



(51) International Patent Classification:

G06N 20/00 (2019.01) G06F 17/40 (2006.01)

(21) International Application Number:

PCT/US2020/050064

(22) International Filing Date:

10 September 2020 (10.09.2020)

(25) Filing Language:

English

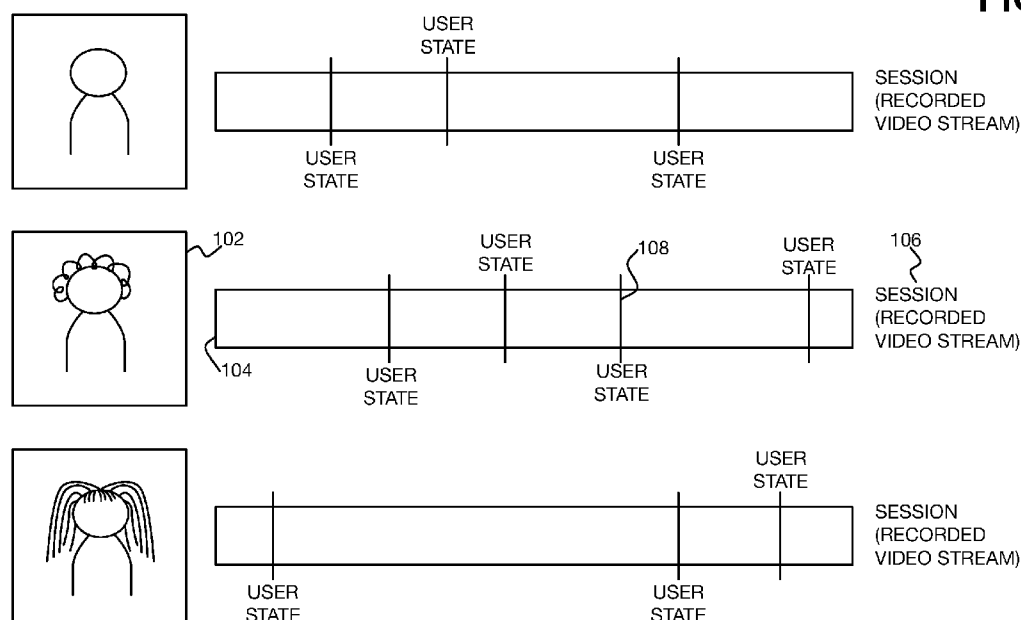
(26) Publication Language:

English

(71) Applicant: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 10300 Energy Dr., Spring, Texas 77389 (US).(72) Inventors: **ZOTTO, Rafael Dal**; Av. Ipiranga, 6681, Bld. 45C, Predios 5-6, 90619-900 Porto Alegre (BR). **SIEGEL, Erika Hansen**; 1501 Page Mill Rd, Palo Alto, California 94304-1100 (US). **BALLAGAS, Rafael**; 1501 Page Mill Rd, Palo Alto, California 94304-1100 (US). **LANDO, Gabriel**; Av. Ipiranga, 6681, Bld. 45C, Predios 5-6, 90619-900 Porto Alegre (BR). **SANTOS DA SILVA, JR., Alexandre**; Av. Ipiranga, 6681, Bld. 45C, Predios 5-6,90619-900 Porto Alegre (BR). **WEI, Jishang**; 286 Tanner Marsh Rd., Guilford, Connecticut 06437 (US). **LIU, Xing**; 1501 Page Mill Rd, Palo Alto, California 94304-1100 (US). **RAMOS, Jose Dirceu Grundler**; Av. Ipiranga, 6681, Bld. 45C, Predios 5-6, 90619-900 Porto Alegre (BR). **HORII, Hiroshi**; 1501 Page Mill Rd., Palo Alto, California 94304-1100 (US). **KUTHURU, Srikanth**; 1501 Page Mill Rd, Palo Alto, California 94304-1100 (US).(74) Agent: **JENNEY, Michael** et al.; HP Inc., Mail Stop 35, 3390 E Harmony Rd, Fort Collins, Colorado 80528 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) Title: TRAINING DATA COLLECTION DURING DATA COLLECTION-UNPOSED SESSIONS

FIG 1

(57) Abstract: While a data collection-unposed session is occurring, the session is recorded and a user is periodically requested at different times to specify labels associated with current user state in the session. The user-specified labels are received from the user and, along with the different times during the session at which the labels were periodically requested and specified, are recorded. The data collection-unposed session and the user-specified labels are collected as training data for a machine learning model that classifies user state via the labels.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

TRAINING DATA COLLECTION DURING DATA COLLECTION-UNPOSED SESSIONS

BACKGROUND

[0001] Machine learning models are a type of artificial intelligence
5 model by which a machine (e.g., a computing device or system) can
adaptively learn over time. A machine learning model can be employed to
classify presented data with labels, and a supervised such machine learning
model is trained with pre-labeled training data. The training data may be
divided into two sets. The model is trained on the labeled data of one set, and
10 then tested on the other set to assess the model's accuracy.

BRIEF DESCRIPTION OF THE DRAWINGS

[0002] FIG. 1 is a diagram of example machine learning model training
data collection during data collection-unposed sessions corresponding to a
videoconference.

15 **[0003]** FIG. 2 is a diagram of an example process by which machine
learning model training data can be collected during a data collection-unposed
session and subsequently used and refined.

[0004] FIG. 3 is a diagram of an example of how machine learning
model training data collected during a data collection-unposed session can be
20 refined.

[0005] FIG. 4 is a diagram of another example of how machine learning
model training data collected during a data collection-unposed session can
be refined.

[0006] FIG. 5 is a diagram of example machine learning model training data collection during a data collection-unposed session corresponding to user interaction with a graphical user interface (GUI) environment.

[0007] FIG. 6 is a diagram of example machine learning model training data collection during a data collection-unposed session corresponding to user interaction with an extended reality (XR) environment.

[0008] FIG. 7 is a diagram of an example computer-readable data storage medium.

[0009] FIG. 8 is a diagram of an example computing system.

10

DETAILED DESCRIPTION

[0010] As noted in the background, a machine learning model can be employed to classify data into labeled categories. For example, a video clip or segment of a person can be labeled to classify user sentiment, such as the evinced mood or emotion of the person in the video. For a longer video recording of the person, the machine learning model can classify changes in user sentiment as they unfold during the recording.

[0011] As also noted in the background, a supervised machine learning model is trained with pre-labeled training data. For example, a machine learning model that classifies the user sentiment of people appearing within video data may be trained using video segments of different people that have been pre-labeled with their user sentiment. The diversity, depth, and sheer amount of the training data can affect the accuracy of the machine learning model.

[0012] Acquiring sufficient training data, in quantity and quality, can be difficult. In the case of user sentiment classification, for instance, training data

25

may be acquired by having users watch previously recorded video clips and assess the user sentiment of the people within the clips. This acquisition technique is laborious and costly because it entails having users watch and label large quantities of video data. That is, the subjects of the video data do not themselves specify their user sentiment.

[0013] Another type of training data for user sentiment classification is posed videos of actors who have been instructed to act out an individual emotion (like fear or disgust). No after-the-fact assessment or labeling of user assessment by users other than the subjects of the video data occurs.

10 However, this technique can be problematic because the user sentiments are posed and forced as opposed to unfolding naturally.

[0014] That is, the scenario in which the training data is being collected is staged and may not represent natural expressions of sentiment as evinced in real life. For instance, a person who is instructed to “look happy” may assume a pose that differs markedly from unposed and unstaged expressions of happiness. Further, there may be significant variation in the expression of sentiment across cultures, age, and gender, limiting the value of posed data for model training.

[0015] Techniques described herein provide for collection of training data for supervised (as well as unsupervised) machine learning models that ameliorate these difficulties. While a data collection-unposed session is occurring, the user appearing within the session is periodically requested to specify labels associated with his or her current user state. The session is not data collection-unposed in that it is not staged for the purposes of collecting machine learning model training data. While the user him or herself is

periodically requested to specify the labels, the session itself is not prompted and does not occur for the purposes of collecting the training data.

[0016] For example, a user may participate in a videoconference in which a video stream of the user is recorded and shared with other

5 participants in the conference. The user may periodically be requested to specify labels associated with the current user sentiment the user is experiencing, evincing, or expressing. The user is not prompted to evince a specific user sentiment, in other words, but rather is periodically requested to specify what his or her current extemporaneous user sentiment is. The user
10 sentiment is thus not forced, staged, or posed.

[0017] Moreover, it is the user him or herself who is specifying what his or her user sentiment is, contemporaneously with when the user is evincing the sentiment, as opposed a different person who attempts after the fact to specify or label the user's sentiment. Because large numbers of

15 videoconferences occur all the time, a large amount of machine learning model training data can be collected without having to request any given user to specify user sentiment labels too often. The videoconferences themselves are conducted for other purposes, unrelated to machine learning model training data collection, but are leveraged to also for training data collection.

20 **[0018]** FIG. 1 shows example machine learning training data collection in the context of a videoconference. A video stream 104 of each of three users 102 participating in the videoconference in the example is captured at a corresponding client computing device and transmitted to client computing devices of the other users 102 for display. Each video stream 104 is also
25 transmitted to another computing device, which records the video stream 104

as a corresponding data collection-unposed session 106. Each session 106 is thus an audiovisual session of a corresponding user 102, and is more specifically a videoconference session of a corresponding user 102. There are, therefore, three audiovisual or videoconference sessions 106 in the example of FIG. 1.

[0019] While each session 106 is occurring, the session 106 (i.e., the corresponding video stream 104) is recorded. Also while each session 106 is occurring, the user 102 of that session 106 (i.e., appearing in the corresponding video stream 104) is periodically requested at different times to specify labels 108 associated with his or her current user state in the session 106. The user state may be user sentiment. The users 102 may be periodically requested at different times, and the number of times each user 102 is prompted to specify current user state labels 108 may likewise differ.

[0020] For example, a user 102 may be presented with a list of labels corresponding to different user states from which the user 102 is requested to specify or select the label 108 that is most indicative of his or her current user state. As to user sentiment, the labels may be “happy,” “angry,” “bored,” “annoyed,” and so on. The user may have the option to not respond to the request. For instance, when a user is periodically prompted to specify his or her current user state, the list of labels may be displayed on the screen of the user’s client computing device, and if the user does not select any label within a specified length of time the list may be removed from display.

[0021] The specified user state labels 108 are thus responsively received and recorded, along with the different times during their corresponding sessions 106 at which the labels 108 were requested and

specified. The recorded sessions 106 and the user-specified labels 108 are collected as training data for a machine learning model that classifies user state via the labels. In the example of FIG. 1, for instance, a machine learning model can be trained from the collected training data for subsequent application to other video streams, such as other videoconferences, to indicate current user sentiment of the people appearing in those streams.

[0022] FIG. 2 shows an example process 200 by which machine learning data can be collected during a data collection-unposed session and subsequently used and refined. The process 200 can be performed by one or multiple computing devices, which can be communicatively connected to one another and/or to a client computing device at which the data collection-unposed session of a user is occurring, such as over a network. The process 200 can be implemented as program code stored on a non-transitory computer-readable data storage medium and executed by the computing device or devices performing the process 200.

[0023] While a data collection-unposed session 202 is occurring, a computing device 200 performing the process 200 can receive and record (204) the session 202 as part of machine learning model training data 206 that is being collected. The computing device can periodically prompt (208) the user of the session 202 at different times to specify labels 210 associated with the user's current user state in the session 202. The computing device can also record (212) the received user state labels 210, along with the different times at which the labels were requested and specified, as part of the machine learning model training data 206.

[0024] A computing device performing the process 200 may also extract (214) metadata 216 regarding the data collection-unposed session 202. The metadata 216 may indicate technical information such as information regarding how the session 202 is being captured at the client computing device of the user, how the session 202 is being recorded at the computing device performing the process 200, the format of the data file of the recorded session 202, and so on. The metadata 216 may include other, non-technical information regarding the session 202 as well, such as the date, time, and location of the session 202, the identity of the user of the session 202, and so on. The computing device can also record (218) the metadata 216 as part of the machine learning training data 206.

[0025] A computing device performing the process 200 can perform machine learning model training (220) using the collected machine learning training data 206, which results in a trained machine learning model 222.

That is, the machine learning model 222 is trained using the recorded data collection-unposed session 202 and the recorded user-specified labels 210 associated with user state during the session 202. The machine learning model 222 is a model that classifies user state via labels, and can be used to label other sessions as they occur or after they have been recorded.

[0026] A computing device performing the process 200 can therefore classify user state of the same or different user of a different session 224 by applying (226) the trained machine learning model 222 to the session 224.

That is, the machine learning model 222 is used to classify user state via labels. The machine learning model 222 receives the session 224 as input

and outputs labels 228 associated with the user state. The machine learning

model 222 may provide one or multiple labels 228 that classifies user state over the entirety of the session 224, or may provide one or multiple labels 228 that classify user state at different times during the session 224.

[0027] A computing device performing the process 200 may also

5 perform different types of post-processing on the collected machine learning model training data 206 to refine the training data 206. The computing device may apply (230) the trained machine learning model 222 against the session 202 for which user-specified labels 210 have been received and recorded.

The machine learning model 222 receives the session 202 as input and

10 outputs model-predicted labels 232 corresponding to the user-specified labels 210. That is, for each different time during the session 202 at which the user specified labels 210, there are corresponding labels 232 generated by the model 222. The model-predicted labels 232 may be the same as or different than the user-specified labels 210.

15 **[0028]** A computing device performing the process 200 may thus

request that the user confirm and correct (234) the user-specified labels 210 on the basis of the model-predicted labels 232. For each different time for which the labels 232 differ from corresponding labels 210, the user may be prompted to confirm that the user-specified labels 210 are correct, or to

20 correct the user-specified labels 210 (e.g., instead select the model-predicted labels 232 as being more accurate of user sentiment). The user that is requested to perform the confirmation or correction can be the same user that previously specified the labels 210. The corrected labels 236 are recorded (212) to update the machine learning model training data 206, and the model
25 222 can be retrained (220).

[0029] As another example of post-processing of the collected machine learning model training data 206, a computing device performing the process 200 may divide (238) the recorded session 202 into session segments 240 in correspondence with the different times during the session 202 at which the user specified the labels 210. For instance, for each different time during the session 202 at which the user specified the labels 210, the computing device may define a corresponding session segment 240 starting at that time and ending a specified length of time later. In another implementation, the trained machine learning model 222 may instead indicate the starting and ending times of each session segment 240.

[0030] A computing device performing the process 200 may thus request that the user confirm and correct (242) the temporal boundaries of the session segments 240 to which the user-specified labels 210 apply.

Specifically, for each different time during the session 202 at which the user specified the labels 210, the user may be prompted to confirm or correct the temporal boundaries of a corresponding session segment 240. The user may adjust the starting and/or ending times of the session segment 240 backwards or forwards so that the segment 240 begins and ends when the user actually began and ceased evincing a corresponding user state. The resulting corrected session segments 244 can be recorded (246) along with the labels 210 for usage as the machine learning model training data 206.

[0031] FIG. 3 shows an example of how collected machine learning model training data can be refined, such as in part 234 of FIG. 2. During a data collection-unposed session 300, a user was prompted at time 302 to provide a user-specified label 304U reflecting the user's current user state at

the time 302. The resultantly trained machine learning model generated a corresponding model-predicted label 304P for the time 302. After the session 300 has been recorded, therefore, the user can be prompted to confirm or correct the user-specified label 304U, such as if the label 304U differs from the model-predicted label 304P. Both user state labels 304P may be displayed to the user during the confirmation and correction process.

[0032] FIG. 4 shows another example of how collected machine learning model training data can be refined, such as in part 242 of FIG. 2. During a data collection-unposed session 400, a session segment 402 of the session 400 may have an initially specified start time 404S corresponding to the time at which the user was prompted for a user-specified label 406 associated with the user's current user state. The session segment 402 may have an initially specified end time 404E that may correspond to a predefined length of time after the start time 404S. After the session 300 has been recorded, the user can thus be prompted to confirm and correct the temporal boundaries of the segment 402 to which the label 406 applies. The user may adjust the start and end times 404S and 404E to more accurately reflect when the user evinced the user state associated with the label 406.

[0033] FIG. 5 shows example machine learning data collection in the context of a scenario 500 in which a user is interacting with a graphical user interface (GUI) environment 502 of a client computing device. For example, the GUI environment 502 may be the displayed GUI desktop, and may include windows and other GUI objects with which the user can interact. A stream 504 of the GUI environment 502 is captured at the client computing device and a corresponding stream 506 of the user's face is captured by a camera as

the user interacts with the GUI environment 502. The streams 504 and 506 are transmitted to another computing device, which records them as a data collection-unposed computing session 508, specifically a GUI computing session.

5 **[0034]** While the session 508 is occurring, the user interacting with the GUI environment 502 is periodically requested at different times to specify labels 510 associated with his or her current user state in the session 508. The user state may be user eye gaze as to which of the different GUI objects of the GUI environment 502 the user's gaze is currently directed (viz., which
10 GUI object the user is currently looking at). The specified user state labels 510 are responsively received and recorded, along with the different times during the session 508 at which the labels 510 were requested.

[0035] The recorded session 508 and the user-specified labels 510 are therefore collected as training data for a machine learning model that
15 classifies user state via the labels. In the example of FIG. 5, for instance, a machine learning model can be trained from the collected training data for subsequent application to other computing sessions, such as computing sessions made up of desktop and face streams. The machine learning model can thus be used to determine the object in the desktop (or other) stream at
20 which the user is currently directing his or her gaze, per the user's face in the face stream.

[0036] FIG. 6 shows example machine learning data collection in the context of a scenario 600 in which a user is wearing extended reality (XR) goggles or glasses to interact with an XR environment. For example, the XR
25 environment may be an augmented reality (AR) environment in which the user

is interacting with physical, real-world objects, such that the goggles or glasses display information regarding the objects to augment the user's reality.

As another example, the XR environment may be a virtual reality (VR) environment in which the user is interacting with virtual objects displayed by the goggles or glasses to immerse the user within a virtual reality.

[0037] A stream 604 of the XR environment 602 – including the real-world or virtual objects of the environment 602 – is captured at the goggles or glasses as the user interacts with the XR environment 602. A corresponding stream 606 of the user's eyes is likewise captured by the goggles or glasses as the user interacts with the XR environment 602. The streams 604 and 606 are transmitted by a client computing device to a computing device that records the streams 604 and 606 as a data collection-unposed computing session, specifically an XR computing session 608.

[0038] While the session 608 is occurring, the user interacting with the XR environment 602 is periodically requested at different times to specify labels 610 associated with his or her current user state in the session 608. The user state may be user eye gaze as to which of the different real-world or virtual objects of the XR environment 602 the user's gaze is currently directed (viz., which object the user is currently looking at). The specified user state labels 610 are responsively received and recorded, along with the different times during the session 608 at which the labels 610 were requested.

[0039] The recorded session 608 and the user-specified labels 610 are therefore collected as training data for a machine learning model that classifies user state via the labels. In the example of FIG. 6, for instance, a machine learning model can be trained from the collected training data for

subsequent application to other computing sessions, such as computing sessions made up of XR and eye streams. The machine learning model can thus be used to determine the object in the XR (or other) stream at which the user is currently directing his or her gaze, per the user's eyes in the eye stream.

[0040] FIG. 7 shows an example non-transitory computer-readable data storage medium 700 storing program code 702 executable by a processor to perform processing. The processing includes, while a data collection-unposed session is occurring, recording the session and periodically requesting at different times that a user specify labels associated with current user state in the session (704). The processing includes responsively receiving the user-specified labels from the user and recording the labels along with the different times during the session at which the labels were periodically requested and specified (706). The session and the labels are collected as training data for a machine learning model that classifies user state via the labels.

[0041] FIG. 8 shows an example computing system 800. The computing system 800 may be implemented as one or multiple computing devices. The computing system 800 includes network hardware to communicatively connect to client devices of users that participate in data collection-unposed sessions. For example, each session may be an audiovisual session like a videoconference session per FIG. 1. As other examples, each session may be a computing session like a GUI computing session per FIG. 5 or an XR computing session per FIG. 6.

[0042] The computing system 800 includes a storage device 804, such as a hard disk drive or a solid-state drive (SSD), to store machine learning model training data 806. The computing system 800 includes a processor 808 and a memory 810 that stores program code 812 executable by the processor 808. The program code 812 is executable by the processor 808 to periodically prompt during the data collection-unposed sessions the users participating in the sessions to specify labels associated with current user state (814). The program code 812 is executable by the processor 808 to record the sessions and the labels along with times at which the labels were specified within the sessions to collect the machine learning model training data 806 (816).

[0043] Techniques have been described for collecting training data for machine learning models. The training data is collected during occurrence of a session that is unposed, unstaged, and unprompted for data collection purposes. Therefore, the training data includes labels associated with current user state specified by the user of the session. Because the user him or herself contemporaneously specifies his or her current user state without being prompted to pose or stage any particular user state, the collected machine learning model training data can result in a more accurate model.

20

We claim:

1. A non-transitory computer-readable data storage medium storing program code executable by a processor to perform processing comprising:
 - while a data collection-unposed session is occurring, recording the
 - 5 session and periodically requesting at different times that a user specify labels associated with current user state in the session; and
 - responsively receiving the user-specified labels from the user and recording the labels along with the different times during the session at which the labels were periodically requested and specified,
 - 10 wherein the data collection-unposed session and the user-specified labels are collected as training data for a machine learning model that classifies user state via the labels.
2. The non-transitory computer-readable data storage medium of claim 1, wherein the session is data collection-unposed in that the session is not
- 15 staged to collect the training data for the machine learning model.
3. The non-transitory computer-readable data storage medium of claim 1, wherein the session is data collection-unposed in that while the user is periodically prompted to specify the labels associated with the current user state in the session, the session itself is unprompted for machine learning
- 20 model training data collection purposes.
4. The non-transitory computer-readable data storage medium of claim 1, wherein the processing further comprises:

training the machine learning model using the recorded data collection-unposed session and the recorded user-specified labels.

5. The non-transitory computer-readable data storage medium of claim 4, wherein the processing further comprises:

5 using the trained machine learning model to classify the user state via the labels during a different session.

6. The non-transitory computer-readable data storage medium of claim 1, wherein the processing further comprises:

applying the machine learning model to the collection-unposed session
10 to generate model-predicted labels for the different times during the session;
for each different time for which the model-predicted labels differ from the user-specified labels, request the user to confirm or correct the user-specified labels; and
training the machine learning model based on the confirmed or
15 corrected user-specified labels to improve accuracy of the model in classifying the user state.

7. The non-transitory computer-readable data storage medium of claim 1, wherein the processing further comprises:

for each different time at which the labels were periodically requested
20 and specified, request that the user confirm or correct temporal boundaries of a corresponding session segment within the session to which the labels apply.

8. The non-transitory computer-readable data storage medium of claim 1, wherein the processing further comprises:

extracting and recording metadata regarding the data collection-unposed session,

5 wherein the training data for the machine learning includes the metadata in addition to the session and the user-specified labels.

9. The non-transitory computer-readable data storage medium of claim 1, wherein the processing further comprises:

dividing the recorded session into session segments in

10 correspondence with the different times during the session at which the user specified the labels associated with the current user state,

wherein the session segments along with the user-specified labels to which the session segments correspond are used as the training data for the machine learning model.

15 10. The non-transitory computer-readable data storage medium of claim 1, wherein the data collection-unposed session comprises an audiovisual session of the user,

wherein the user state comprises user sentiment during the session.

11. The non-transitory computer-readable data storage medium of claim 10,

20 wherein the audiovisual session is one of a plurality of videoconference sessions of a plurality of videoconference participants including the user.

12. The non-transitory computer-readable data storage medium of claim 1,
wherein the data collection-unposed session comprises a computing session
in which the user is interacting with different objects,
and wherein the user state comprises user eye gaze in relation to the
5 different objects during the session.

13. The non-transitory computer-readable data storage medium of claim 12,
wherein the computing session comprises a graphical user interface (GUI)
computing session in which the user is interacting with different GUI objects of
a GUI environment,
10 and wherein the user state comprises the user eye gaze as to which of
the different GUI objects the user eye gaze is currently directed.

14. The non-transitory computer-readable data storage medium of claim 12,
wherein the computing session comprises an extended-reality (XR) computing
session in which the user is interacting with different virtual or real-world
15 objects within an XR environment,
and wherein the user state comprises the user eye gaze as which of
the different virtual or real-world objects the user eye gaze is currently
directed.

15. A computing system comprising:
20 network hardware to communicatively connect to a plurality of client
devices of users that participate in data collection-unposed sessions;
a storage device to store machine learning model training data;
a processor; and

a memory storing program code executable by the processor to:

periodically prompt during the data collection-unposed sessions

that the users participating in the sessions specify labels associated with

current user state; and

5 record the sessions and the user-specified labels along with

times at which the labels were specified within the sessions to collect the

machine learning model training data.

FIG 1

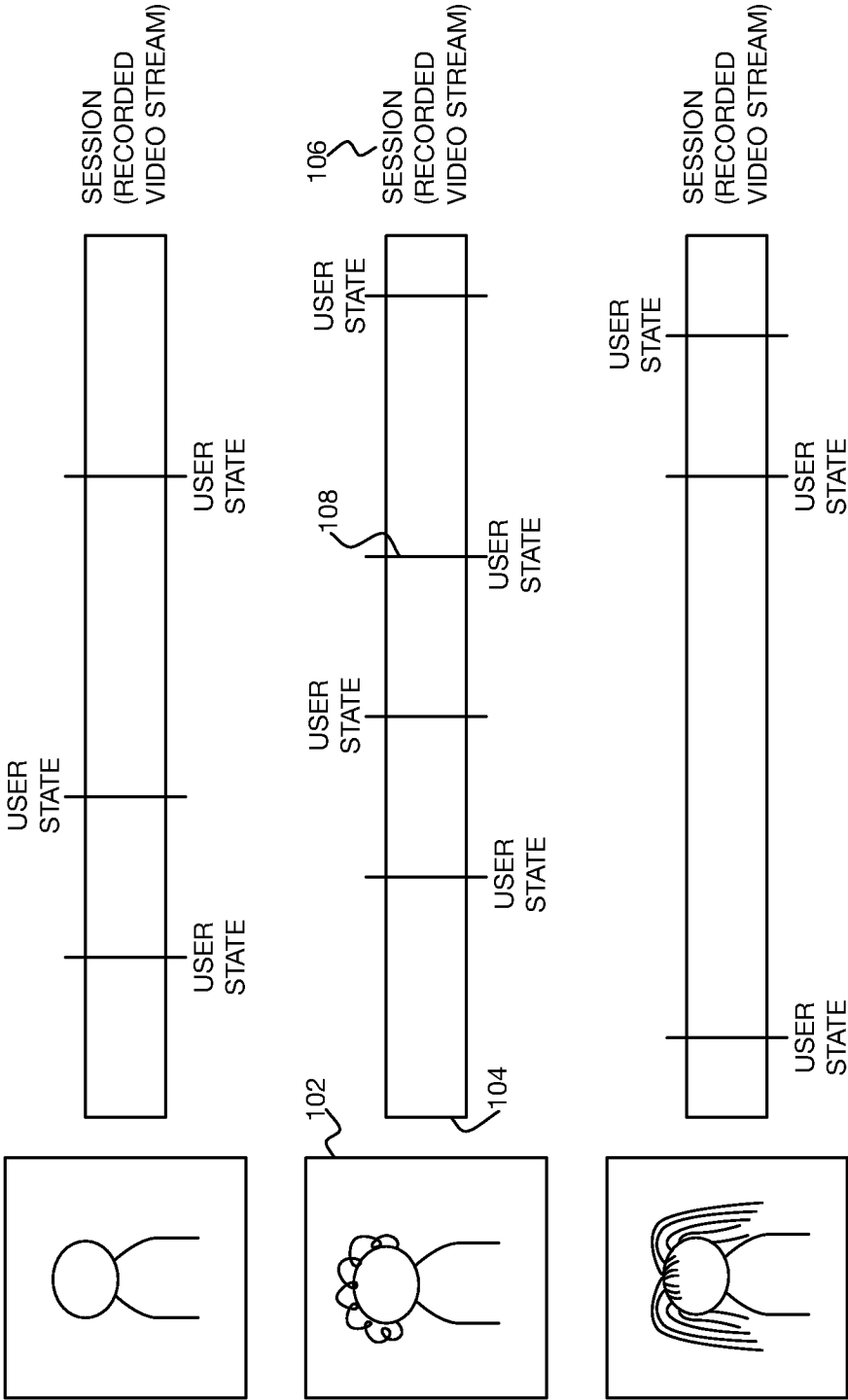


FIG 2

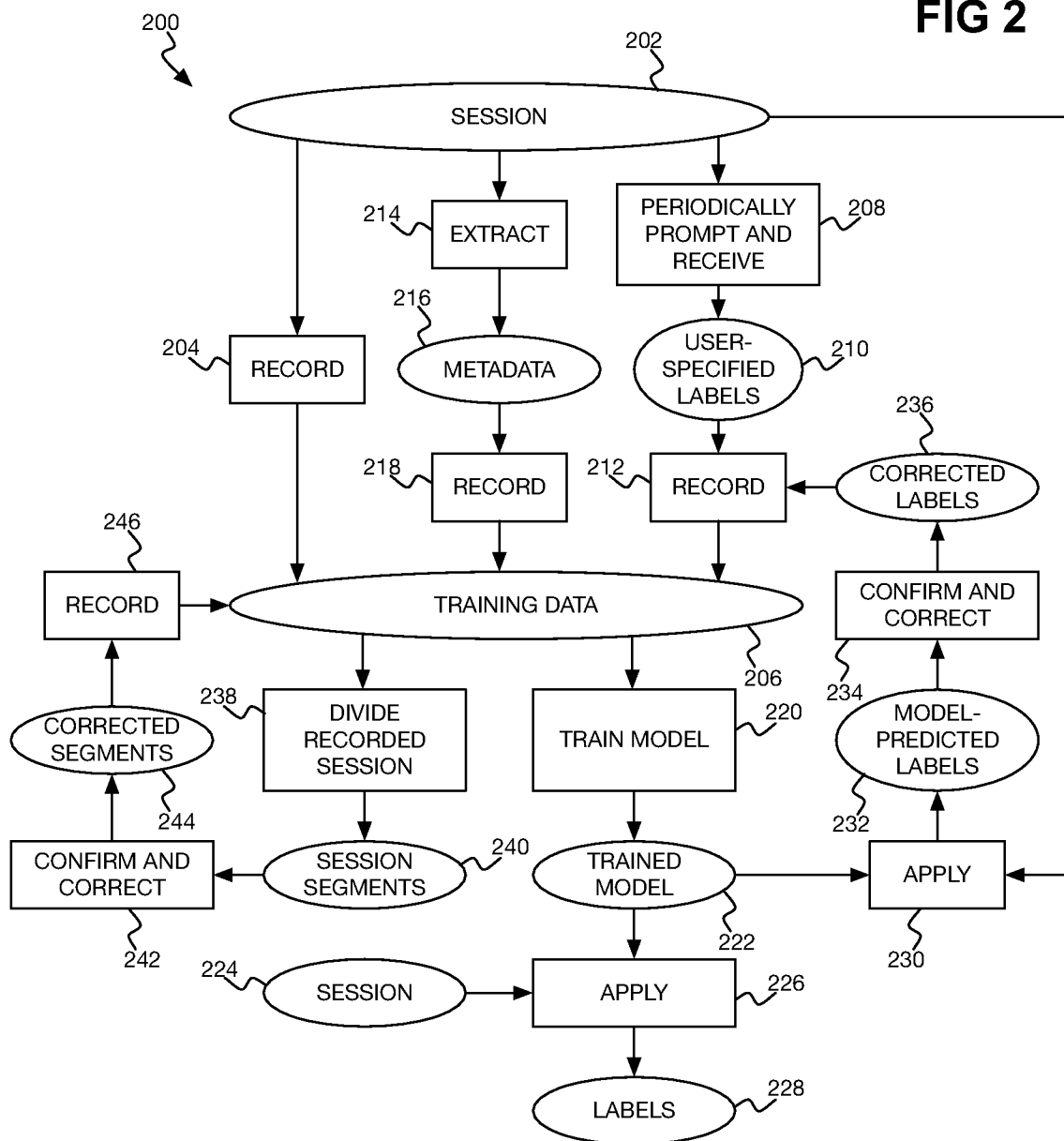


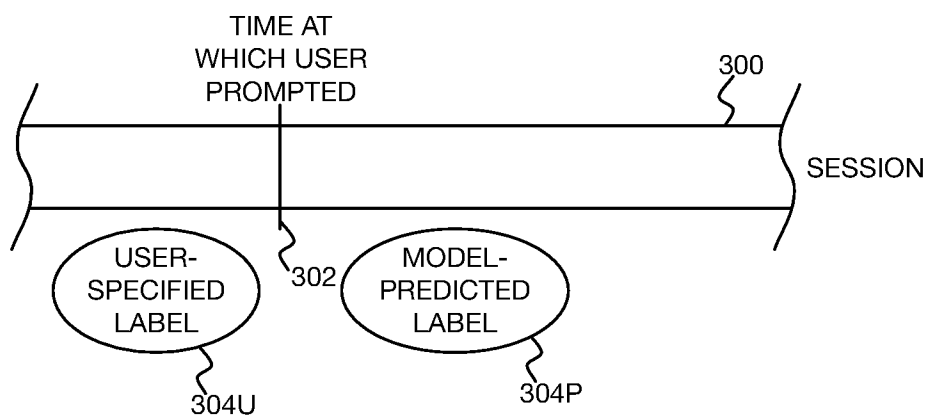
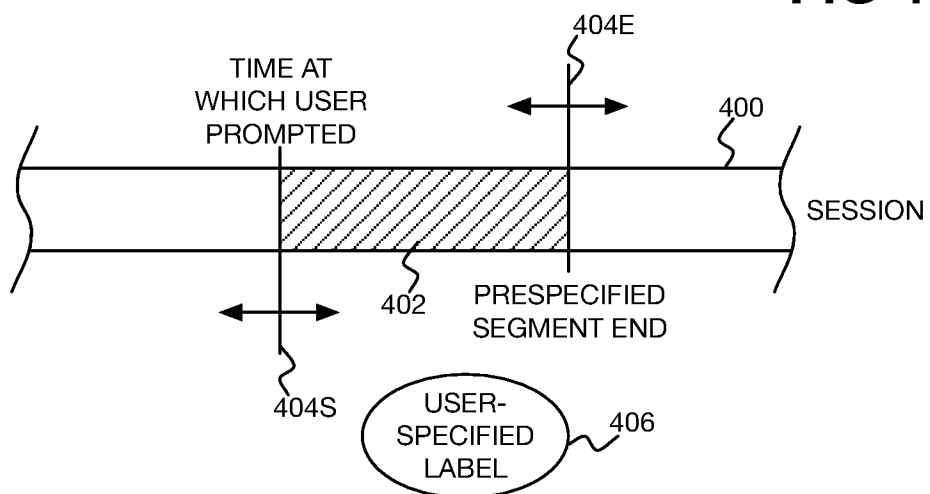
FIG 3**FIG 4**

FIG 5

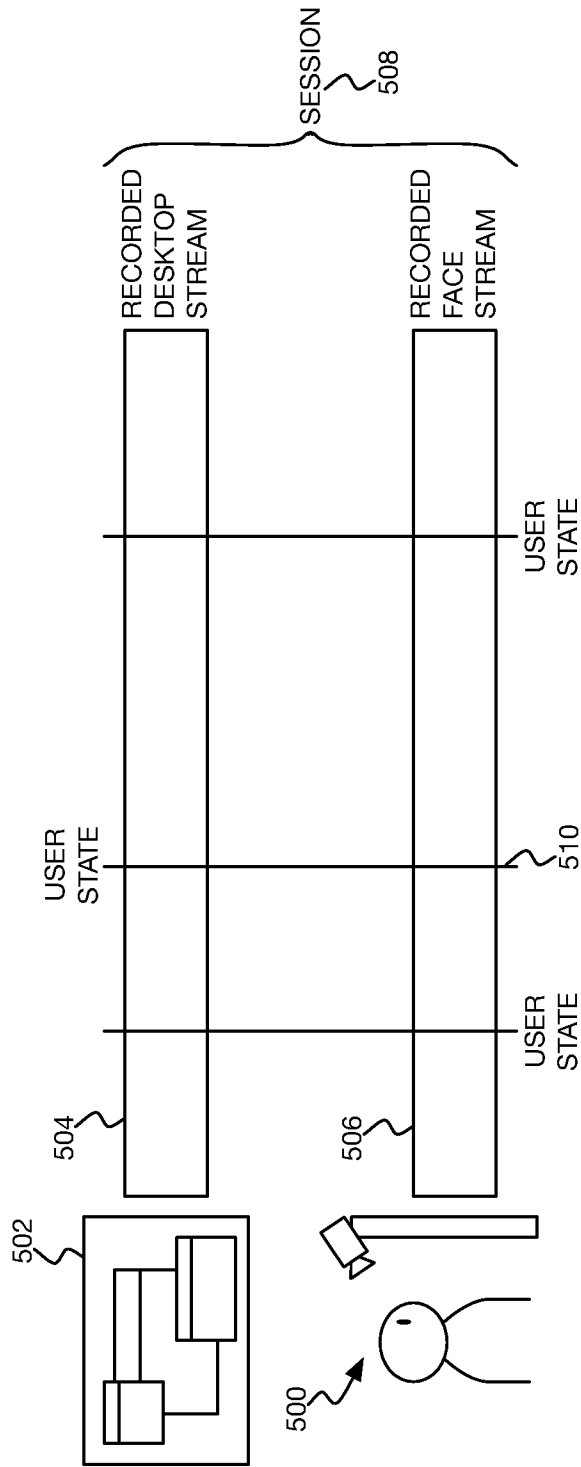


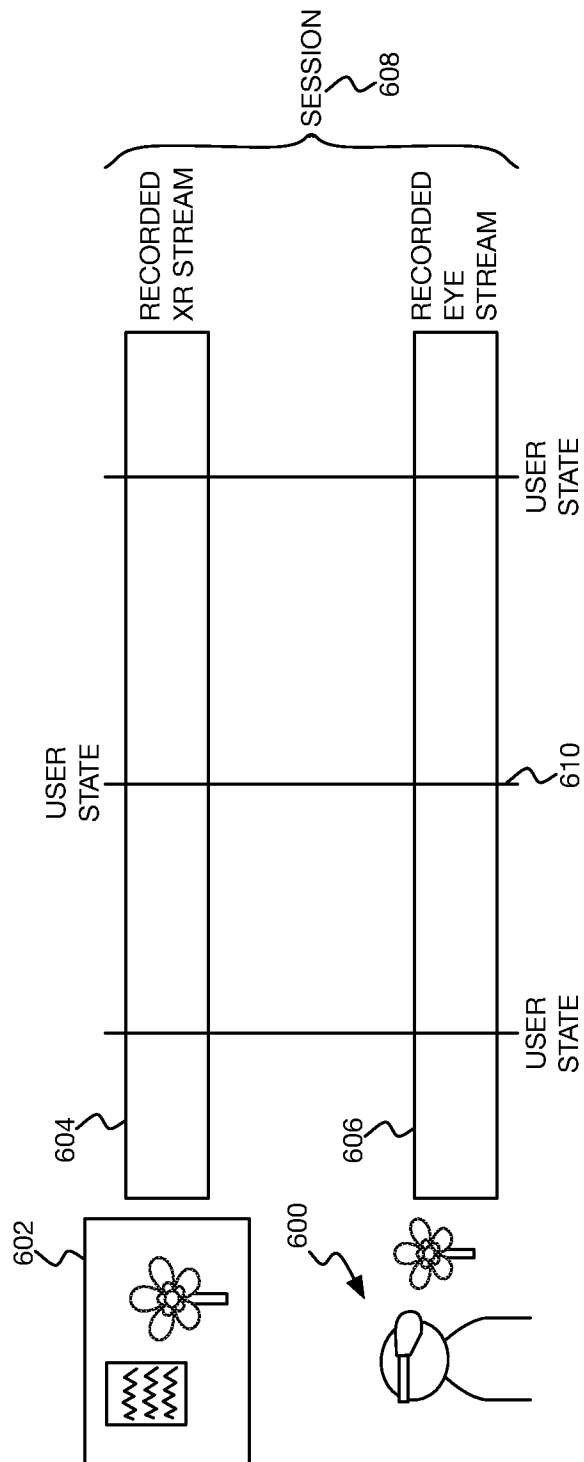
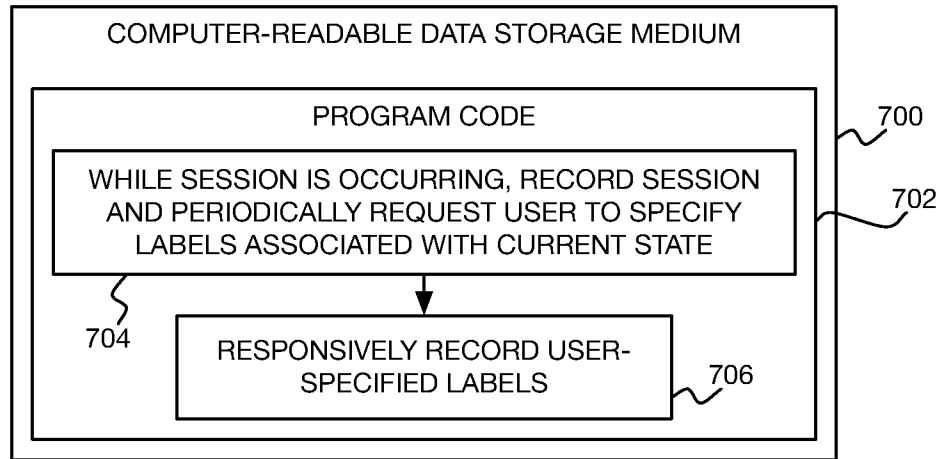
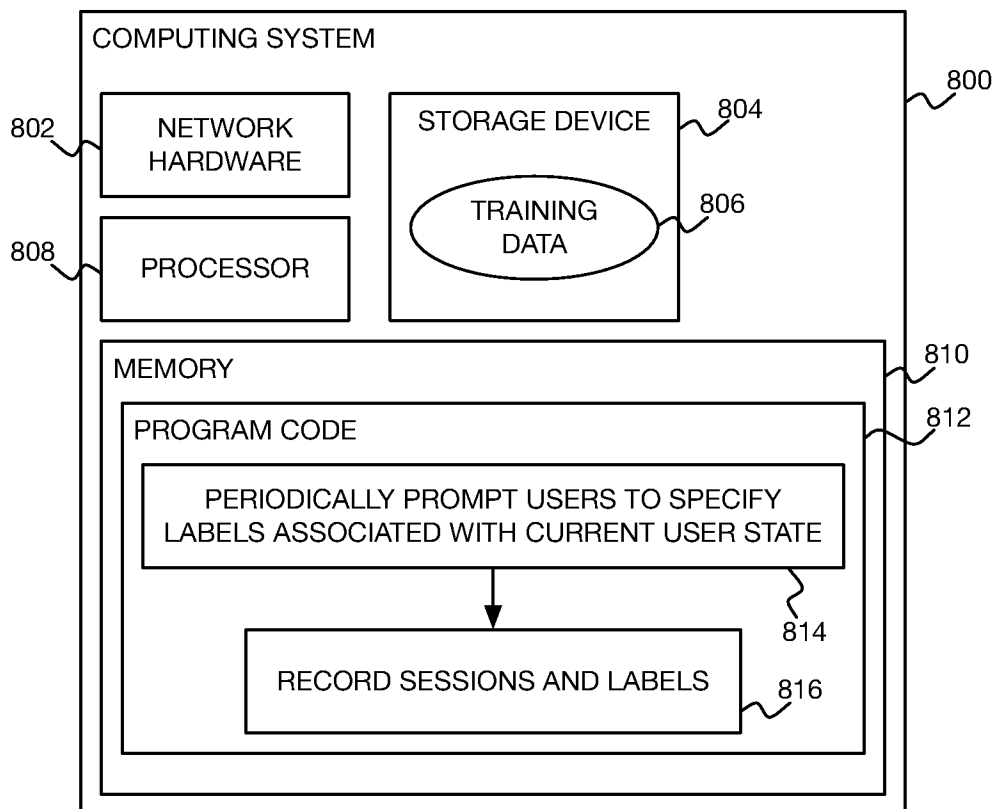
FIG 6

FIG 7**FIG 8**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 2020/050064

A. CLASSIFICATION OF SUBJECT MATTER <i>G06N 20/00 (2019.01)</i> <i>G06F 17/40 (2006.01)</i> According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06N 20/00, G06F 7/00, 17/00-17/40, G06K 9/00-9/72, G06Q 10/00, 30/00, 30/02, 50/00, G10L 15/00-15/30 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) PatSearch (RUPTO Internal), USPTO, PAJ, Espacenet, Information Retrieval System of FIPS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2015/0039405 A1 (ARI M. FRANK et al.) 05.02.2015, paragraphs [0004], [0012], [0037]-[0049], [0056]-[0058], [0070], [0114], [0395], [0397], [0432]-[0434], claims 1, 20, figures 2, 3, 5	1-15
A	US 2020/0160385 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 21.05.2020	1-15
A	US 10216850 B2 (FACEBOOK, INC.) 26.02.2019	1-15
A	US 2020/0219495 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 09.07.2020	1-15
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents: “A” document defining the general state of the art which is not considered to be of particular relevance “D” document cited by the applicant in the international application “E” earlier document but published on or after the international filing date “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) “O” document referring to an oral disclosure, use, exhibition or other means “P” document published prior to the international filing date but later than the priority date claimed	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art “&” document member of the same patent family	
Date of the actual completion of the international search 11 May 2021 (11.05.2021)	Date of mailing of the international search report 20 May 2021 (20.05.2021)	
Name and mailing address of the ISA/RU: Federal Institute of Industrial Property, Berezhkovskaya nab., 30-1, Moscow, G-59, GSP-3, Russia, 125993 Facsimile No: (8-495) 531-63-18, (8-499) 243-33-37	Authorized officer D. Starshinov Telephone No. 8(495)531-65-15	