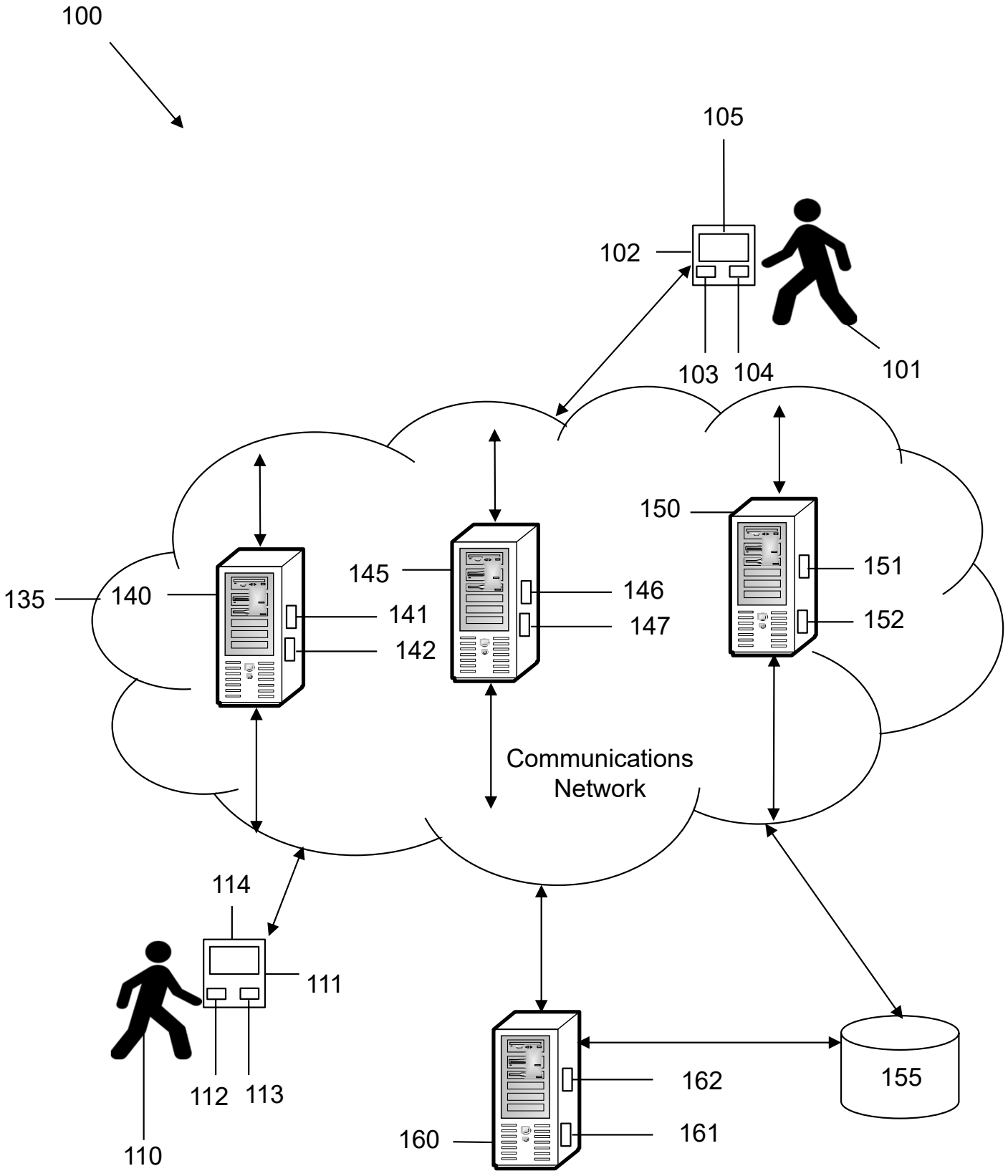

100
105
102
103
104
101
150
151
152
145
146
147
135
140
141
142
Communications Network
114
111
112
113
110
162
161
160
155

FIG. 1

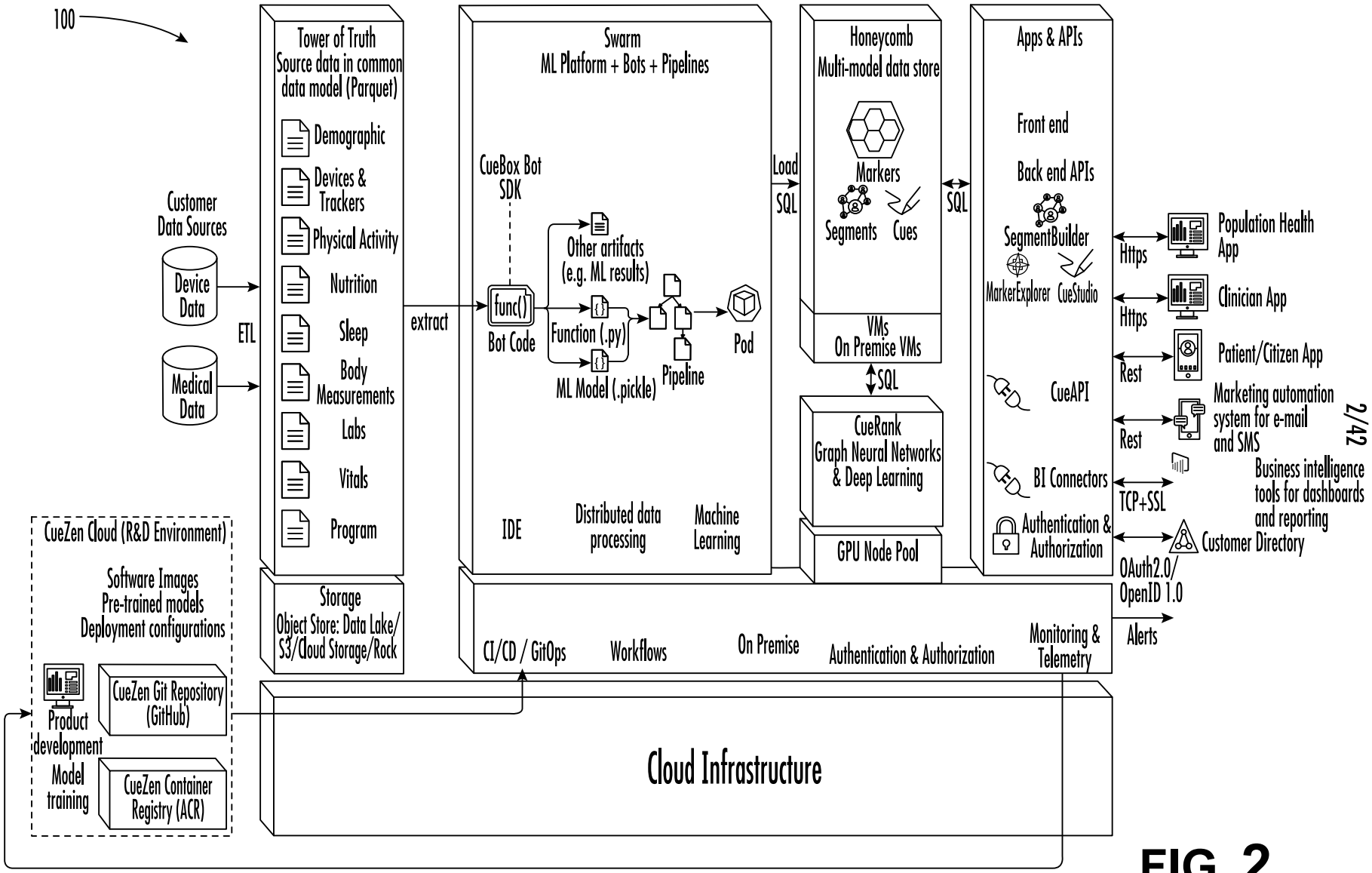

100
Tower of Truth
Source data in common data model (Parquet)
Demographic
Devices & Trackers
Physical Activity
Nutrition
Sleep
Body Measurements
Labs
Vitals
Program
Customer Data Sources
Device Data
Medical Data
ETL
Storage
Object Store: Data Lake/ S3/Cloud Storage/Rock
extract
Swarm
ML Platform + Bots + Pipelines
CueBox Bot SDK
func()
Bot Code
Other artifacts (e.g. ML results)
Function (.py)
ML Model (.pickle)
Pipeline
Pod
IDE
Distributed data processing
Machine Learning
Load
SQL
Honeycomb
Multi-model data store
Markers
Segments
Cues
VMs
On Premise VMs
SQL
CueRank
Graph Neural Networks & Deep Learning
GPU Node Pool
SQL
Apps & APIs
Front end
Back end APIs
SegmentBuilder
MarkerExplorer
CueStudio
CueAPI
BI Connectors
Authentication & Authorization
Https
Population Health App
Https
Clinician App
Rest
Patient/Citizen App
Rest
Marketing automation system for e-mail and SMS
TCP+SSL
Business intelligence tools for dashboards and reporting
Customer Directory
OAuth2.0/ OpenID 1.0
Alerts
CueZen Cloud (R&D Environment)
Software Images
Pre-trained models
Deployment configurations
Product development
CueZen Git Repository (GitHub)
Model training
CueZen Container Registry (ACR)
CI/CD / GitOps
Workflows
On Premise
Authentication & Authorization
Monitoring & Telemetry
Cloud Infrastructure
2/42

FIG. 2

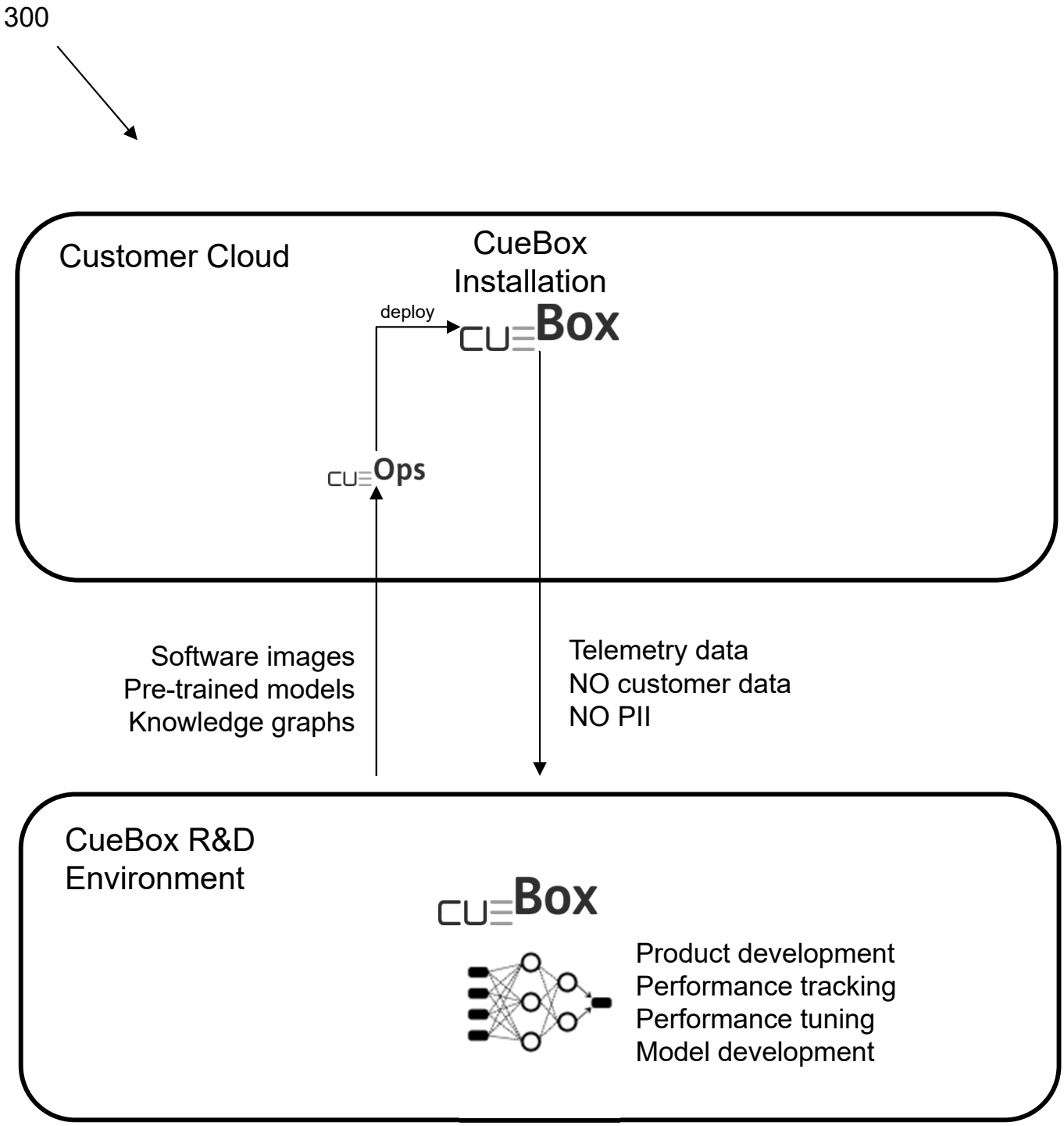

300
Customer Cloud
CueBox
Installation
deploy
CUEBox
CUEOps
Software images
Pre-trained models
Knowledge graphs
Telemetry data
NO customer data
NO PII
CueBox R&D
Environment
CUEBox
Product development
Performance tracking
Performance tuning
Model development

FIG. 3

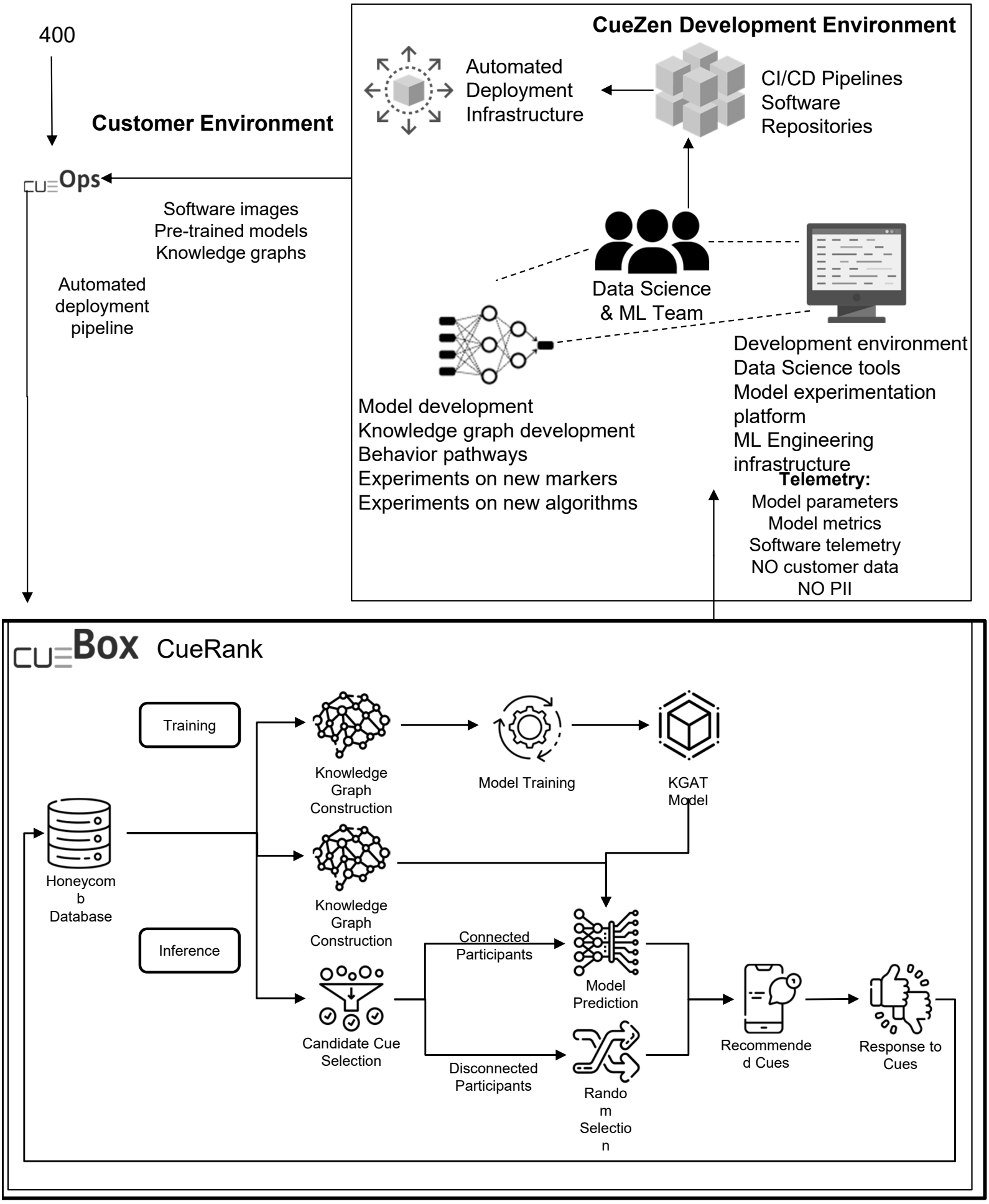

400
Customer Environment
CueZen Development Environment
Automated Deployment Infrastructure
CI/CD Pipelines
Software Repositories
CueOps
Software images
Pre-trained models
Knowledge graphs
Automated deployment pipeline
Data Science & ML Team
Development environment
Data Science tools
Model experimentation platform
ML Engineering infrastructure
Model development
Knowledge graph development
Behavior pathways
Experiments on new markers
Experiments on new algorithms
Telemetry:
Model parameters
Model metrics
Software telemetry
NO customer data
NO PII
CueBox CueRank
Training
Knowledge Graph Construction
Model Training
KGAT Model
Honeycomb Database
Knowledge Graph Construction
Inference
Connected Participants
Model Prediction
Candidate Cue Selection
Disconnected Participants
Random Selection
Recommended Cues
Response to Cues

FIG. 4

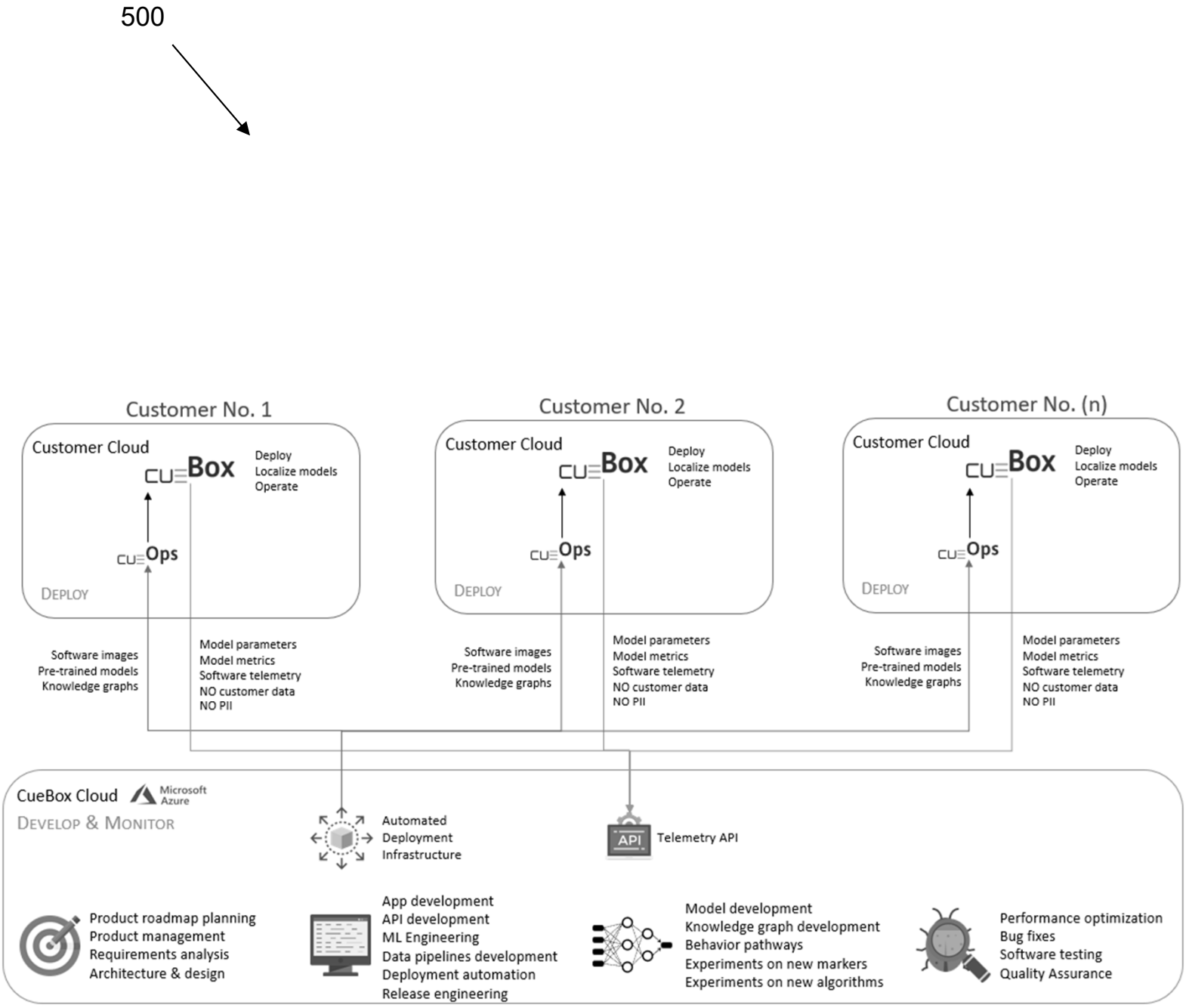

500
Customer No. 1
Customer Cloud
CUBox
Deploy
Localize models
Operate
CUOps
DEPLOY
Software images
Pre-trained models
Knowledge graphs
Model parameters
Model metrics
Software telemetry
NO customer data
NO PII
Customer No. 2
Customer Cloud
CUBox
Deploy
Localize models
Operate
CUOps
DEPLOY
Software images
Pre-trained models
Knowledge graphs
Model parameters
Model metrics
Software telemetry
NO customer data
NO PII
Customer No. (n)
Customer Cloud
CUBox
Deploy
Localize models
Operate
CUOps
DEPLOY
Software images
Pre-trained models
Knowledge graphs
Model parameters
Model metrics
Software telemetry
NO customer data
NO PII
CueBox Cloud
Microsoft Azure
DEVELOP & MONITOR
Automated
Deployment
Infrastructure
API
Telemetry API
Product roadmap planning
Product management
Requirements analysis
Architecture & design
App development
API development
ML Engineering
Data pipelines development
Deployment automation
Release engineering
Model development
Knowledge graph development
Behavior pathways
Experiments on new markers
Experiments on new algorithms
Performance optimization
Bug fixes
Software testing
Quality Assurance

FIG. 5

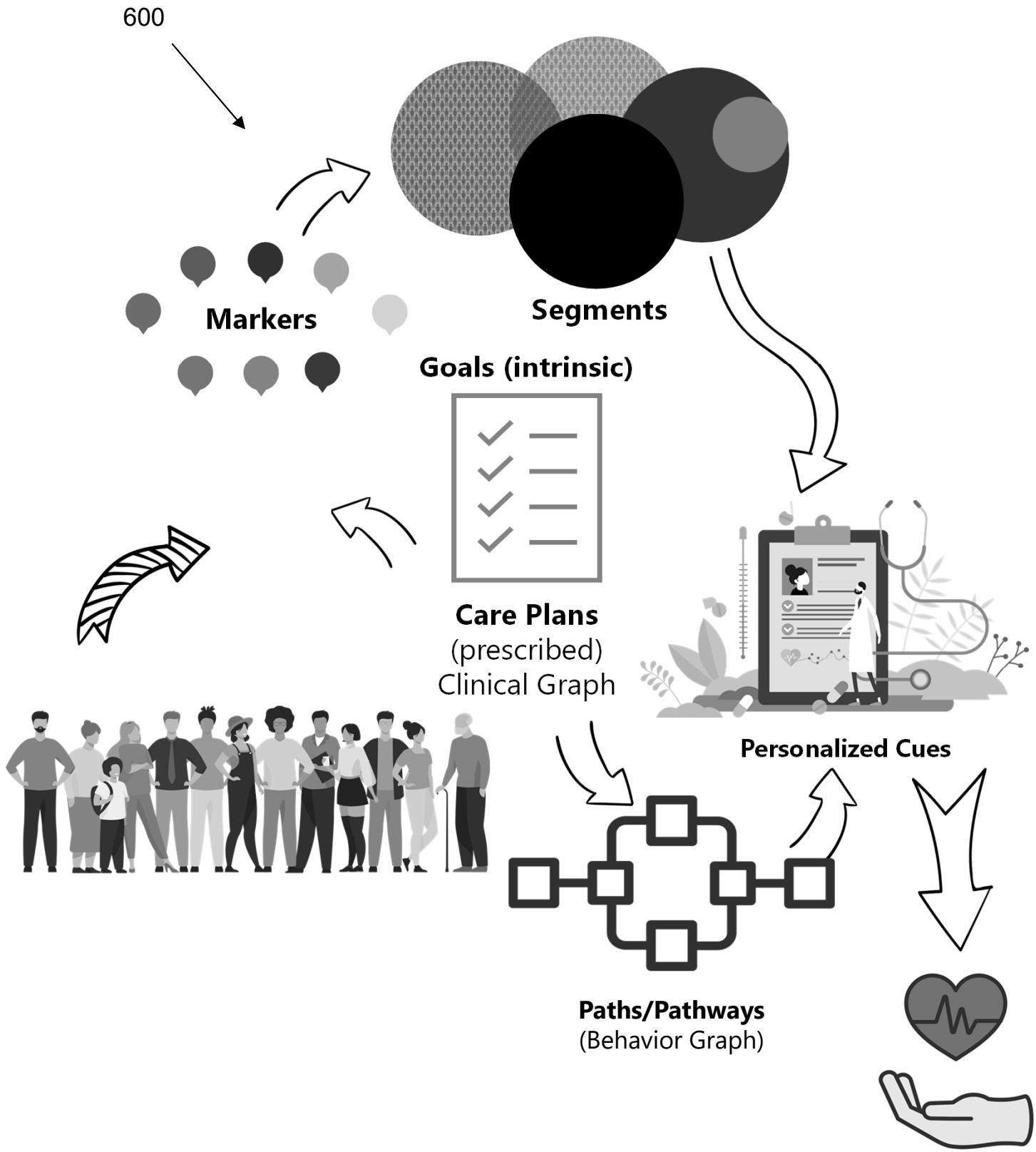
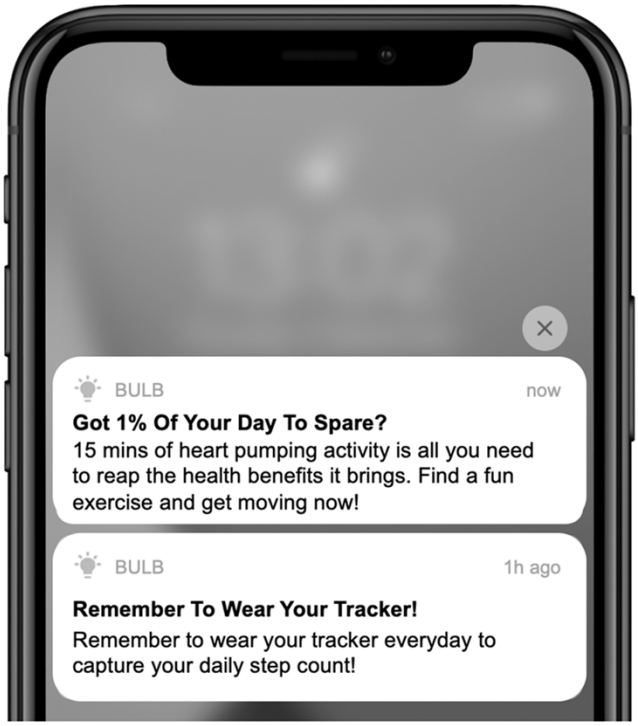

600
Markers
Segments
Goals (intrinsic)
Care Plans
(prescribed)
Clinical Graph
Personalized Cues
Paths/Pathways
(Behavior Graph)

FIG. 6

700



Batch scoring

uuid uuid	cue_id integer	participant_id text	date date	time time with time zone	content text
37ea2b26-16c5-4128-...	308	04LRVJ	2021-07-01	08:00:00+08:00	In the last 7 days, you clocked a total of 128 exercise minutes and 5,450 steps. Well done!
0976e495-a0e4-41a8-...	296	0CEPQK	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
da2607e6-ce70-4e4f-a...	296	12HNNJ	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
eb8b1a7b-fb7f-4381-a...	275	1C9UNH	2021-07-01	08:00:00+08:00	A high BMI is a sure way to become diabetic. Your BMI is currently 16.8 and very close to be...
d2f49e36-f138-44e0-8...	296	1C9UNH	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
8022815a-1e02-4bda-...	334	25JTP1	2021-07-01	08:00:00+08:00	In the past week, you had an average of 6.8 hours of sleep and you made 0 healthy purchase...
b724aac5-a529-4407-...	308	25JTP1	2021-07-01	08:00:00+08:00	In the last 7 days, you clocked a total of 220 exercise minutes and 4,675 steps. Well done!
9c9086e1-e412-4ab1-...	275	2E466W	2021-07-01	08:00:00+08:00	A high BMI is a sure way to become diabetic. Your BMI is currently 24.0 and very close to be...
0324bd02-89c2-4455-...	296	2E466W	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
1fcb2c0e-dc23-4399-8...	308	2NWEWT	2021-07-01	08:00:00+08:00	In the last 7 days, you clocked a total of 257 exercise minutes and 1,245 steps. Well done!
460c2897-4544-4fea-9...	275	2NWEWT	2021-07-01	08:00:00+08:00	A high BMI is a sure way to become diabetic. Your BMI is currently 23.5 and very close to be...
13a13f11-b69d-49e8-...	296	2ONGTQ	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
f8a1b8be-4749-4cac-8...	296	2TNASC	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
494fdcbf-5409-48c7-8...	334	30DT7N	2021-07-01	08:00:00+08:00	In the past week, you had an average of 5.0 hours of sleep and you made 0 healthy purchase...
a3719ae1-7c97-4c80-...	296	30DT7N	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
7cb08084-d162-4da4-...	296	3QYMDH	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
68d33346-a4f9-4d6d-...	275	3XGMG6	2021-07-01	08:00:00+08:00	A high BMI is a sure way to become diabetic. Your BMI is currently 24.3 and very close to be...
e793ebd3-f831-4b57-...	296	3XGMG6	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
47f03785-e336-47b6-...	334	40USAM	2021-07-01	08:00:00+08:00	In the past week, you had an average of 7.0 hours of sleep and you made 2 healthy purchase...
6f796870-eeda-413f-a...	296	40USAM	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
9b0ac399-a719-4620-...	347	4B4QY8	2021-07-01	08:00:00+08:00	Got 1% of your day to spare? 15mins of heart pumping activity is all you need to reap the he...
6c196716-eae0-417b-...	308	4RJ7ZE	2021-07-01	08:00:00+08:00	In the last 7 days, you clocked a total of 184 exercise minutes and 11,683 steps. Well done!
92057c6b-957b-4be2-...	296	4W9XUX	2021-07-01	08:00:00+08:00	Eat Good, Feel Good! Get healthier by opting for food products and beverages that carry the ...
e7143de9-f5fd-4f5c-b...	275	53XB3F	2021-07-01	08:00:00+08:00	A high BMI is a sure way to become diabetic. Your BMI is currently 15.3 and very close to be...

FIG. 7

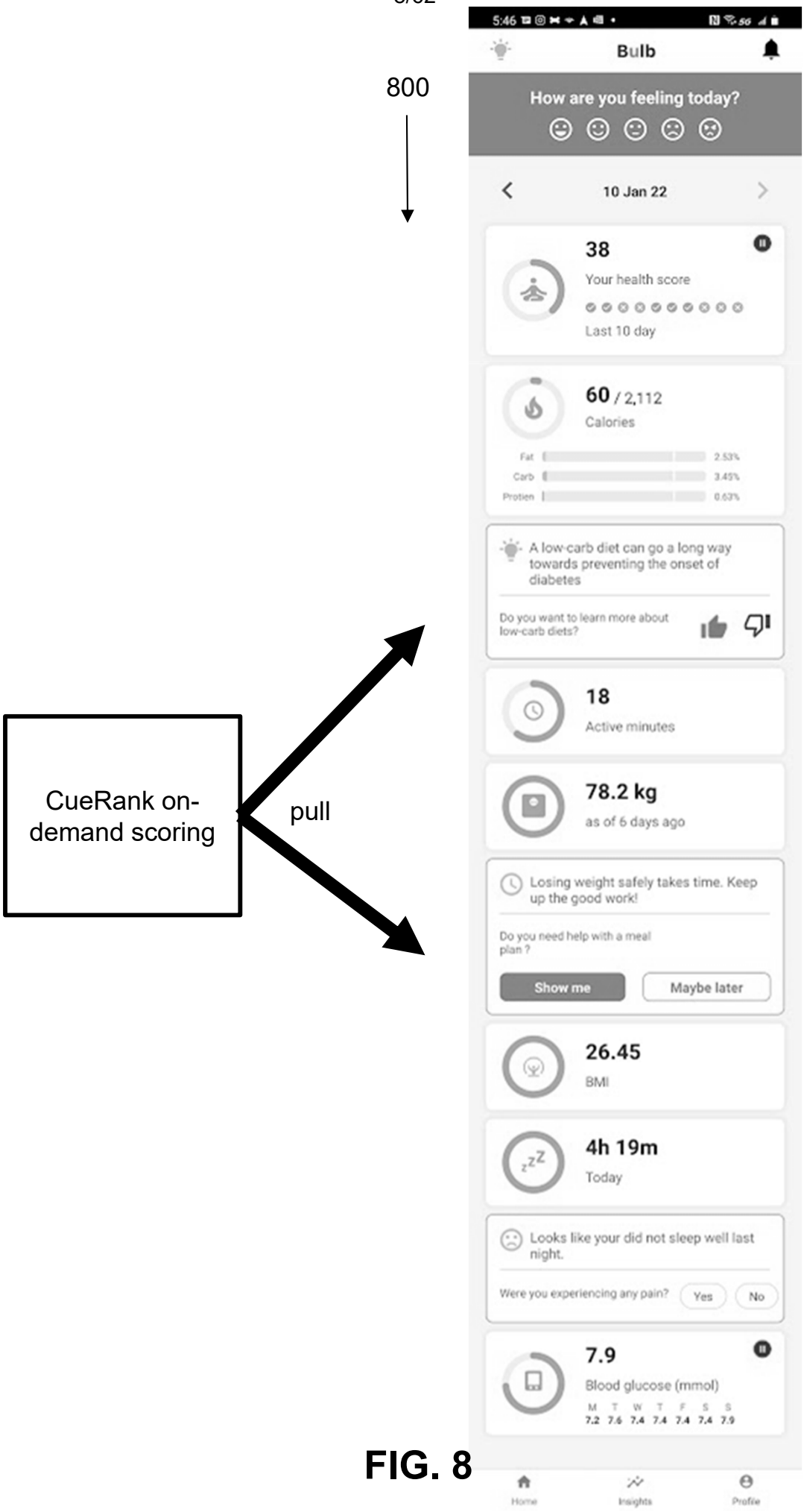

800
CueRank on-demand scoring
pull
Bulb
How are you feeling today?
10 Jan 22
38
Your health score
Last 10 day
60 / 2,112
Calories
Fat
Carb
Protien
A low-carb diet can go a long way towards preventing the onset of diabetes
Do you want to learn more about low-carb diets?
18
Active minutes
78.2 kg
as of 6 days ago
Losing weight safely takes time. Keep up the good work!
Do you need help with a meal plan ?
Show me
Maybe later
26.45
BMI
4h 19m
Today
Looks like your did not sleep well last night.
Were you experiencing any pain?
Yes
No
7.9
Blood glucose (mmol)
M T W T F S S
7.2 7.6 7.4 7.4 7.4 7.4 7.9
Home
Insights
Profile

FIG. 8

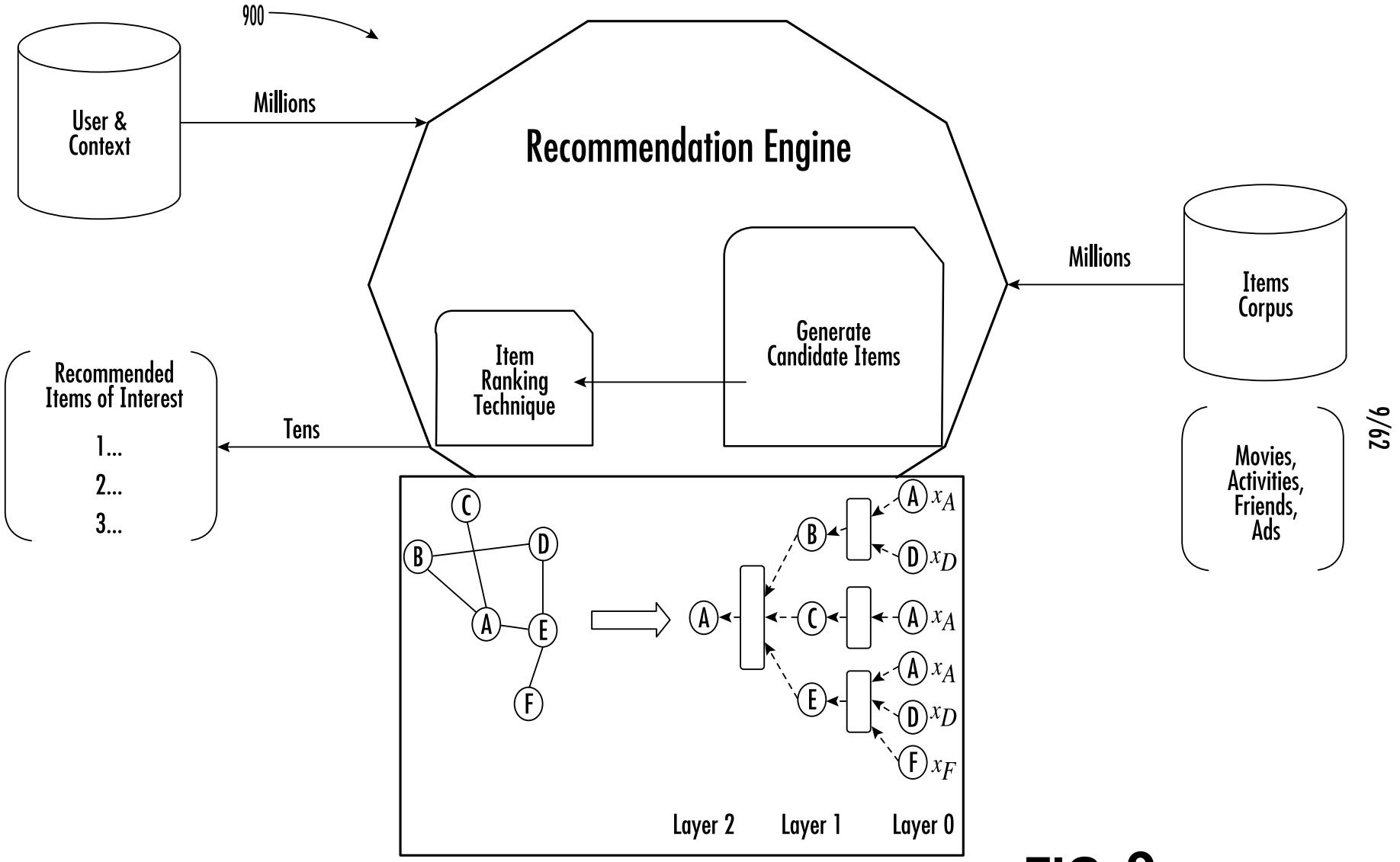

900
User &
Context
Millions
Recommendation Engine
Millions
Items
Corpus
Generate
Candidate Items
Item
Ranking
Technique
Recommended
Items of Interest
1...
2...
3...
Tens
Movies,
Activities,
Friends,
Ads
Layer 2
Layer 1
Layer 0

FIG. 9

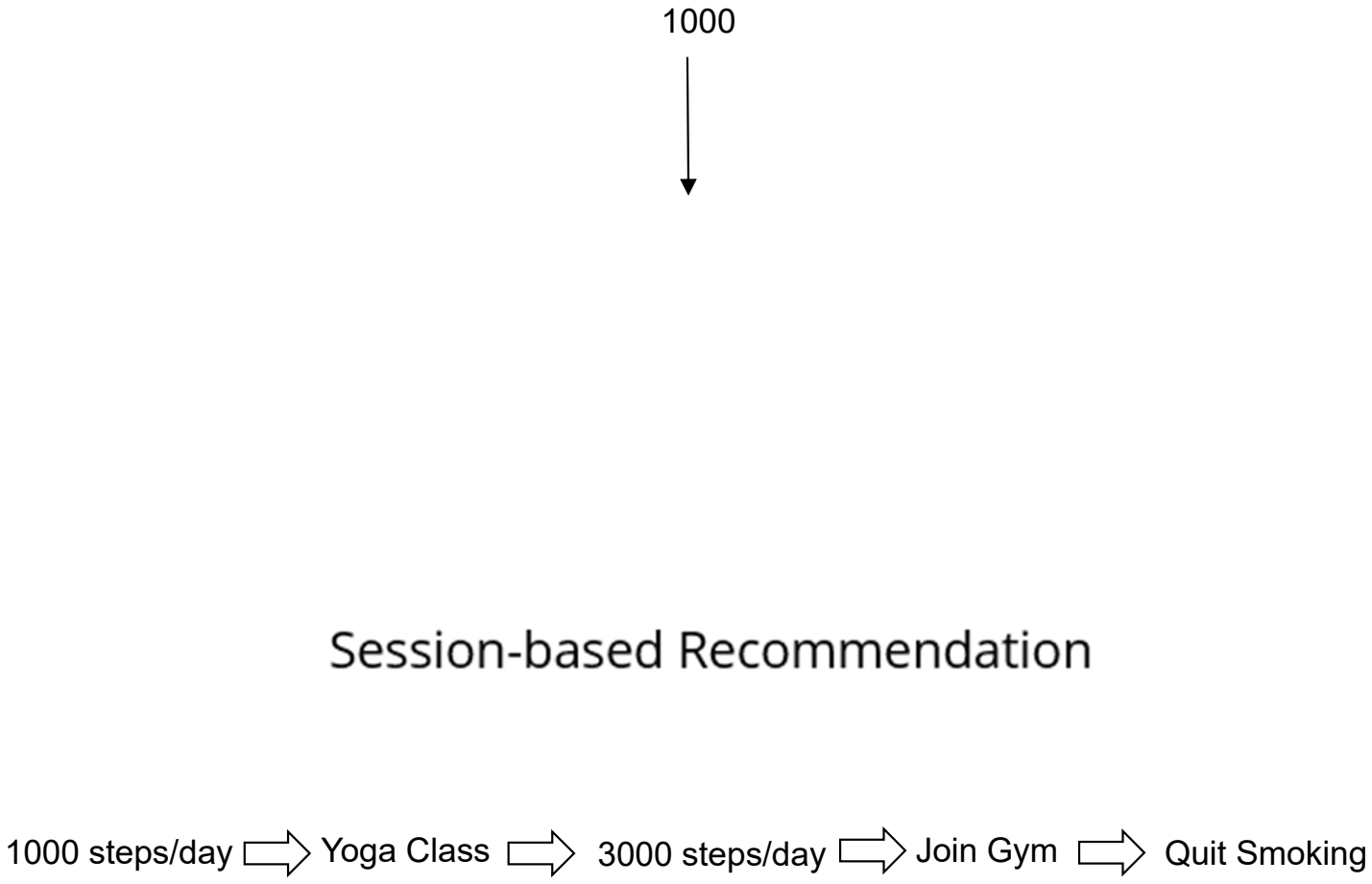

1000
Session-based Recommendation
1000 steps/day
Yoga Class
3000 steps/day
Join Gym
Quit Smoking

FIG. 10

1100

Loss Functions

Pointwise Loss (e.g., log loss)

- Treats the recommendation task as a **classification** problem
- Optimizes for the prediction scores to be close to the target scores
- Applications: Rating Prediction, CTR Prediction

e.g., segment -> Propensity P(i)

$$\mathcal{L} = \sum_{(u,i,X) \in \mathcal{O}} -y_{ui} \log \sigma(\hat{y}_{ui}) - (1 - y_{ui}) \, log(1 - \sigma(\hat{y}_{ui}))$$

Pairwise Loss (e.g., Bayesian Personalized Ranking (BPR) loss

- Treats the recommendation task as a **ranking** problem
- Optimizes for the relative order between observed and unobserved interactions
- Applications: Top-N Recommendation, Preference Ranking

e.g., Marker -> Next Best Activity $A_1, A_2, \ldots, A_k$

$$\mathcal{L} = \sum_{(u,i,j) \in \mathcal{O}} -\ln \sigma \left(\hat{y}_{ui} - \hat{y}_{uj} \right)$$

FIG. 11

1200



Candidate Item Generation: General Idea

Step 1: Generate Data Instance
Take each participant (user) and their context (e.g., markers, segments) and generate a user-item pair (u, i) or if side-information (c) is available, then a trio (u, i, c). Transform each pair or trio into a separate data instance.
Step 2: Feature Representation
Represent each data instance using hand crafted features or features learnt via representation learning (e.g., deep Learning).
Step 3: Prediction on Data Instance
Perform predictions or rankings based on user-item-knowledge interaction modeling using matrices or graphs.

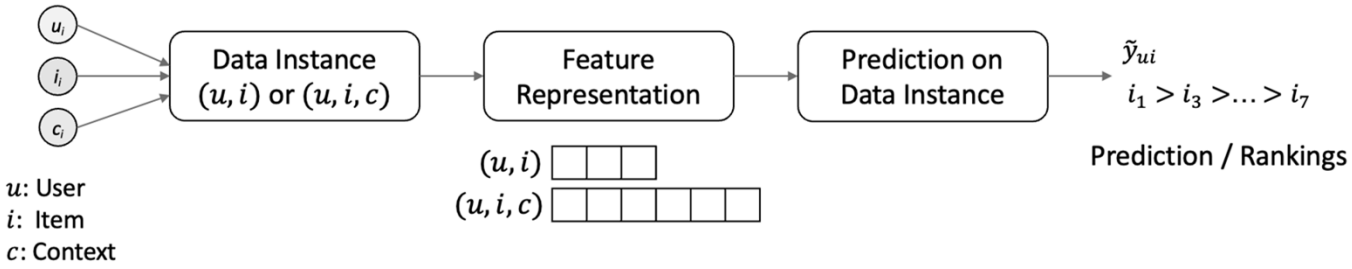


FIG. 12

1300

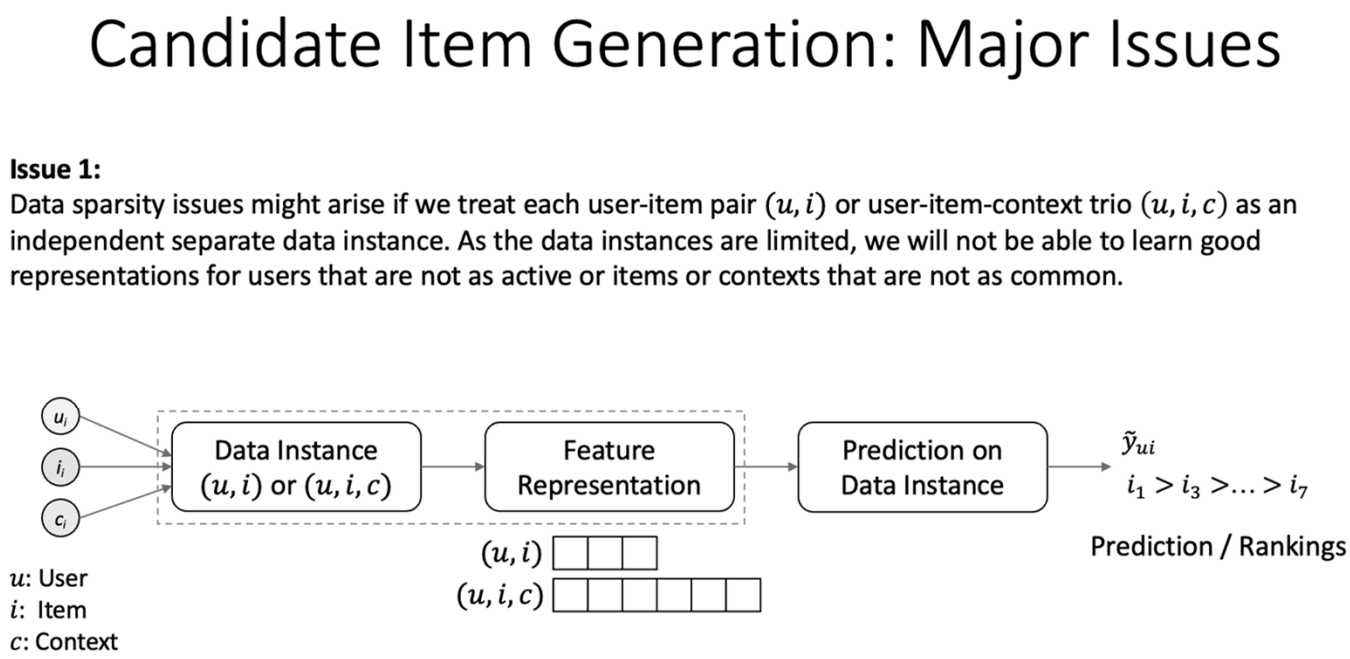

Candidate Item Generation: Major Issues
Issue 1:
Data sparsity issues might arise if we treat each user-item pair (u, i) or user-item-context trio (u, i, c) as an independent separate data instance. As the data instances are limited, we will not be able to learn good representations for users that are not as active or items or contexts that are not as common.
u_i
i_i
c_i
Data Instance (u, i) or (u, i, c)
Feature Representation
Prediction on Data Instance
$\tilde{y}_{ui}$
$i_1 > i_3 > \ldots > i_7$
Prediction / Rankings
(u, i)
(u, i, c)
u: User
i: Item
c: Context

FIG. 13

1400



Candidate Item Generation: Major Issues

Issue 2:
Suboptimal representations of unseen or infrequent user-item / user-item-context interactions would result in poor prediction performance and hence bad interaction models.

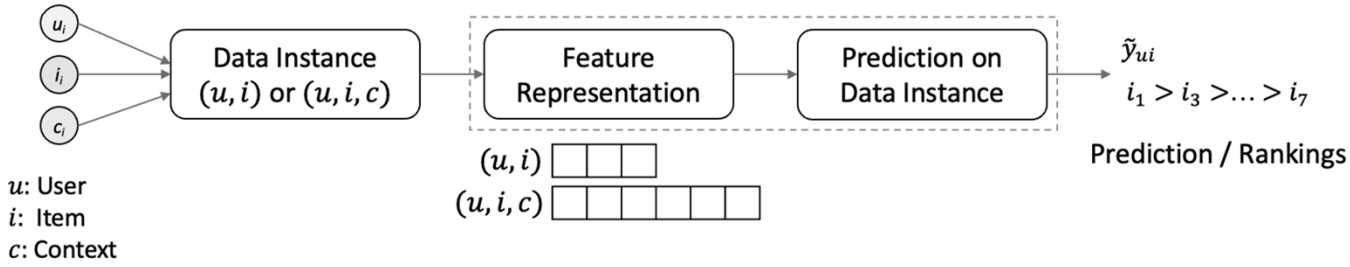


FIG. 14

1500

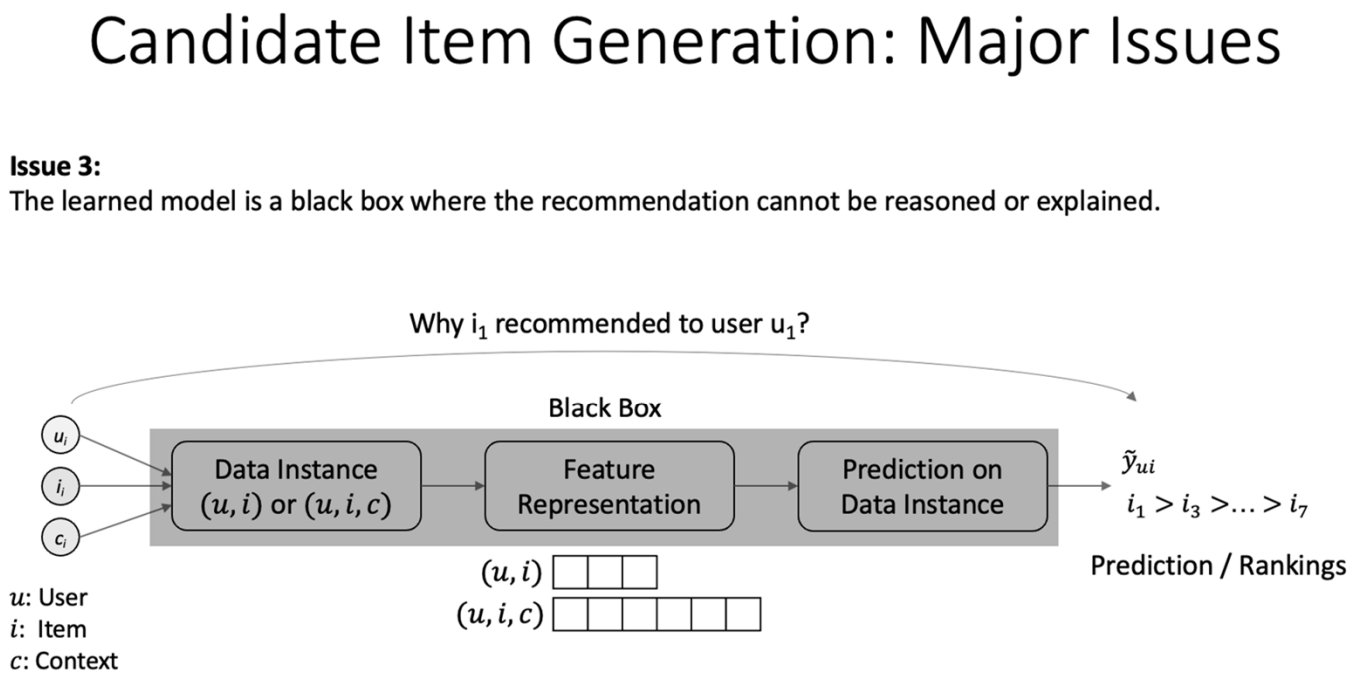
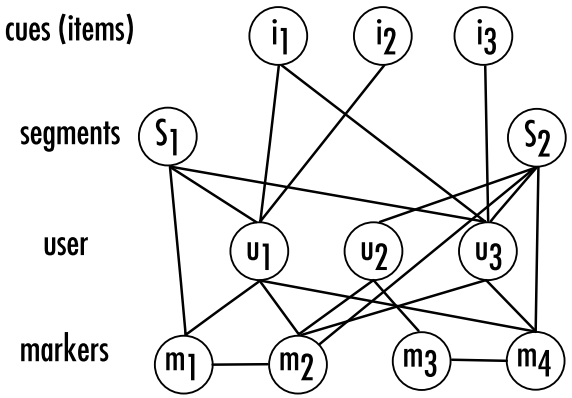

Candidate Item Generation: Major Issues
Issue 3:
The learned model is a black box where the recommendation cannot be reasoned or explained.
Why i1 recommended to user u1?
Black Box
Data Instance
(u, i) or (u, i, c)
Feature
Representation
Prediction on
Data Instance
ỹui
i1 > i3 >... > i7
Prediction / Rankings
(u, i)
(u, i, c)
u: User
i: Item
c: Context

FIG. 15

1600

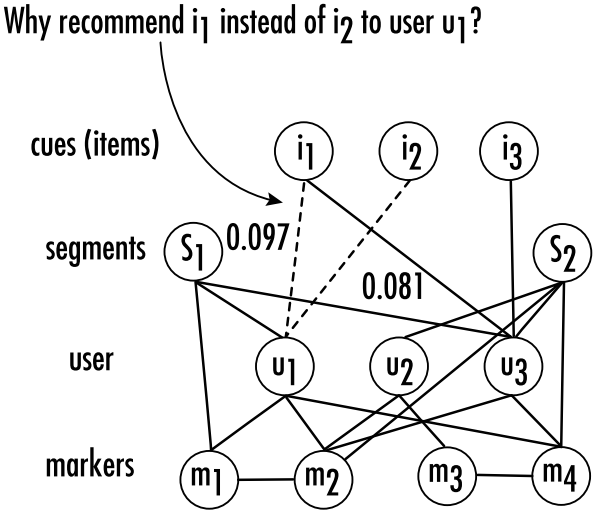
Candidate Item Generation: Addressing the Issues

Representing the users, items, and their contexts within a graph enables exploration and exploitation of high-order relations betwen them.



Addresses the data sparsity issue and may lead to improved model performance.

Inspection of learned model weights may offer explainations on why a recommendation was made.

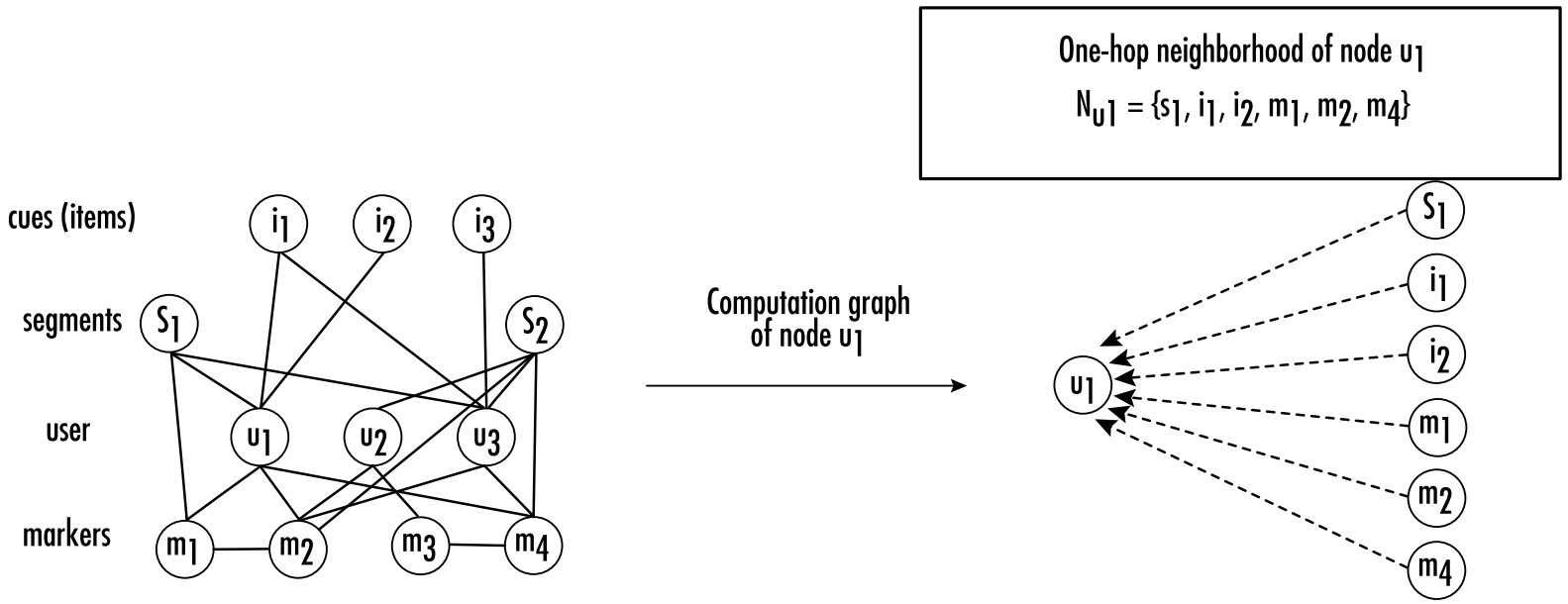


Addresses the black box issue.

FIG. 16

1700

Implementation of RS as a Graph w Deep Learning

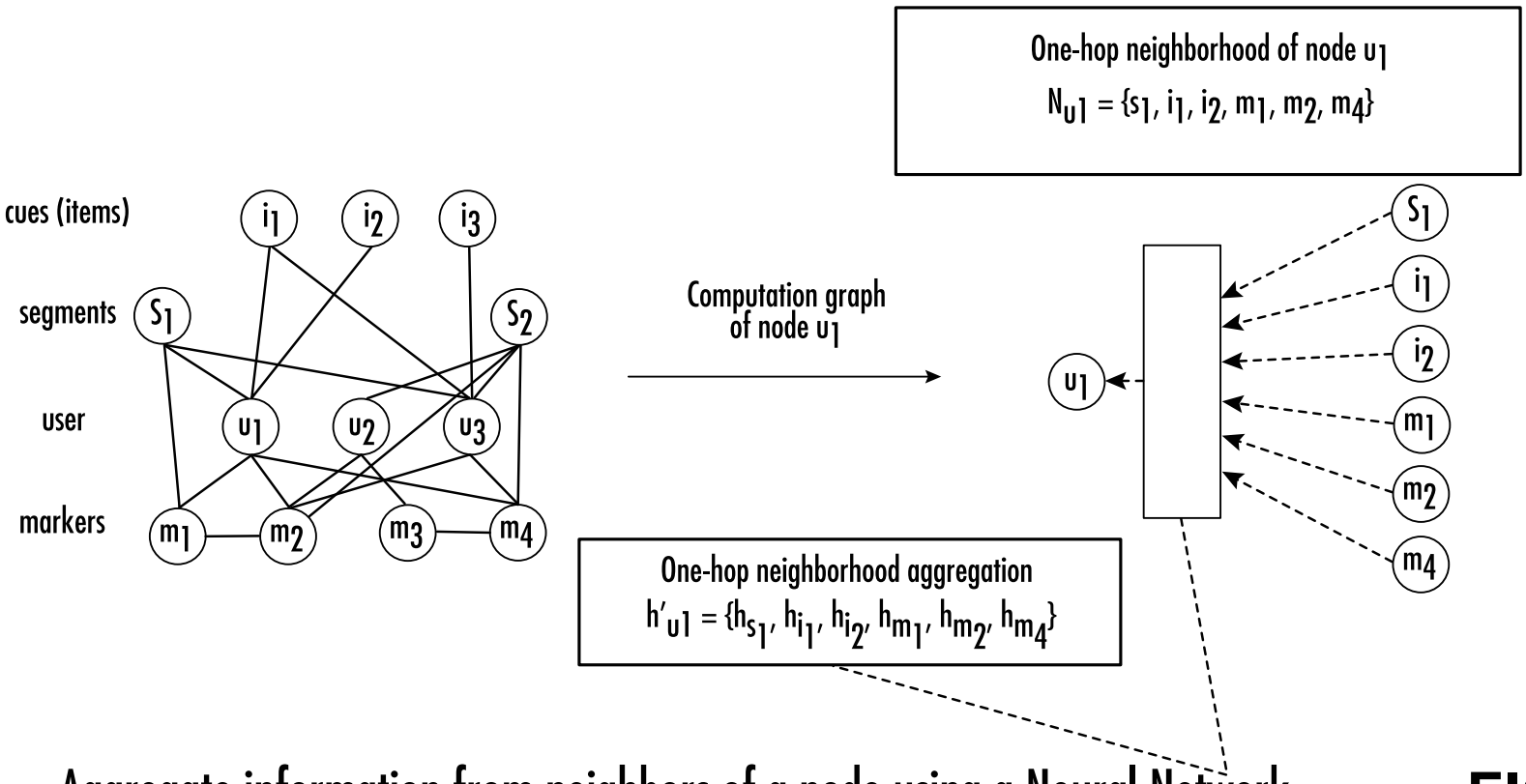


Use local neighbors to "learn" the representation of a node.

FIG. 17

1800

Implementation of RS as a Graph w Deep Learning

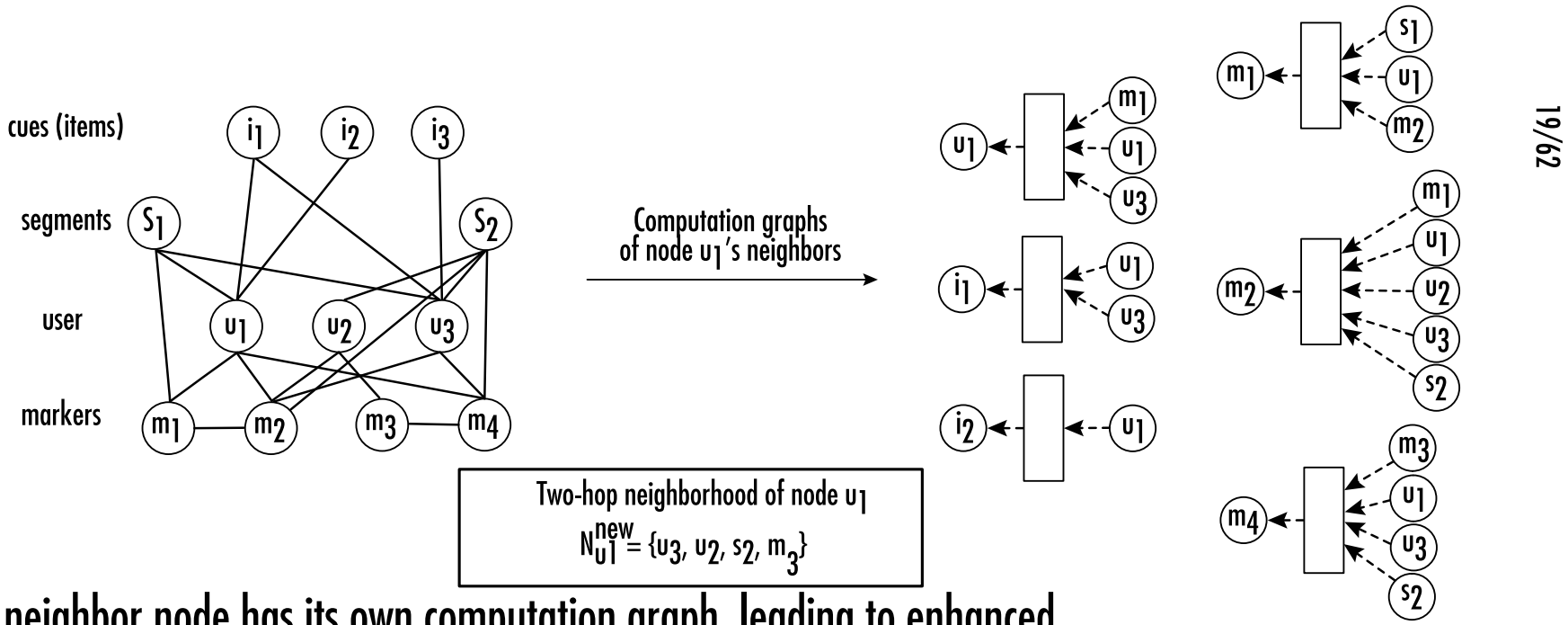


Aggregate information from neighbors of a node using a Neural Network.

FIG. 18

1900

Implementation of RS as a Graph w Deep Learning

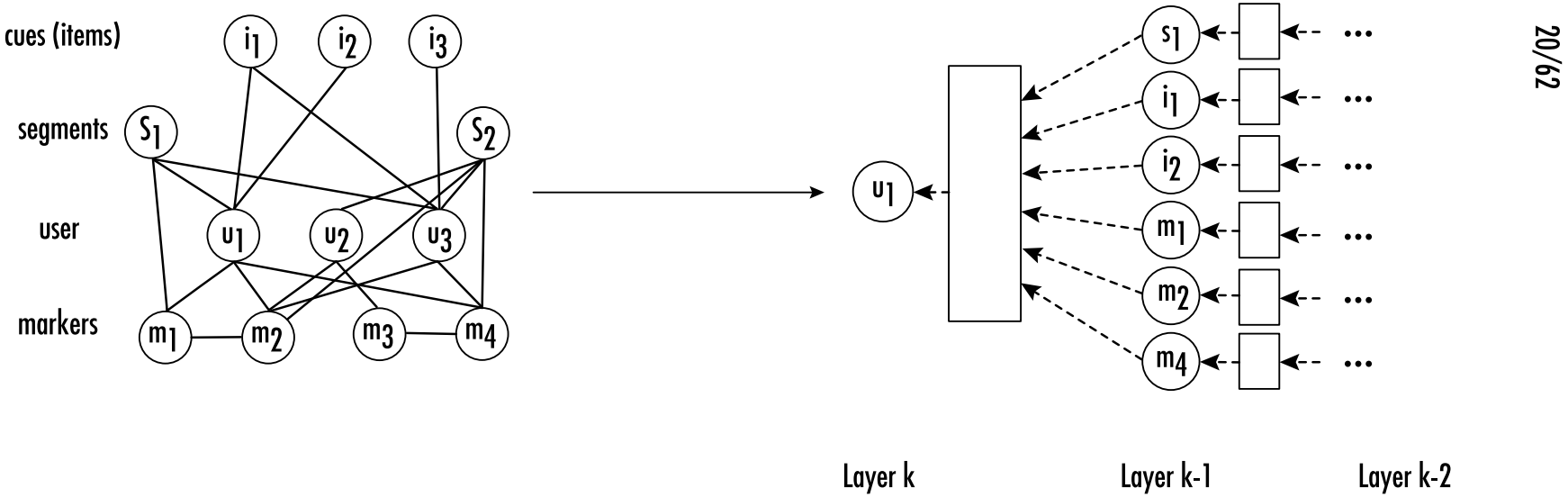


Each neighbor node has its own computation graph, leading to enhanced information propagation and feature embeddings.

FIG. 19

2000

Implementation of RS as a Graph w Deep Learning



Each neighborhood hop level (layer) is stackable up to arbitrary depth, leading to rich information embedding for each node and its neighborhood.

FIG. 20



Session-based Recommendation

- aims to **predict which item a user will click next**
- solely based on the user's current sequential session data without accessing to the long-term preference profile.
- user identification is unknown and only the user behavior history during an ongoing session is available

FIG. 21

2200



Challenges

Limitations of previous work:

1. Difficulty in estimating user representations without adequate user behavior in one session
2. Only modeling single-way transitions between consecutive items and neglect the transitions among contexts
3. Assumes each user only accesses one item at each timestep

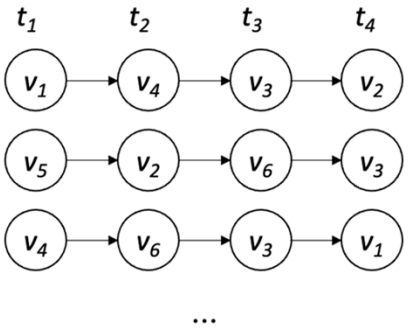
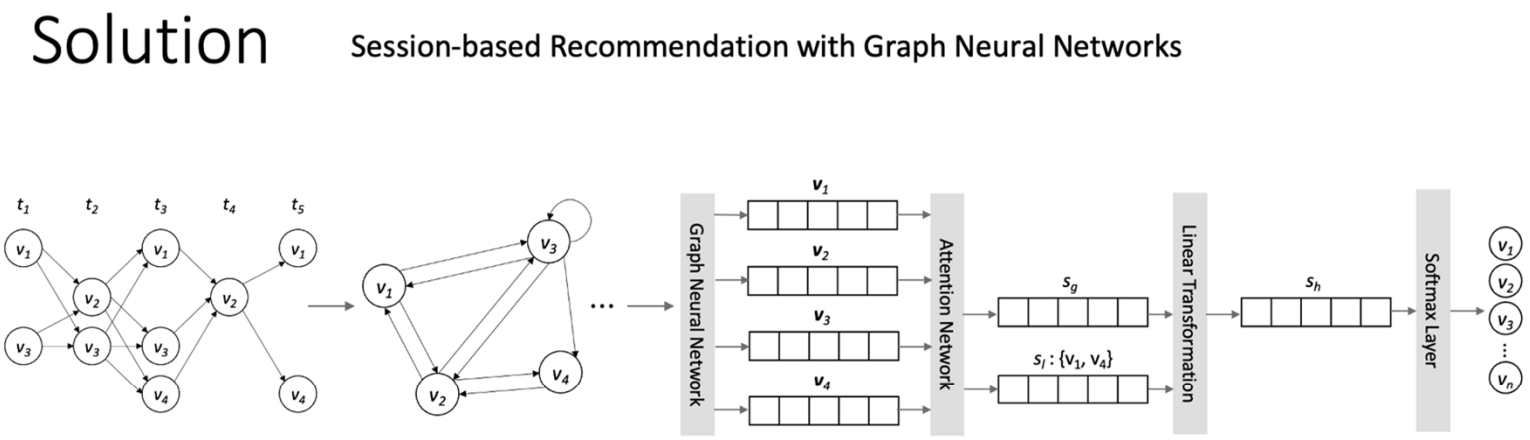


FIG. 22

2300



1. Model all session sequences as session graphs
2. Latent vectors for all nodes can be obtain through graph neural network
3. Each session is represented as the combination of the global preference and local interests of this session, which is composed by latent vectors of nodes
4. Predict the probability that each item will appear in the next itemset of the session

FIG. 23

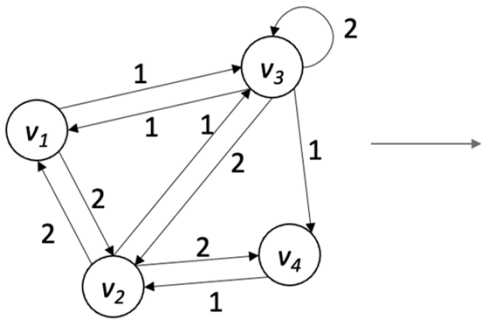
2400

1. Constructing Session Graphs

Each session sequence modeled as a directed graph

- Node represents item
- Edge represents successive clicks

Assign each edge with a normalized weight, calculated as the occurrence of the edge divided by the outdegree of start node



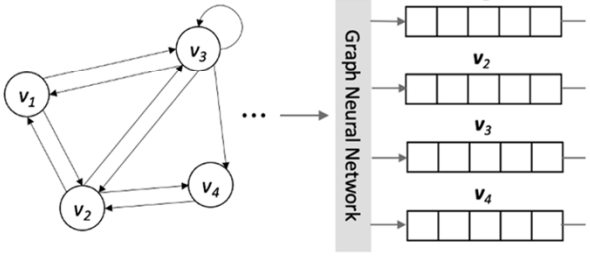
	v_1	v_2	v_3	v_4	v_1	v_2	v_3	v_4
v_1	0	2/3	1/3	0	0	2/3	1/3	0
v_2	2/5	0	1/5	2/5	2/5	0	2/5	1/5
v_3	1/6	2/6	2/6	1/6	1/4	1/4	2/4	0
v_4	0	1	0	0	0	2/3	1/3	0

FIG. 24

2500

2. Learning Item Embeddings on Session Graphs

Obtain latent vectors for each node through **gated graph neural network**



$$\mathbf{a}_{s,i}^{t} = \mathbf{A}_{s,i:}\left[\mathbf{v}_1^{t-1}, \ldots, \mathbf{v}_n^{t-1}\right]^{\top}\mathbf{H} + \mathbf{b}, \quad (1)$$

$$\mathbf{z}_{s,i}^{t} = \sigma\left(\mathbf{W}_z\mathbf{a}_{s,i}^{t} + \mathbf{U}_z\mathbf{v}_i^{t-1}\right), \quad (2)$$

$$\mathbf{r}_{s,i}^{t} = \sigma\left(\mathbf{W}_r\mathbf{a}_{s,i}^{t} + \mathbf{U}_r\mathbf{v}_i^{t-1}\right), \quad (3)$$

$$\widetilde{\mathbf{v}}_i^{t} = \tanh\left(\mathbf{W}_o\mathbf{a}_{s,i}^{t} + \mathbf{U}_o\left(\mathbf{r}_{s,i}^{t} \odot \mathbf{v}_i^{t-1}\right)\right), \quad (4)$$

$$\mathbf{v}_i^{t} = \left(1 - \mathbf{z}_{s,i}^{t}\right) \odot \mathbf{v}_i^{t-1} + \mathbf{z}_{s,i}^{t} \odot \widetilde{\mathbf{v}}_i^{t}, \quad (5)$$

$$\mathbf{i}_{s,j}^{(l+1)} = \mathbf{GRU}(\mathbf{i}_{s,j}^{(l)}, \mathbf{a}_{s,j}^{(l+1)}) \quad (6)$$

- vectors of neighborhoods propagates to current node
- candidate state constructed by the previous state and the current state. Reset gate as described in Eq. (4).
- The final state is combination of the previous hidden state and the candidate state, controlled by update gate.

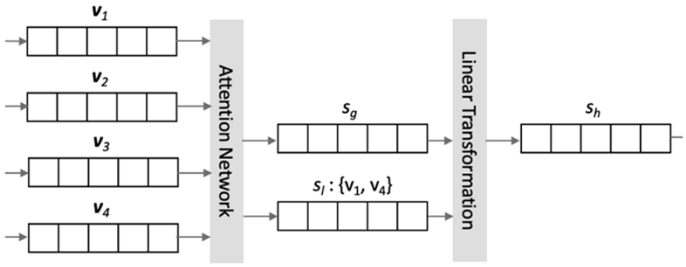
FIG. 25

2600

3.Generating Session Embeddings

Session embedding: combine **global preference** and **current interests** of the session

- **local embedding** s_l is defined as the average of the node vectors of the last-clicked itemset
- **global embedding** s_g is obtained by aggregating all node vectors within each itemset using soft-attention mechanism and averaging across all itemsets.
- **hybrid embedding** s_h is obtained by taking linear transformation over the concatenation of the local and global embedding vectors



$$\alpha_i = \mathbf{q}^\top \sigma(\mathbf{W}_1 \mathbf{v}_n + \mathbf{W}_2 \mathbf{v}_i + \mathbf{c}),$$

$$\boldsymbol{s}_g = \frac{1}{m} \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \alpha_i v_i \right) \tag{6}$$

$$\mathbf{s}_\mathrm{h} = \mathbf{W}_3 \left[\mathbf{s}_\mathrm{l}; \mathbf{s}_\mathrm{g} \right], \tag{7}$$

FIG. 26

2700

4. Making Recommendation and Model Training

- compute the score z_i for each candidate item v_i by multiplying its embedding v_i by session representation s_h

$$\hat{\mathbf{z}}_i = \mathbf{s}_{\mathrm{h}}^{\top} \mathbf{v}_i. \qquad (8)$$

$$\hat{\mathbf{y}} = \mathrm{softmax}\left(\hat{\mathbf{z}}\right), \qquad (9)$$

- For each session graph, the loss function is defined as the cross-entropy of the prediction and the ground truth

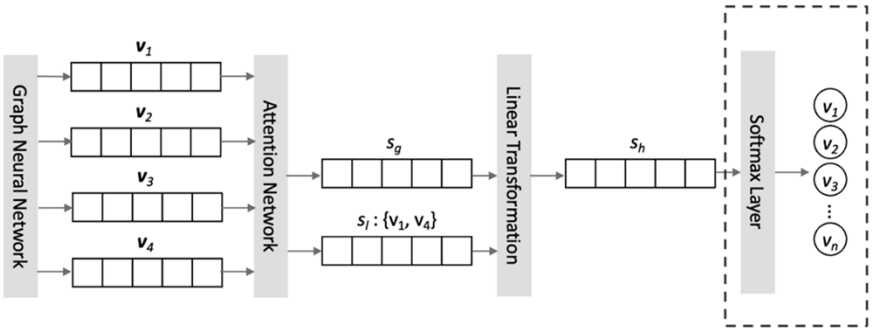


FIG. 27

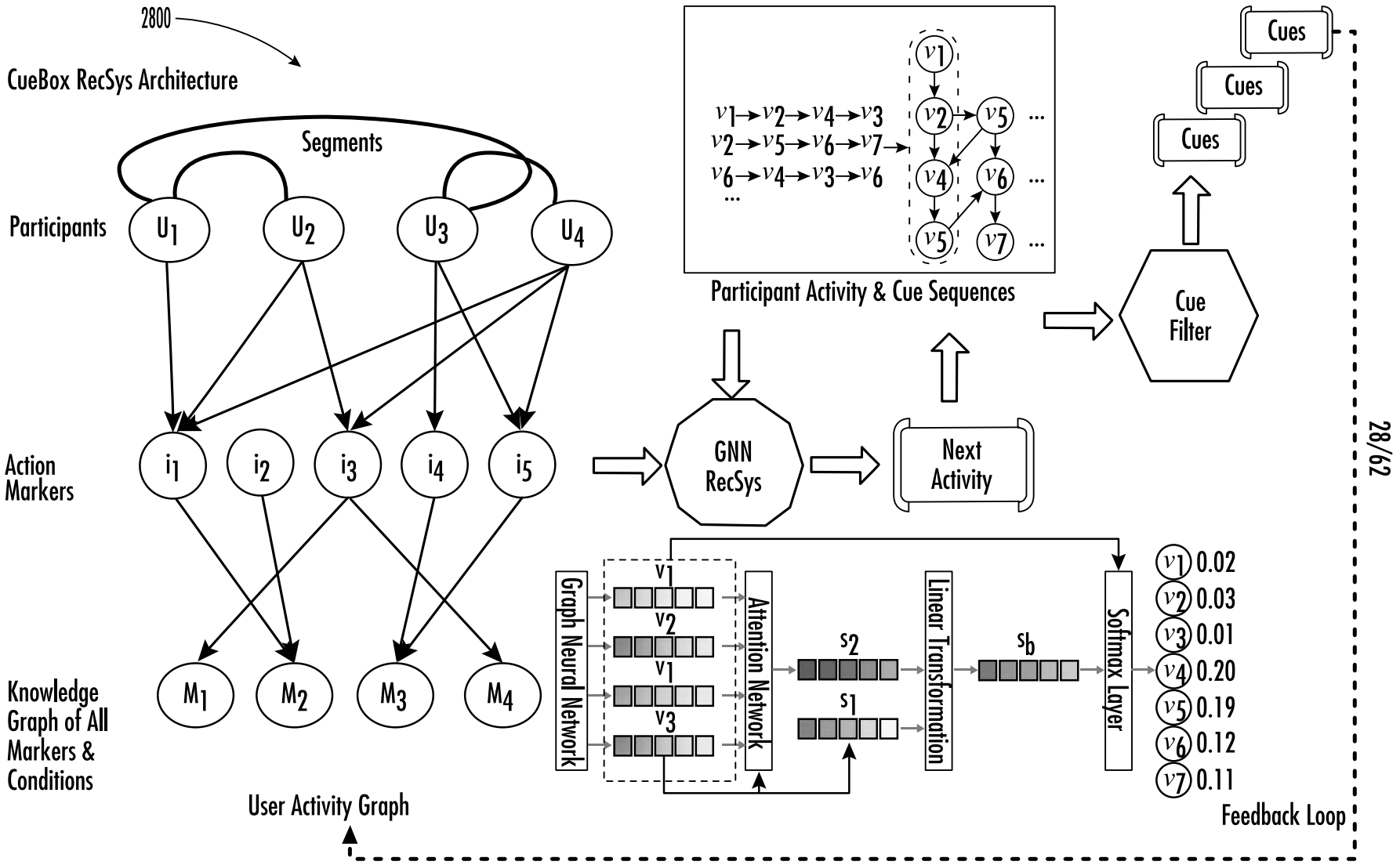

2800
CueBox RecSys Architecture
Segments
Participants
U1
U2
U3
U4
Action Markers
i1
i2
i3
i4
i5
Knowledge Graph of All Markers & Conditions
M1
M2
M3
M4
User Activity Graph
v1→v2→v4→v3
v2→v5→v6→v7→
v6→v4→v3→v6
Participant Activity & Cue Sequences
Cues
Cues
Cues
Cue Filter
GNN RecSys
Next Activity
Graph Neural Network
Attention Network
Linear Transformation
Softmax Layer
s2
s1
sb
v1 0.02
v2 0.03
v3 0.01
v4 0.20
v5 0.19
v6 0.12
v7 0.11
Feedback Loop
28/62

FIG. 28

2900

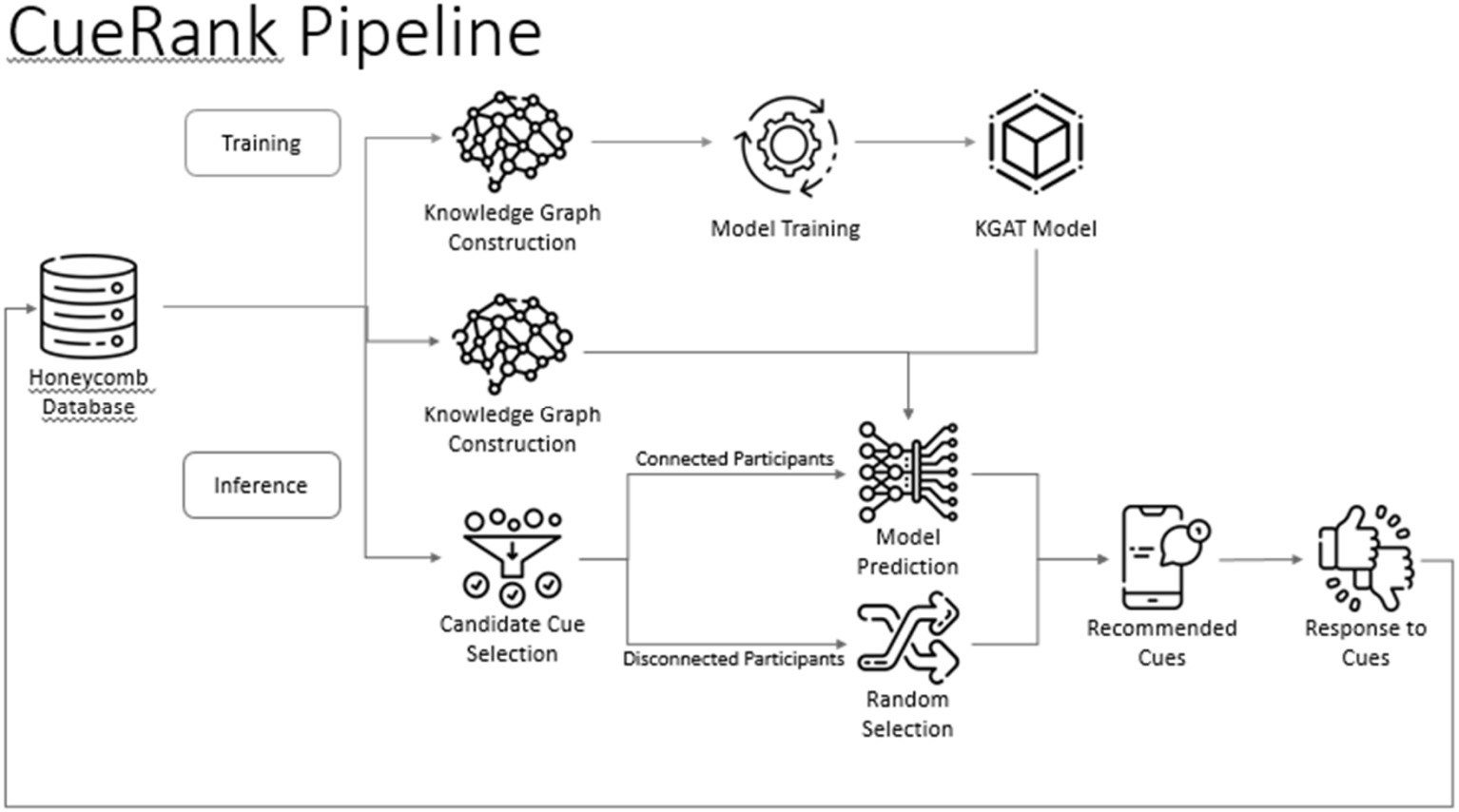

CueRank Pipeline
Training
Knowledge Graph Construction
Model Training
KGAT Model
Honeycomb Database
Knowledge Graph Construction
Inference
Connected Participants
Model Prediction
Candidate Cue Selection
Disconnected Participants
Random Selection
Recommended Cues
Response to Cues

FIG. 29

2900

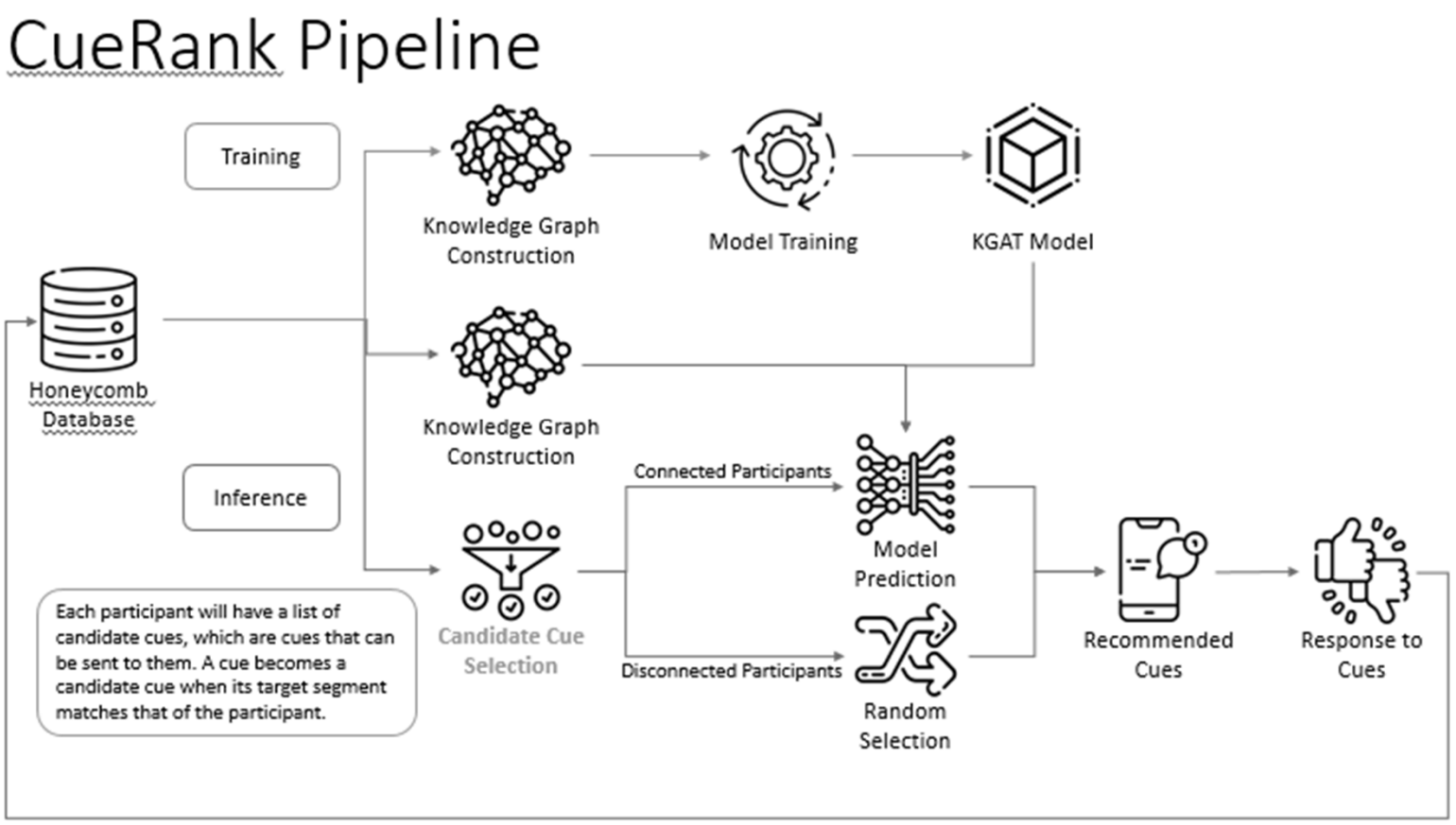

CueRank Pipeline
Training
Knowledge Graph Construction
Model Training
KGAT Model
Honeycomb Database
Knowledge Graph Construction
Inference
Connected Participants
Model Prediction
Each participant will have a list of candidate cues, which are cues that can be sent to them. A cue becomes a candidate cue when its target segment matches that of the participant.
Candidate Cue Selection
Disconnected Participants
Random Selection
Recommended Cues
Response to Cues

FIG. 30

2900

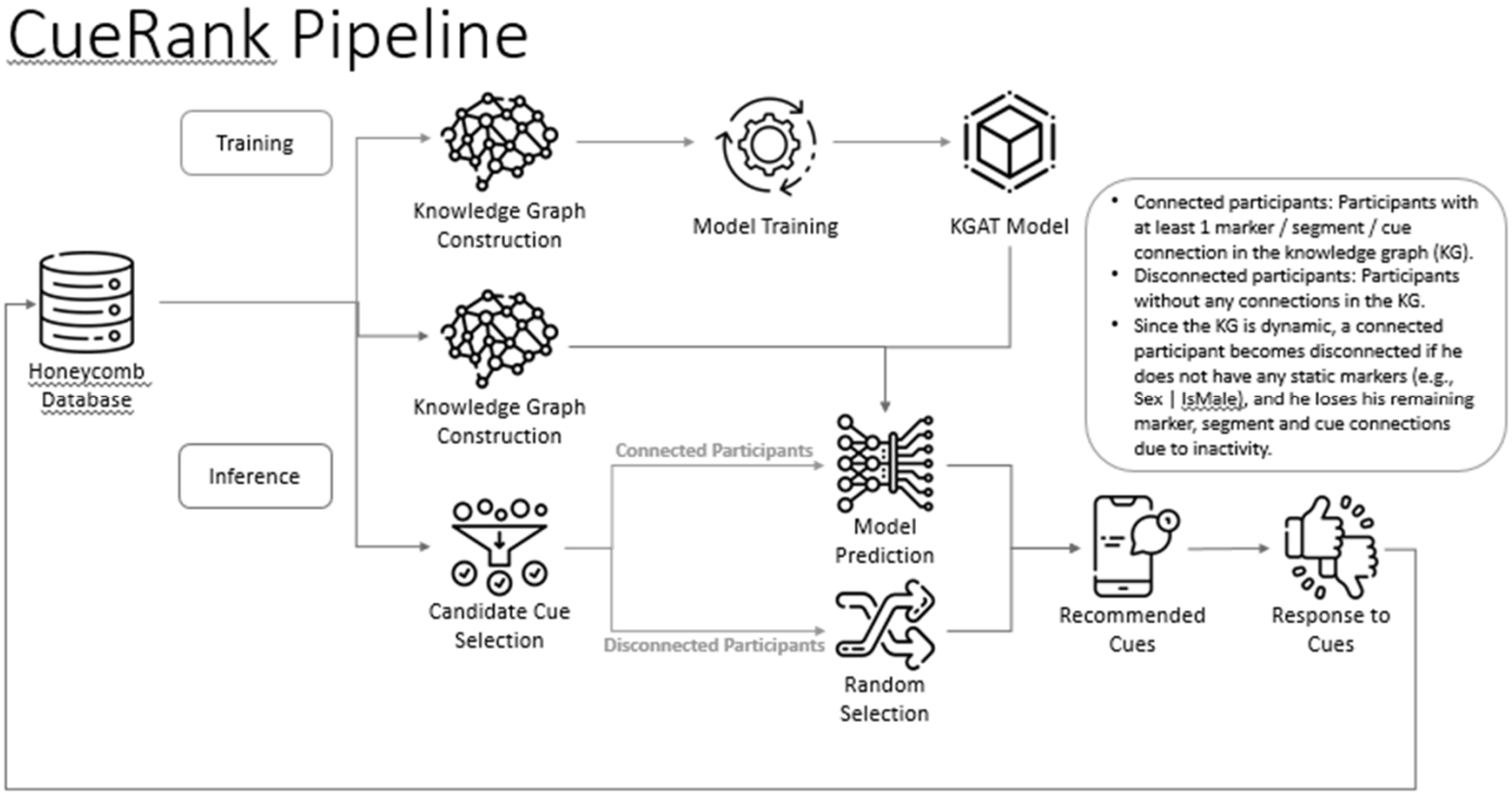

CueRank Pipeline
Training
Knowledge Graph Construction
Model Training
KGAT Model
Honeycomb Database
Knowledge Graph Construction
Inference
Connected Participants
Model Prediction
Candidate Cue Selection
Disconnected Participants
Random Selection
Recommended Cues
Response to Cues
Connected participants: Participants with at least 1 marker / segment / cue connection in the knowledge graph (KG).
Disconnected participants: Participants without any connections in the KG.
Since the KG is dynamic, a connected participant becomes disconnected if he does not have any static markers (e.g., Sex | IsMale), and he loses his remaining marker, segment and cue connections due to inactivity.

FIG. 31

2900

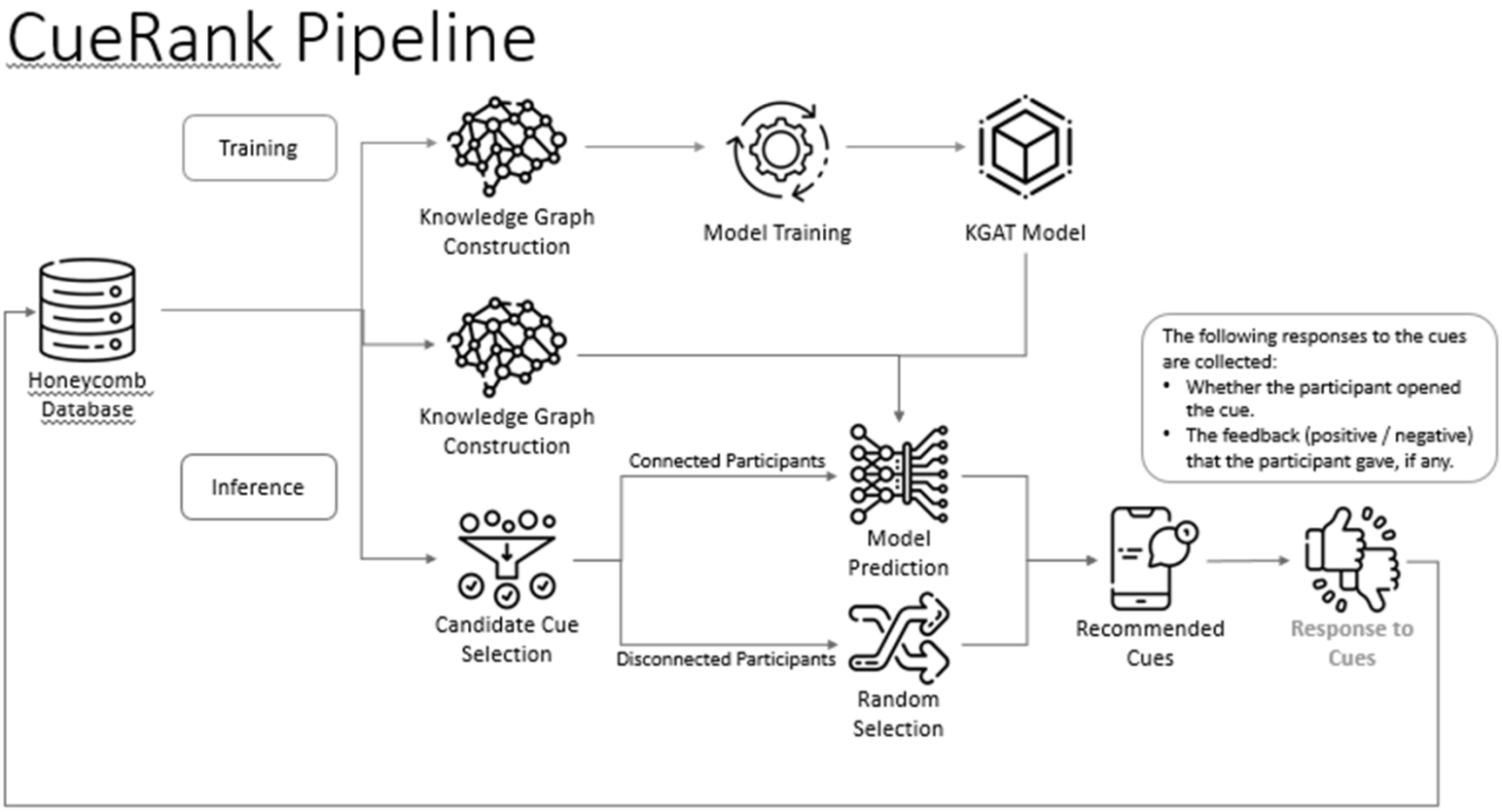
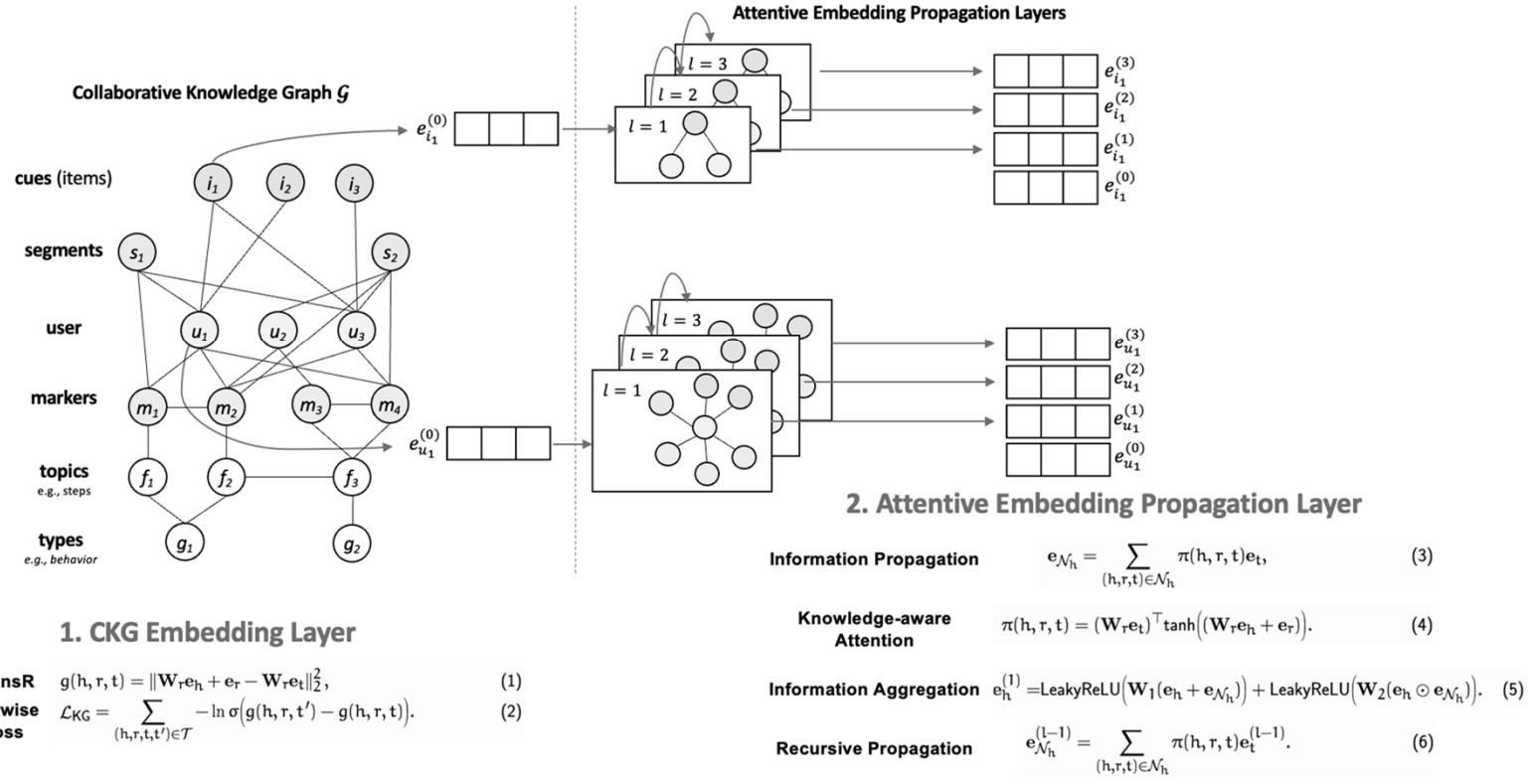

CueRank Pipeline
Training
Knowledge Graph Construction
Model Training
KGAT Model
Honeycomb Database
Knowledge Graph Construction
Inference
Connected Participants
Model Prediction
Candidate Cue Selection
Disconnected Participants
Random Selection
Recommended Cues
Response to Cues
The following responses to the cues are collected:
• Whether the participant opened the cue.
• The feedback (positive / negative) that the participant gave, if any.

FIG. 32

3300



$$g(h, r, t) = \|W_r e_h + e_r - W_r e_t\|_2^2, \quad (1)$$

$$\mathcal{L}_{KG} = \sum_{(h,r,t,t') \in \mathcal{T}} -\ln \sigma\big(g(h, r, t') - g(h, r, t)\big). \quad (2)$$

$$e_{\mathcal{N}_h} = \sum_{(h,r,t) \in \mathcal{N}_h} \pi(h, r, t) e_t, \quad (3)$$

$$\pi(h, r, t) = (W_r e_t)^\top \tanh\big((W_r e_h + e_r)\big). \quad (4)$$

$$e_h^{(1)} = \text{LeakyReLU}\big(W_1(e_h + e_{\mathcal{N}_h})\big) + \text{LeakyReLU}\big(W_2(e_h \odot e_{\mathcal{N}_h})\big). \quad (5)$$

$$e_{\mathcal{N}_h}^{(l-1)} = \sum_{(h,r,t) \in \mathcal{N}_h} \pi(h, r, t) e_t^{(l-1)}. \quad (6)$$

FIG. 33

3300

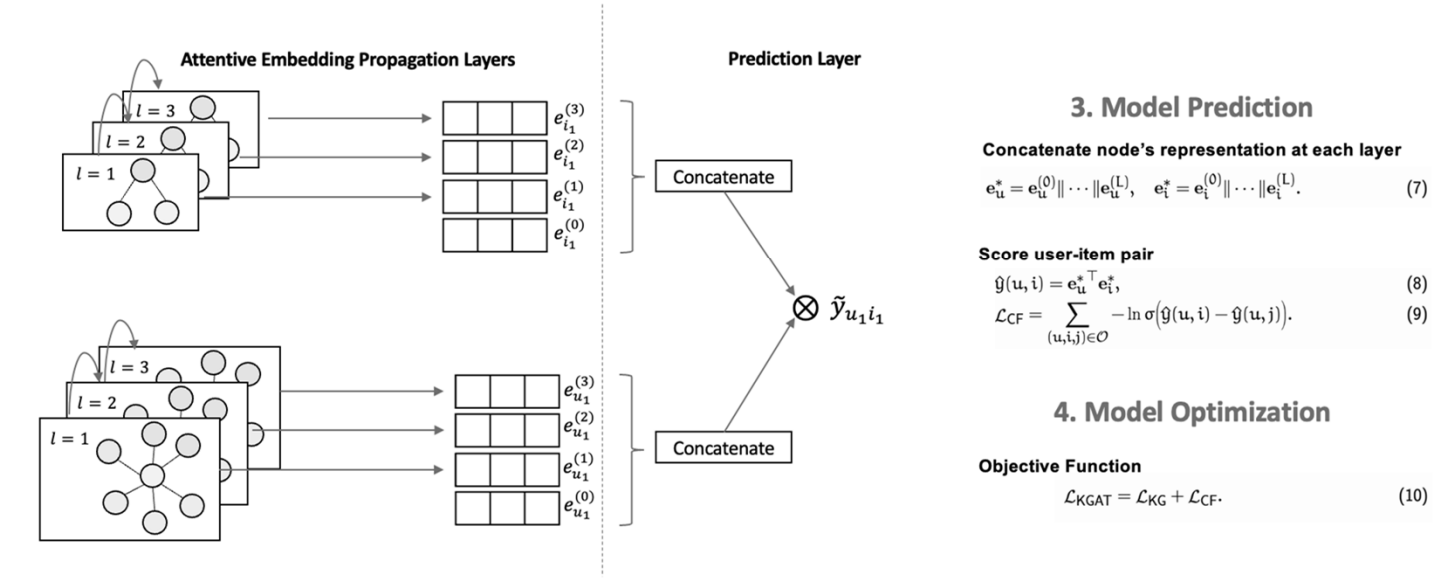

Knowledge Graph Attention Network (KGAT)
Attentive Embedding Propagation Layers
Prediction Layer
l = 3
l = 2
l = 1
Concatenate
Concatenate
3. Model Prediction
Concatenate node's representation at each layer
$\mathbf{e}_u^* = \mathbf{e}_u^{(0)} \| \cdots \| \mathbf{e}_u^{(L)}, \quad \mathbf{e}_i^* = \mathbf{e}_i^{(0)} \| \cdots \| \mathbf{e}_i^{(L)}.$ (7)
Score user-item pair
$\hat{y}(u,i) = \mathbf{e}_u^{*\top} \mathbf{e}_i^*,$ (8)
$\mathcal{L}_{CF} = \sum_{(u,i,j) \in \mathcal{O}} -\ln \sigma\big(\hat{y}(u,i) - \hat{y}(u,j)\big).$ (9)
4. Model Optimization
Objective Function
$\mathcal{L}_{KGAT} = \mathcal{L}_{KG} + \mathcal{L}_{CF}.$ (10)

FIG. 34

3500



KGAT to CueRank Knowledge Graph Attention Network (CRKGAT)

	KGAT	CRKGAT
User connections	Users are only connected to items.	Users are connected to items (cues) and other entities (markers and segments).
Cold start scenario	KGAT is not able to handle a cold start scenario and is only able to make predictions on entities that were seen during training.	CRKGAT can easily handle a cold start scenario and make predictions on unseen entities without the need to retrain the model.
Model Parameters	7 model parameters.	9 model parameters, 7 from KGAT and 2 are new parameters: *neg_slope* and *train_kg*.
Model Outputs	Returns a prediction probability on whether a user would like an item.	Returns a ranked list of items (cues) for a user in order of user preference.

FIG. 35

3600

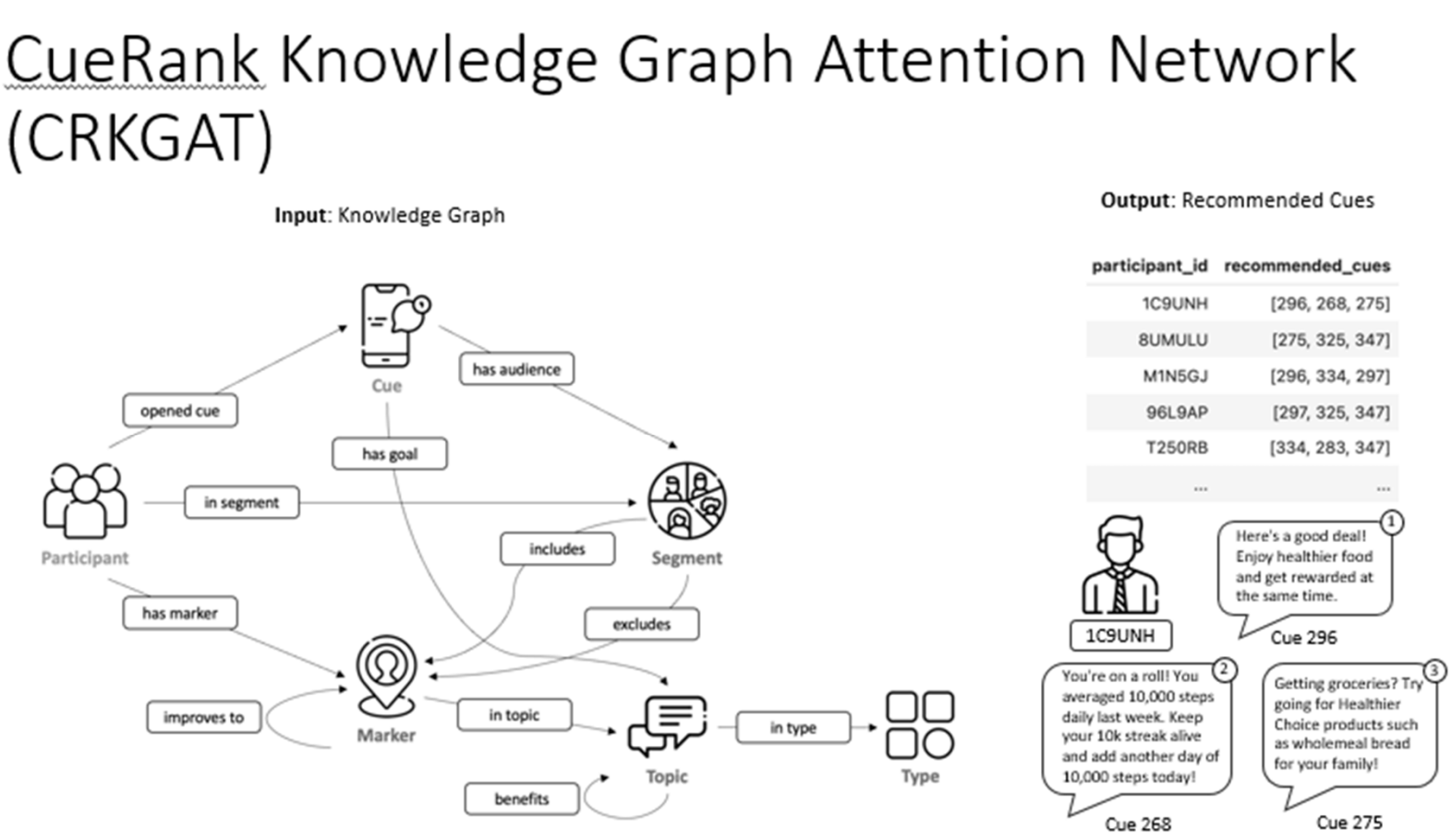

CueRank Knowledge Graph Attention Network (CRKGAT)
Input: Knowledge Graph
Cue
has audience
opened cue
has goal
Participant
in segment
Segment
includes
has marker
excludes
improves to
Marker
in topic
Topic
in type
Type
benefits
Output: Recommended Cues
participant_id
recommended_cues
1C9UNH
[296, 268, 275]
8UMULU
[275, 325, 347]
M1N5GJ
[296, 334, 297]
96L9AP
[297, 325, 347]
T250RB
[334, 283, 347]
1C9UNH
Here's a good deal! Enjoy healthier food and get rewarded at the same time.
Cue 296
You're on a roll! You averaged 10,000 steps daily last week. Keep your 10k streak alive and add another day of 10,000 steps today!
Cue 268
Getting groceries? Try going for Healthier Choice products such as wholemeal bread for your family!
Cue 275

FIG. 36

3700

CueRank Knowledge Graph Attention Network (CRKGAT)

No.	Model Parameters	Description	Example
1	entity_dim	The entity embedding size.	100
2	relation_dim	The relation embedding size.	50
3	conv_layers	The output sizes of every aggregation layer. The length of conv_layers represents the number of layers used.	[100, 50]
4	dropout	The probability of dropout in each layer.	0.1
5	agg_type	The type of aggregation to use from {gcn, graphsage, bi-interaction}.	gcn
6	neg_slope	The negative slope of the LeakyReLU.	0.4
7	LR	The learning rate.	1e-4
8	L2	The L2 regularization.	1e-5
9	train_kg	0: Train model only on CF loss, 1: Train model on both CF and KG loss.	0

FIG. 37

3800



CueRank Knowledge Graph Attention Network (CRKGAT)

Initial Experiment Results

model id	entity_dim	relation_dim	conv_layers	dropout	agg_type	neg_slope	LR	L2	train_kg	cf_loss	kg_loss	train_auc	test_auc
1	50	50	[50]	0.1	gcn	0.01	1E-04	1E-05	1	0.5606	1.13E-03	0.6675	0.6778
2	50	50	[50]	0.1	gcn	0.1	1E-04	1E-05	1	0.5411	1.15E-03	0.6692	0.6771
3	50	50	[50]	0.1	gcn	0.3	1E-04	1E-05	1	0.5237	1.14E-03	0.6786	0.6908
4	50	50	[50]	0.1	gcn	0.5	1E-04	1E-05	1	0.518	1.13E-03	0.6778	0.6903
5	50	50	[50]	0.1	gcn	0.2	1E-04	1E-05	1	0.5291	1.14E-03	0.6791	0.6904
6	50	50	[50]	0.1	gcn	0.4	1E-04	1E-05	1	0.5201	1.13E-03	0.6794	0.6919
7	100	100	[100]	0.1	gcn	0.4	1E-04	1E-05	1	0.4846	1.08E-03	0.681	0.6921
8	100	50	[100]	0.1	gcn	0.4	1E-04	1E-05	1	0.485	1.09E-03	0.6848	0.6953
9	100	50	[100, 100]	0.1	gcn	0.4	1E-04	1E-05	1	0.4114	1.11E-03	0.6893	0.7004
10	100	50	[100, 50]	0.1	gcn	0.4	1E-04	1E-05	1	0.4105	1.10E-03	0.6866	0.6983
11	100	50	[100, 100, 100]	0.1	gcn	0.4	1E-04	1E-05	1	0.3482	1.11E-03	0.6794	0.6895
12	100	50	[100, 50, 25]	0.1	gcn	0.4	1E-04	1E-05	1	0.3522	1.12E-03	0.6825	0.6923
13	100	50	[100, 50]	0.1	graphsage	0.4	1E-04	1E-05	1	0.415	1.11E-03	0.6857	0.6938
14	100	50	[100, 50]	0.1	bi-interaction	0.4	1E-04	1E-05	1	0.4072	1.11E-03	0.6834	0.6935
15	100	50	[100, 50]	0.3	gcn	0.4	1E-04	1E-05	1	0.4618	1.11E-03	0.6766	0.6893
16	100	50	[100, 50]	0.1	gcn	0.4	5E-04	1E-05	1	0.08268	9.96E-04	0.6607	0.6631
17	100	50	[100, 50]	0.1	gcn	0.4	1E-04	1E-05	0	0.4206	-	0.6995	0.7108

FIG. 38

3900

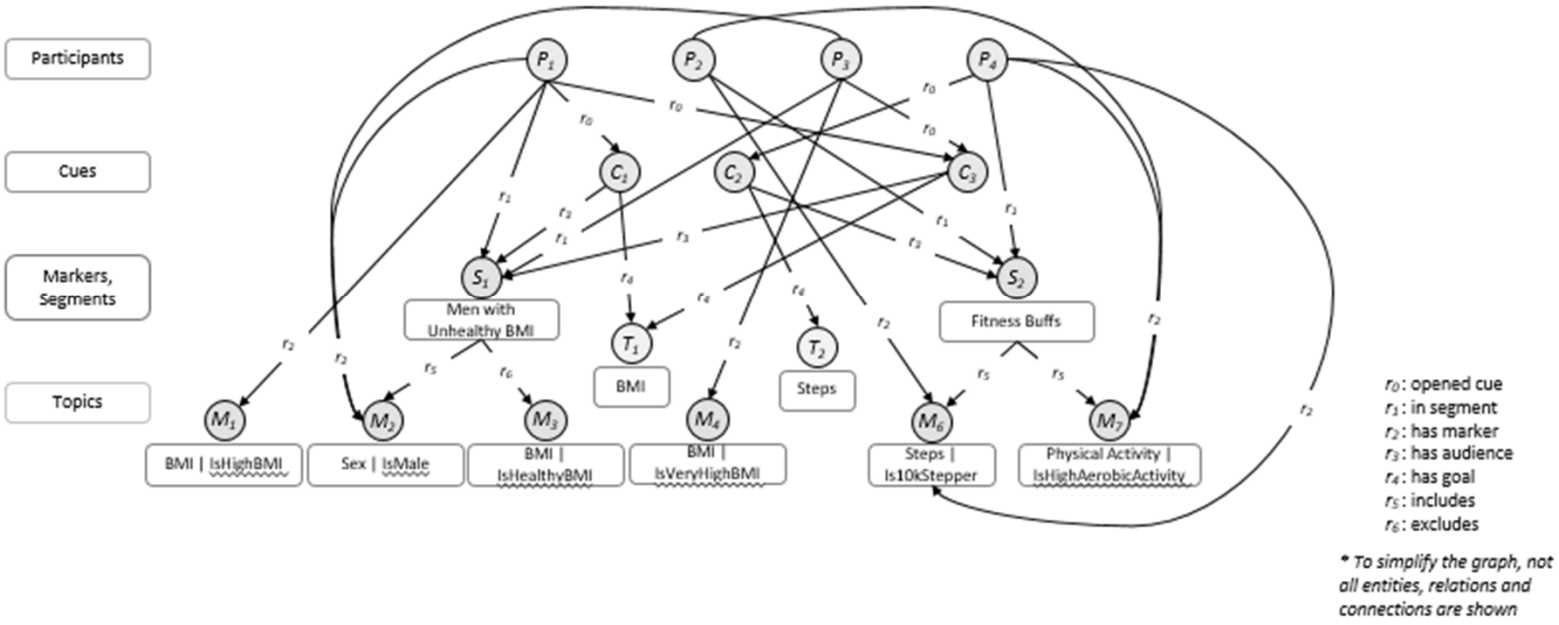

Participants
Cues
Markers, Segments
Topics
P1
P2
P3
P4
C1
C2
C3
S1
S2
Men with Unhealthy BMI
Fitness Buffs
T1
BMI
T2
Steps
M1
M2
M3
M4
M6
M7
BMI | IsHighBMI
Sex | IsMale
BMI | IsHealthyBMI
BMI | IsVeryHighBMI
Steps | Is10kStepper
Physical Activity | IsHighAerobicActivity
r0: opened cue
r1: in segment
r2: has marker
r3: has audience
r4: has goal
r5: includes
r6: excludes
* To simplify the graph, not all entities, relations and connections are shown

FIG. 39

3900

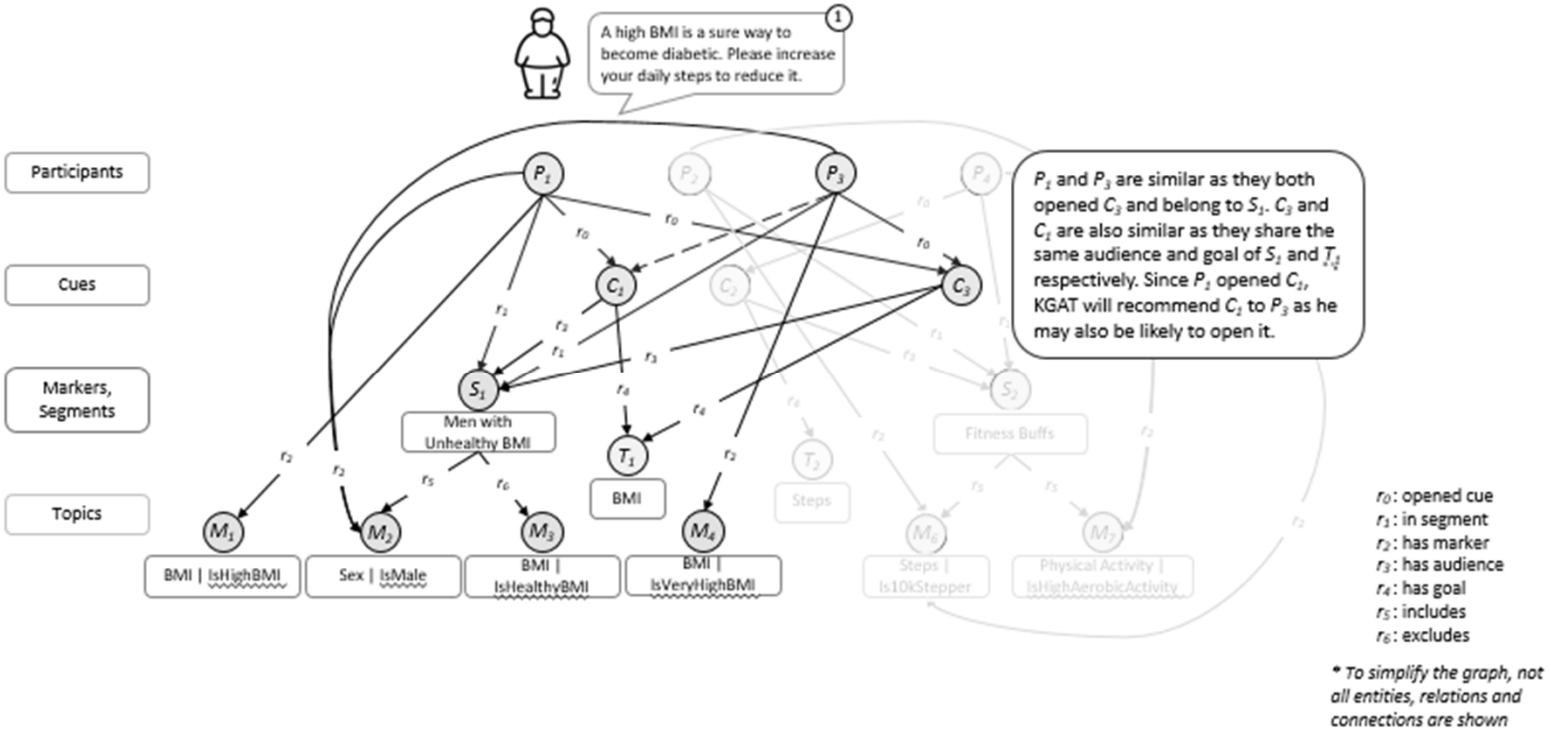

A high BMI is a sure way to become diabetic. Please increase your daily steps to reduce it.
1
Participants
Cues
Markers, Segments
Topics
P_1 and P_3 are similar as they both opened C_3 and belong to S_1. C_3 and C_1 are also similar as they share the same audience and goal of S_1 and T_1 respectively. Since P_1 opened C_1, KGAT will recommend C_1 to P_3 as he may also be likely to open it.
Men with Unhealthy BMI
BMI
Fitness Buffs
Steps
BMI | IsHighBMI
Sex | IsMale
BMI | IsHealthyBMI
BMI | IsVeryHighBMI
Steps | Is10kStepper
Physical Activity | IsHighAerobicActivity
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
* To simplify the graph, not all entities, relations and connections are shown

FIG. 40

3900

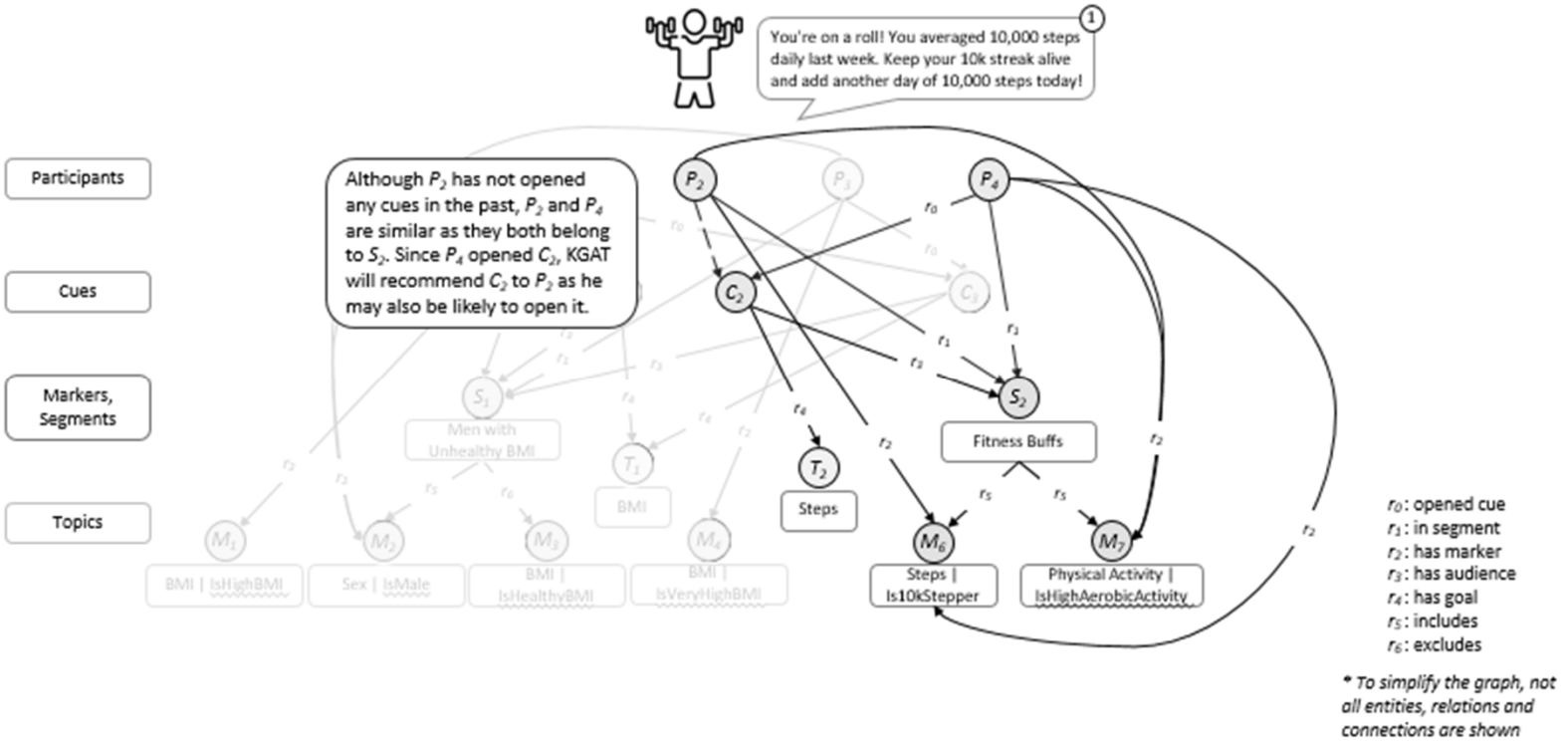

You're on a roll! You averaged 10,000 steps daily last week. Keep your 10k streak alive and add another day of 10,000 steps today!
1
Participants
Cues
Markers, Segments
Topics
Although P_2 has not opened any cues in the past, P_2 and P_4 are similar as they both belong to S_2. Since P_4 opened C_2, KGAT will recommend C_2 to P_2 as he may also be likely to open it.
P_2
P_3
P_4
C_2
C_3
S_1
S_2
Men with Unhealthy BMI
Fitness Buffs
T_1
BMI
T_2
Steps
M_1
M_2
M_3
M_4
M_6
M_7
BMI | IsHighBMI
Sex | IsMale
BMI | IsHealthyBMI
BMI | IsVeryHighBMI
Steps | Is10kStepper
Physical Activity | IsHighAerobicActivity
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
* To simplify the graph, not all entities, relations and connections are shown

FIG. 41

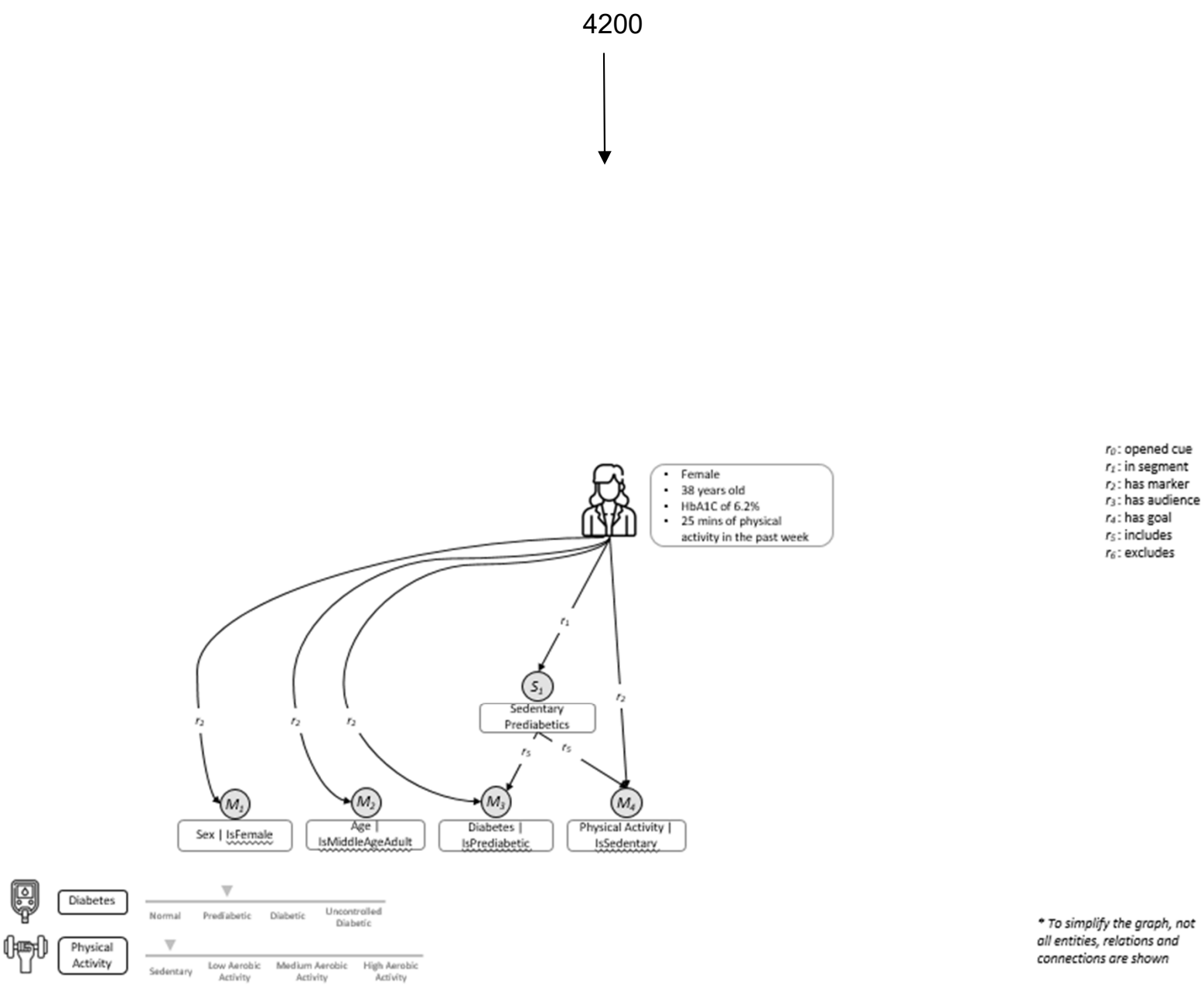

4200
Female
38 years old
HbA1C of 6.2%
25 mins of physical activity in the past week
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
S_1
Sedentary Prediabetics
M_1
Sex | IsFemale
M_2
Age | IsMiddleAgeAdult
M_3
Diabetes | IsPrediabetic
M_4
Physical Activity | IsSedentary
Diabetes
Normal
Prediabetic
Diabetic
Uncontrolled Diabetic
Physical Activity
Sedentary
Low Aerobic Activity
Medium Aerobic Activity
High Aerobic Activity
* To simplify the graph, not all entities, relations and connections are shown

FIG. 42

4300

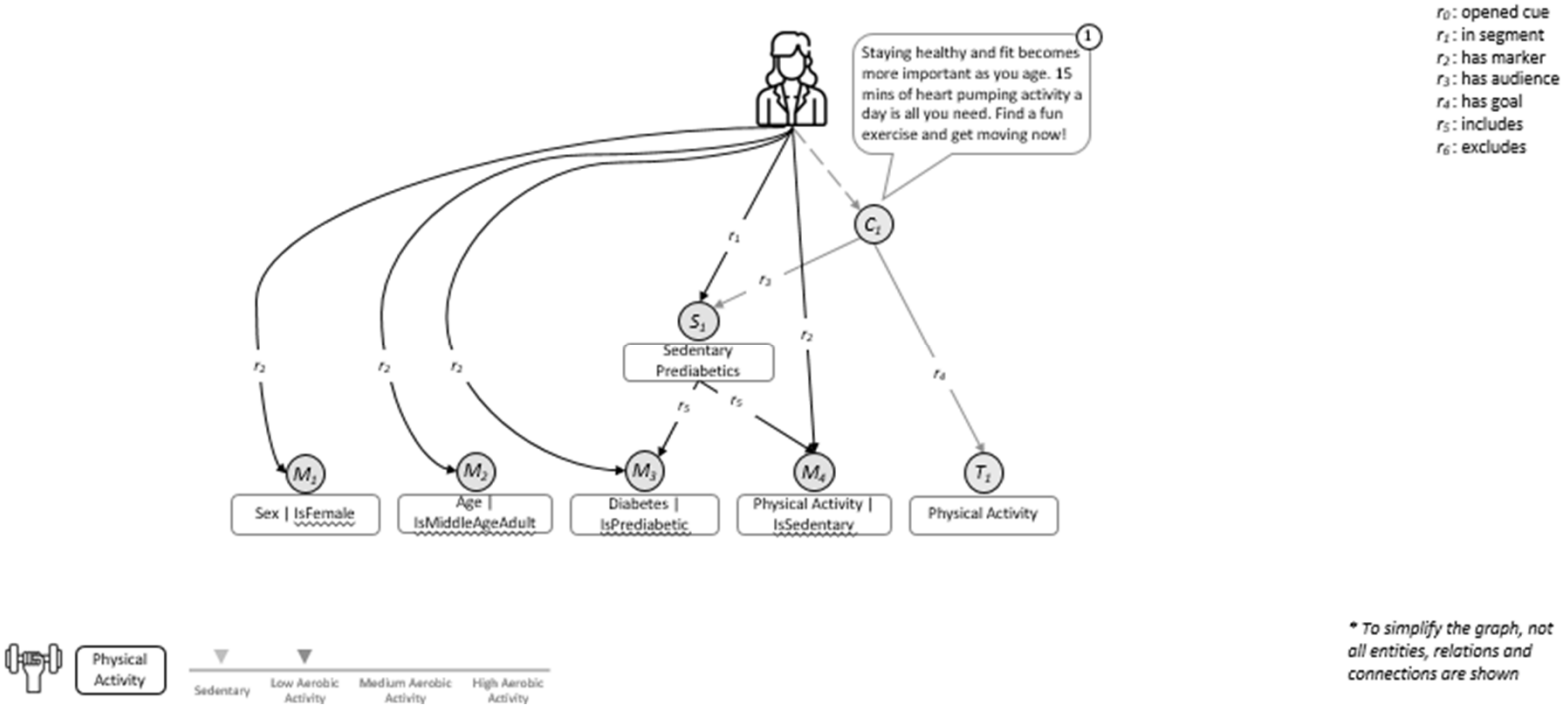

Staying healthy and fit becomes more important as you age. 15 mins of heart pumping activity a day is all you need. Find a fun exercise and get moving now!
1
C_1
S_1
Sedentary Prediabetics
M_1
Sex | IsFemale
M_2
Age | IsMiddleAgeAdult
M_3
Diabetes | IsPrediabetic
M_4
Physical Activity | IsSedentary
T_1
Physical Activity
r_1
r_2
r_3
r_4
r_5
r_6
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
Physical Activity
Sedentary
Low Aerobic Activity
Medium Aerobic Activity
High Aerobic Activity
* To simplify the graph, not all entities, relations and connections are shown

FIG. 43

4400

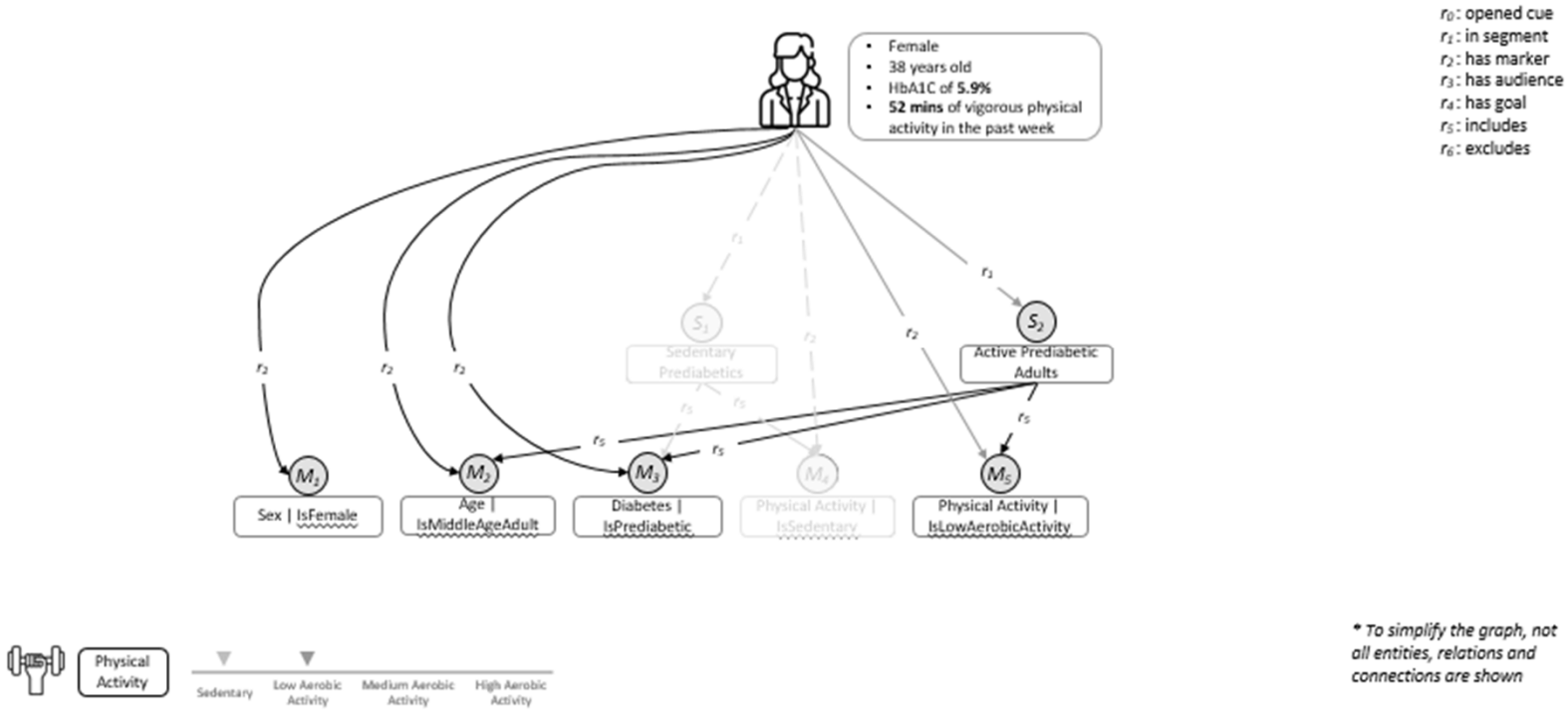

Female
38 years old
HbA1C of 5.9%
52 mins of vigorous physical activity in the past week
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
S_1
Sedentary Prediabetics
S_2
Active Prediabetic Adults
M_1
Sex | IsFemale
M_2
Age | IsMiddleAgeAdult
M_3
Diabetes | IsPrediabetic
M_4
Physical Activity | IsSedentary
M_5
Physical Activity | IsLowAerobicActivity
Physical Activity
Sedentary
Low Aerobic Activity
Medium Aerobic Activity
High Aerobic Activity
* To simplify the graph, not all entities, relations and connections are shown

FIG. 44

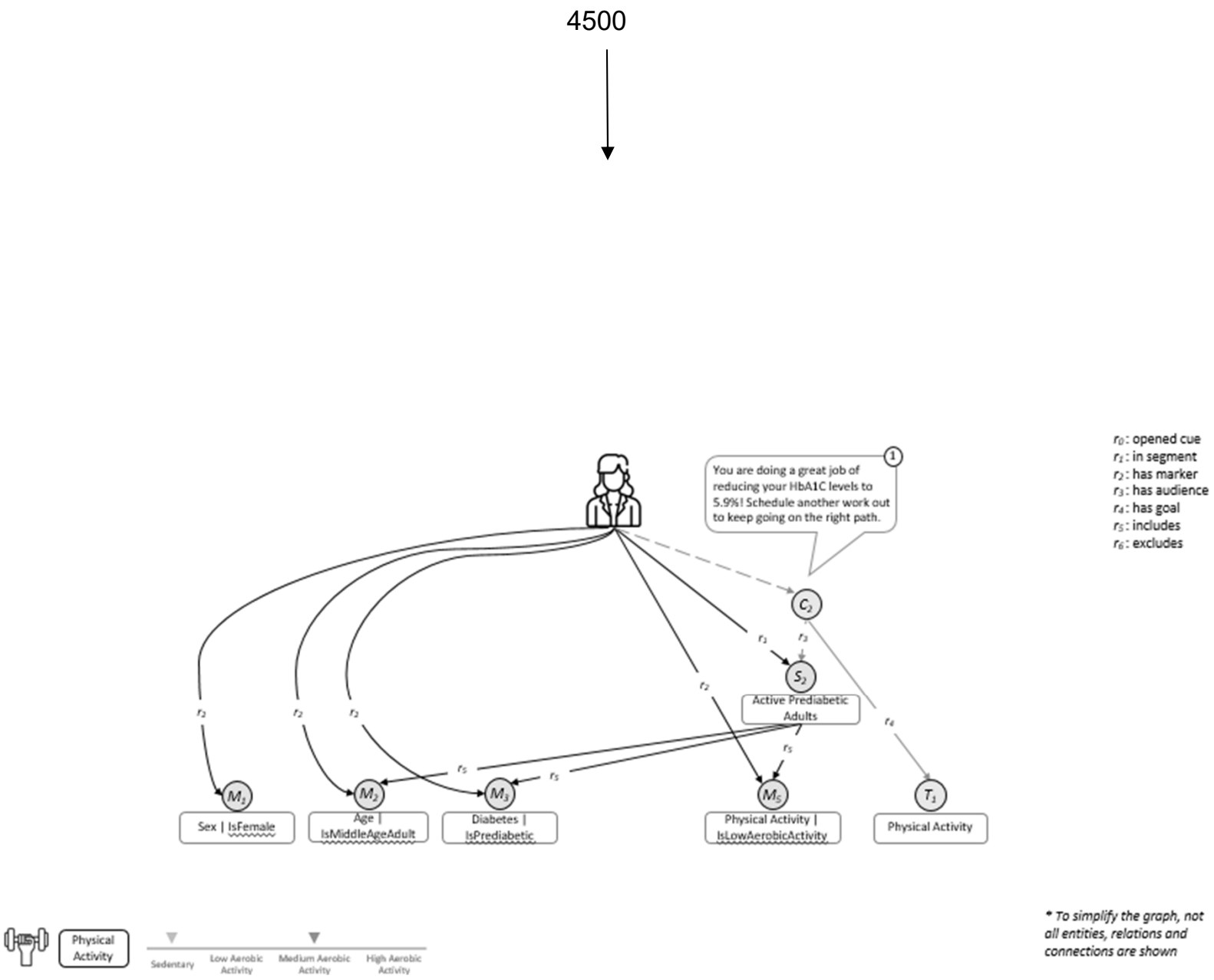

4500
You are doing a great job of reducing your HbA1C levels to 5.9%! Schedule another work out to keep going on the right path.
1
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
C_2
S_2
Active Prediabetic Adults
M_1
Sex | IsFemale
M_2
Age | IsMiddleAgeAdult
M_3
Diabetes | IsPrediabetic
M_5
Physical Activity | IsLowAerobicActivity
T_1
Physical Activity
Physical Activity
Sedentary
Low Aerobic Activity
Medium Aerobic Activity
High Aerobic Activity
* To simplify the graph, not all entities, relations and connections are shown

FIG. 45

4600

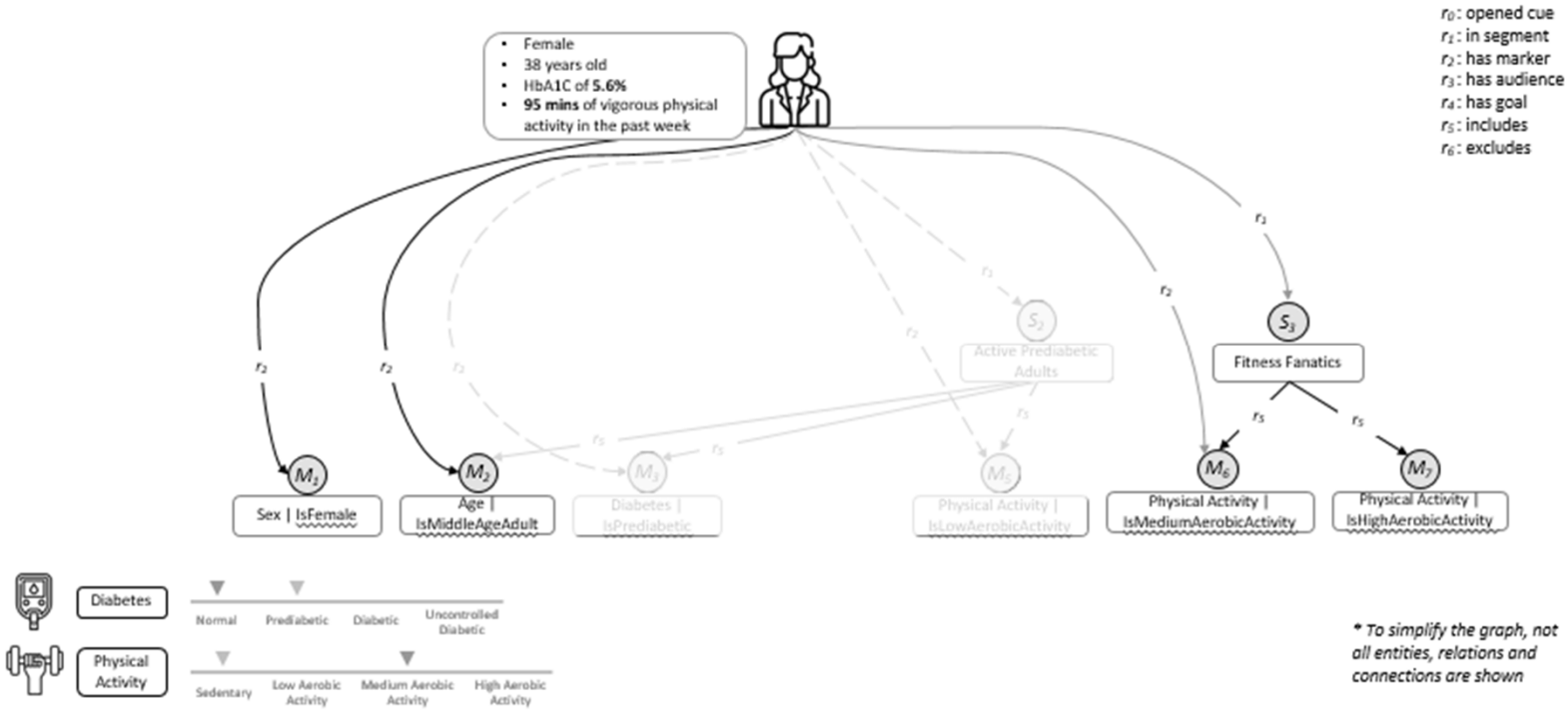

Female
38 years old
HbA1C of 5.6%
95 mins of vigorous physical activity in the past week
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
S_2
Active Prediabetic Adults
S_3
Fitness Fanatics
M_1
Sex | IsFemale
M_2
Age | IsMiddleAgeAdult
M_3
Diabetes | IsPrediabetic
M_5
Physical Activity | IsLowAerobicActivity
M_6
Physical Activity | IsMediumAerobicActivity
M_7
Physical Activity | IsHighAerobicActivity
Diabetes
Normal
Prediabetic
Diabetic
Uncontrolled Diabetic
Physical Activity
Sedentary
Low Aerobic Activity
Medium Aerobic Activity
High Aerobic Activity
* To simplify the graph, not all entities, relations and connections are shown

FIG. 46

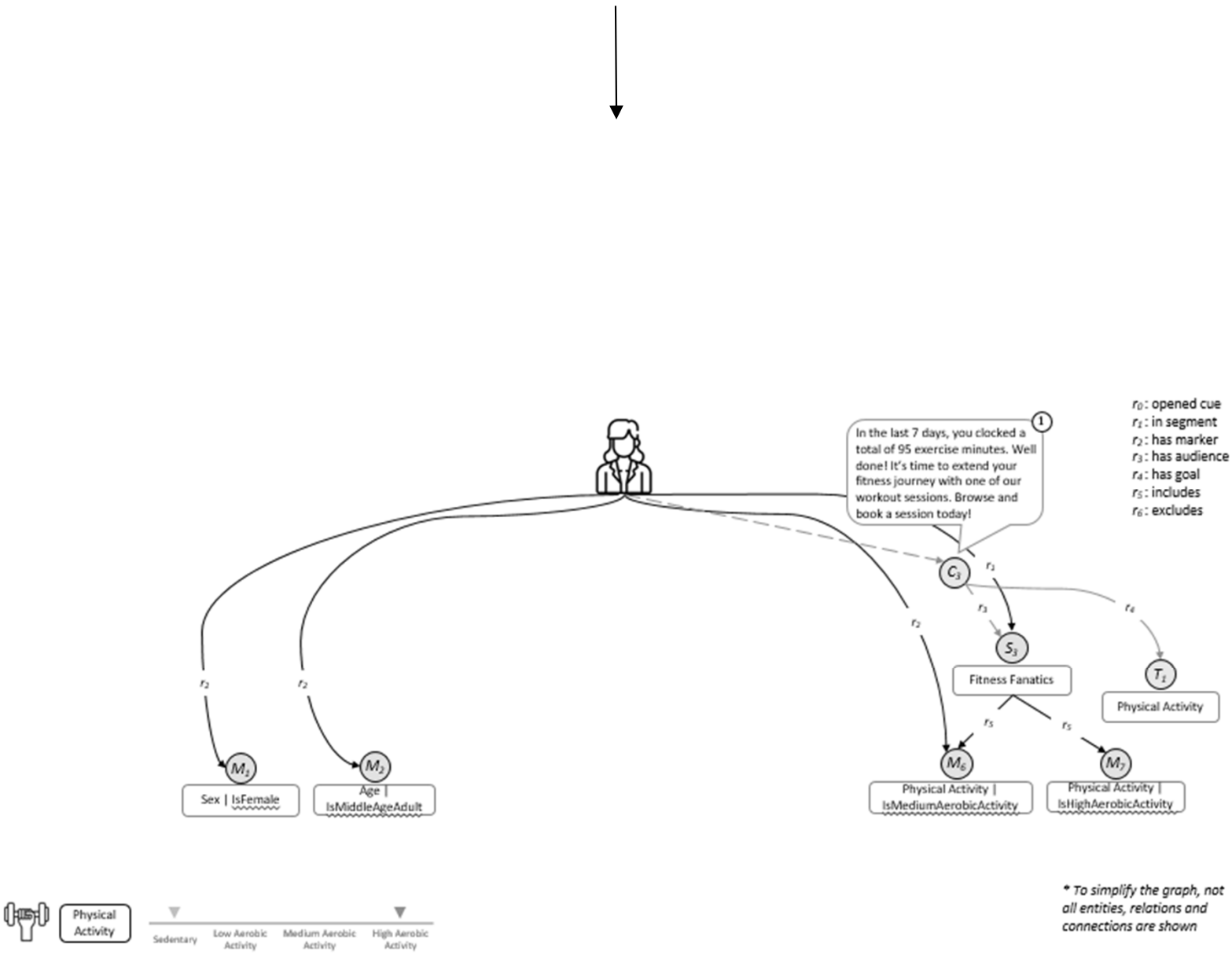

4700
In the last 7 days, you clocked a total of 95 exercise minutes. Well done! It's time to extend your fitness journey with one of our workout sessions. Browse and book a session today!
1
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
C_3
S_3
Fitness Fanatics
T_1
Physical Activity
M_1
Sex | IsFemale
M_2
Age | IsMiddleAgeAdult
M_6
Physical Activity | IsMediumAerobicActivity
M_7
Physical Activity | IsHighAerobicActivity
Physical Activity
Sedentary
Low Aerobic Activity
Medium Aerobic Activity
High Aerobic Activity
* To simplify the graph, not all entities, relations and connections are shown

FIG. 47

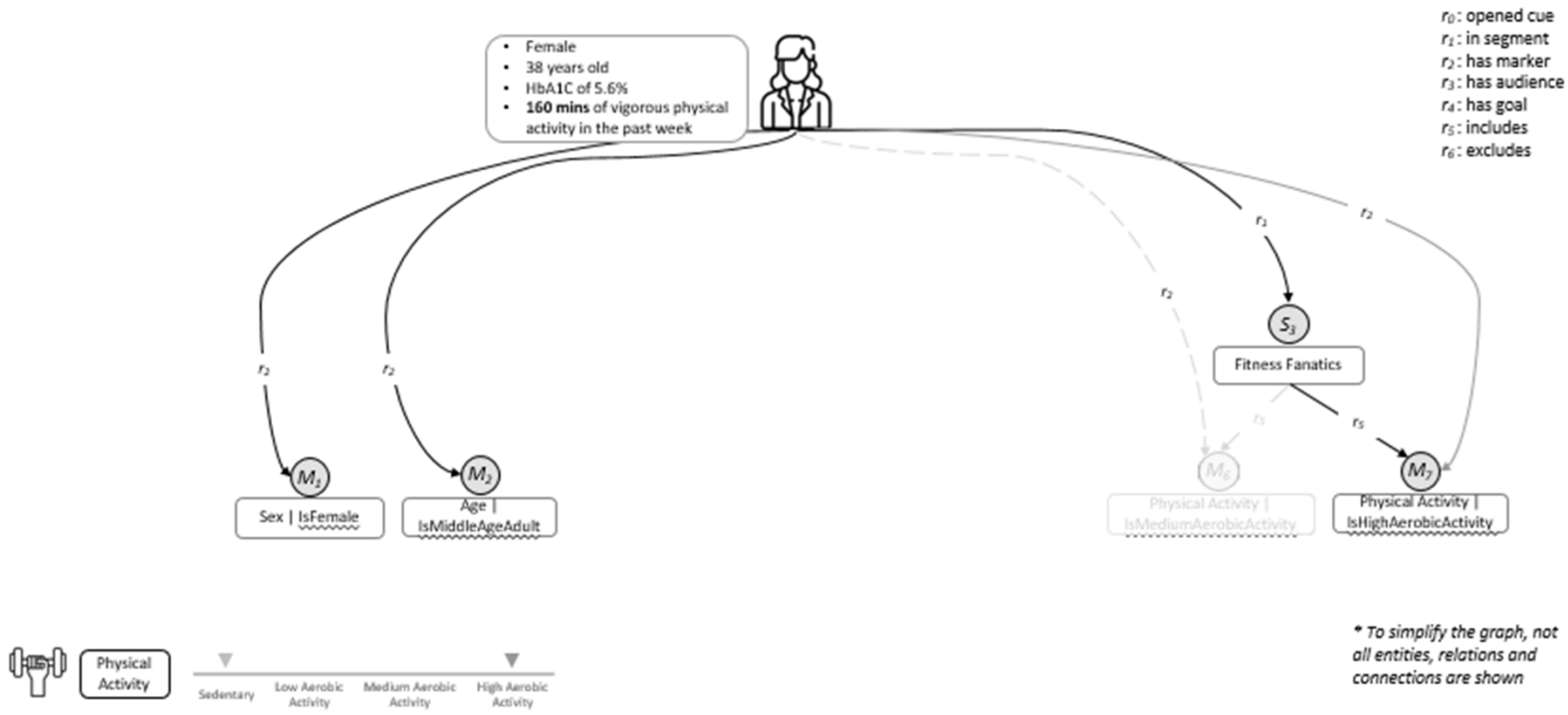

4800
Female
38 years old
HbA1C of 5.6%
160 mins of vigorous physical activity in the past week
r_0: opened cue
r_1: in segment
r_2: has marker
r_3: has audience
r_4: has goal
r_5: includes
r_6: excludes
S_3
Fitness Fanatics
M_1
Sex | IsFemale
M_2
Age | IsMiddleAgeAdult
M_6
Physical Activity | IsMediumAerobicActivity
M_7
Physical Activity | IsHighAerobicActivity
Physical Activity
Sedentary
Low Aerobic Activity
Medium Aerobic Activity
High Aerobic Activity
* To simplify the graph, not all entities, relations and connections are shown

FIG. 48

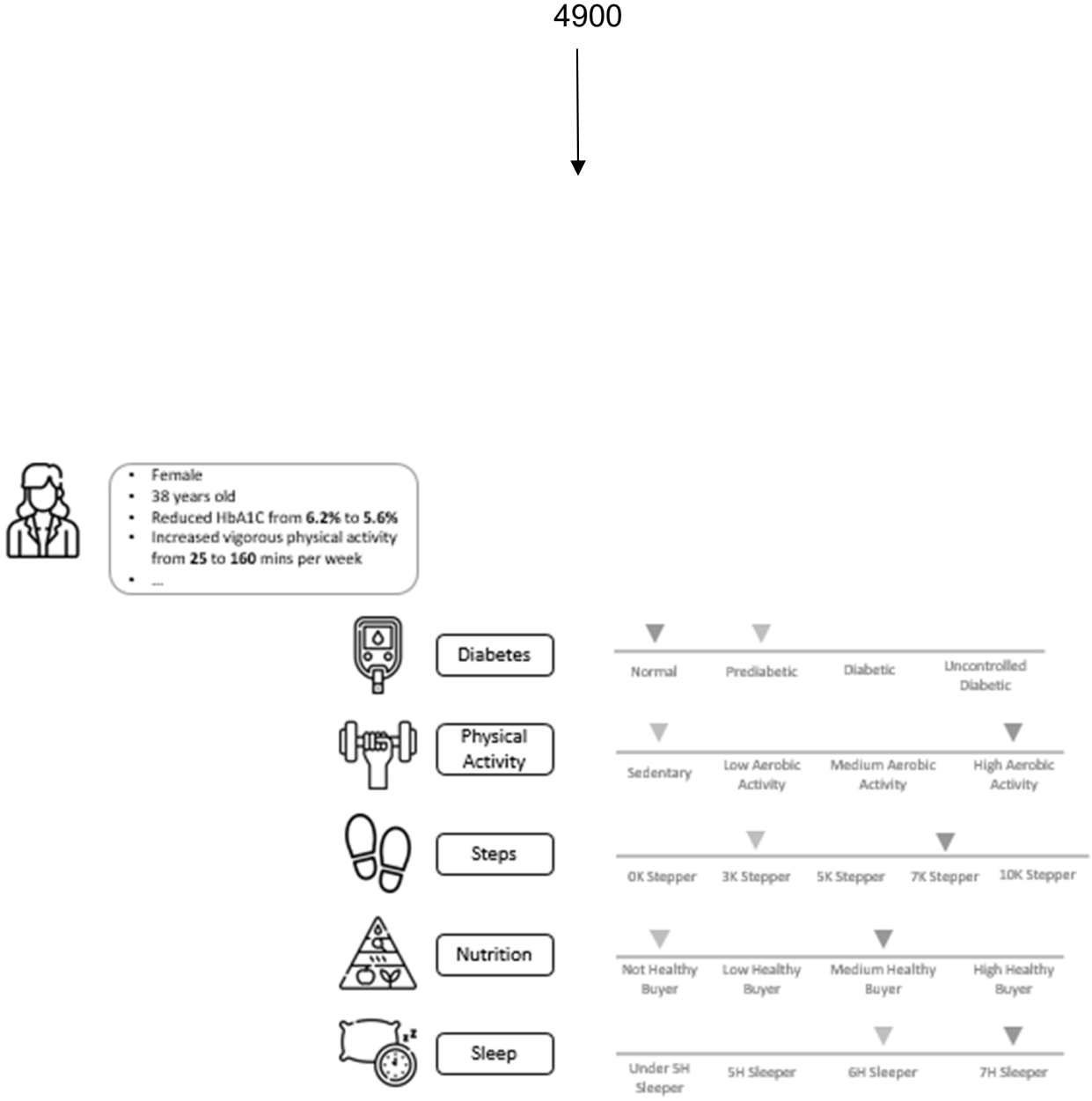

4900
Female
38 years old
Reduced HbA1C from 6.2% to 5.6%
Increased vigorous physical activity from 25 to 160 mins per week
...
Diabetes
Normal
Prediabetic
Diabetic
Uncontrolled Diabetic
Physical Activity
Sedentary
Low Aerobic Activity
Medium Aerobic Activity
High Aerobic Activity
Steps
0K Stepper
3K Stepper
5K Stepper
7K Stepper
10K Stepper
Nutrition
Not Healthy Buyer
Low Healthy Buyer
Medium Healthy Buyer
High Healthy Buyer
Sleep
Under 5H Sleeper
5H Sleeper
6H Sleeper
7H Sleeper

FIG. 49

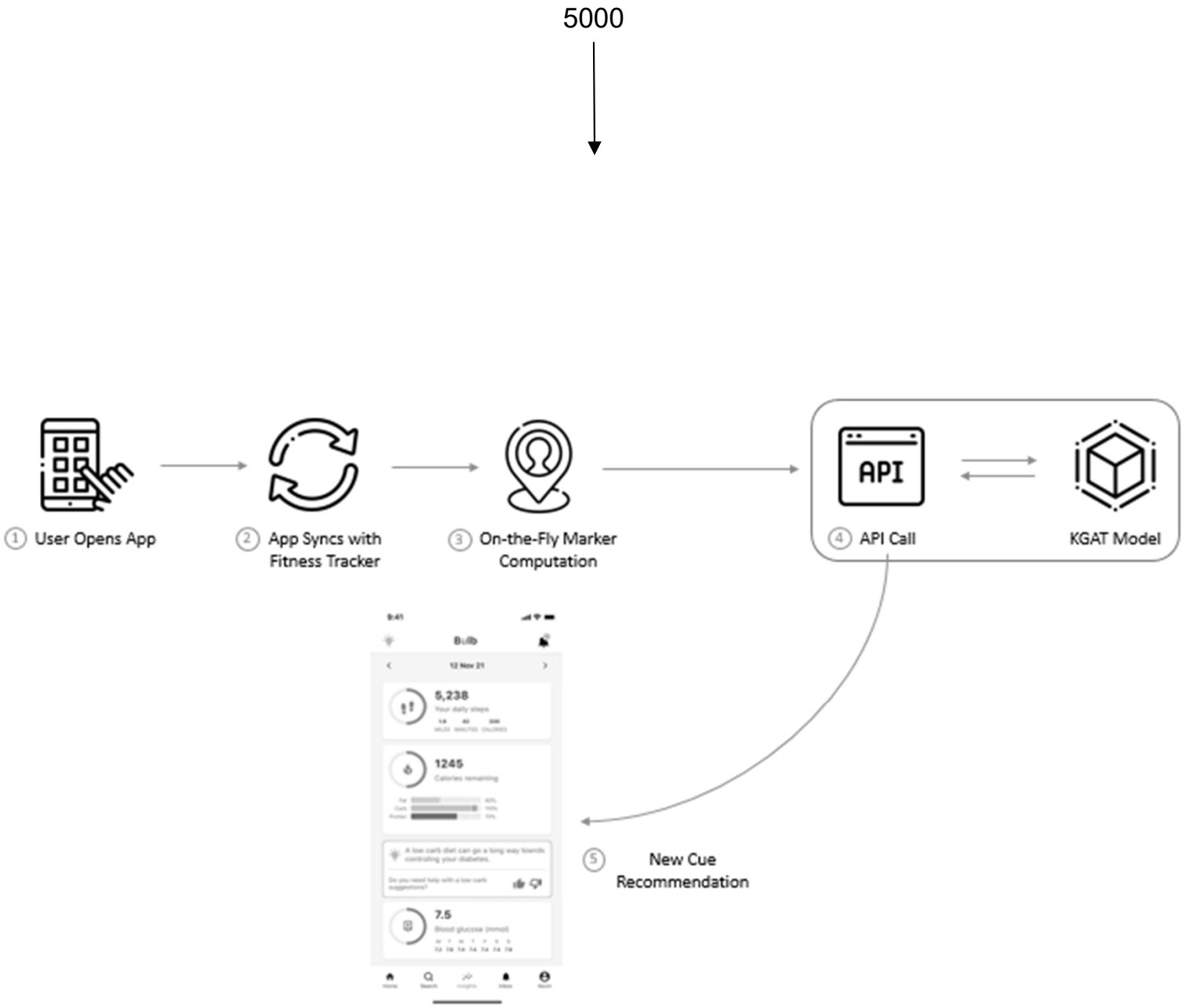

5000
1 User Opens App
2 App Syncs with Fitness Tracker
3 On-the-Fly Marker Computation
API
4 API Call
KGAT Model
5 New Cue Recommendation
5,238
1245
7.5

FIG. 50

5100

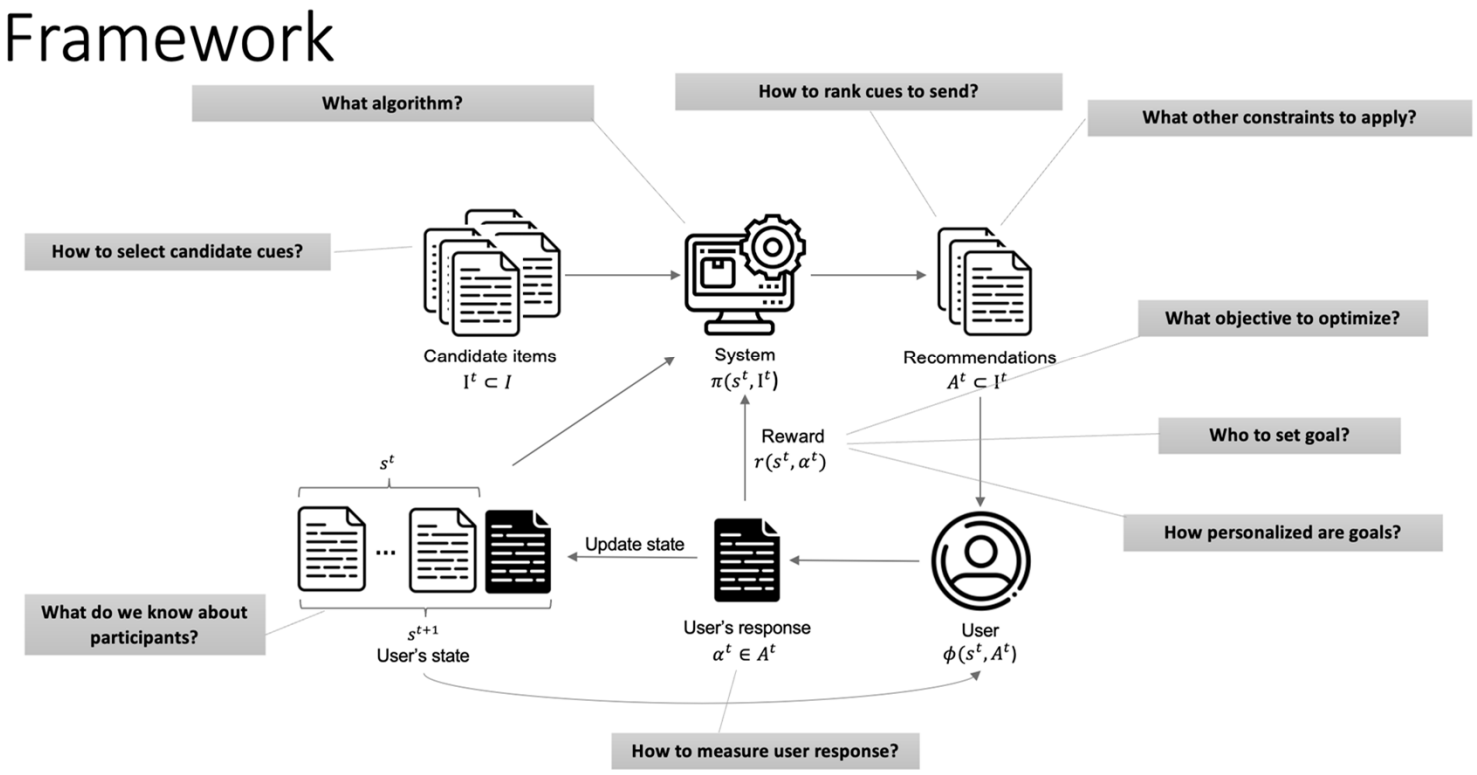

Framework
What algorithm?
How to rank cues to send?
What other constraints to apply?
How to select candidate cues?
Candidate items
$I^t \subset I$
System
$\pi(s^t, I^t)$
Recommendations
$A^t \subset I^t$
What objective to optimize?
Reward
$r(s^t, a^t)$
Who to set goal?
How personalized are goals?
s^t
Update state
What do we know about participants?
s^{t+1}
User's state
User's response
$a^t \in A^t$
User
$\phi(s^t, A^t)$
How to measure user response?

FIG. 51

5200

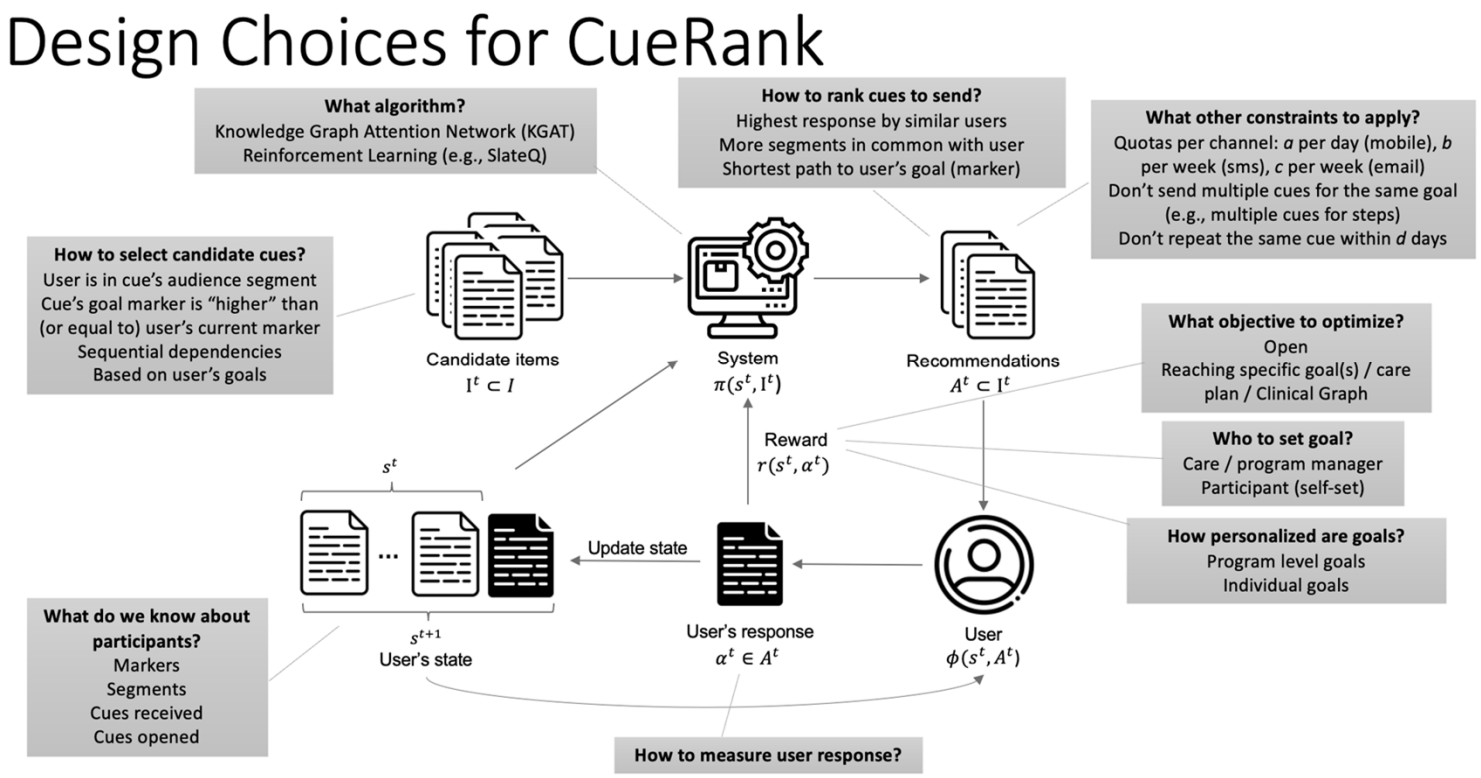
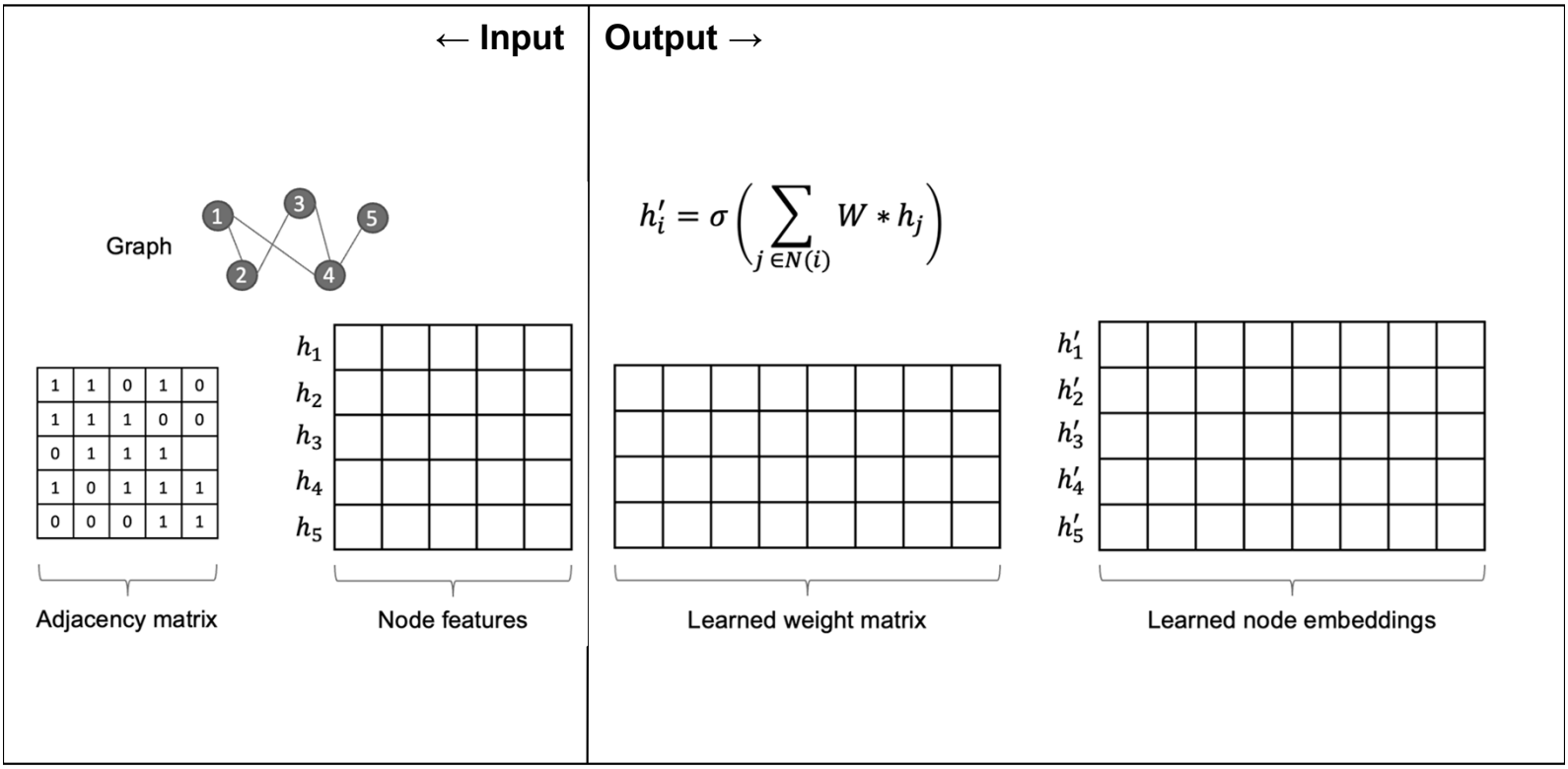

Design Choices for CueRank
What algorithm?
Knowledge Graph Attention Network (KGAT)
Reinforcement Learning (e.g., SlateQ)
How to rank cues to send?
Highest response by similar users
More segments in common with user
Shortest path to user's goal (marker)
What other constraints to apply?
Quotas per channel: a per day (mobile), b per week (sms), c per week (email)
Don't send multiple cues for the same goal (e.g., multiple cues for steps)
Don't repeat the same cue within d days
How to select candidate cues?
User is in cue's audience segment
Cue's goal marker is "higher" than (or equal to) user's current marker
Sequential dependencies
Based on user's goals
Candidate items
$I^t \subset I$
System
$\pi(s^t, I^t)$
Recommendations
$A^t \subset I^t$
What objective to optimize?
Open
Reaching specific goal(s) / care plan / Clinical Graph
Reward
$r(s^t, a^t)$
Who to set goal?
Care / program manager
Participant (self-set)
How personalized are goals?
Program level goals
Individual goals
s^t
Update state
s^{t+1}
User's state
User's response
$a^t \in A^t$
User
$\phi(s^t, A^t)$
What do we know about participants?
Markers
Segments
Cues received
Cues opened
How to measure user response?

FIG. 52

5300



Transferrable:

1. Weight matrix
2. Subset of embeddings:
   - Markers
   - Topics (Goals)
   - Cue Signatures
   - Segment Signatures
   - Participant Signatures

Cannot / not useful to transfer:

1. Adjacency matrix
2. Node features, if any
3. Embeddings for actual participants and cues

FIG. 53

5400

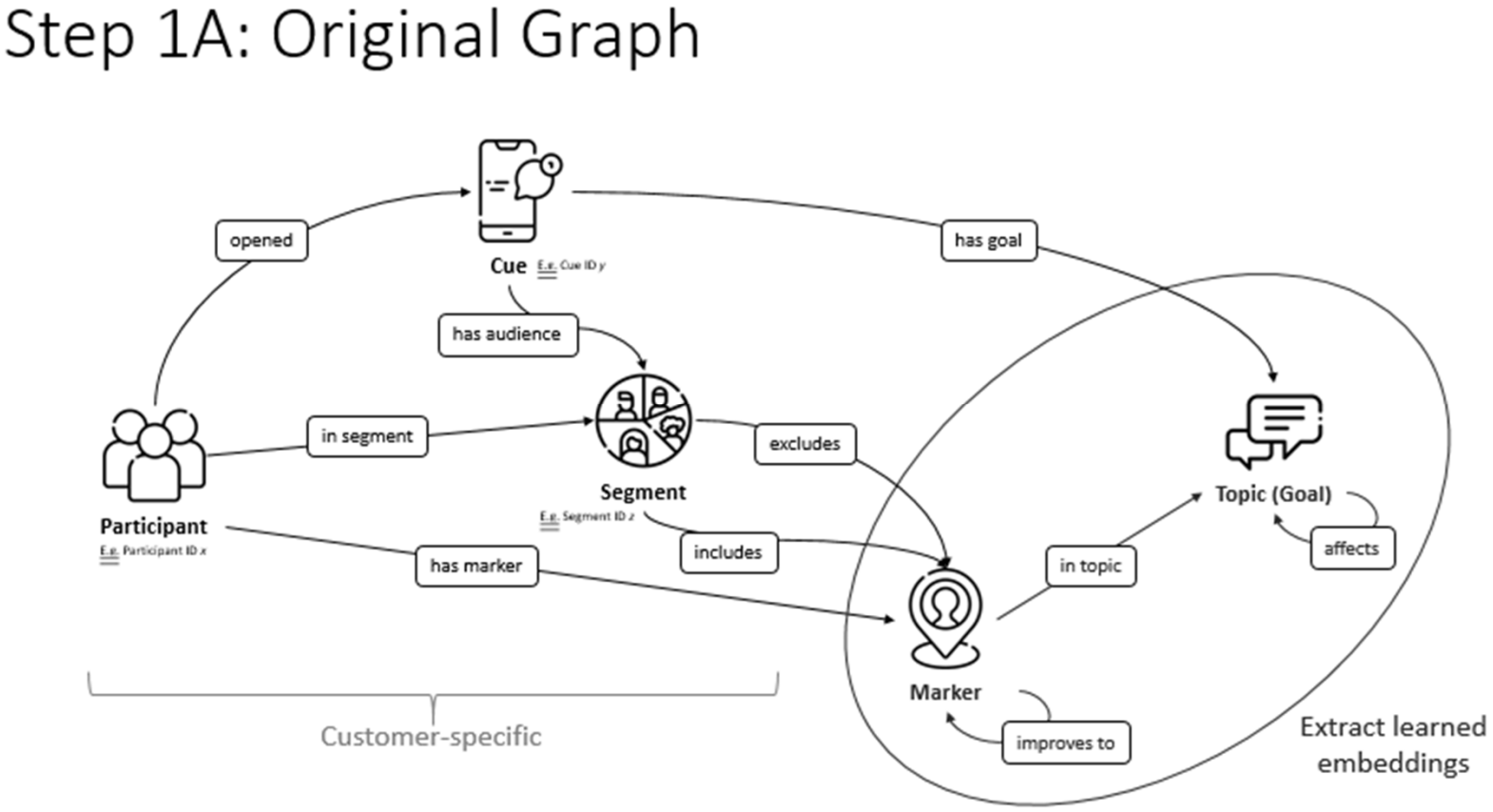

Step 1A: Original Graph
opened
Cue
E.g. Cue ID y
has goal
has audience
in segment
Segment
E.g. Segment ID z
excludes
Participant
E.g. Participant ID x
has marker
includes
in topic
Topic (Goal)
affects
Marker
improves to
Customer-specific
Extract learned embeddings

FIG. 54

5500

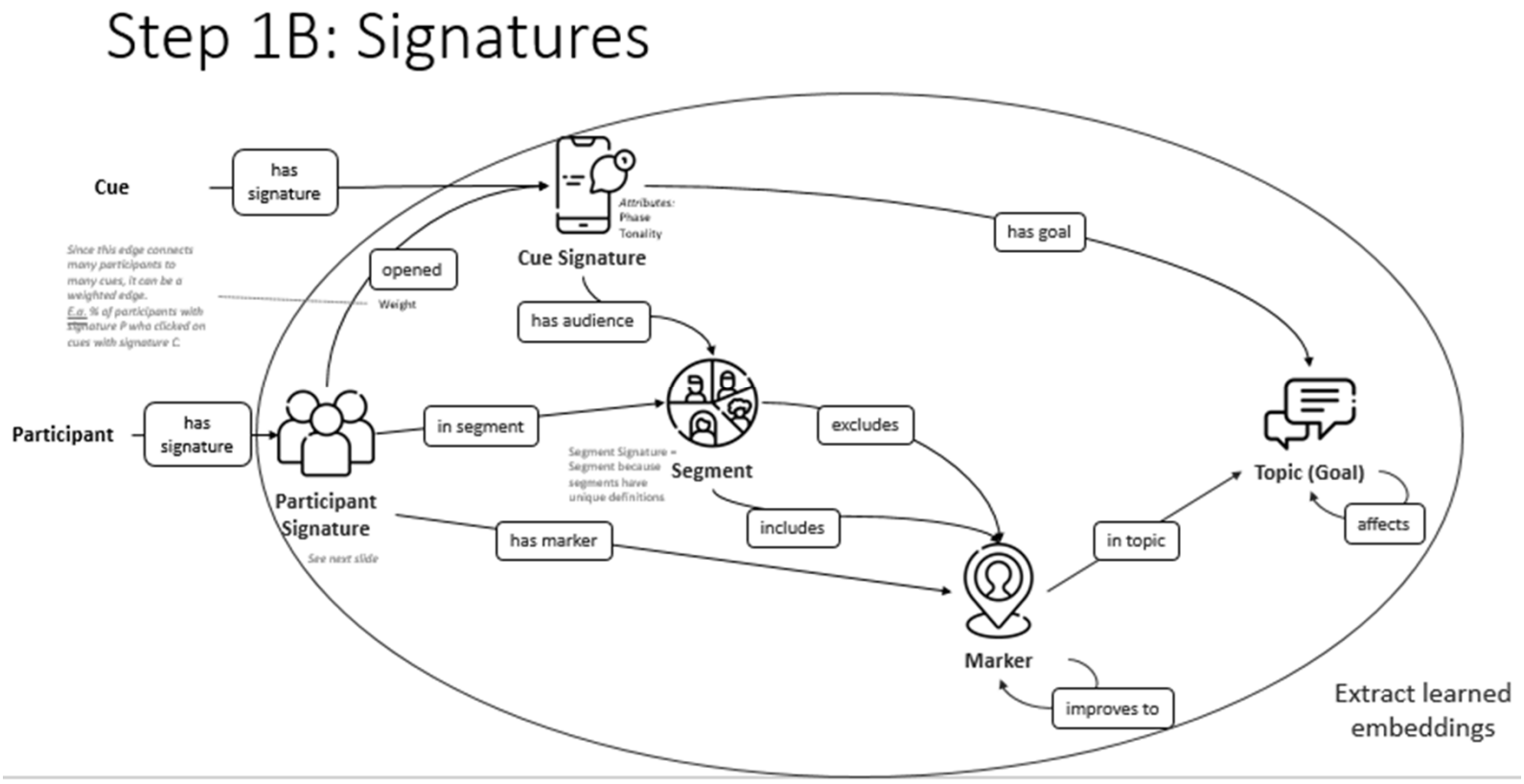

Step 1B: Signatures
Cue
has signature
Attributes:
Phase
Tonality
Cue Signature
has goal
Since this edge connects many participants to many cues, it can be a weighted edge.
E.g. % of participants with signature P who clicked on cues with signature C.
opened
Weight
has audience
Participant
has signature
in segment
Segment Signature = Segment because segments have unique definitions
Segment
excludes
Topic (Goal)
Participant Signature
See next slide
has marker
includes
in topic
affects
Marker
improves to
Extract learned embeddings

FIG. 55

5600

No. of possible Participant Signatures is determined by the markers in the signature

Example 1:

Age	isInfant isToddler isChild isAdolescent isYoungAdult	isAdult isMiddleAge isSenior Unknown	9
Sex	isMale isFemale	Unknown	3
MVPA	isLow isMedium isRecommended	isHigh Unknown	5
Steps	isUnder1k is1k is3k is5k	is7k is10k is15k Unknown	8
Nutrition	isLow isMedium	isHigh Unknown	4
Sleep	isUnder6h is6h is7h	is8h is9h Unknown	6
BMI	isUnderweight isNormal isOverweight	isObeseClassI isObeseClassII isObeseClassIII Unknown	7
Markers	**35**	**Unique Signatures**	**181,440**

Example 2:

Age	isChild isAdolescent isYoungAdult	isMiddleAge isSenior Unknown	6
Sex	isMale isFemale	Unknown	3
MVPA	isLow isMedium	isHigh Unknown	4
Steps	isUnder3k is3k is5k	is7k is10k Unknown	6
Nutrition	isLow isMedium	isHigh Unknown	4
Sleep	isLow isMedium	isHigh Unknown	4
BMI	isUnderweight isNormal isOverweight	isObese Unknown	5
Markers	**25**	**Unique Signatures**	**34,560**

FIG. 56

5700

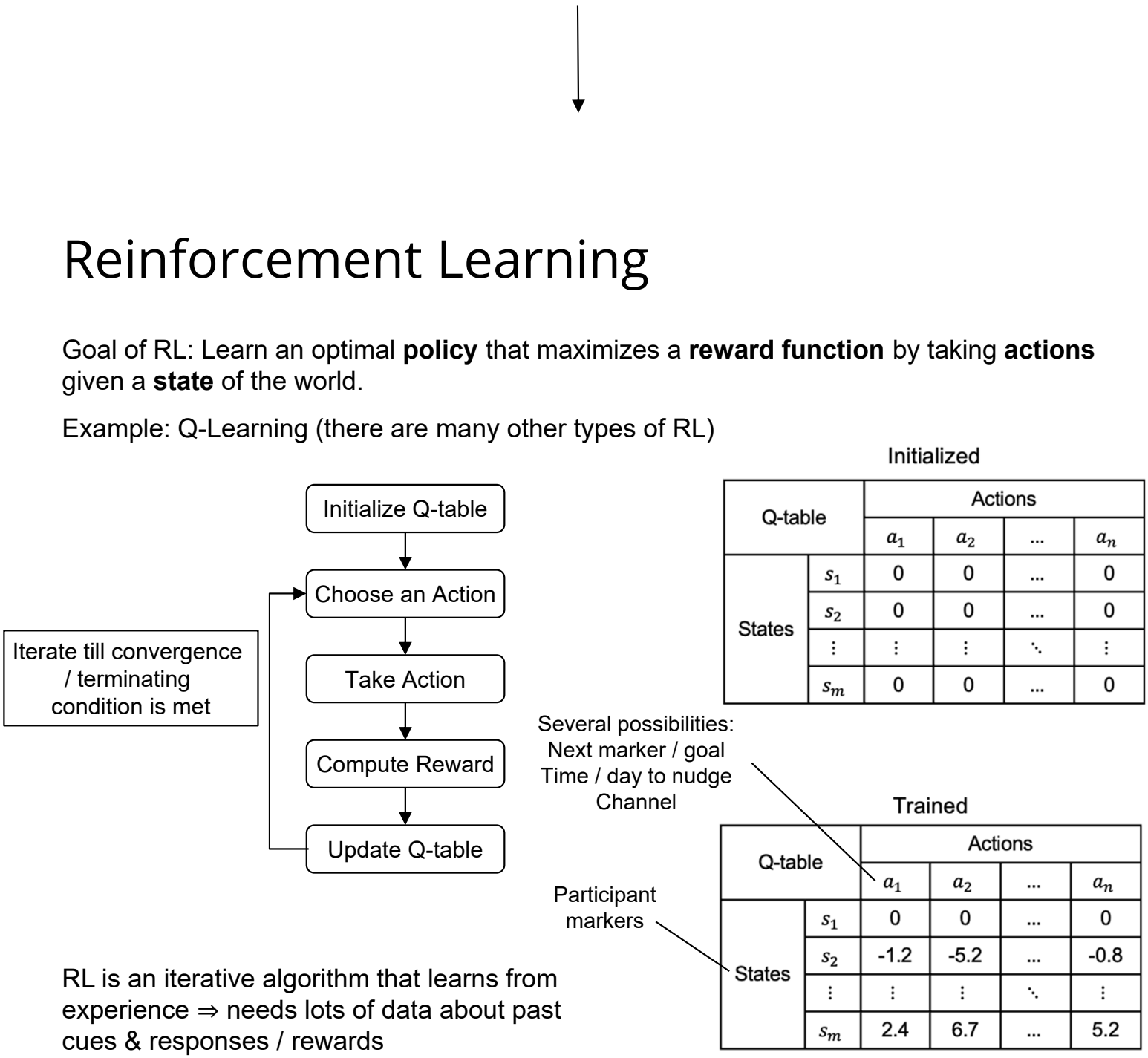

Reinforcement Learning
Goal of RL: Learn an optimal **policy** that maximizes a **reward function** by taking **actions** given a **state** of the world.
Example: Q-Learning (there are many other types of RL)
Initialize Q-table
Choose an Action
Take Action
Compute Reward
Update Q-table
Iterate till convergence / terminating condition is met
Initialized
Q-table
Actions
a_1 a_2 ... a_n
States
s_1 0 0 ... 0
s_2 0 0 ... 0
s_m 0 0 ... 0
Several possibilities:
Next marker / goal
Time / day to nudge
Channel
Trained
Participant markers
s_2 -1.2 -5.2 ... -0.8
s_m 2.4 6.7 ... 5.2
RL is an iterative algorithm that learns from experience ⇒ needs lots of data about past cues & responses / rewards

FIG. 57

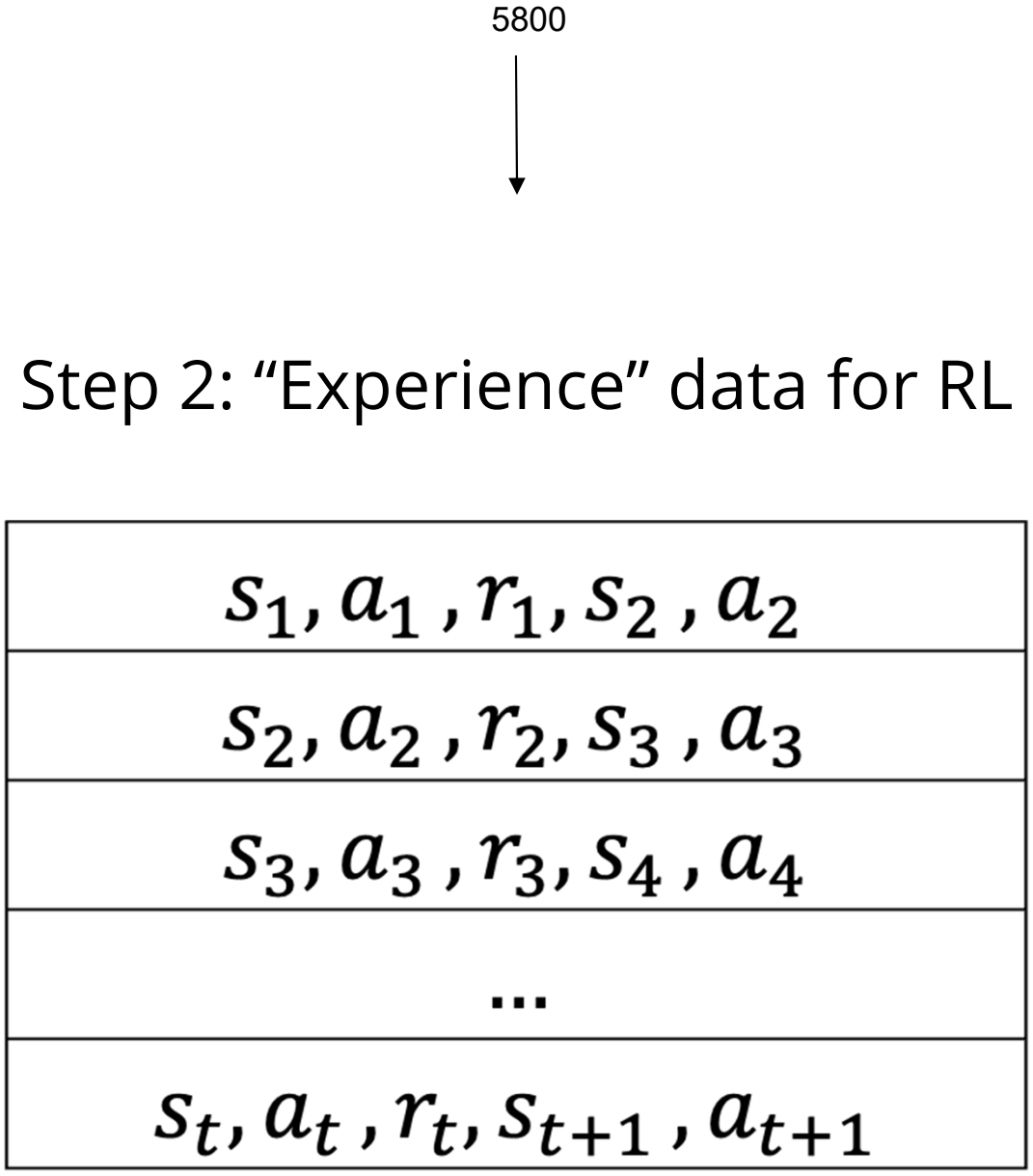

5800
Step 2: “Experience” data for RL
s_1, a_1, r_1, s_2, a_2
s_2, a_2, r_2, s_3, a_3
s_3, a_3, r_3, s_4, a_4
...
$s_t, a_t, r_t, s_{t+1}, a_{t+1}$

FIG. 58

5900

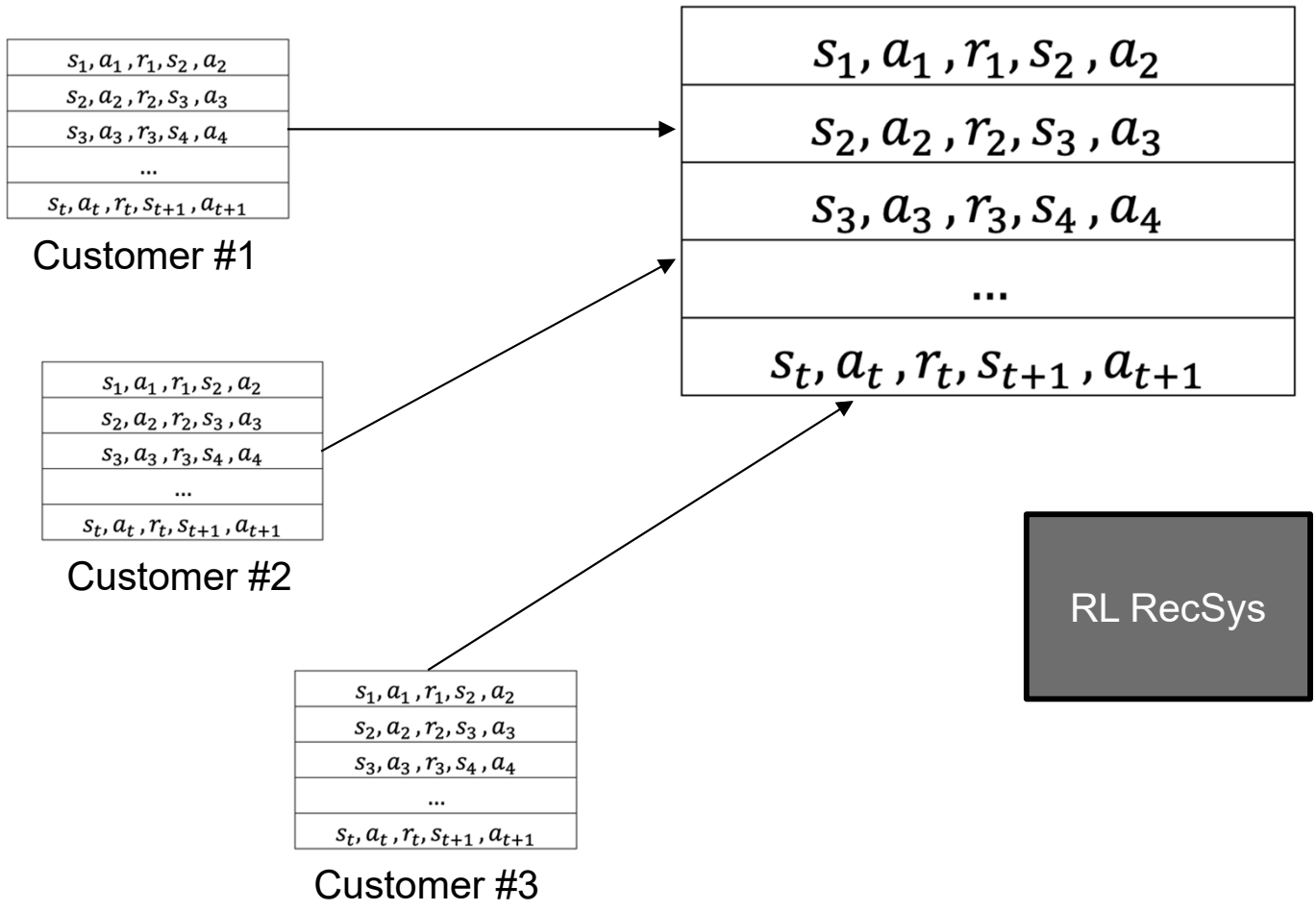

Step 2: “Experience” data for RL
s_1, a_1, r_1, s_2, a_2
s_2, a_2, r_2, s_3, a_3
s_3, a_3, r_3, s_4, a_4
...
$s_t, a_t, r_t, s_{t+1}, a_{t+1}$
Customer #1
Customer #2
Customer #3
RL RecSys

FIG. 59

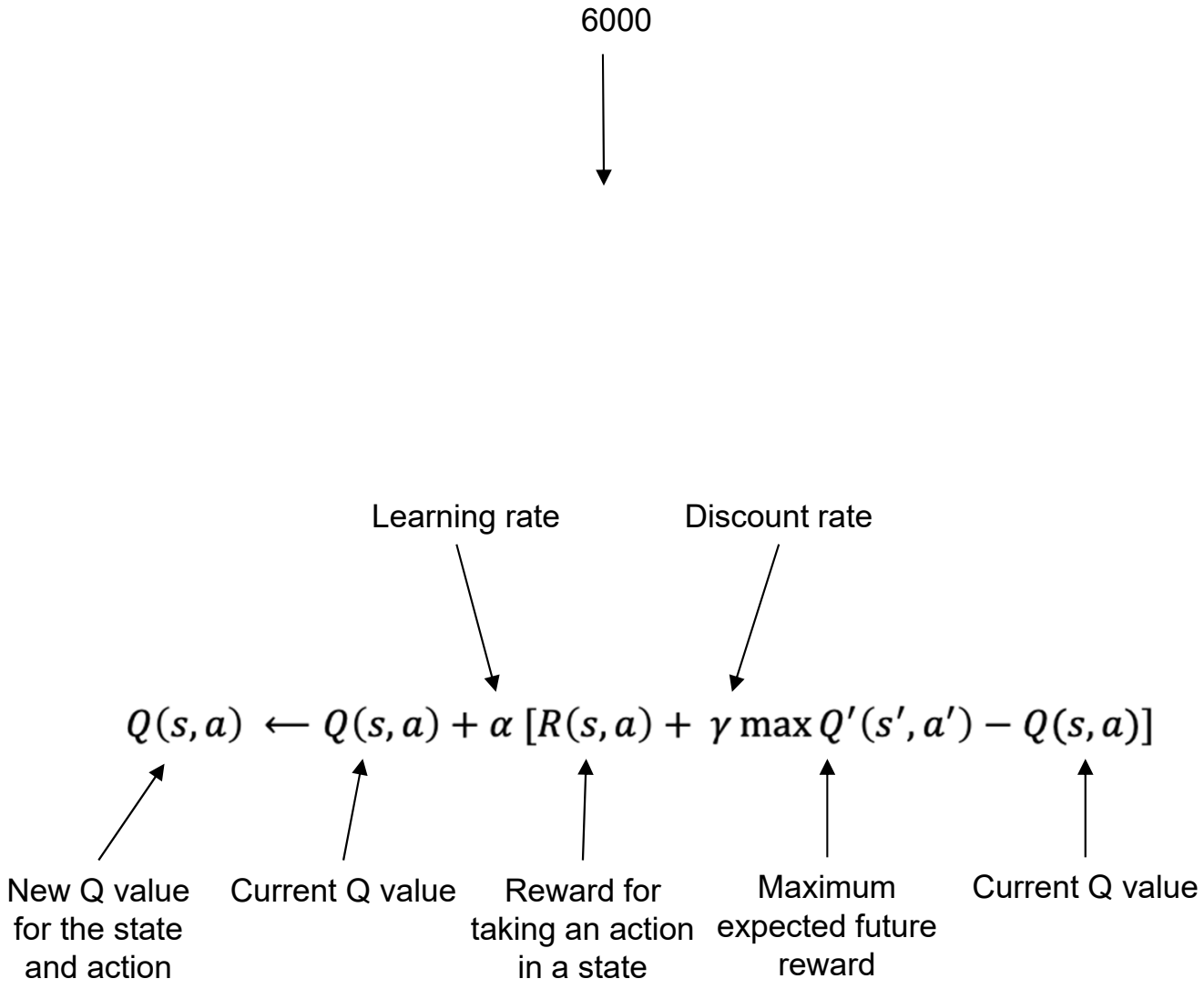

6000
Learning rate
Discount rate
$Q(s,a) \leftarrow Q(s,a) + \alpha\,[R(s,a) + \gamma \max Q'(s',a') - Q(s,a)]$
New Q value for the state and action
Current Q value
Reward for taking an action in a state
Maximum expected future reward
Current Q value

FIG. 60

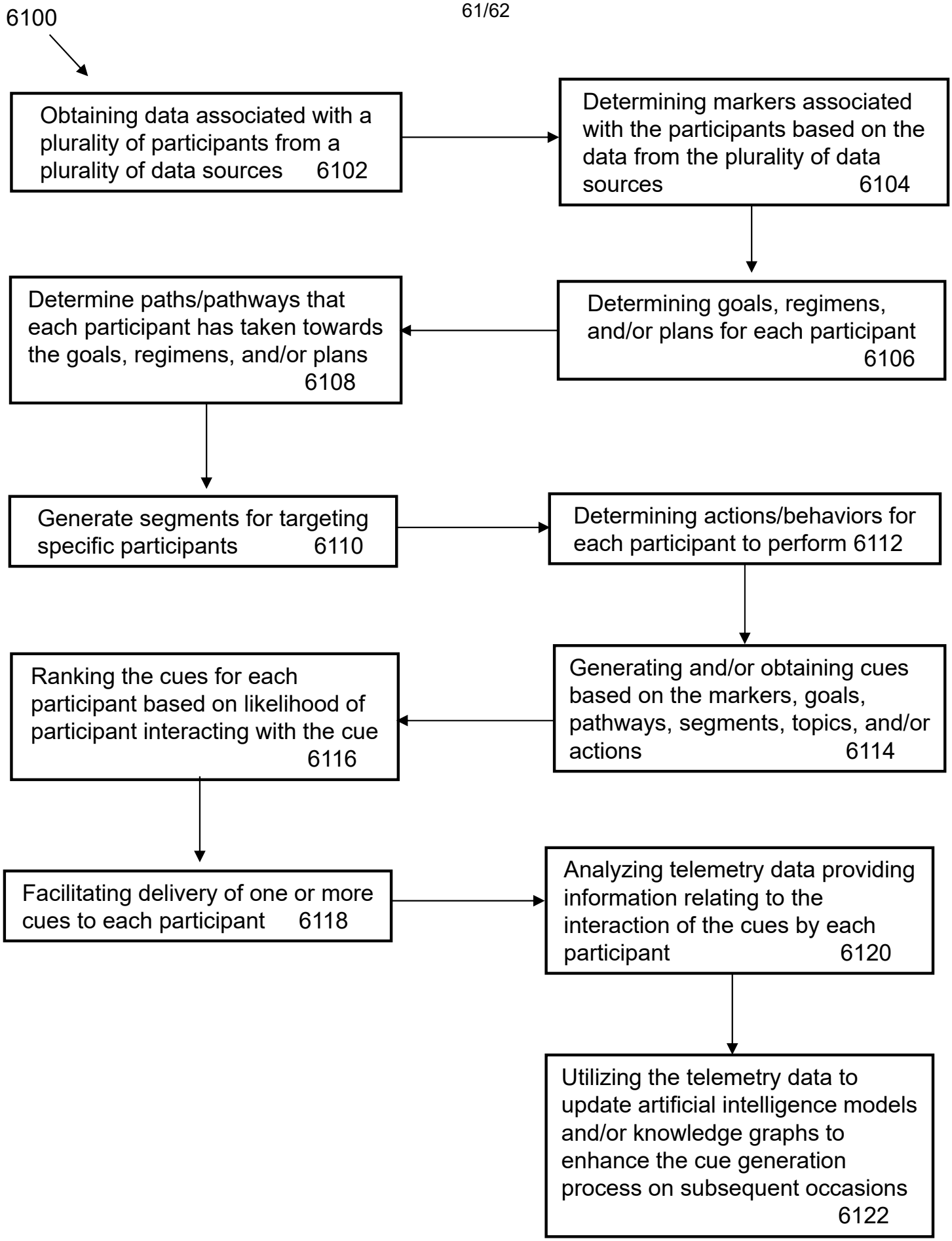

6100
Obtaining data associated with a plurality of participants from a plurality of data sources 6102
Determining markers associated with the participants based on the data from the plurality of data sources 6104
Determining goals, regimens, and/or plans for each participant 6106
Determine paths/pathways that each participant has taken towards the goals, regimens, and/or plans 6108
Generate segments for targeting specific participants 6110
Determining actions/behaviors for each participant to perform 6112
Generating and/or obtaining cues based on the markers, goals, pathways, segments, topics, and/or actions 6114
Ranking the cues for each participant based on likelihood of participant interacting with the cue 6116
Facilitating delivery of one or more cues to each participant 6118
Analyzing telemetry data providing information relating to the interaction of the cues by each participant 6120
Utilizing the telemetry data to update artificial intelligence models and/or knowledge graphs to enhance the cue generation process on subsequent occasions 6122

FIG. 61

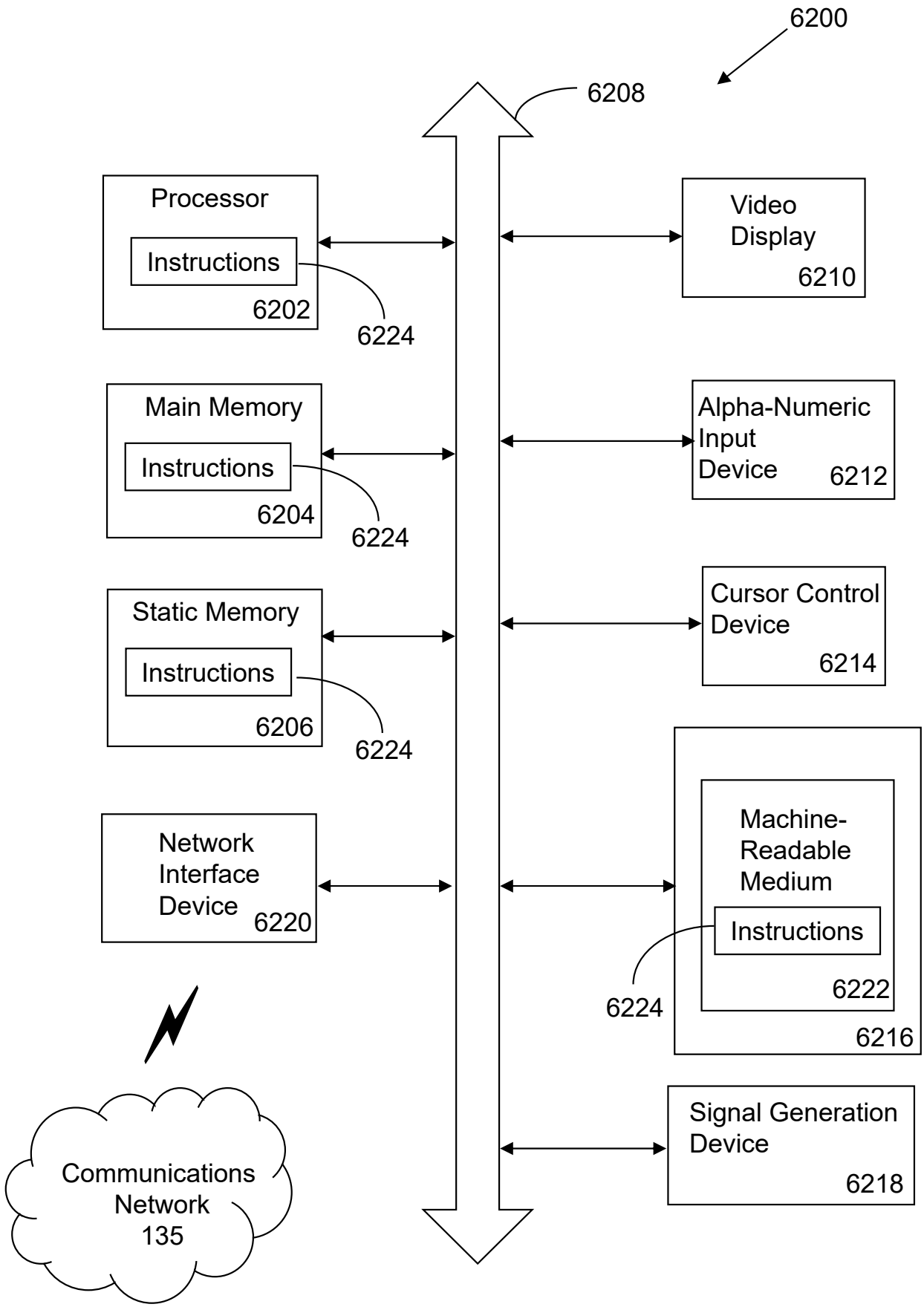

6200
6208
Processor
Instructions
6202
6224
Main Memory
Instructions
6204
6224
Static Memory
Instructions
6206
6224
Network Interface Device
6220
Communications Network
135
Video Display
6210
Alpha-Numeric Input Device
6212
Cursor Control Device
6214
Machine-Readable Medium
Instructions
6222
6224
6216
Signal Generation Device
6218

FIG. 62