, ¶0013. Thus, Syed discloses **associating the virtual clip** (video clips) with **one or more taxonomy tags** (tags/labels).

337. Giunio-Zorkin discloses “[u]sers are also associated with a subset of one or more of the keywords selected from [sic] the plurality of keyword sets wherein the keywords are selected, explicitly or implicitly by user and/or system 100, based on their interests.” Ex-1013, 12:47-50. “As depicted in FIG. 14, **keywords, for association with a user or with a content item**, may be selected from a user interface 1400 provided by the system 100....” Ex-1013, 12:63-65. Giunio-Zorkin further discloses “the keyword sets are stored in an interest database or data structure stored in the data store 124 wherein each keyword is associated with a unique identifier. **Associations between users or content items and keywords are implemented by storing the associated identifiers of the selected keywords in association therewith.**” Ex-1013, 13:30-35. Thus, Giunio-Zorkin discloses **associating, in an account of the user (user), the virtual clip** (content

item) with **each of the one or more taxonomy tags** (keywords from keyword sets).

338. As set forth above under the “Rationale (Motivation) to Combine” Section, in my opinion, it would have been obvious for a PHOSITA to modify Syed based on the teachings of Giunio-Zorkin to store tags associated with video clips in an account of the user in order to provide media content to the user that is more aligned with their interests. As explained under claim element [10e], Giunio-Zorkin teaches “content items are associated with one or more interest classifications” and its system “maps or otherwise correlates **the interests of each user against the interest classification(s) of content items to determine which content items to transmit to a user.**” Ex-1013, 11:3-13. Storing tags by a user in an account of the user would have required nothing more than standard computer programming.

CLAIM 11

“11. The method of claim 10, wherein associating the virtual clip with the one or more taxonomy tags comprises arranging the one or more taxonomy tags according to the sequence in the user input of the one or more taxonomy tags.”

339. As a preliminary matter, it is unclear whether the limitation “arranging the one or more taxonomy tags according to the sequence in the user input of the one or more taxonomy tags” is intended to imply a hierarchical relationship exists between the taxonomy tags in the user input of the one or more

taxonomy tags when associating the taxonomy tags with the virtual clip. I note the '442 patent discloses “the system may require that the taxonomy tags of the query appear in the same sequence of the stored taxonomy tags, in order to improve accuracy and relevance of the search results,” which suggests the above limitation refers only to the sequence of the taxonomy tags without implying any hierarchical relationship between the taxonomy tags. Ex-1001, 14:27-36. However, I also note this discussion of the sequence of taxonomy tags in the '442 patent is presented in connection with the portion of the specification that recites the language of claim 9 as prose and, in particular, the term “directed relationship,” which is discussed under claim element [9a] as possibly implicating a hierarchical relationship exists between the taxonomy tags. Ex-1001, 14:27-36.

340. Since this claim element only refers to arranging one or more taxonomy tags based on a sequence of one or more taxonomy tags in the user input, I interpret this claim element as not imposing a hierarchical relationship between the taxonomy tags. Rather, “arranging the one or more taxonomy tags” is interpreted to mean the virtual clip is associated with the sequence of the one or more taxonomy tags in the user input such that the virtual clip can be later retrieved if a user’s query contains all the taxonomy tags in the one or more taxonomy tags in the exact same sequence. As discussed under claim element [9a], this interpretation is analogous to a conventional phrase search where a sequence

of keywords enclosed by quotation marks causes a search engine to search for results containing only that exact sequence of keywords.

341. Under this interpretation, Syed in view of Giunio-Zorkin, and Kwon renders obvious this claim element. Kwon further teaches “receiving, from a user terminal, content and a user tag” (wherein “the receiving comprises **receiving a plurality of user tags**”) and “registering the content and the received user tag” (wherein “registering comprises, when a number of the received plurality of user tags is higher than a threshold number, **assigning the threshold number of user tags**, from among the plurality of user tags, as general tags of the content **in an order in which the received plurality of tags are received**”). Ex-1012, cl. 1, 5. Thus, Kwon discloses **arranging one or more taxonomy tags** (assignment of the sequence of user tags) according to **a sequence in the user input of the one or more taxonomy tags** (order in which the tags are received).

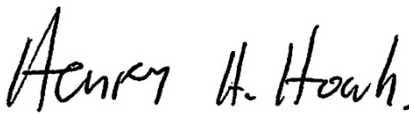
342. In my opinion, it would have been obvious for a PHOSITA to modify Syed’s system to associate a virtual clip with one or more taxonomy tags arranged in the same sequence presented in the user input according to Kwon’s teachings because this would allow users to produce search results that are more relevant to their query. Moreover, arranging one or more taxonomy tags in the order in which they are presented in the user input is one of the simplest approaches of incorporating data on sequences of tags to improve content retrieval.

343. The modifications represent nothing more than the predictable combination of elements taught by Syed and Kwon. In my opinion, a PHOSITA would have been able to combine the teachings of Syed and Kwon using standard computer programming. Furthermore, the respective functions of the elements taught by Syed and Kwon would not change when combined together: users would be able to propose a new label for a portion of a video via a label text entry field according to Syed's teachings, and users would be able propose sequences of labels for association with content according to Kwon's teachings.

X. OATH

344. This declaration and my opinions herein are made to the best of my knowledge and understanding, and based on the material available to me, at the time of signing this declaration. I declare under penalty of perjury under the laws of the United States of America that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true. I understand that willful false statements and the like are punishable by fine or imprisonment, or both (18 U.S.C. 1001) and may jeopardize the validity of the application or any patent issuing thereon.

August 12, 2023
Date



Dr. Henry Houh