

Provisional Patent Application

System and Method for Video Processing, Behavior Monitoring, Modeling, and Interaction

Justin Wilson
21566 Iredell Terrace
Broadlands, VA 20148

Seunghye Jung Wilson
21566 Iredell Terrace
Broadlands, VA 20148

Prior Art.

US Patent Documents

US Pat. US20100306138A1	Jun., 2009	Scott Hotes et al.	G06Q30/0206
US Pat. US20070100595A1	Oct., 2005	Earles Alison C et al.	G06Q30/0207
US Pat. US20040210458A1	Apr., 2003	Junius Evans et al.	G06Q10/10
US Pat. US20070156585A1	Jan., 2006	International Business Machines Corp.	G06Q20/3829
US Pat. US20080097550A1	Oct., 2006	Kent Dicks et al.	A61N1/37282
US Pat. US20080097793A1	Oct., 2006	Kent Dicks et al.	G06Q10/10
US Pat. US8165893B1	Feb., 2005	Ideal Life Inc.	G06Q40/08
US Pat. US5997476A	Mar., 1997	Health Hero Network, Inc.	A61B5/4839
US Pat. US6024699A	Mar., 1998	Healthware Corporation	G16H40/63
US Pat. US6168563B1	Nov., 1992	Health Hero Network, Inc.	A61B5/4839
US Pat. US6221010B1	Jul., 1999	Donald A. Lucas	A61B5/0008
US Pat. US20030153836A1	May, 2000	Claude Gagnandre et al.	G08B29/183
US Pat. WO2010009361A1	Jul., 2008	Indiana Univ. Research & Tech. Corp.	H04N7/181
US Pat. US20100033332A1	Aug., 2008	The Quantum Group, Inc.	G06Q10/10
Pat. WO2018007738A1	Jul., 2016	Novia Search	G06Q50/22
US Pat. US20040019259A1	Nov., 1992	Stephen Brown et al.	A61B5/0002
US Pat. US6039688A	Nov., 1996	Peter Douglas et al.	A61B5/4833
US Pat. US5722418A	Aug., 1993	L. William Bro	G09B7/04

Advantages.

Various aspects of the system and method for video processing, behavior monitoring, modeling, and interaction may have one or more of the following advantages:

- Provide families a more expedient alternative for access to healthcare professionals given it is not uncommon for wait lists of a year or more to see a developmental pediatrician, for instance
- Provide labeled data and a more universal behavior mapping and tagging within and across organizations, industries, & academia
- Provide an interactive historical archive from users' prior recordings to identify trends and other analysis metrics based on video data and behavior modeling

- Provide a complete and comprehensive, end-to-end behavior monitoring system by processing a variety of device and video formats along with collaborative interactions of a community between users such as parents and healthcare professionals, for example
- Provide a more integrated approach for video processing, behavior monitoring, modeling, and interaction, etc. from multiple sources
- Provide an automated system for identifying developmental behaviors using machine learning
- Provide an active community that offers opportunities and availability of healthcare entities
- Provide a more dynamic system that learns & adapts to user experience and usage

Drawings.

FIG. 1 illustrates an exemplary digital behavior monitoring network environment using modeling features and interactions among users;

FIG. 2 is a block diagram of an exemplary system and method for video processing, modeling, and interaction for behavior monitoring;

FIG. 3 is a block diagram of exemplary system components to view recordings, access calendar views, automate behavior tag prediction, and access member resources, to name a few;

FIG. 4 is an exemplary screen shot of a home page with links to view recordings and access member resources, to name a few;

FIG. 5 is a block diagram of an exemplary system component enabling a user to access member resources such as prior video recordings, uploading new videos, or viewing calendar of uploads and tags;

FIG. 5A is an exemplary screen shot of a registration and login page for users;

FIG. 5B is an exemplary screen shot to view information on previous recordings with video filename, link title, and behavior tags either manually or automatically associated with each video using a trained machine learning classification model based on one second audio segments;

FIG. 5C is an exemplary screen shot of a user's own recording;

FIG. 5D is an exemplary screen shot of a setup for automatic recordings using a Raspberry Pi (or Google Nest, Amazon Echo, or some other recording device) and a user manually uploading a new video recording from a computer or mobile device;

FIG. 5E is an exemplary screen shot of a user calendar view;

FIG. 5F is an exemplary screen shot of a summary video and animated GIF comprised of only tagged segments which enables quick review of daily videos in a snapshot summary video while still allowing access to the individual videos that make up the summary to watch in entirety for more detail;

FIG. 6 is an exemplary screen shot of sortable, filterable, and/or searchable blog posts;

FIG. 6A is an exemplary screen shot of a blog post where users have accessed a specific article allowing them to view statistics, hyperlink to the blog post, and/or comment;

FIG. 7 is an exemplary screen shot of multimodal convolutional neural networks that are trained as behavior classifiers with audio only and audio-visual data.

Detailed Description.

The inventors listed invented system and method for video processing, behavior monitoring, modeling, and interaction using data capture techniques and machine learning algorithms such as cloud storage, elastic compute instances, classification for automatic behavior tagging, and data visualization, to name a few. Behavior monitoring can be defined as, but not limited to, a system and method that can identify

patterns by tagging times within a video of behavior (e.g. yelling, crying, laughing) that differ against a baseline of non-tagged video segments. These patterns can be analyzed over time and accessed for review by the owner or users whom the owner shared the video with such as healthcare professionals. Relevant prior art includes remote viewing for in-home patient monitoring and collaboration systems for feedback, data sharing, and collaboration. Primary differentiators of this invention are that it allows the user to record and upload video, access recordings, and grant access to any others who can view and comment on a per video or user basis as well as automatic tagging of behavior classifications using machine learning multi-class classification based on sound or audio-visual data from video recordings. As a result, the system and method for video processing, behavior monitoring, modeling, and interaction **130** provides a complete and comprehensive, end-to-end behavior monitoring system by processing a variety of device and video formats along with enabling collaborative interactions from a community of users such as parents and healthcare professionals. This invention also offers a more end-to-end behavior monitoring experience from automatic video capture **526** and user uploads **496** from a computer **522** or mobile device **524** for cloud storage and access **120**, behavior classification **495**, and trending **550** to name a few, enabling users **101** to view recordings **206**, parents **101A** and healthcare professionals **101B** members to access member resources **208** and calendar views and tags **370**, and sponsors **101C** to advertise sponsored content items. The more data captured about the user's **101A** history of behaviors **550**, more accurate behavior models **490** and predictions **495** can be made for future recordings and users. Manual behavior tagging is another way to provide more information to improve classification. Calendar views **550** display counts of behavior tags of second segments for yelling **554**, crying **556**, and laughing **558** based on audio and visual data from the video as well as a total count of upload videos **552** by day. Behavior tagged video recordings **512** save time for reviewing users such as healthcare professionals by automatically highlighting parts of the video with different behavior tags that may require attention and follow-up. In addition, these tags can serve as the basis for summary videos **562** or animated GIFs **564** comprised of only tagged segments **560**. This enables quick review of daily videos in a snapshot summary and access to the individual videos that make up the summary to watch in entirety for more detail. Comments **545** to videos can be added by users **101** such as parents **101A** and healthcare professionals **101B** for additional feedback on specific behaviors, trends over time, advice, and recommendations.

The system and method for video processing, behavior monitoring, modeling, and interaction **130** may host a website that allows one or more users **101**, e.g., healthcare professional users **101B** to advise parents and caregivers **101A** for behavior monitoring **130** or the sponsor user **101C** to advertise, at one or more user devices **110** via a communications network **120** to the website and/or mobile application. User devices **110** include a computer terminal, a laptop, a personal digital assistant (PDA), a wireless telephone, a smart phone, a smart phone application, and/or the like. Communications network **120** includes a local area network (LAN), a wide area network (WAN), a wireless network, a cellular network, an intranet, an internet, and/or the like.

Terms of use **202** are accessible via a terms of use link **306**. The terms of use **202** apply to all visitors, users, and others who access the service. Non-registered visitors **102** may apply and register **204** by accessing an online application via registration section **502** on the web page to access member resources **208** or advertise resources **210**. If deemed appropriate for membership and if membership is being limited, non-registered visitors **102** will receive a confirmation code to use during login **504**. Otherwise, they are given read-only access to access social media pages **360**, join the mailing list **350**, share pages via LinkedIn **355** or other social media, and/or view blog posts **345** until they login **504**. Users **101** unable to login may use forgot password functionality **506**. Users **101** may login by entering required login

information **504**, such as registered email address and password. Confirmation code may also be required when logging in from a communication network **120** for the first time. Once logged in, users **101** can access their member resources **208** such as their own video recordings **402** where they can playback the video **510** and jump to tagged behavior segments **512** in the video or view shared videos from other users **406** which they have been granted access. Alternative ways to add additional experiences **505** may be integration with healthcare systems, social media such as LinkedIn, and/or other company APIs. System activity is stored in several databases, schemas, and tables referred to as a resource database **212** and data encrypted on transfer and securely stored in private cloud storage **125**.

The web pages for the system and method for video processing, behavior monitoring, modeling, and interaction **330** consists of a few sections. Along the top, a navigation bar **310** is available to navigate between internal pages. External links to social media **302** are available below the navigation bar. Also along the top right, non-registered visitors **102** can enter their email address to join the mailing list. Along the right side, filter options **408** are available to filter the list of video links being displayed and any other filterable information relevant to the selected recording. Each video displayed **402** displays the video filename **401**, video description **403**, a link to delete the video **404**, and number of behavior tags **405** associated with the video. Click on the link of the video filename **401** to playback the video **510** and see the list of behavior tags **512**, if any. In this exemplary screen shot **510**, there is a single yelling tag at 105 seconds into the video and a single crying tag at 176 seconds into the video. Clicking the seconds link will jump to that time in the video. Tags can be added manually or automatically based on predictions from a trained classification model **495** based on sound alone **572** or multiple modes **574**. Both manual and automatic tags are associated to the video in resource database **212**. Sound from video can be classified **495** based on a trained model **490** and displayed in video recording **512** and used to generate summary video **560**. Table 1 below shows example binned mel-scaled spectrograms for different behavior classifications. Note that each sample of a behavior class may have slightly varied mel-scaled spectrograms (shifted start time of behavior, different intensities, etc.) and the model is trained to be able to learn and distinguish these features given the training and test datasets used. Data augmentation and additional training may occur to continue to improve classification accuracy and expand to predict more behavior types (e.g. coughing, clapping). Models may use audio **572**, visual, and audio-visual **574** data.

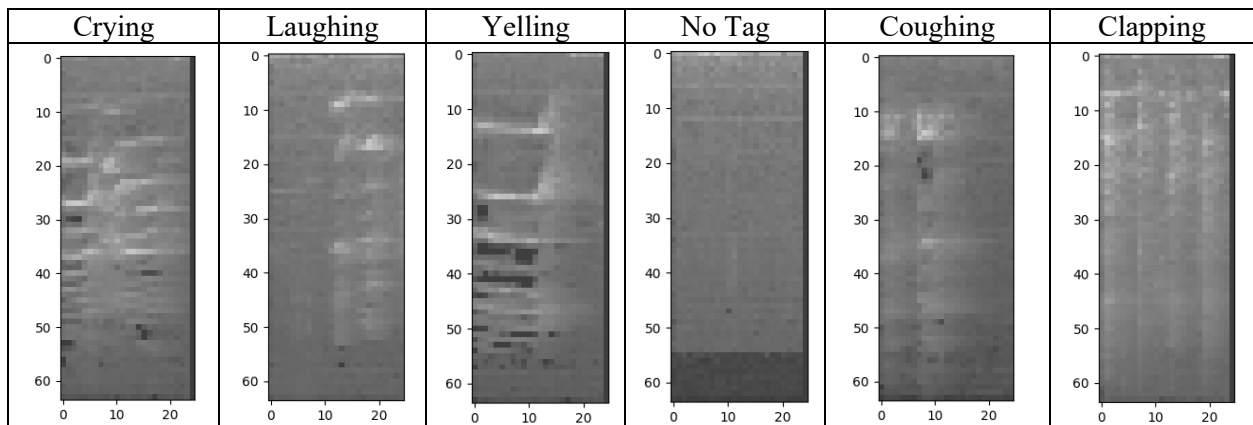


Table 1: example binned mel-scaled spectrograms for different behavior classifications

To access member resources **340** of the system, users **101** must have successfully logged in **480** and have confirmed **476** their email address. To confirm, users **101** may obtain a confirmation code from the email that they used to register **502** and enter the code along with email address and password during login **504**. If a user **101** has access to the email address, then he or she will be able to retrieve the confirmation code and enter it into the appropriate account confirmation field. If the confirmation code entered by the user **101** matches what is stored in the resource database **212**, then the account has been confirmed and will successfully login if the email and password are also correct. If the confirmation code entered is incorrect, an error message will display and the user **101** may try again **480**.

After successfully logging in **594**, user **101** is directed to his or her list of video recordings **507**. Non-users **102** have read-only access to view blogs **345**, access social media pages **360**, view terms of use **365**, and join the mailing list **350**. Parents and caregivers **101A** as well as healthcare professionals **101B** can add new videos **520** and sponsor **101C** users can advertise by adding promotional videos and other types of content. Note that sponsor items are labeled as such. Healthcare professionals **101B** and sponsor users **101C** may not view user data unless it has been explicitly shared with them by the video owner **101A**. Videos can be entered at any time and in any order. Videos do not have to be entered in the sequence that they occurred because the user can manage the metadata associated with the content and arrange in the appropriate chronological order based on the start and end dates of each video, not necessarily the order by which they are entered or created. The list of user owned videos **402** and shared videos **406** are clickable links that result in the video **510** being selected. If the behavior tags are associated with the video, they will be marked and listed **512**. If there are comments associated with the video from parent **101A** or healthcare professional **101B**, they will also be listed, and new ones may be added **545**.

To share a video with another user **406**, parent user **101A** that uploaded the video can click on the share video link **407** and then specify which user or users have access to view the video **510**. When sharing a video, the owner can grant read-only access without comments or read-only access with comments **545**. For privacy and security, access to view the video **510** as well as a login session are time limited. Table 2 below is an example of how shared video information may be stored in resource database **212**.

User Video ID	User ID	Video ID	Access	Comments
1	1	1	Owner	User 1 has access to Video 1
2	1	2	Owner	User 1 has access to Video 2
3	2	2	w/ Comments	User 2 also has access to Video 2

Table 2: example of single or multiple user access based on video sharing controlled by video owners

Drop-down list of pre-populated users along with any metadata or reviews associated with the users may be enabled to assist the user to select the appropriate user with whom to share. If the text field does not auto-complete because it is a new entry or user, the system may allow a user to enter free text and the entry will be validated and verified after submission. Finally, start date and end dates for video sharing and commenting can be enabled if temporary access is desired by users.

By default, all blog posts are sorted by publish date ascending and limited to 10 articles to display. Specific blog posts **544** and general blog posts **534** are displayed via hyperlinked thumbnail images. Articles may be filtered via the filtering drop-down, sorted via the sorting drop-down, and limited via the results per page drop-down in section **532**. Search for an article using the search functionality **538** by selecting the search criteria and typing the keyword to search. Each article has its own section of

information and statistics **540** displaying article name, image, author, publish date, and description in section **544**. Users **101**, parents **101A**, healthcare professionals **101B**, and sponsors **101C** may publish articles as well as comment on articles **409**. To read an article, click on the hyperlinked article name **539**. To comment on an article, click on the article thumbnail **536** and access the specific page for selected article **540**. On this page, notice that the filtering drop-down in section **532** displays “selected” to notify that a specific post has been selected. As a result, you can then find statistics, information, and comments **409** specific to the blog post selected. Users **101** can view and publish blog posts, where sponsor user **101C** blog posts will be flagged as posted by a sponsor. To submit a blog post, users **101** must have successfully logged in **480** and have confirmed **476** their email address.

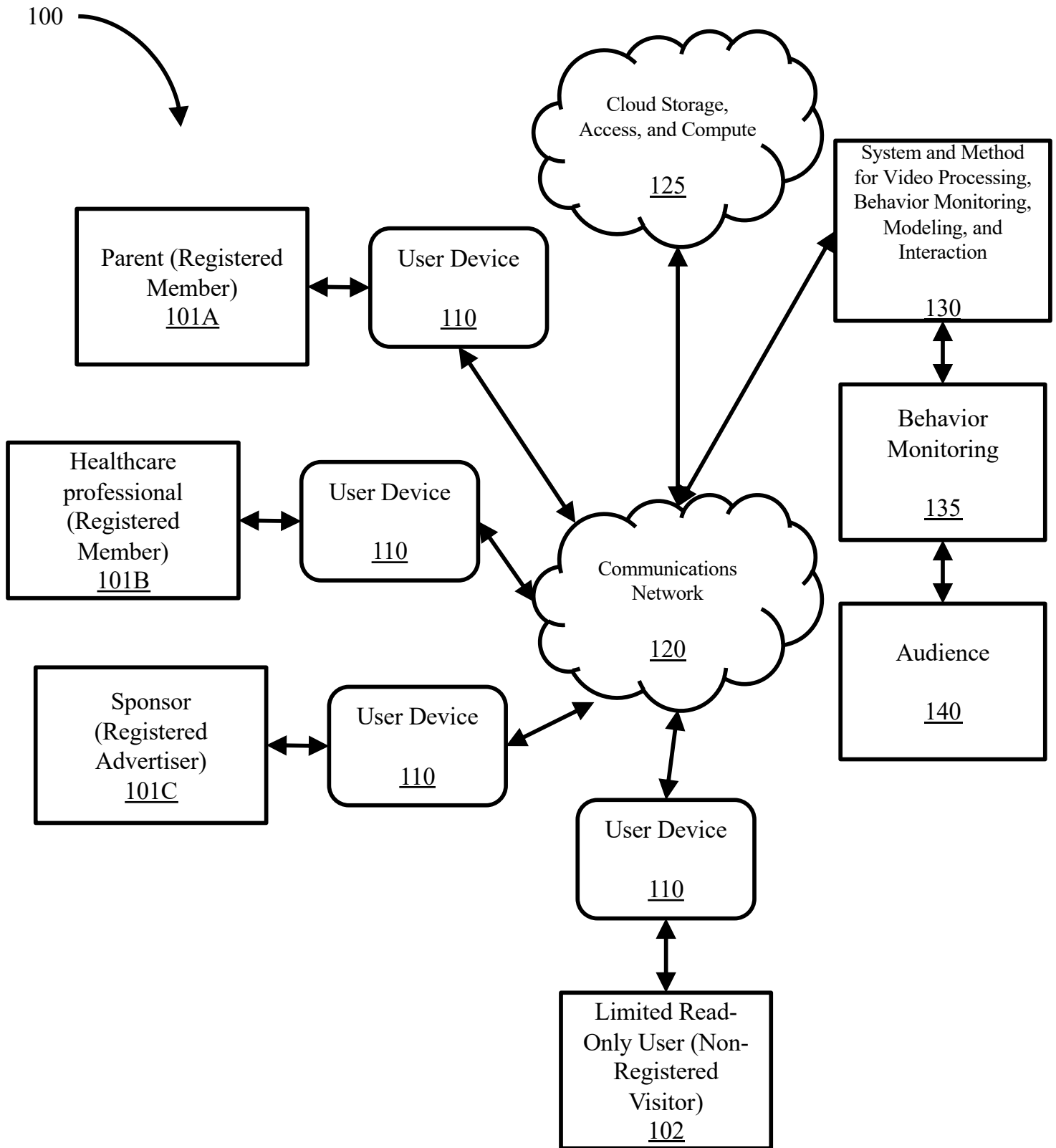
There are many alternative ways that the system and method for video processing, behavior monitoring, modeling, and interaction can be implemented:

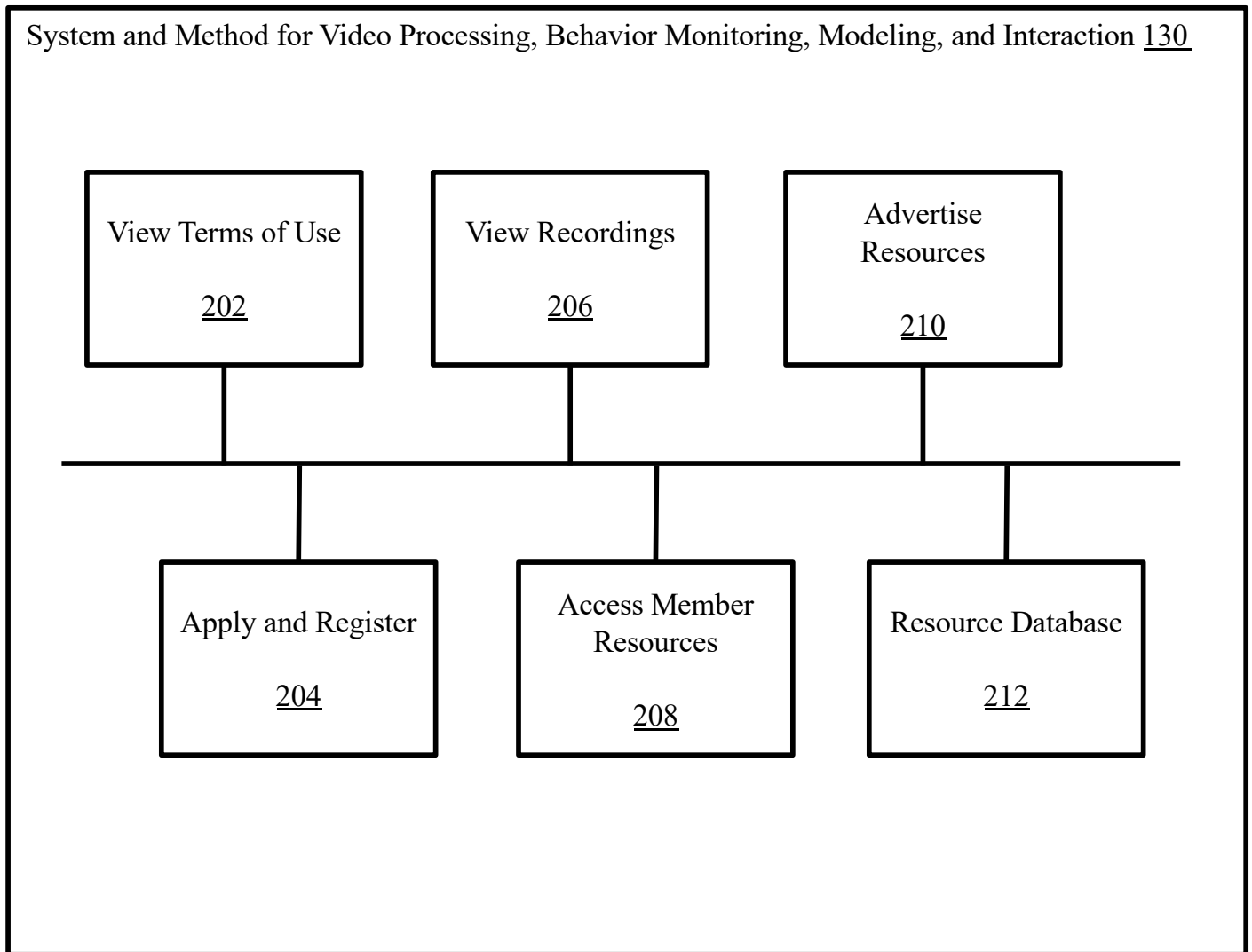
- Additional required and optional fields can be added to apply and/or register
- Non-users can register automatically without having to apply or entering a confirmation code
- Non-users can login using their social media accounts via integrations such as Sign in with LinkedIn, Google, Amazon, Facebook Connect, and/or the like
- Additional video, user, and behavior information can be displayed and in more places throughout the system
- Behavior modeling and tagging can be based on alternative methods using existing and/or additional data and/or metrics along with different time lengths for data segments
- Suggested users to share a video can be based on additional and/or alternative information
- Filtering and sorting can be based on alternative calculations using existing and/or additional metrics; for example, geographic information such as country, CBSA, city, state, and/or the like
- Additional search criteria can be added to the search functionality
- Additional filtering criteria can be added to the filtering functionality
- Additional social networks and sponsor applications can be integrated and connected
- The system and method can be integrated within a healthcare organization or other application
- Information can be further customized by soliciting user input and/or retrieving information from one or more other sources containing such information
- Video recordings from additional and multiple recording devices can be added
- Resources, courses, and blogs from more online and offline educational platforms can be added
- Sponsor healthcare organizations and/or professionals associated with the system can be refreshed automatically on a periodic basis and/or initiated by a company, sponsor, and/or user
- Additional content can be added on a continuous basis by advertisers, other users, inventors, and/or others associated with this invention
- For interactions, in addition to comments, system may allow for telemedicine, phone call, or other communication scheduling
- Users can view percentage likelihood of a predicted behavior and/or the like based on previous data and/or obtained from sources containing such information
- Sponsor users can filter and/or sort how the advertising users and their posts are displayed
- Eligibility can be defined based on sponsor needs and/or advertiser user demographics, previous postings, and/or the like to permit and/or prevent eligibility to advertise specific information
- Additional required and optional fields can be added to user information

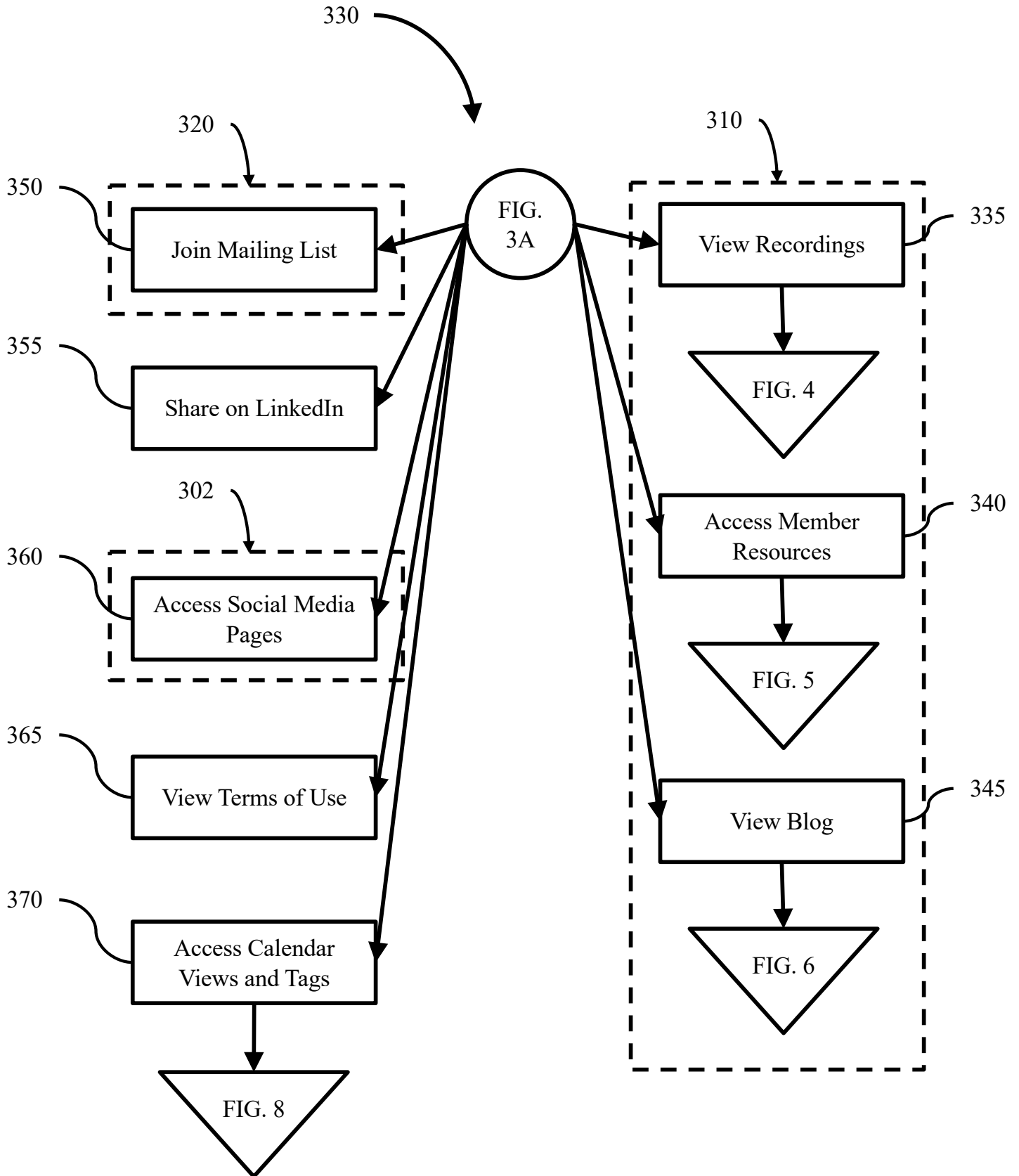
- Users can be alerted when new information is submitted, advertised, and/or the like
- Users can connect within and across users, companies, industries, etc. for resources and additional information
- Users can submit questions via text message, email, and/or other messaging channel and they will be directed to the appropriate user, healthcare professional, and/or experts to answer, using same or different messaging channel from how the question was submitted
- Questions and/or communications submitted via Google, Amazon, Facebook, Twitter, and/or other social media networks can be directed to users and/or the appropriate representation for comment and/or follow-up
- Additional criteria can be incorporated for selecting similar users and/or behaviors
- Users can find other users that are in similar circumstances, geographical area, etc. and/or in a similar state of system use
- Users can save videos, specific or summary, locally and/or share on social media
- Summaries could be based on varying date ranges, besides daily, and generated in other formats
- Users can view their resources, save information, and/or receive information to a profile page specific for each user
- Sites such as WebMD and other healthcare related sites can be integrated to further enhance the end-to-end experience from monitoring to interaction
- Faces and/or other identifiable information can be blurred in videos using face detection, object detection, and other related computer vision and machine learning techniques
- Electrical medical record (EMR) and/or electronic health record (HER) integration can allow users to connect with their healthcare providers and organizations to share/receive information
- Ability to virtually shadow users to experience the system and method for information from users or from other sources containing such information (e.g. prior to joining)
- Since the invention spans across industries, companies, etc., it may offer features like an umbrella or parent organization; for example, behavior workshops, best practice sharing, etc. at various companies but are members of this invention's community
- Expanded behavior tags based on transcribed speech to text and keywords and/or visual data, such as identifying autism symptoms, tics, etc.
- Users of like situations or healthcare professionals supporting similar parents can receive notifications about each other in the community and are given the opportunity to share resources such as classes, training, and/or comment/message with other users
- Messaging between users online and offline
- External articles and resources visualized, grouped together, and filterable for relevant users
- How content, video, publications, etc. can be produced and displayed about this invention, inventors, and/or others associated with this invention
- Data, algorithms, clustering techniques, machine learning, etc. may be modified and enhanced
- Blog posts, articles, publications, and/or the like added to a video, company, industry, etc. and accessible in alternative ways using alternative filtering and visualizations
- Functionality to schedule time with others can be incorporated
- Highlight healthcare organizations, professionals, resources, etc. regarded as relevant for users

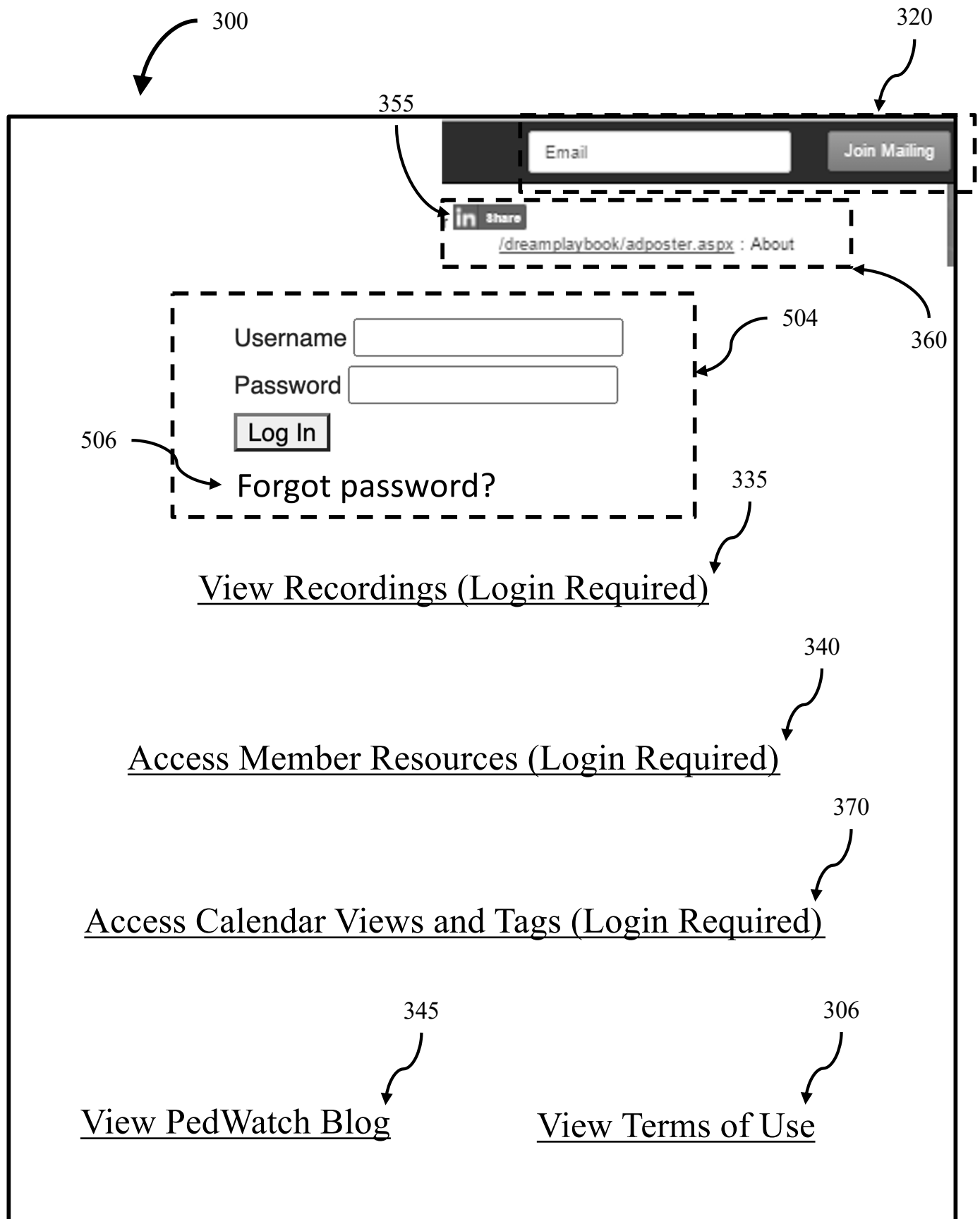
In summary, the system and method for video processing, behavior monitoring, modeling, and interaction **130** of this invention provide a complete and comprehensive, end-to-end behavior monitoring system by processing a variety of device and video formats along with collaborative interactions of a community between users such as parents **101A** and healthcare professionals **101B**. Other alternative ways this invention could be implemented may involve additional functionality, data, algorithms, machine learning, and benefits to support the advantages and claims of this invention.

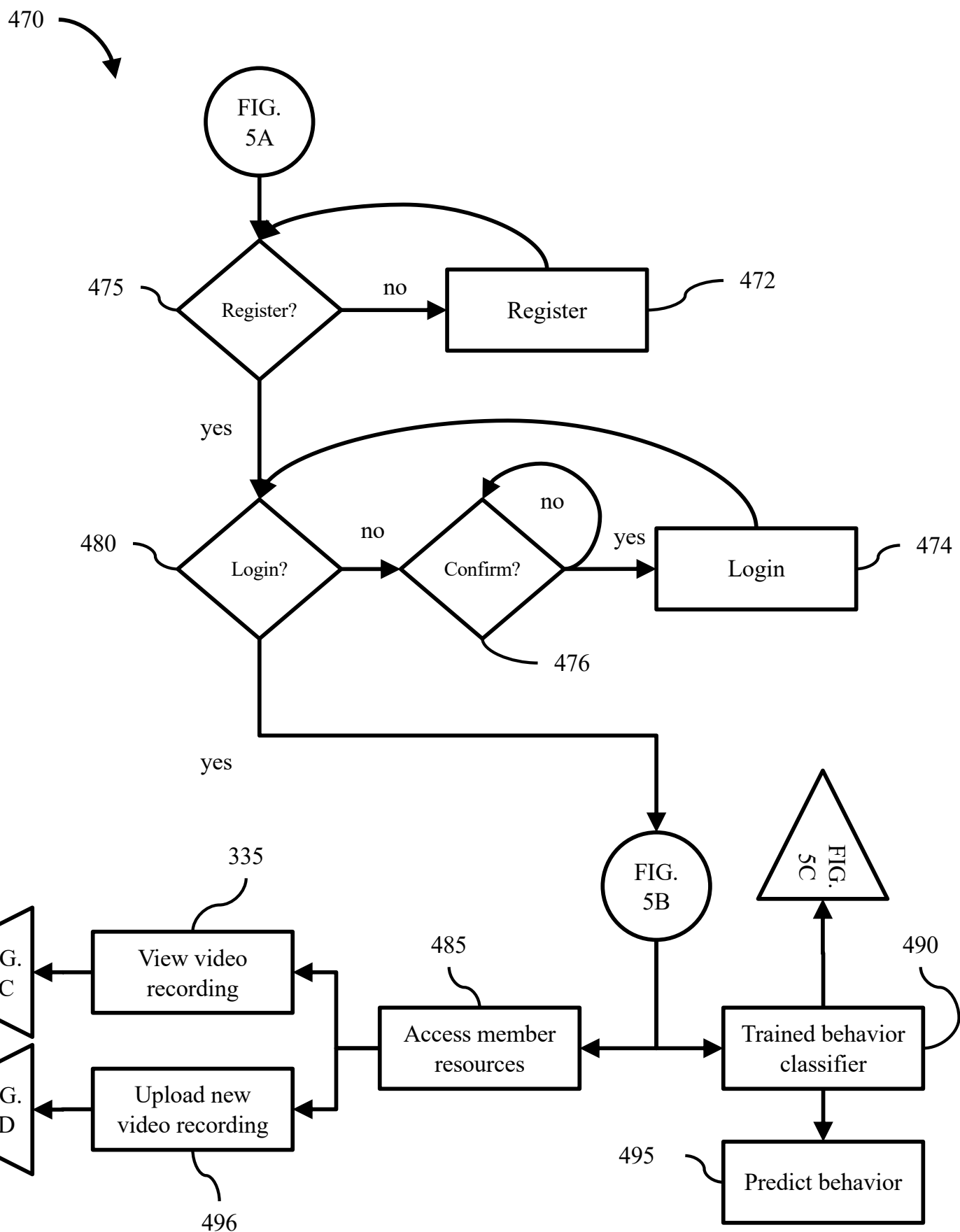
While the invention has been illustrated and described in detail in the drawings and foregoing description, the same is to be considered as illustrative and not restrictive in character, it being understood that only the preferred embodiment has been shown and described and that all equivalents, changes, and modifications that come within the spirit of the inventions as described herein are desired to be protected.

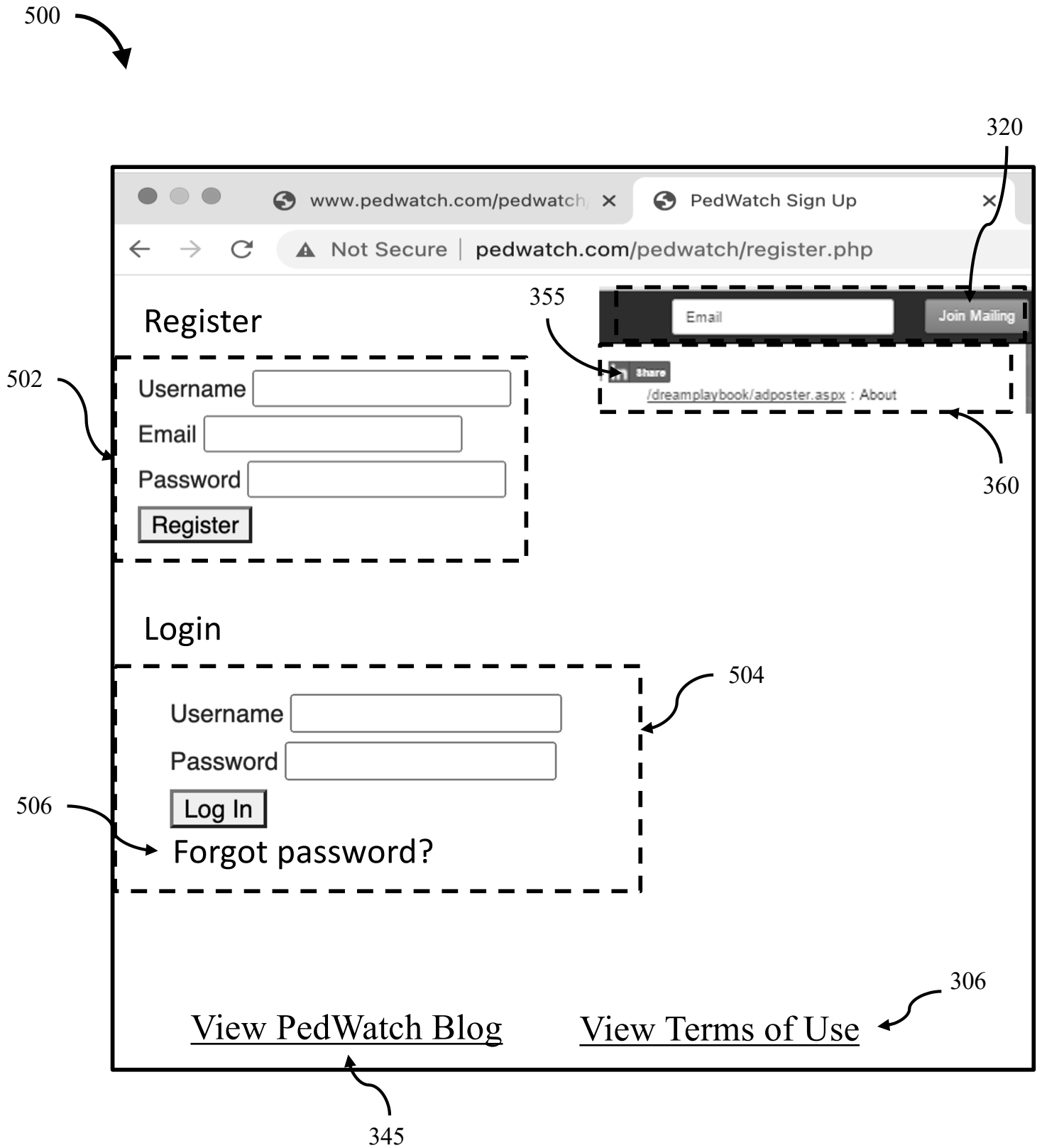


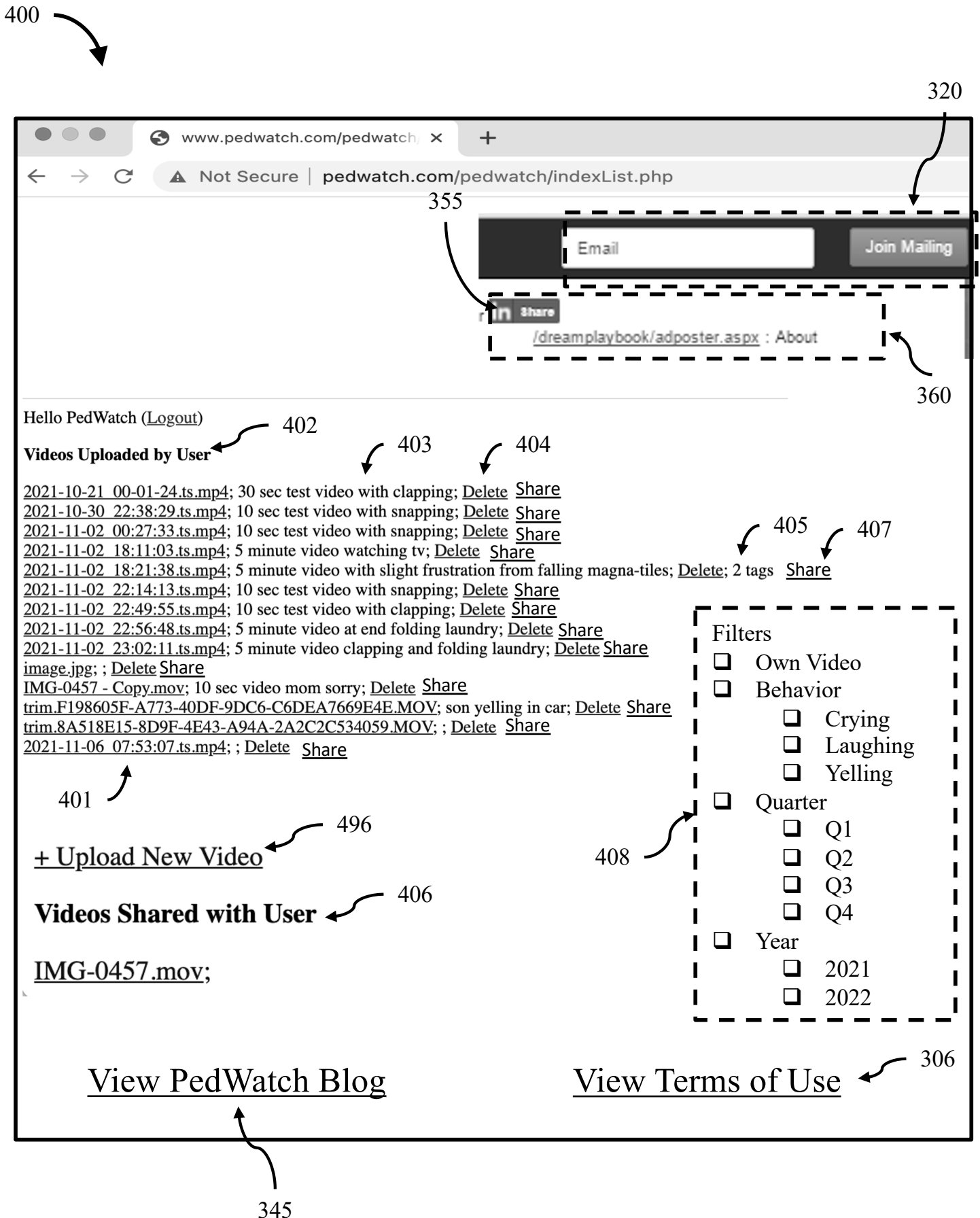




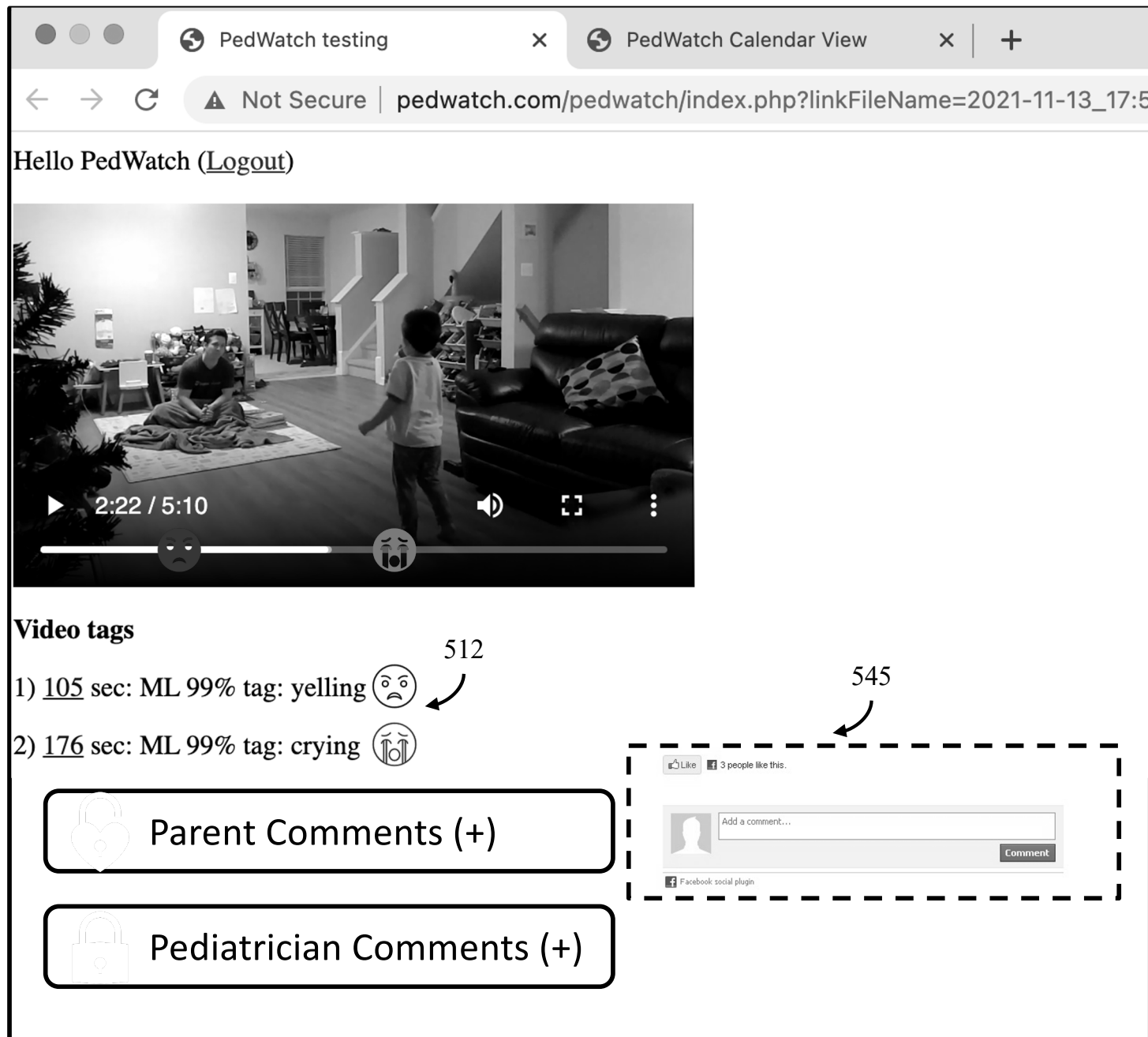








510



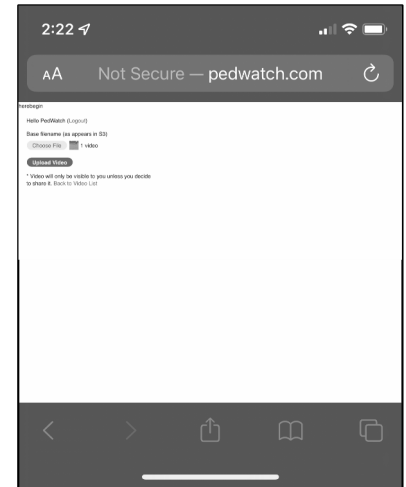
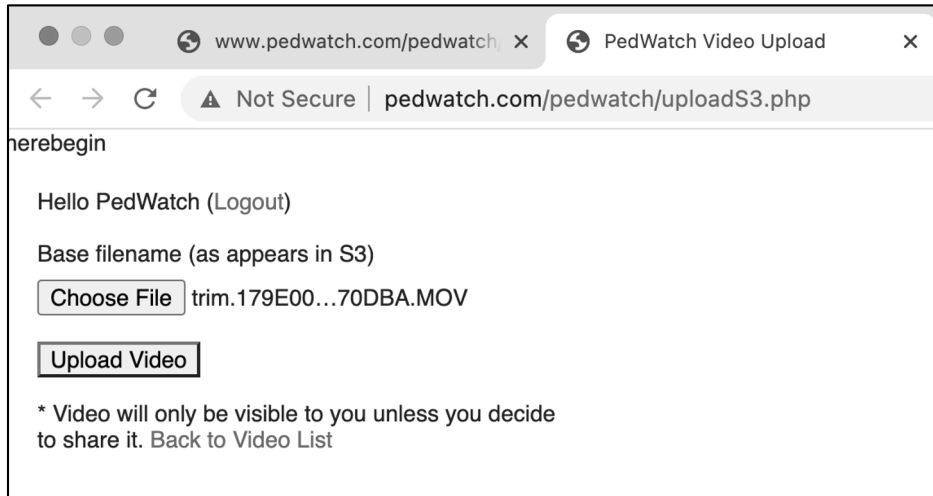
520

Upload from computer

522

Upload from mobile

524



Automatic upload from device (e.g. Raspberry Pi)

526



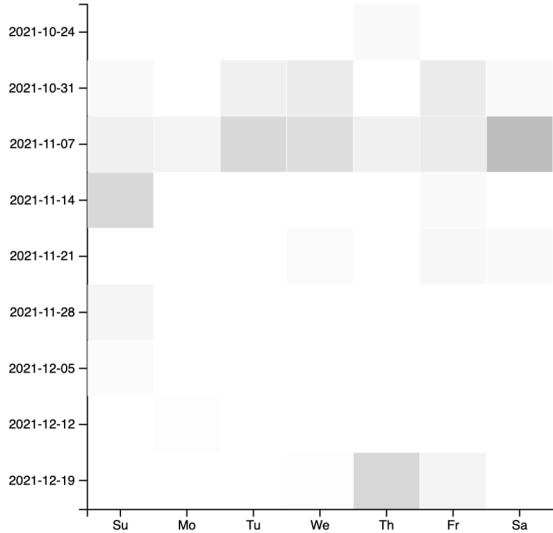
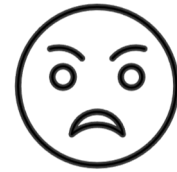
550

Videos Uploaded

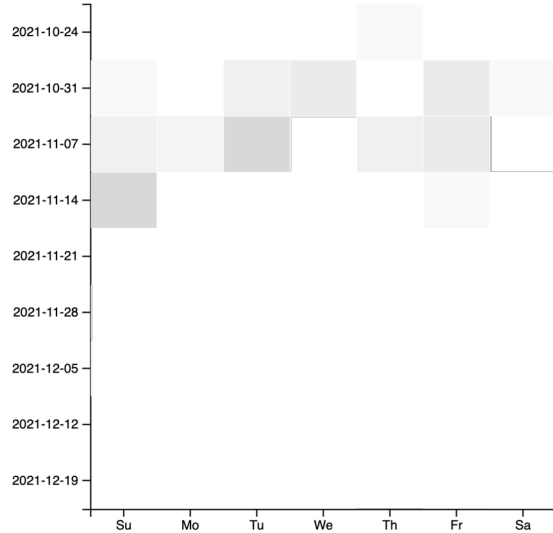
552

554

Yelling Tags



The exact value of this cell is: 29



The exact value of this cell is: 29

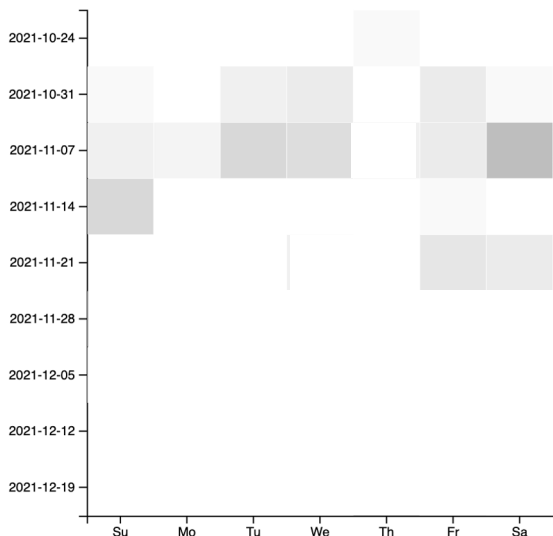
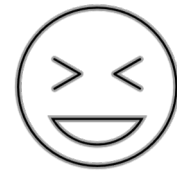
556

Crying Tags

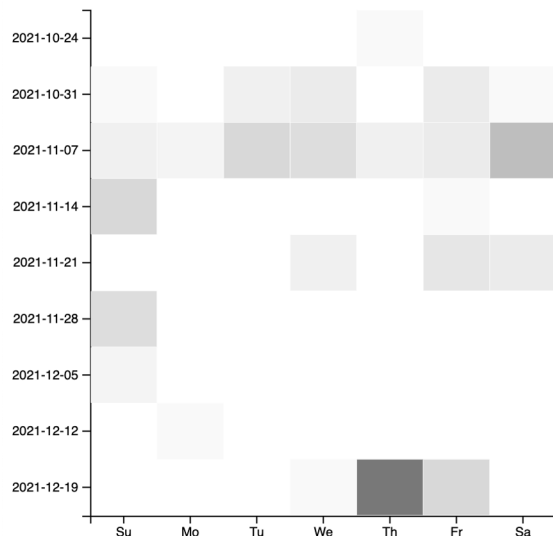


558

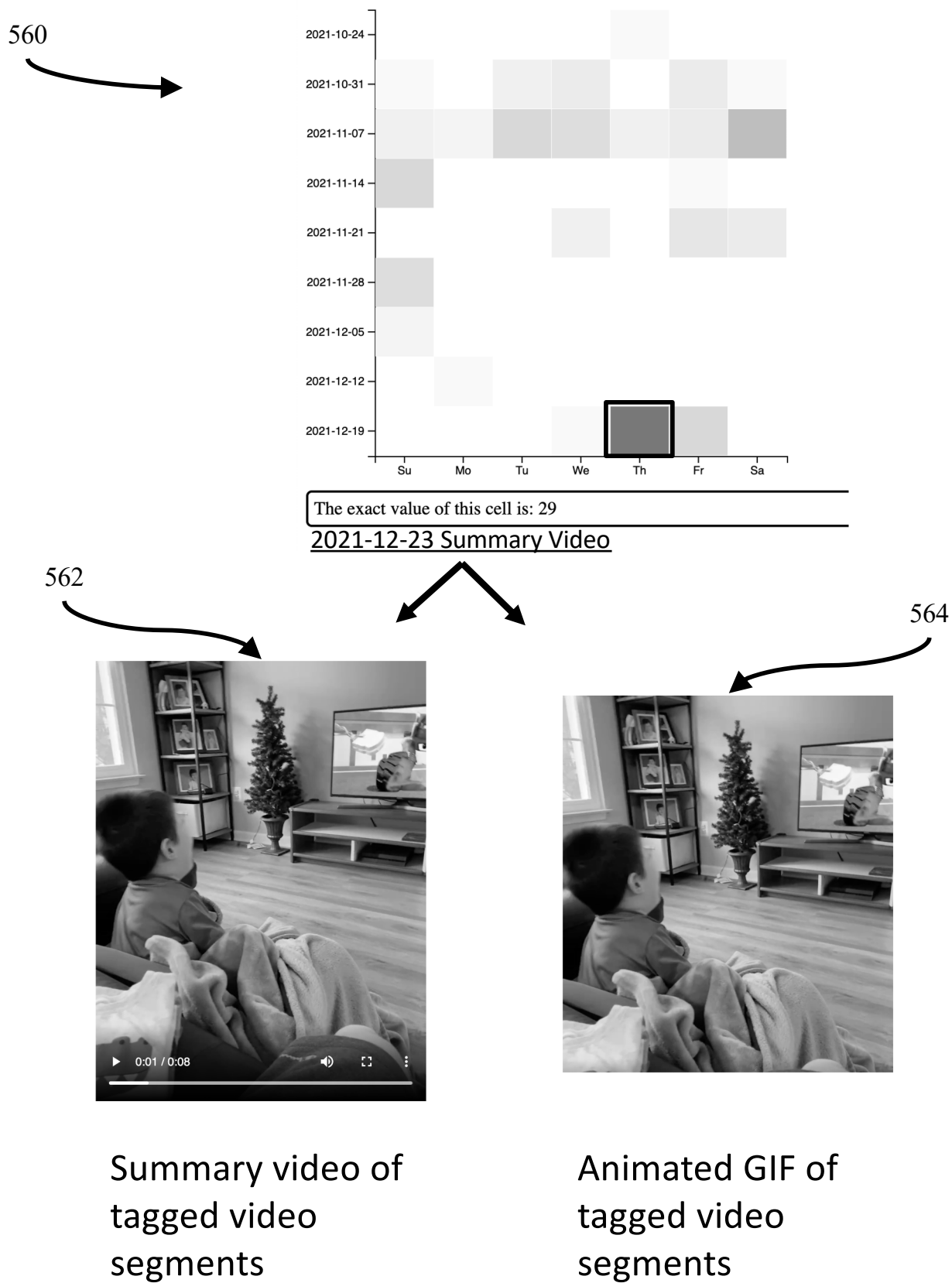
Laughing Tags

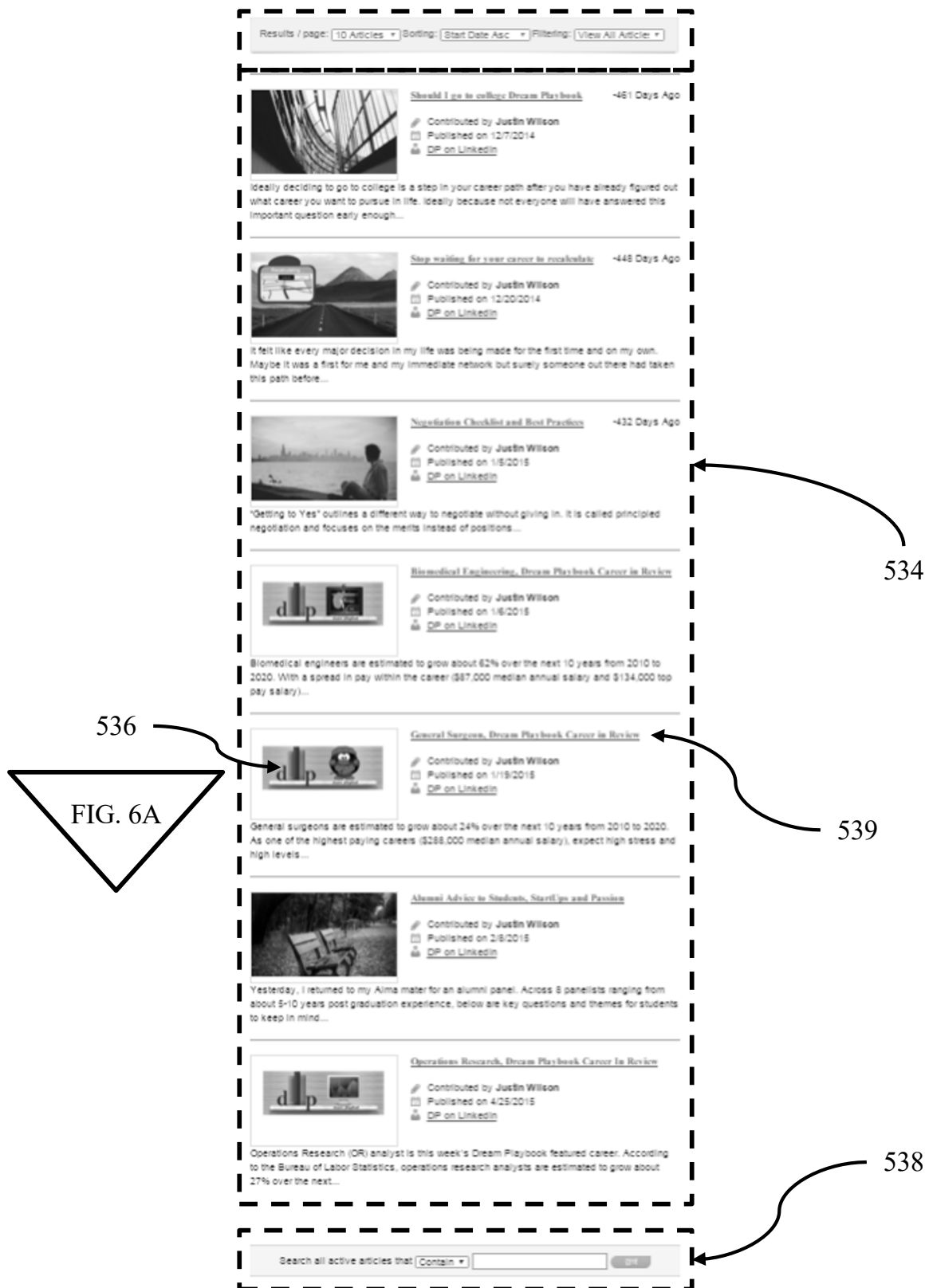


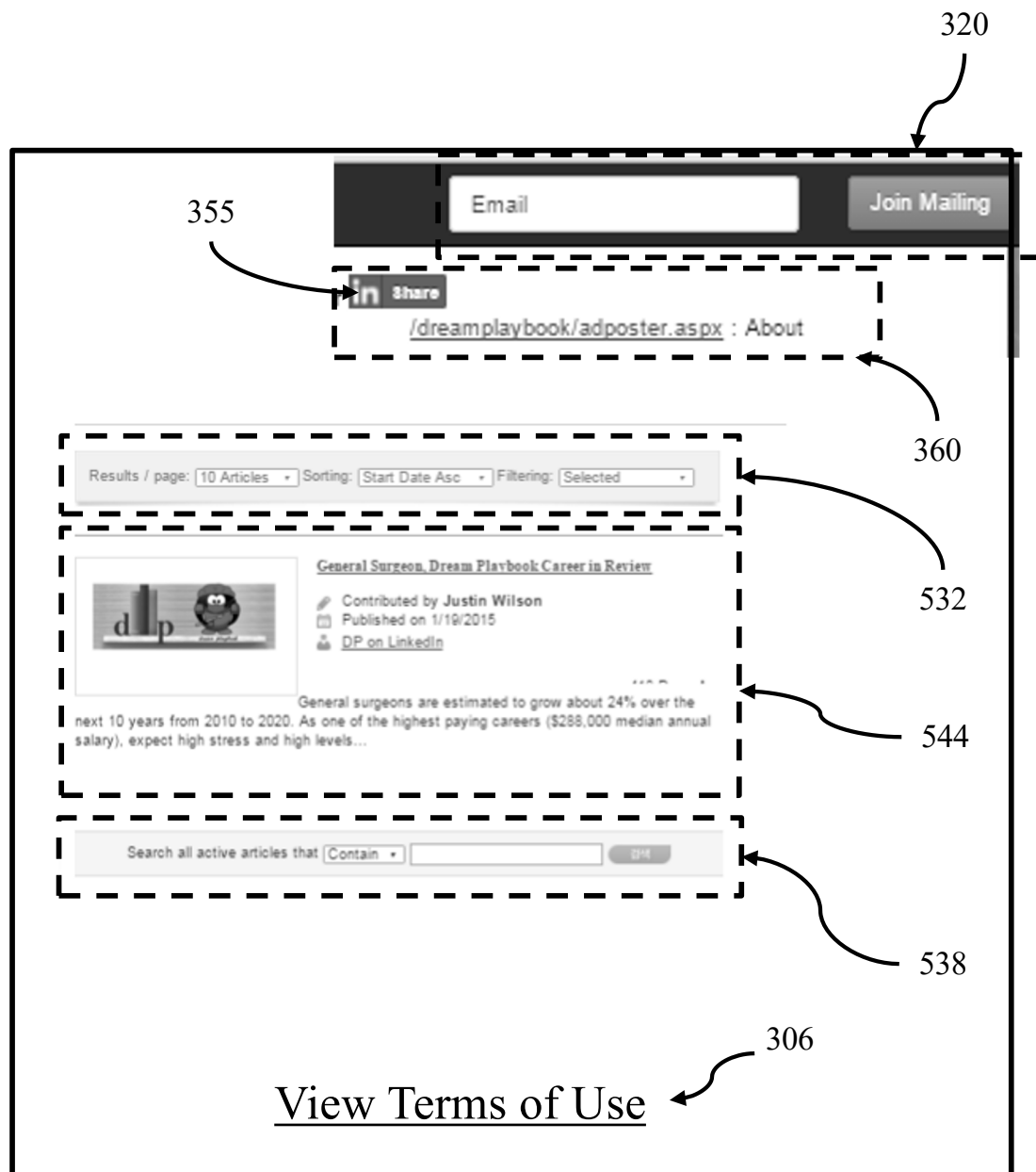
The exact value of this cell is: 29



The exact value of this cell is: 29







570

