Expert Systems With Applications 186 (2021) 115731

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A hierarchical multi-agent architecture based on virtual identities to explain black-box personalization policies

Elvira Amador-Domínguez [a], Emilio Serrano [a,*], Daniel Manrique [b]

[a] Ontology Engineering Group, Department of Artificial Intelligence, ETSI Informáticos, Universidad Politécnica de Madrid, 28660 Madrid, Spain
[b] Artificial Intelligence Lab, Department of Artificial Intelligence, ETSI Informáticos, Universidad Politécnica de Madrid, 28660 Madrid, Spain

ARTICLE INFO

Keywords:
Multi-agent system
Virtual identities
Personalization

ABSTRACT

Hyper-personalization policies entail a considerable improvement regarding previous personalization approaches. However, they present several issues that need to be addressed, such as minimal explainability and privacy invasion. A hierarchical Multi-Agent System (MAS) is presented in this work to provide a solution to these concerns. The system is formulated as a hybrid approach, where some of the agents work autonomously, while the user input triggers the remaining. At the autonomous level, a set of Virtual Identities (VIs) representing different user profiles interact with Black-Box Hyper-Personalization Online Systems (BBHOS), gathering a set of targeted responses. Associative patterns and profile aggregations can then be inferred from the analysis of these responses. In the user-triggered level, the real user is virtualized as an identity that represents their features. The virtual identity serves as an intermediary between the personalization system and the real user. This virtualization hinders the personalization service from extracting sensitive contextual information about the real user, protecting their privacy. The results obtained by the user identity on its interaction with the personalization service are then analyzed, adjusting the content of the response to fit the user's requests instead of their features. A use case on the functioning of the analysis of search engines is presented to illustrate the complete behavior of the proposed architecture.

1. Introduction

Hyper-personalization is on the basis of current marketing approaches. While classic personalization techniques group users by general and static traits, such as age or location, hyper-personalization also considers dynamic and particular information, such as navigation data or contextual information. Therefore, the marketing experience provided by hyper-personalization techniques gives the user a sense of uniqueness, as the displayed content is specifically tailored for the customer's necessities and interests. According to Salesforce annual report (SalesForce, 2017), more than half of the customers are likely to switch brands if the company does not personalize communications to them. Moreover, the report states that efficient user personalization and segmented campaigns can increase the revenue of a company by over 760%.

Amazon is one of the prime examples of hyper-personalization. This service not only tailors its main page for the user (e.g., showing recommendations based on liked items or discounts on products that the user has previously visited), but it also feeds the user with real-time personalized notifications. With this approach, the user feels more comfortable using the service, subsequently feeling more prone to make purchases.

However, maintaining a favorable balance on the level of personalization provided to the user is not an easy task. Whereas a sound level of hyper-personalization is engaging to the user and can significantly increase the profits of a company, it is very easy to overstep the user's comfort boundaries. If the number of notifications is very elevated, or the content of these notifications is too specific, the customer may feel a sense of being under control, therefore retrieving from using the service. Hyper-personalization policies also present issues when dealing with changes, e.g., the development of a new interest or lifestyle change. Thus, if the customer suddenly develops a new interest, the personalization policy would be oblivious to this for a long time, or until enough data about the subject is generated to change the characterization of the user.

According to European Commission (2019), data protection awareness has increased during recent times. Current hyper-personalization policies are supported by the permanent extraction and exploitation of contextual data, to which the user is generally oblivious. It is essential to maintain the user informed on how their data is obtained and used

* Corresponding author.
E-mail addresses: eamador@fi.upm.es (E. Amador-Domínguez), emilioserra@fi.upm.es (E. Serrano), dmanrique@fi.upm.es (D. Manrique).

https://doi.org/10.1016/j.eswa.2021.115731
Received 15 January 2020; Received in revised form 29 June 2021; Accepted 4 August 2021
Available online 10 August 2021
0957-4174/

to raise awareness and be aligned with the existing data protection ethics. Therefore, a solution is needed that protects the user from contextual data exploitation and provides an insight into the inference process of hyper-personalization policies while still benefiting from their functionality.

This work presents a Multi-Agent System (MAS) based on virtual identities to deal with these shortcomings. Opposite to hyper-personalization systems, which usually rely on deep learning techniques and their behavior is opaque to the user, multi-agent systems are interpretable and provide insight into their inference process. The MAS framework presented in this work is capable of extracting behavior patterns from hyper-personalization systems. The benefits of this proposal are twofold. Firstly, the framework serves as an intermediary between the personalization system and the real user, thus improving its privacy. Secondly, the extraction of patterns from the hyper-personalization policy may provide an insight into its behavior, subsequently reducing its opacity.

In the proposed framework, *Virtual Identities* (VIs), representing diverse user profiles, enable the extraction of hyper-personalized content tailored to each of the proposed profiles. Thus, based on the knowledge of the different profiles and the content presented to each of them, general behavior patterns can be inferred to explain the functioning of these policies. This information can be used not only to endure the privacy over the real user but also to exploit these policies to their benefit. A specific-to-general approach is considered to present the content, leaving the user in charge of deciding whether the selected content is relevant or not. Multi-agent systems are selected for this task due to their capability of orchestrating and coordinating the interactions between agents and VIs in an autonomous way, while selecting, processing, and summarizing the most important information.

Static predefined data is combined with information extracted from social networks to obtain accurate and human-like virtual identities, giving them unique features. A set of deep learning approaches are employed to extract and analyze the retrieved content. *Deep Convolutional Neural Networks* are considered for image classification (Rawat & Wang, 2017), while for text processing tasks, such as sentiment analysis (Graves et al., 2013; Tang et al., 2015), *Recurrent Neural Networks* can be employed. VIs are then aggregated based on the similarity of the response content returned by hyper-personalization services, providing an overview of their behavior. An operation log summarizing the interactions between the VIs and the hyper-personalization system is provided to the user.

The content is organized as follows: Section 2 depicts the related works, whereas Section 3 presents the developed multi-agent system. Section 4 provides a study case for the system. Conclusions and future works are drawn in Section 5.

2. Related works

Although the idea of hyper-personalization is relatively recent, the analytical study of customer behavior, as well as the simulation of the effectiveness of marketing strategies, has been lingering for a long time. Multi-Agent Systems (Ferber, 1999) are one of the most extended paradigms for the simulation of scenarios involving complex entities, such as humans (García-Valverde et al., 2012). In the context of marketing, we can distinguish between two main objectives regarding human simulation: (i) User profile generation or user modeling (Bhowmick et al., 2010; Gao et al., 2016; Kim et al., 2014; Singh & Sharma, 2019; Xu et al., 2021) and (ii) Customer behavior simulation (Chang et al., 2018; Dehuri et al., 2008; Duarte et al., 2018; Ennaji et al., 2016; Li et al., 2019; Yang & Li, 2017; Zhang & Zhang, 2007). In the first case, the goal of the MAS is to generate user profiles that accurately characterize them to be subsequently used for different tasks. In the second case, the focus is not on the design of realistic user profiles, but on simulating and studying the responses that these profiles provide when facing certain situations, e.g., marketing campaigns or product placement strategies.

The design of user profiles that are accurate and representative, while still capable of responding to stimuli in a human-like manner, poses a considerable challenge. Ideally, a user profile should be composed of two different components: static information and context. Some traits, such as age, gender, or nationality, are generally stable throughout time, representing the core values of the user that represents its identity. However, this information is not enough to accurately define a user; hence, contextual information has to be considered for the generation of user profiles. Navigation cookies, as well as browsing histories, are examples of contextual information that are widely used for this purpose. Kim et al. (2014) propose a context-aware MAS that considers the textual information generated by the user and visual information to generate a user preference profile. Xu et al. (2021) propose an adaptive system to the user's preferences based on multi-round recommendation, while Gao et al. (2016) and Bhowmick et al. (2010) rely on ontologies to model contextual information. Gao et al. (2016), consider mobile sensor information for context, updating the generated user profile when it interacts with other profiles of similar preferences. Abu Sulayman and Ouda (2020) focus on anomaly behavior detection to automatically update the user's profile. Singh and Sharma (2019) propose a MAS framework for dynamic user profiling for web personalization that, in addition to contextual information, takes into consideration the user's interests, both short and long term. The same criteria are applied in this work to construct the VIs, where a combination of static features and contextual information is used. Social media data is employed to build the user's contextual information, as in Trifa et al. (2017).

Virtual Identities are a particular case of user modeling. While user modeling approaches aim to accurately portray the user's traits and behavior, virtual identities focus on masking certain aspects of the user's profile. Gomaa et al. (2017) study the applications of virtual identities in distributed virtual environments for network security improvement. In their work, virtual identities are created automatically and are not related to real-life users. Similarly, in Gomaa and Abd-Elrahman (2015), a new paradigm for virtual identity creation on the cloud is presented, preventing access to the user's real data. Finally, Jin (2012) studies the so-called 'virtual identity discrepancy', which explores the interactions established by users when represented by a virtual self.

MAS, beyond simulation purposes, are a widely employed paradigm to solve engineering problems that are very difficult to cover by monolithic systems. For example, in the context of marketing, they can be used for e-commerce simulation (Duarte et al., 2018; Serrano & Iglesias, 2016), development of customized product services (Chang et al., 2018), predicting the success of a particular campaign (Dehuri et al., 2008), or analyzing pricing approaches (Li et al., 2019). Other works, such as Zhang and Zhang (2007), are less goal-oriented and aim to study and analyze the customer decision-making process when a purchase is made. Similarly, Yang and Li (2017) employ a MAS to determine the ideal product arrangement depending on the particular user, while Doniec et al. (2020) focus on the behavior of in-store shoppers, considering several parameters such as the customer's walking time inside the shop. Ennaji et al. (2016) present a MAS framework for customer relationship management to extract and analyze real social media opinions.

The mentioned above approaches center their attention on the customer and their interactions. Some other works focus on the improvement and development of personalized systems based on the input and responses provided by the user. Hence, user input can be used to enhance recommendation systems (Yusof et al., 2016) and to provide personalized web content (Kazienko & Adamski, 2007).

Despite being the core element of these approaches, the direct impact of the users on their refinement and development is very minimal. Therefore, a hybrid approach is considered in this work, where human behavior simulation is conducted through VIs, but the user input directly impacts the working of the proposed system.

3. A multi-agent system based on virtual identities to explain black-box personalization online systems

Although hyper-personalization policies entail an improvement regarding previous marketing approaches, some flaws can be identified from the user's perspective:

- *Opacity:* User data is at the core of these hyper-personalization policies. With the growth of social media usage, users are now prone to post and publish personal information that can be subsequently used for their characterization. In this case, the user is fully conscious of the existence and delivering of this information. However, the majority of the data employed by hyper-personalization systems comes from mobile sensors and online activities, instead of being explicitly provided by the user. Therefore, when the user faces content selected entirely from these data sources, it can create a sense of rejection as it can be perceived as an invasion of privacy. As stated in previous sections, openness is essential as a scaffold for the user's trust in a hyper-personalization system. In the case of e-commerce services, the personalized content should be split into different categories, each one based on different types of data, for example, previous visits, purchase history, or items usually bought together. With this approach, the user has a general idea of how their data is used, giving a sense of clarity.
- *Limited adaptability:* One of the most desirable traits of hyper-personalization systems is their ability to adapt to the user's necessities dynamically. These changes are usually detected when there is a sudden shift in the data generated by the user about a particular topic. For example, if the user initiates a new activity, it must provoke an update to the profile previously established. However, there is usually a delay between the time when the user develops a new interest and when the system detects it. Hence, if the time elapsed is elevated, it will deter the user. Letting the user indicate whether a particular content is relevant or not can be a potential solution to this issue, reducing the time required by the system to adapt.
- *Lack of privacy:* Although hyper-personalization policies are based on giving a sense of uniqueness, it can be very challenging to determine the optimal level of specificity for each particular user. What it is acceptable for some users, like email notifications, personalized messages, or purchase reminders, may not be appropriate for others. Thus, what is thought to provide a sense of specialty can be felt as controlling, making the user withdraw from using that particular service. Therefore, ensuring that the user feels comfortable with the level of personalization is critical. This is also directly linked to openness since it is closely related to trust.
- *Absence of user input:* Users are the core of every hyper-personalization system. Accurately predicting their necessities and successfully fulfilling them is the objective of these policies. However, even though users are the most relevant component of these systems, the level of influence they have over the content they receive is minimal. By taking a hybrid approach, where the user actively participates in determining the content they receive, not only the user feels more comfortable with the system, but also helps increasing the accuracy of the predictions.

In order to tackle the aforementioned issues found within hyper-personalization services, a hybrid MAS based on virtual identities with direct user input is proposed. In this approach, in contrast to the works presented in Section 2, the focus is not on the user, but on the hyper-personalization services themselves. Thus, the goal is to analyze these *Black-Box Hyper-Personalization Online Systems* (BBHOS) behavior based on their responses to a set of artificially generated user profiles. These *Generated Virtual Identities* (GVIDs) emulate real users, each unique and composed of a combination of previously defined behaviors with contextual information extracted from social media data. The proposed MAS coordinates the interactions between the GVIDs and the BBHOS, gathering a set of responses that feed each of the proposed profiles. Filtering and annotating operations are then performed to extract and analyze the content from the collected untreated data. Then, hierarchical profile clustering is applied to group the GVIDs based on the similarity of the content provided by the BBHOS for each of them. The generated aggregations are oblivious to the features of the proposed user profiles. Although BBHOS are opaque, general associative patterns between the defined profiles and their associated content can be inferred. These patterns can explain, in a superficial manner, the functioning of the BBHOS.

The knowledge extracted from these patterns can then be used for two different purposes: protecting the privacy of the user and providing relevant content. In the proposed system, the real user is represented by an agent, referred to as *User Virtualized Identity*, or UVID. Instead of having the real user interact with the BBHOS, it is the UVID who makes the requests, hindering the BBHOS from extracting specific contextual information about the real user. This ensures a high degree of protection of the user's privacy. While GVIDs take a passive role, the UVID is active, as it serves as an intermediary between the system and the real user. The UVID is the only virtual identity with the capability of making certain transactions with the BBHOS.

As stated previously in this Section, even though users are at the core of BBHOS, they lack control over the content they are presented. A potential solution for this issue would be to leave the user in charge of accepting or discarding the received content without any previous analysis. However, this is not optimal, as it would hinder the adaptability of the system by relying entirely on the user's decisions. The MAS proposed in this work adopts a hybrid solution since it combines user interaction with autonomous behavior to achieve a balance between adaptability and user reliability. In addition to triggering certain transactions, the user's input also determines what content is selected and presented to them. Fig. 1 shows a high-level instantiation of the procedure followed by the proposed system for the task of airplane ticket purchasing.

This procedure comprises the following steps:

1. The real user inputs the objective, which in this example is to book a flight to a particular destination. Additional parameters, such as preferred schedules, airline, or cost, can also be specified.
2. A UVID is then generated representing the features and goals of the real user, masking its real identity, and hindering the extraction of specific contextual information.
3. The UVID makes a request to the BBHOS, which in this context is an online flight booking portal.
4. The BBHOS makes an offer based on the information known about the UVID.

    4.1 If the offer meets the real user's requirements (e.g., fits their specified budget or it is aligned with their preferred departure time), it is sent to the real user. If approved by the user, the UVID accepts the provided offer, finishing the purchase of the tickets.

    4.2 If the offer does not meet the requirements, an iterative procedure is followed until the user accepts an offer:

        4.2.1 A set of GVIDs is selected, considering the cluster with the highest similarity to the UVID as the initial set. As the aggregations are obtained through hierarchical clustering, in the initial iterations, those clusters containing elements that are more affine to the UVID are selected. The selected aggregations become more distant from the UVID as the iterations move forward.

        4.2.2 The selected GVIDs interact with the online flight booking platform, retrieving new offers.

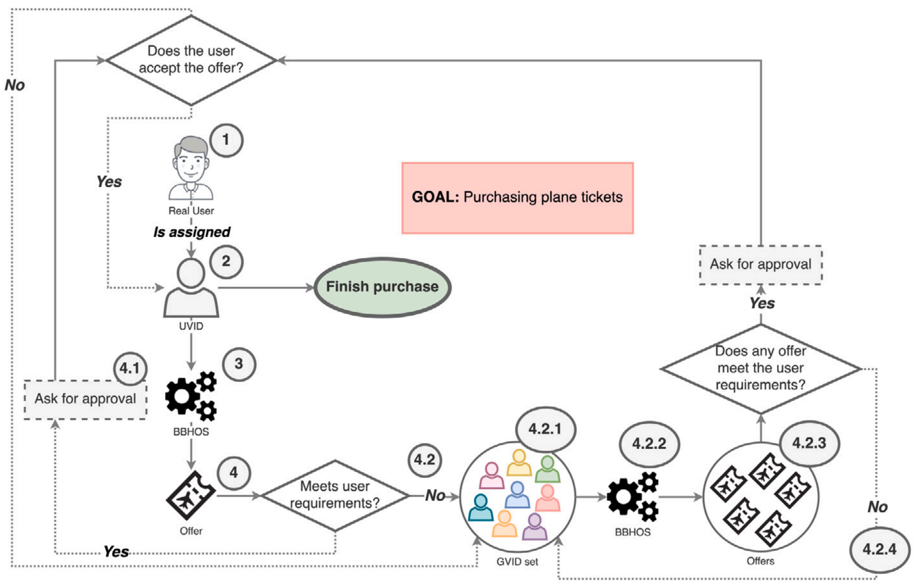

Fig. 1. Example of the procedure conducted by the system to purchase airline tickets.

4.2.3 Those offers meeting the user specifications are summarized and sent to the real user for approval.
4.2.4 If none of the offers meet the requirements, or the user declines them, the system returns to the step 4.2.1 until it reaches its finishing state.

3.1. Design of the architecture

A four-layered multi-agent system is presented to implement the proposal, as shown in Fig. 2. Each layer represents a set of functionalities conducted by the agents contained within. These agents can communicate with each other, both from the same and adjacent layers. Aside from the functionality layers, a storage layer is also considered, where the collected data can be stored for further operations. Two primary levels can be distinguished in the proposed system: an autonomous and a user-triggered level. The autonomous level of the system corresponds to layers one and two, which execute without any user interaction and prepare the system for its posterior usage. The user-triggered level, associated with layers three and four, is reliant on the user to begin its execution.

Layer 1: Data retrieval. The first layer of the proposed system comprises all operations related to the creation of the GVIDs, as well as the data extraction from the BBHOS. This layer sets the purpose and the specific objective of the system, defining how the GVIDs and the BBHOS behave. This layer involves three different agents:

1. *GVID Generating Agent:* Two different elements are considered by the agent to generate the required GVIDs to provide them with unique features and to emulate the behavior of real users realistically. First, the core elements about each of the designed GVIDs, which are stored in the system. These elements include information of mid-term static nature, such as gender, age segment, location, interests, and goals.
   Second, since this information is not enough to realistically define a user, data available from social networks is also gathered to further characterize each identity, bestowing them with the information provided directly by social-network users. This data is extracted *ad-hoc* from external data sources, and is not stored in the system for further usage. *Natural Language Processing* (NLP) techniques can be employed to extract information from this data, such as topic modeling (Alghamdi & Alfalqi, 2015) or sentiment analysis (Tang et al., 2015). Preferences, opinions, or interests can be extracted from this data, which can be then manually or randomly assigned to the specified GVIDs. The combination of these two different types of information enables GVIDs to have clearly defined features and goals while achieving a higher level of complexity with the inclusion of real, user-provided data. These GVIDs are stored for its usage on the superior layers of the system. The behavior of the GVIDs is predefined and static throughout the execution. If the GVIDs were updated, they would need to be rebuilt using new, available online data.
2. *Data Retrieval Agent:* The GVIDs interact with the selected BBHOS making requests and receiving responses aligned with their features and behavior. This agent orchestrates and supervises the interactions between the identities and services, collecting the data generated within the process. This data can have different formats depending on the selected BBHOS, ranging from plain text to video files. This data is stored for further analysis, as it can be considered to redefine the features of the GVIDs from the obtained responses.
   As BBHOS are online systems and, therefore, prone to errors, it is necessary to deal with these issues before transmitting the data to the next agent. Hence, this agent also performs operations such as duplicate or corrupted data removal.
3. *Filtering Agent:* This agent receives the data extracted from the interactions between the GVIDs and the BBHOS, and classifies it according to its format into categories like image, text, audio... Filtering and homogenization operations are required to extract the relevant content of the received data and to facilitate further data understanding procedures. Depending on the format of the data, suitable filtering operations are required, like normalization, noise filtering, or text preprocessing. This agent only takes into consideration the format of the data, without delving into its content.

Layer 2: Data understanding. Once the outputs provided by the BBHOS when interacting with the GVIDs have been collected, separated and filtered, data understanding processes are needed to extract knowledge from the gathered data. Thus, the second layer of the architecture deals with the procedures related to this task. General knowledge on the behavior of BBHOS can be inferred by understanding and analyzing the content fed to the considered profiles. Furthermore, GVDIs can be grouped based on the extracted BBHOS associative patterns. Two different agents comprise this layer:

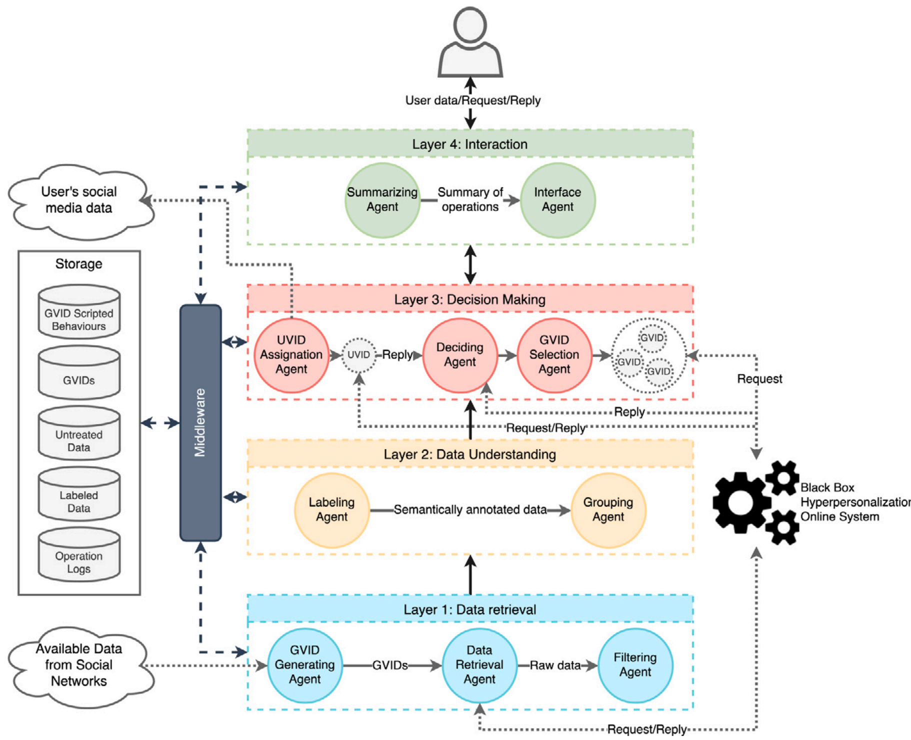

Fig. 2. Overview of the proposed Multi-Agent System. Dotted lines represent Interactions between the Agents of the system with external elements and with the storage layer. Clouds are used to represent external elements.

1. *Labeling Agent:* This agent annotates, classifies, and extracts knowledge from the collected data. As several formats may appear, suitable approaches are required to extract knowledge effectively from the gathered data. In the case of images, multi-label classification employing Deep Convolutional Neural Networks (DCNNs) can be used to determine the elements composing the images extracted from the BBHOS. State-of-the-art architectures such as ResNet (He et al., 2015) or CNN-RNN (Wang et al., 2016) can be employed for this task. DCNNs are also suitable for audio and video classification (Hershey et al., 2017). However, in some tasks, the multi-label classification may not be enough to understand the content of audio data. Speech recognition techniques (Chiu et al., 2018) can be used to convert audio to text, which enables higher-complexity operations.
   As most of the responses generated by BBHOS are textual, a more in-depth analysis is needed for this data compared to previously considered formats. Several NLP techniques can be considered to analyze the data from different perspectives. As in the formats mentioned above, text classification techniques (Mirończuk & Protasiewicz, 2018) may be suitable to categorize the retrieved content. Topic modeling approaches (Alghamdi & Alfalqi, 2015; Nallapati et al., 2011) enable finer-grained classification and analysis of the content. Finally, document clustering, as well as text summarization techniques, (Neto et al., 2000; Oikonomakou & Vazirgiannis, 2005; Yousefi-Azar & Hamey, 2017) simplify and group the retrieved data, facilitating its comprehension.
   The resulting annotated data is then stored in the system. Since most of the current state-of-the-art approaches for the techniques mentioned above rely on deep learning, periodical re-trains to consider newly retrieved data are required to maintain the models updated.
2. *Grouping Agent:* Despite the opacity of BBHOS, specific patterns can be inferred based on the content of the outputs provided for each GVID. Furthermore, relations between the GVIDs can be established based on the similarity of their displayed content. This agent establishes GVID aggregations based on this information. Hierarchical clustering is performed to enable a specific-to-general approach. This procedure ensures heterogeneity between the content of the GVID clusters in the lower levels while converging to a single aggregation in the top level. The obtained aggregations are stored for its usage by the subsequent layers of the system.

Layer 3: Decision making. At this layer, the user's data input is collected and, after being processed by the different agents, the output is sent to the interaction layer to be presented back to the user. While the previous layers followed a feed-forward structure, where data was transmitted unidirectionally to the superior level, this layer has a bidirectional flow, receiving data from both adjacent layers. This layer can be therefore considered the core element of the system, and its execution is triggered directly by the real user. This layer involves three agents:

1. *UVID Assignation Agent:* The lack of privacy is one of the most significant issues currently existing within hyper-personalization policies. This agent creates and deploys a virtual identity that represents the real user within the system, referred to as the *User Virtual Identity* (UVID) to tackle this issue, eliminating the direct interaction between the real user and the BBHOS. The same procedure conducted by the *GVID Generating Agent* is followed to obtain the UVID to enable comparison with the existing GVIDs. Therefore, in addition to the static information about the user, its publicly available social media data is considered to further analyze their interests and preferences.
2. *Deciding Agent:* Once the initial UVID is set, it interacts with the selected BBHOS. If it meets the users' specifications, the received response is then transferred onto the superior layer, where the user can specify whether it is relevant or not. If it is not, then this agent passes this information onto the *GVID Selection Agent*, which would return a set of new responses. This agent is aware of the requirements of the real user, and decides

which, if any, of the returned responses are suitable. If there are no valid responses, this agent solicits a new set of responses. This procedure is iteratively performed until the real user indicates that at least one of the provided responses is valid, or until no GVID aggregations remain.
This agent orchestrates the interactions between the GVIDs and the BBHOS, as well as collecting the information provided by the real user to send it to the following agent. The data generated by the operations performed is also stored for analysis.

3. *GVID Selection Agent:* If the response generated by the UVID is not satisfactory to the user, then new responses are required that fit the specified requirements, as shown in Fig. 1. Once this agent receives a negative response from the *Deciding Agent*, it selects a new GVID set. Since a specific to general approach is considered, similar GVID aggregations are favored in the initial iterations of the decision process. In case the user professes a strong rejection on the previous BBHOS response, the agent could choose a dissonant GVID set in the following iteration.
Several approaches could be considered to select the initial GVID set, based on measuring the similarity between the UVID with the existing GVID clusters (Rawashdeh & Ralescu, 2015; Zhang et al., 2015). These approaches range from minimal cosine distance between feature vectors to regular string comparison. However, as GVIDs are aggregated using a hierarchical clustering approach, a simplistic and efficient way to select the initial set is to apply the nearest neighbor criterion. Thus, the aggregation containing the nearest neighbor of the UVID is selected as the initial set.
Fig. 3 shows the general selection procedure. The GVID with the highest similarity with the UVID, represented inside the red dashed line box, belongs to *Cluster 1*. Thus, *Cluster 1* would be selected as the initial set. Then, the superior level of the hierarchy is explored, represented by *Cluster 4*, which is composed of *Cluster 1* and *Cluster 2*. As *Cluster 1* has already been selected, *Cluster 2* would be chosen as the next GVID aggregation. If the user does not specify a strong disagreement with the provided responses, the system follows the specified order to select the aggregations.

Layer 4: Interaction. Finally, an interaction layer is required to present the data gathered by the performed operations to the user in a simplified and precise format. The data flow in this layer is also bidirectional as the user not only receives, but introduces information. The user's input is crucial, as it directly affects the behavior of the system. The agents of this layer determine what and how the content is displayed to the user. They also collect input from the real user to send it to the lower layers.

1. *Summarizing Agent:* For the sake of openness, to develop the trust of the user in a given system, they not only need to know the output, but the operations performed by the system for its generation. This agent extracts the logs of the iterations required by the *Deciding Agent* to reach its final reply and summarizes the most relevant information. Alongside with the final response, this agent generates a simple, user-readable record on which virtual identity obtained each response, and how many iterations were required. Thus, the real user can be conscious of whether the responses have been obtained employing their virtual identity (UVID) or whether a different profile has obtained them.
2. *Interface Agent:* User input is key to the behavior of the system on its superior level. Thus, appropriate mechanisms are required to interact with the real user to fulfill their requirements effectively. The communication with the user is bidirectional, implying that the system not only has to collect the user's input and make it readable by the system but also has to present the output in a clear, readable manner. Therefore, this agent serves as an intermediary between the real user and the system. It collects the user's input and transmits it to the lower layer for processing. Besides, this agent receives the output generated by the *Summarizing Agent*, containing the final BBHOS response and the operation logs of the procedure. This content is displayed to the real user in a simplified and visual format, employing adequate layout techniques.

Storage units. As shown in Fig. 2, storage units serve as a sharing point for all layers of the architecture. The data generated during the execution of the system has a heterogeneous nature, coming from different sources in different formats. Five databases are needed to suit the particularities and different types of data generated within the system:

1. *GVID Scripted Behaviors:* A set of unique features are required to create a set of diverse user profiles so that the behavior of BBHOS can be analyzed. These features refer to mid-term information, such as interests, age, ..., which barely changes throughout the execution of the system.
2. *Untreated Data:* This unit stores data extracted from the interactions between the GVIDs and the BBHOS in the first stages of execution. The responses obtained from the BBHOS system can be used to redefine the behavior of the GVIDs.
3. *Labeled Data:* Once the untreated data has been separated according to its format, filtered, and annotated; it is stored in this database. This data is used to periodically retrain the models within the *Labeling Agent* to maintain them updated.
4. *GVIDs:* As GVIDs are at the core of the system and, therefore, used by multiple agents of the system across different layers, it is critical to store them and ensure their persistence and accessibility. As well as the single GVIDs generated within the first layer of the system, the aggregations obtained employing hierarchical clustering are also stored within this storage unit. If there is an update on the GVIDs, it triggers a reaction to the *GVID Generating Agent*, initiating the execution of the autonomous level.
5. *Operation logs:* A log of the communications between the GVIDs and the selected BBHOS is saved. The data generated from these interactions allows the extraction of associative patterns that enable understanding over the BBHOS. These logs are created in the third layer and read in the fourth for their subsequent summarizing.

There is a *middleware* between the databases and the agents that orchestrates and supervises the introduction and extraction of data from the corresponding databases. The middleware ensures the persistence of the data and avoids concurrency issues.

3.2. Data flow

Fig. 4 presents the data flow associated with the proposed architecture. In the initial stage, the specified GVIDs are generated from a combination of scripted behaviors and existing social media data from users. The set of generated GVIDs is stored in its corresponding database for its subsequent usage. These GVIDs then interact with the selected BBHOS, generating a set of untreated, or "raw" data. This data comes in different formats, such as audio, video, or HTML text. Untreated data is also stored in its corresponding database. This data is then analyzed to adjust the behavior of the GVIDs to improve the performance of the system. For example, if two GVIDs receive the same responses, it may imply that there are redundancies within the predefined behaviors. Thus, a revision is required.

At this stage, the data gathered is untreated, containing both relevant and irrelevant information. Since relevant information must be extracted for its subsequent annotation and analysis, the data is first

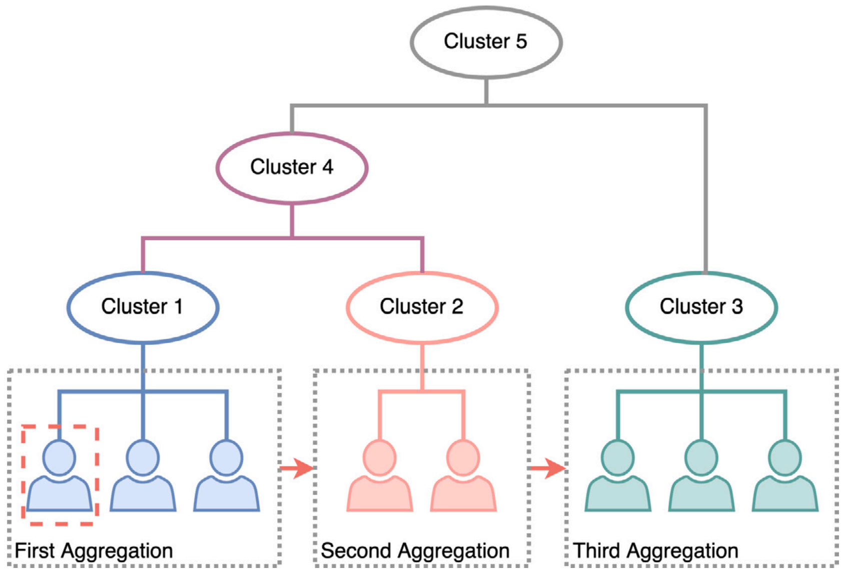

Fig. 3. Example of the general GVID selection procedure. Dotted lines delimit the clusters. Different colors are used to denote the elements of each aggregation. The red dashed line box indicates the GVID with the highest similarity to the UVID.

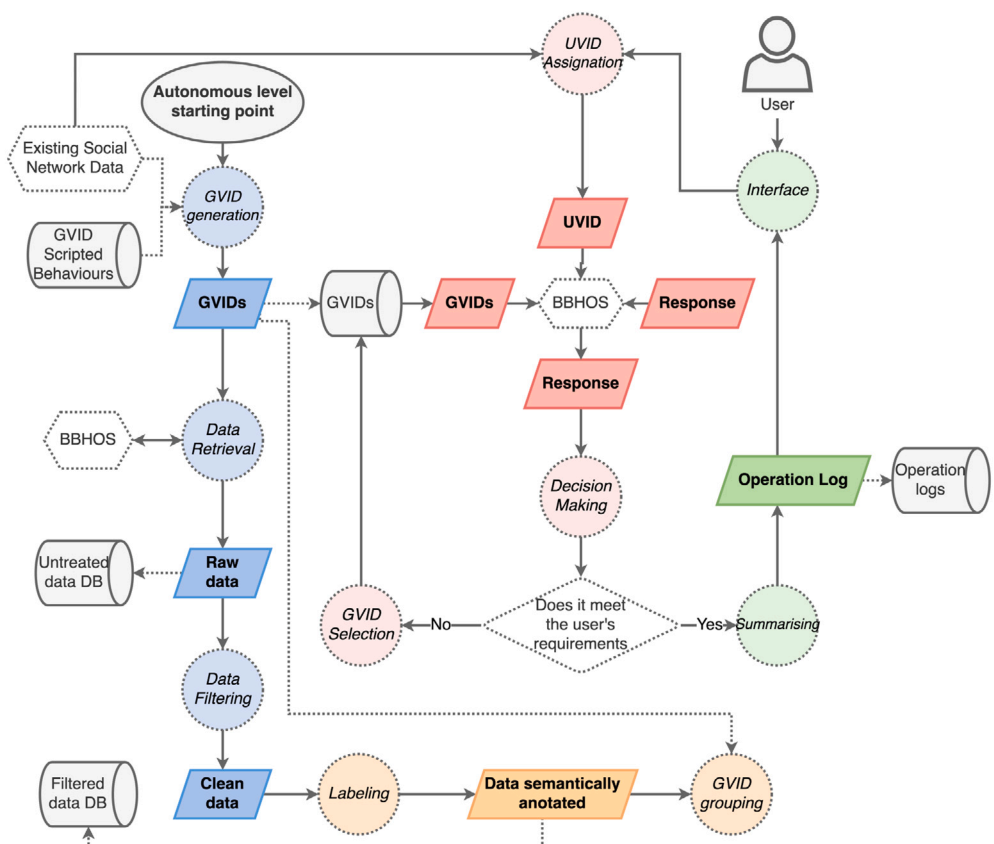

Fig. 4. Data flow of the presented system. Elements in dotted lines represent procedures (circles) and external data sources (hexagons). Elements in continuous lines represent internal data transfers (trapezoids) and data storage units (cylinders). The following color code is used to associate each stage of the data flow with its corresponding layer on the MAS: blue for the first layer, orange for the second layer, red for the third layer, and green for the fourth layer.

classified according to its format into video, audio, image, or text. Then, appropriate filtering and homogenization operations are performed over the data depending on its format: in the case of images, normalization and resizing operations are conducted while in the case of text, stopword removal or markup text elimination is performed. Filtering and cleaning the data not only facilitates its subsequent annotation by labeling models but also considerably reduces its size, which positively affects the performance. The models in the *Labeling Agent* then annotate the filtered data and subsequently store it in its corresponding database. Then, hierarchical clustering is performed to aggregate the GVIDs according to the content of the responses provided to each of them. The obtained aggregations are stored into the GVID database for its posterior usage.

As mentioned in Section 3.1, the system includes an automatic and a user-triggered level. The operations mentioned above, corresponding to the first two layers, are executed autonomously. Then, the system remains suspended until the real user reignites it. The remaining operations, corresponding to layers three and four, have the user as the starting point. The real user accesses the system via an interface with a particular goal. Once the real user is logged into the system, its UVID is generated, acting as their intermediary. The UVID interacts with the targeted BBHOS, generating a tailored response. If this response meets the user requirements, it is then summarized and sent back to the real user via the interface. On the contrary, if the user rejects the provided response, or it is not aligned with their requirements, a set of GVIDs is selected, repeating the same request to the BBHOS and obtaining a new set of responses. Once the user agrees with the provided response, the system is suspended again.

4. Case study

A real-life case study is provided to illustrate the proposed architecture. The presented MAS is composed of two differentiated levels: an autonomous and a user-triggered level. Fig. 1 illustrates an example of usage of the system focusing on the procedures involved in the superior level. The case study presented in this Section provides a clearer and more-in-depth insight on the interactions and the data transfer between the agents to give a complete overview on the behavior of the system both at the autonomous and the user-triggered level.

The proposed case study focuses on the analysis of the behavior of search engines, more specifically on the relevance of the entries on the result page. From a user perspective, the results that best fit the input query should be featured on the initial result page, independently from the employed search engine. However, this statement is rarely met, as the results for the same search not only vary between engines but also between users. Moreover, most search engines feature advertising content amongst the top results, even though this content is generally unrelated to the given query. Thus, this content needs to be detected and replaced with valid result entries.

The considered case scenario meets the following specifications:

- The goal of the system is to detect, analyze, and optimize result entries in search engine results.
- Multiple search engines are selected, which play the role of the BBHOS.
- A reduced set of GVIDs is considered. Each GVID has unique features for the generated results to be heterogeneous.
- The system determines which results are related to the real user's request, discarding those entries that contain unrelated content, such as advertisements.

As the proposed architecture is composed of two disjoint levels, the interactions between the agents on each level are presented separately. Fig. 5 illustrates the presented example. Operations related to the autonomous layer are numbered using Roman numerals, while Arabic numerals are used for the user-triggered level. As the operations of the *Deciding Agent* can lead to two different scenarios, letters *a* and *b* are used to designate the order in which the agents are executed.

Autonomous level. In this level, the system gathers data from the selected BBHOS, which the agents subsequently process. The workflow is as follows:

1. The *GVID Generating Agent* collects the available social network user data dumps and extracts a set of N GVID behaviors from the database. The collected social network data is shuffled and randomly distributed into N different sets, each containing information such as tweets, reviews, or opinions. Sentiment analysis (Devlin et al., 2019) and topic modeling (Moody, 2016) are then performed over the sets to extract interests and opinions. This unique information is combined with the GVID scripted behavior, generating the final GVID set. The generated GVIDs are stored in the database.
2. Once the GVIDs have been generated, the *Data Retrieval Agent* selects a set of search engines to retrieve information from, as well as defining the browsing terms to be browsed. The browsing term set is composed of the topics extracted in the previous stage. Each GVID does the same number of searches, and multiple GVIDs consult the same term across the different engines. This enables the analysis of the results returned when the same term is browsed on different search engines and those obtained by distinct GVIDs when both the search term and the engine are the same. The collected responses are subsequently stored in the database in the form *(entry, GVID_id, search term, search engine)*.
3. The result pages gathered by the GVIDs are then passed to the *Filtering Agent*, which performs cleaning operations on the data. The collected data is in the form of HTML webpages, thus containing non-significant elements such as markup text, references, or symbols. The following operations are performed to extract the fresh content of the gathered responses: entry extraction, HTML and JavaScript markup text removal, empty word removal, and non-alphanumeric character removal. A set of clean result entries is obtained as an outcome of these procedures.
4. As stated early in this Section, search engines tend to merge advertising and real result entries. Hence, these invalid entries need to be detected and discarded to present the real user only with those entries representing valid search results. The *Labeling Agent* receives the result entry set from the previous agent and classifies them into two disjoint categories: advertisements and real results. This process can be seen as a binary text classification problem, thus can be approached with state-of-the-art models for the task such as FastText (Joulin et al., 2017) or BERT (Devlin et al., 2019). The combination of the previous agent filtering procedures with the *FastText* classification module achieves a classification accuracy of almost 90% (Amador-Domínguez et al., 2019). Labeled result entries are stored in the database following the format *(entry, label, confidence, GVID_id, search term, search engine)*.
5. Finally, GVIDs are hierarchically aggregated by their result entries to analyze the behavior of the studied search engines. The *Grouping Agent* receives the labeled entries in the format mentioned above and performs hierarchical clustering over them. As a result, GVIDs are grouped based on the result entries returned for each of them by the different engines. Relevant information about the considered search engines can then be inferred, such as which profiles are the most prone to receive advertising content, which engine includes the least amount of irrelevant content amongst its entries, or whether the results provided by the same engine on the same search vary between different GVIDs. The aggregations are stored in the GVIDs database.

User-triggered level. Once the autonomous level finishes its execution, the system remains suspended until a user makes a request. In this example, the request is a search query that the real user types in to a search engine. When the user makes the request, the user-triggered level of the system initiates its execution:

1. The *Interface Agent* collects the real user's search query, alongside with the necessary user information. The input is sent to the following agent for further processing.

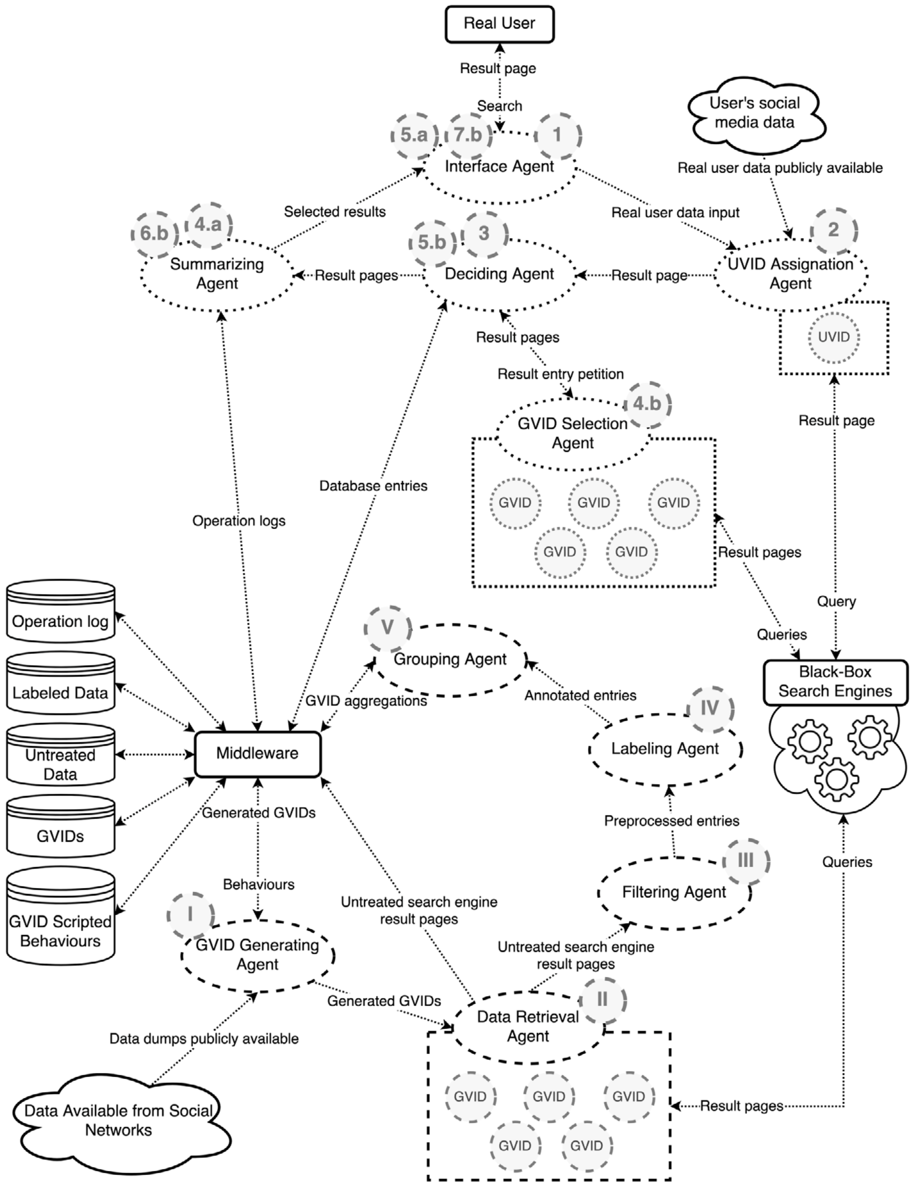

Fig. 5. Use case of the presented architecture. Agents in dashed lines belong to the autonomous level, while dotted lines represent agents in the user-triggered level.

2. As presented in Section 3, one of the goals of the system is to protect the user's sensitive information, for which a virtual artificial identity is employed. The virtual identity, or UVID, is generated from the contextual information about the real user, alongside with their publicly available social media data. The procedure followed to generate the UVID is identical to the one employed by the *GVID Generating Agent* to build the GVIDs. The UVID is the element making the required search to hinder the search engine from extracting contextual information about the user that could lead to inappropriate responses.
3. The UVID performs the search requested by the real user on the search engine. Then, the *Deciding Agent* receives the result page and determines whether the entries returned by the engine fit the real user's search. Result entries are extracted from the result page to analyze their validity and filtered following the same procedure from Stage III. Then, the stored binary classification model is employed to determine whether there is advertising content in any of the collected entries. At this point, two different scenarios can occur based on whether the result entries are real results (*Scenario A*) or not (*Scenario B*).

Scenario A.

4.a If all results provided by the engine to the UVID search are relevant, implying that no commercial content has been detected, the *Summarizing Agent* compiles them for the user, e.g., removes duplicate entries or result entries that refer to different paths on the same site.

5.a Finally, the *Interface Agent* builds a result page from the final result entry set and displays it to the real user. The constructed result page is formatted employing an adequate layout to present the information to the real user in a clear, understandable manner.

Scenario B.

4.b If there is unrelated content amongst the results, the *Deciding Agent* discards those entries and makes a petition to the *GVID Selection Agent* to obtain new, valid entries that replace the discarded ones. A set of GVIDs is selected following the selection criteria exposed in Section 3, which repeat the same query over the same search engine than the UVID. The obtained result pages are sent back to the *Deciding Agent*.

5.b The new set of result entries obtained by the GVIDs is analyzed by the *Deciding Agent*, following the procedure presented in Stage 3. Then, those results with the highest similarity to the UVID are selected to replace the discarded entries. If no relevant results are found, the agent sends a new petition to the *GVID Selection Agent*, repeating the process until valid result entries are obtained.
6.b The constructed result entry set, composed by the results obtained by the UVID and the GVID set, is transferred to the *Summarizing Agent*. In addition to summarizing the result entries, this agent also generates a log comprising the operations required to obtain the final result entry set, identifying the virtual identity that originated each result.
7.b The *Interface Agent* composes a result page from the result entry set and displays it to the real user, finishing the execution of the system.

5. Conclusions

This work presents a Multi-Agent System (MAS) architecture with virtual identities to analyze and explain black-box personalization services. Although Black-Box Hyper-Personalization Online Systems (BBHOS) entail an improvement regarding previous personalization services, they still present significant flaws that need to be appointed. A diminished sensation of privacy and the absence of user input are some of the most prominent concerns existing within these services.

The proposed MAS architecture extracts behavior patterns from black-box personalization policies that can explain their functioning and can be exploited to solve the aforementioned issues. A set of virtual identities (VIs) representing different user profiles is employed for pattern extraction. A particular virtual identity conveying the user (UVID) is deployed to protect the user's privacy. The UVID serves as an intermediary and a firewall between the BBHOS and the user, executing the operations required to fulfill the request. In the proposed architecture, VIs encode both static user traits, such as age or location, alongside dynamic features, such as user's interests. VIs interact with hyper-personalization services, receiving tailored responses. The content of the responses is analyzed and semantically annotated with *Natural Language Processing* techniques that best fit the considered BBHOS. VIs are then hierarchically aggregated based on the similarity of the content displayed on each of them, subsequently extracting associative patterns between the profiles and the responses they receive. These patterns can then be exploited to protect the user's sensitive data while still fulfilling their requirements. Moreover, the extracted patterns can also provide an insight into the criteria followed by the BBHOS when displaying certain content to a user, reducing the opacity of these systems.

A use case is presented to instantiate the presented approach, providing an example of the system for the analysis of search engines. In the proposed use case, the MAS analyzes how the results for the same search query vary across different search engines, as well as whether or not the results change depending on the user profile. The information extracted on the subject is then used to provide optimized search results, detecting and substituting the result entries that contain inappropriate information, such as advertising content or duplicates.

Our main future work is the implementation of the architecture proposed by instantiating its main components with some of the alternative technologies detailed in this work. In some of our previous works (Amador-Domínguez et al., 2019), we implement agents with simple VIs which are capable of detecting and retrieving semantics from online advertising by several methods such as Deep Learning and Expert Systems. Another important future work is the generation of agents with more complex Generated Virtual Identities (GVIDs), capable of automatically tuning their predefined behaviors from their gathered responses. Thus, these agents would be capable of detecting whether the responses obtained are heterogeneous and fitting enough for the designated goal. For example, if agents with two different GVIDs detect that their responses are the same, one of them could request a redefinition to the GVID Generating Agent, receiving new features that would lead to different responses.

CRediT authorship contribution statement

Elvira Amador-Domínguez: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Emilio Serrano:** Conceptualization, Methodology, Supervision, Writing – reviewing & editing. **Daniel Manrique:** Conceptualization, Methodology, Supervision, Writing – reviewing & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the reviewers and editors for their valuable comments and suggestions, which have improved this paper. This research work is supported by the "Comunidad de Madrid", under the programs "Ayudas para el fomento de la innovación tecnológica en sectores tecnológicos de la Comunidad de Madrid, en el marco de la Estrategia Regional de Investigación e Innovación para una Especialización Inteligente (RIS3)"; and, by the "Universidad Politécnica de Madrid" under the program "Ayudas para Contratos Predoctorales para la Realización del Doctorado".

References

Abu Sulayman, I. I. M., & Ouda, A. (2020). Designing security user profiles via anomaly detection for user authentication. In *2020 International symposium on networks, computers and communications* (pp. 1–6). http://dx.doi.org/10.1109/ISNCC49221.2020.9297252.

Alghamdi, R., & Alfalqi, K. (2015). A survey of topic modeling in text mining. *International Journal of Advanced Computer Science and Applications*, *6*, http://dx.doi.org/10.14569/IJACSA.2015.060121.

Amador-Domínguez, E., Serrano, E., Mateos-Nobre, J. D., & Ayala-Muñoz, A. (2019). An intelligent and autoadaptive system of virtual identities based on deep learning for the analysis of online advertising networks. In *Highlights of practical applications of survivable agents and multi-agent systems. the PAAMS collection - international workshops of proceedings* (pp. 26–28). http://dx.doi.org/10.1007/978-3-030-24299-2_26.

Bhowmick, P. K., Sarkar, S., & Basu, A. (2010). Ontology based user modeling for personalized information access. *International Journal on Computational Science and Applications*, *7*, 1–22.

Chang, C.-Y., Tsai, C.-H., & Chu, M.-C. (2018). Construct a customized product service system utilizing multi-agent system. In I. Moon, G.M. Lee, J. Park, D. Kiritsis, & G. von Cieminski (Eds.), Advances in production management systems. production management for data-driven, intelligent, collaborative, and sustainable manufacturing (pp. 193–200).

Chiu, C.-C., Sainath, T. N., Wu, Y., Prabhavalkar, R., Nguyen, P., Chen, Z., Kannan, A., Weiss, R. J., Rao, K., Gonina, E., Jaitly, N., Li, B., Chorowski, J., & Bacchiani, M. (2018). State-of-the-art speech recognition with sequence-to-sequence models. In *2018 IEEE international conference on acoustics, speech and signal processing* (pp. 4774–4778). http://dx.doi.org/10.1109/ICASSP.2018.8462105.

Dehuri, S., Cho, S., & Jagadev, A. K. (2008). Honey bee behavior: A multi-agent approach for multiple campaigns assignment problem. In *2008 international conference on information technology* (pp. 24–29). http://dx.doi.org/10.1109/ICIT.2008.14.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the north American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171–4186). Minneapolis, Minnesota: Association for Computational Linguistics, http://dx.doi.org/10.18653/V1/N19-1423.

Doniec, A., Lecoeuche, S., Mandiau, R., & Sylvain, A. (2020). Purchase intention-based agent for customer behaviours. *Information Sciences*, *521*, 380–397. http://dx.doi.org/10.1016/j.ins.2020.02.054, URL https://www.sciencedirect.com/science/article/pii/S0020025520301377.

Duarte, D., Ferreira, H. S., Dias, J. P., & Kokkinogenis, Z. (2018). Towards a framework for agent-based simulation of user behaviour in e-commerce context. In *Trends in cyber-physical multi-agent systems. the PAAMS collection - 15th international conference* (pp. 30–38).

Ennaji, F. Z., El Fazziki, A., El Alaouiel Abdallaoui, H., Sadiq, A., Sadgal, M., & Benslimane, D. (2016). Multi-agent framework for social crm: Extracting and analyzing opinions. In *2016 IEEE/ACS 13th international conference of computer systems and applications* (pp. 1–8). http://dx.doi.org/10.1109/AICCSA.2016.7945700.

European Commission (2019). *Communication from the commission to the european parliament and the council*. The European Commission.

Ferber, J. (1999). *Multi-agent systems: an introduction to distributed artificial intelligence* (1st ed.). Addison-Wesley Longman Publishing Co. Inc.

Gao, Q., Fu, D., & Dong, X. (2016). A context-aware mobile user behavior-based neighbor finding approach for preference profile construction. *Sensors*, *16*(1230), http://dx.doi.org/10.3390/s16081230.

García-Valverde, T., Campuzano, F., Serrano, E., Villa, A., & Botía, J. A. (2012). Simulation of human behaviours for the validation of ambient intelligence services: A methodological approach. *Journal of Ambient Intelligence and Smart Environments*, *4*, 163–181. http://dx.doi.org/10.3233/AIS-2012-0147.

Gomaa, I. A., & Abd-Elrahman, E. (2015). A novel virtual identity implementation for anonymous communication in cloud environments. *Procedia Computer Science*, *63*, 32–39. http://dx.doi.org/10.1016/j.procs.2015.08.309.

Gomaa, I., Said, A. M., Abd-Elrahman, E., Hamdy, A., & Saad, E. M. (2017). Performance evaluation of virtual identity approaches for anonymous communication in distributed environments. *Procedia Computer Science*, *109*, 710–717. http://dx.doi.org/10.1016/j.procs.2017.05.382.

Graves, A., Mohamed, A., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In *2013 IEEE international conference on acoustics, speech and signal processing* (pp. 6645–6649). http://dx.doi.org/10.1109/ICASSP.2013.6638947.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. In *IEEE conference on computer vision and pattern recognition* (pp. 770–778).

Hershey, S., Chaudhuri, S., Ellis, D. P. W., Gemmeke, J. F., Jansen, A., Moore, R. C., Plakal, M., Platt, D., Saurous, R. A., Seybold, B., Slaney, M., Weiss, R. J., & Wilson, K. (2017). Cnn architectures for large-scale audio classification. In *2017 IEEE international conference on acoustics, speech and signal processing* (pp. 131–135). http://dx.doi.org/10.1109/ICASSP.2017.7952132.

Jin, S.-A. A. (2012). The virtual malleable self and the virtual identity discrepancy model: Investigative frameworks for virtual possible selves and others in avatar-based identity construction and social interaction. *Computers in Human Behavior*, *28*, 2160–2168. http://dx.doi.org/10.1016/j.chb.2012.06.022, URL https://www.sciencedirect.com/science/article/pii/S0747563212001744.

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2017). Bag of tricks for efficient text classification. In *Proceedings of the 15th conference of the European chapter of the association for computational linguistics: volume 2, short papers* (pp. 427–431). Valencia, Spain: Association for Computational Linguistics.

Kazienko, P., & Adamski, M. (2007). Adrosa—adaptive personalization of web advertising. *Information Sciences*, *177*, 2269–2295. http://dx.doi.org/10.1016/j.ins.2007.01.002.

Kim, J., Gao, Q., & Cho, Y. (2014). A context-awareness modeling user profile construction method for personalized information retrieval system. *International Journal of Fuzzy Logic and Intelligent Systems*, *14*, 122–129. http://dx.doi.org/10.5391/IJFIS.2014.14.2.122.

Li, F., ting Du, T. C., & Wei, Y. (2019). Offensive pricing strategies for online platforms. *International Journal of Production Economics*, *216*, 287–304. http://dx.doi.org/10.1016/j.ijpe.2019.06.009.

Mirończuk, M. M., & Protasiewicz, J. (2018). A recent overview of the state-of-the-art elements of text classification. *Expert Systems with Applications*, *106*, 36–54. http://dx.doi.org/10.1016/j.eswa.2018.03.058.

Moody, C. E. (2016). Mixing dirichlet topic models and word embeddings to make lda2vec. arXiv:1605.02019, CoRR abs/1605.02019. URL http://arxiv.org/abs/1605.02019.

Nallapati, R., McFarland, D., & Manning, C. (2011). Topicflow model: Unsupervised learning of topic-specific influences of hyperlinked documents. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics (vol. 15)* (pp. 543–551). Lauderdale, FL, USA: PMLR.

Neto, J. L., Santos, A. D., Kaestner, C. A., Alexandre, N., Santos, D., A, C. A., Alex, K., Freitas, A. A., & Parana, C. (2000). Document clustering and text summarization.

Oikonomakou, N., & Vazirgiannis, M. (2005). A review of web document clustering approaches. In O. Maimon, & L. Rokach (Eds.), *Data mining and knowledge discovery handbook* (pp. 921–943). Boston, MA: Springer US, http://dx.doi.org/10.1007/0-387-25465-X_43.

Rawashdeh, A., & Ralescu, A. L. (2015). Similarity measure for social networks - a brief survey. In *Modern artificial intelligence and cognitive sciences conference*.

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, *29*, 2352–2449.

SalesForce (2017). Annual report. Retrieved from https://c1.sfdcstatic.com/content/dam/web/en_us/www/assets/pdf/datasheets/salesforce-research-fourth-annual-state-of-marketing.pdf. (Accessed 30 July 2019).

Serrano, E., & Iglesias, C. A. (2016). Validating viral marketing strategies in twitter via agent-based social simulation. *Expert Systems with Applications*, *50*, 140–150. http://dx.doi.org/10.1016/j.eswa.2015.12.021.

Singh, A., & Sharma, A. (2019). A multi-agent framework for context-aware dynamic user profiling for web personalization. In M. N. Hoda, N. Chauhan, S. M. K. Quadri, & P. R. Srivastava (Eds.), *Software engineering* (pp. 1–16). Singapore: Springer Singapore.

Tang, D., Qin, B., & Liu, T. (2015). Document modeling with gated recurrent neural network for sentiment classification. In *Proceedings of the 2015 conference on empirical methods in natural language processing* (pp. 1422–1432). http://dx.doi.org/10.18653/v1/D15-1167.

Trifa, A., Sbai, A. H., & Chaari, W. L. (2017). Enhancing assessment of personalized multi-agent system through convlstm. In *Knowledge-based and intelligent information & engineering systems: proceedings of the 21st international conference Procedia Computer Science*, *112*, 249–259. http://dx.doi.org/10.1016/j.procs.2017.08.239.

Wang, J., Yang, Y., Mao, J., Huang, Z., Huang, C., & Xu, W. (2016). Cnn-rnn: A unified framework for multi-label image classification. In *2016 IEEE conference on computer vision and pattern recognition* (pp. 2285–2294). http://dx.doi.org/10.1109/CVPR.2016.251.

Xu, K., Yang, J., Xu, J., Gao, S., Guo, J., & Wen, J.-R. (2021). Adapting user preference to online feedback in multi-round conversational recommendation. In *Proceedings of the 14th ACM international conference on web search and data mining* (pp. 364–372). New York, NY, USA: Association for Computing Machinery, http://dx.doi.org/10.1145/3437963.3441791.

Yang, S., & Li, L. (2017). Personalization in dynamic assortment planning: An analysis based on multi-agent simulation method. In *2017 5th international conference on enterprise systems* (pp. 157–162). http://dx.doi.org/10.1109/ES.2017.32.

Yousefi-Azar, M., & Hamey, L. (2017). Text summarization using unsupervised deep learning. *Expert Systems with Applications*, *68*, 93–105. http://dx.doi.org/10.1016/j.eswa.2016.10.017.

Yusof, A., Mahmoud, M. A., & Ahmad, M. S. (2016). A conceptual multi-agent semantic web model of a self-adaptive website for intelligent strategic marketing in learning institutions. In *2016 2nd international symposium on agent, multi-agent systems and robotics* (pp. 133–138). http://dx.doi.org/10.1109/ISAMSR.2016.7810016.

Zhang, T., & Zhang, D. (2007). Agent-based simulation of consumer purchase decision-making and the decoy effect. *Journal of Business Research*, *60*, 912–922. http://dx.doi.org/10.1016/j.jbusres.2007.02.006.

Zhang, S., Zheng, X., & Hu, C. (2015). A survey of semantic similarity and its application to social network analysis. In *2015 IEEE international conference on big data* (pp. 2362–2367).