Fuzzy-Based Conversational Recommender for Data-intensive Science Gateway Applications

Arjun Ankathatti Chandrashekara, Radha Krishna Murthy Talluri, Sai Swathi Sivarathri, Reshmi Mitra, Prasad Calyam, Kerk Kee, and Satish Nair
University of Missouri-Columbia, USA; Webster University, USA; Chapman University, USA
{*aacwb, rtm4x, sss26x*} *@mail.missouri.edu, reshmimitra25@webster.edu,*
calyamp@missouri.edu, kee@chapman.edu, nairs@missouri.edu

***Abstract*—Neuro-scientists are increasingly relying on parallel and distributed computing resources for analysis and visualization of their neuron simulations. Although science gateways have democratized relevant high performance/throughput resources, users require expert knowledge about programming and infrastructure configuration that is beyond the repertoire of most neuroscience programs. These factors become deterrents for the successful adoption and the ultimate diffusion (i.e., systemic spread) of science gateways in the neuroscience community. In this paper, we present a novel intuitionistic fuzzy logic based conversational recommender that can provide guidance to users when using science gateways for research and education workflows. The users interact with a context-aware chatbot that is embedded within custom web-portals to obtain simulation tools/resources to accomplish their goals. In order to ensure user goals are met, the chatbot profiles a user's cyberinfrastructure and neuroscience domain proficiency level using a 'usability quadrant' approach. Simulation of user queries for an exemplary neuroscience use case demonstrates that our chatbot can provide step-by-step navigational support and generate distinct responses based on user proficiency.**

***Index Terms*—Conversational Recommenders, Intutionistic Fuzzy Logic, Mamdani Inference, Neuroscience Workflows, Science Gateways, Virtual Agents, Guided User Interfaces.**

I. Introduction

Research and training in neural science and engineering are increasingly dealing with the analysis of voluminous multi-parameter data [1]. The NSF iNeuro report [2] states these issues as access to multi-omics data archives, heterogeneous software and computing resources, and multi-site interdisciplinary expertise, e.g., engineering, biology, and psychology. Existing parallel and distributed computing resources available in the form of cyberinfrastructures (CIs) enable multiple overlapping steps: access, collect, clean, model, analyze, visualize and communicate such data [3]. However, to completely utilize their full capabilities, the neuroscientists with limited CI skills are required to invest in training with High-Performance Computing (HPC)/High Throughput Computing (HTC), Big data and related tools, instead of focusing on bold experiments for knowledge discovery in neuroscience.

This material is based upon work supported by the National Science Foundation under awards: ACI-1440582, OAC-1730655, OAC-1322305 and ACI-1453864. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

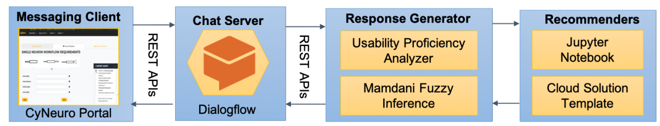

Fig. 1. Overview of Web-Portal and fuzzy based Chatbot Integration with the Underlying recommender systems

In addition to these challenges, other factors such as usability, reproducibility, automation, and distributed cognition [4] have been significant barriers for use of science gateways (SG) in computational neuroscience. There is a bigger challenge of usability and user support since all users cannot fall on the same domain proficiency level. In fact, the productivity of novice or even expert user may suffer as they begin to explore the gateway capabilities with the wide variety of tools and their features. SG user support is time-intensive, often requiring one-on-one help and training [5], [6]. New users often have research needs and/or usage assumptions that are different from the owners and developers of the CI/SG tools. Therefore, new users usually require personalized support to become proficient with the SG, before they are able to use it to conduct their own research independently. Voluntary user support by SG projects when the technologies are in early diffusion cycles may be feasible. Once the SG begins to gain traction in the user community, personalized user support becomes practically impossible. Scaling to keep up with adoption and providing expert service support has been difficult in practice.

In this paper, we address the above challenges by proposing an intuitionistic fuzzy logic based conversational recommender for the exemplar neuroscience use case in a next-generation SG as shown in Fig. 1. The conversational recommender supplements SG by providing support across novice/expert scientific community to conduct and debug repeatable computational workflow experiments. We name our context-aware chatbot "Vidura", after the wise adviser in Indian mythology, as a metaphor of the chatbot is an expert guide for CI/SG users to navigate through complex tasks. The user will interact with the "Vidura" to communicate their intent and proficiency level on the SG web portal. Based on their interactions, the chatbot uses usability quadrant analyzer to profile the user HPC and neuroscience domain proficiency and uses Mamdani inference system to generate recommendations. These recommendations include either guided user interface [5] or a self-service training platform supported by Jupyter Notebooks within HPC/HTC infrastructures [7] or a CIPRES workflow manager [8].

The novelty of our chatbot approach is the ability to use relevant Jupyter Notebook or CI template recommender systems as shown in Fig. 1 to provide user support for diverse SG actions. For instance, the chatbot uses an underlying resource configuration template recommender [11] for all the CI related queries, to suggest template options along with configuration information such as compute, network and memory resources. Similarly, we also use an underlying Jupyter Notebook Recommender and also a Publication Recommender which suggest appropriate Jupyter Notebooks and related research publications, respectively. Our chatbot implementation is built using Dialogflow, which is a conversational artificial intelligence service from Google. It performs the natural language processing to find the intent and entities in the user queries. We use Dialogflow API feature called 'fulfillment' to invoke external REST API calls to our respective conversational recommender modules which in turn uses underlying recommendation systems.

We evaluate our conversational recommender with simulated user interactions for a neuron use case, which involves building computational models for neuron single cell using the NEURON Python module [2]. To measure the reliability and consistency of our user quadrant questionnaire, we calculate its Cronbach's alpha score. Once the user proficiency is established, we simulate a series of user queries for the four different usability quadrants related to the aforementioned use case and record the responses from our Vidura chatbot. We describe experiments to study how Vidura can improve the ability of novice/expert users to effectively express their intents and subsequently access appropriate CI/SG resources and tools. Thus, our work can improve the usability, automation, and productivity for the novice/expert users, while increasing diffusion and adoption of CI/SG.

The remainder of the paper is organized as follows: Section II presents related work. Section III details the design of our fuzzy logic based conversational recommender. Section IV presents the chatbot architecture evaluation of responses generated by chatbot for the neuron simulation use case. Section V concludes the paper.

II. Related Work

A. User Profile and Recommendation System

The quality of recommendations improves with the advancement in artificial intelligence (AI) and storing user-specific profile information. Many e-commerce websites solicit user interest data to match user needs with their products to advance their sales. Studies such as [13] and [14] demonstrate this point and provide insights into the contents of the user profile. In e-learning platforms, the user proficiency is one of the most significant factors that influences the absorption of learning content [15]. In [15], a questionnaire related to domain knowledge is provided to the student to assess their knowledge level and Mamdani approach is used to decide how much content should be presented to a student according to his/her intellectual level. In this work, we leverage the user profile to provide sustainable and scalable support in SG for neuroscience workflows. In our case, the profile information is best represented by their HPC and (neuroscience) domain proficiency.

B. Automating Data Analysis with Chatbots

Recently, natural language processing has advanced to an extent that industry is widely enabling users to interact with virtual assistants such as Apple Siri, Amazon Alexa, and Google Personal Assistant to accomplish day-to-day life chores (e.g., playing music, appointment scheduling). With artificial intelligence platforms such as Google Dialogflow, IBM Watson, AWS Lex, Microsoft LUIS, developing a chatbot has been made simple. Recent bibliometric study on chatbots [16] reveals the scope of contributions in the state-of-the-art for chatbot applications in many domains with a high number of alternatives. This also confirms the novelty of our contribution to SG, since there is no prior work in this domain to the best of our knowledge.

In recent development, a new category of chatbot enables natural language processing to map suitable commands for their execution to provide convenience for users' requirements. An example is implemented in [17] for a conversational visualization service to extract keywords from the conversation and to apply filters to the visualizations. Chatbots for data analysis workflow management such as AVA [9] and more recently IRIS [10] are promising solutions that motivate our work in this paper. The chatbot interface AVA [9] is on the Jupyter Notebook, where users can interact with a bot to execute Python code and build their data model. IRIS [10] uses linguistic theory to handle complex requests interactively by combining commands through a nested conversation.

Inspired by the above works, we have developed our Vidura chatbot for the science gateway serving neuroscience researchers and educators. Instead of just allowing users to do few pre-configured data workflow actions, we guide users in their research experiments by providing varying resources such as text, simulation tools and recommending related publications via a chatbot interface. Users can not only get appropriate tools but we also guide users through execution of the workflow/simulation steps by answering queries related to the steps. The Vidura chatbot acts as a virtual expert service to support users on SG. Students, teachers and faculty members with a varying level of expertise can interact with our bot. A major novel contribution in our approach is in our approach to analyze user proficiency and to generate custom responses using the Mamdani approach as well as rule-based intelligence. We use the user quadrant questionnaire to estimate users' proficiency in HPC systems and neuroscience domain as explained later (see Section III-A). This approach is beneficial because it helps in generating contextually correct responses to the users based on their proficiency level.

III. Intuitionistic Fuzzy Based Conversational Recommender

In this section, we detail the user quadrant analysis using an intuitionistic fuzzy logic, and the Mamdani inference system which uses underlying resource recommenders to generate custom responses to user queries.

A. User Profiling and Quadrant Analysis

We use a questionnaire related to HPC/CI and neuroscience domain to analyze the user proficiency. Although in a traditional questionnaire, a user is expected to select one option from many alternatives, it is usually hard to concretely describe human thinking using binary logic. Instead, we rely on the benefits of many-valued logic in the fuzzy approach [18] for our user quadrant questionnaire. Fuzzy logic can capture an accurate and detailed representation of the questionnaire information, thereby providing a realistic analysis of the real-time user data.

Fuzziness represents vagueness and imprecision in human understanding and thinking. Fuzzy logic is an approach to handle this vagueness with "many-valued" logic instead of traditional true or false Boolean logic. Fuzzy logic is a part of the Fuzzy set theory. A fuzzy set is a class of items with a continuity in membership function. A membership function defines how each point in the input space is mapped to a degree of membership between 0 and 1. If $X = \{x\}$ is a universe of discourse which is the range of all possible values for an input to a fuzzy system, then $A \subseteq X$ denotes fuzzy subset where A belongs to X. The membership function associated with A is a mapping of

$$\mu : X \mapsto \{0, 1\} \tag{1}$$

For any element $x \in X$, if $\mu(x) = 1$ then x is a member of fuzzy set A, if $\mu(x) = 0$ then x is not a member of fuzzy set A.

Intuitionistic fuzzy logic [15] extends fuzzy logic and deals with uncertainty. It assigns each element a membership degree and a non-membership degree. It allows assessment of the elements by membership and non-membership functions that belong to the real unit interval [0, 1] with the sum also belonging to the same interval. It helps to move closer to human thinking and work with linguistic variables and terms.

We have developed our Vidura chatbot based on SG challenges and user requirements for multi-disciplinary neuroscience applications from the findings of our prior work [5]. By leveraging the proficiency and prior knowledge of the users in the neuroscience domain and HPC resources, we have categorized them into four user quadrants, as shown in Fig. 2.

- *Quadrant 1*: 'Novice users' with limited knowledge of neuroscience domain and the computational resources. For example, typically IT/CS freshmen.
- *Quadrant 2*: 'Domain Experts' have a high domain or subject knowledge, but have limited knowledge about SGs and HPC resources. For example, typical neuroscience Master students.
- *Quadrant 3*: 'HPC Experts' have low domain or subject knowledge but have high computational resource usage proficiency. For example, IT/CS Master Students.
- *Quadrant 4*: 'Domain and HPC Experts' have expertise in domain/subject as well as the usage of computational resources. For example, neuroscience faculty and PhD students.

Our user quadrant questionnaire has a total of 14 multiple choice questions to evaluate the proficiency of user in HPC and neuroscience domains. This questionnaire is divided into two sections: (1) HPC/CI and (2) neuroscience domain, with 7 questions in each of the section.

Example neuroscience questions involve "Rate your familiarity/understanding about" each of the following items:

- Concept of time constant as it relates to Neuron cell membranes
- Ionic currents give rise to an Action Potential
- Neuron modeling using the Hodgkin-Huxley formulation

Furthermore, example HPC questions includes "Rate your familiarity/understanding about" for each of the following items:

- National high-performance computing (HPC) resources such as XSEDE (Extreme Science & Engineering Discovery Environment)
- Using Linux Operating System like (Ubuntu, Fedora, Redhat)
- Proficiency in installing required software for your research simulations on your local machine

Each question was measured on a four-point Likert scale, ranging from (1) little to none, (2) below average, (3) above average, and (4) mastery.

While answering the questionnaire, the user can assign a weight between 0 and 1 for every option in each question to convey their respective degree of vagueness in the option selection. The sum of weights given by a user of all options for each question should not exceed 1. After the user has completed the questionnaire, we calculate the sum of weights of all options from all questions belonging to a specific category. For example, we add weights of option 1 and option 2 to find the 'Novice' membership value. Similarly, we add weights of option 3 and option 4 to find 'Expert' membership value. We have provided details about this procedure of categorizing users into the four quadrants in Algorithm 1.

Algorithm 1 User Quadrant Analyzer using fuzzy logic

```
/*calculate sum of weights of all options for all questions*/
for each option k = 1, ..., K, weights w:
for each hpc_question t = 1, ..., T, neuro_question s = 1, ..., S:
Compute r = Σ_w k;
Select maximum value, val = max(r)
Categorize user, proficiency(val)
```

B. Mamdani Inference System

After obtaining the user proficiency quadrant in Section III-A, we apply the Mamdani-type fuzzy inference model

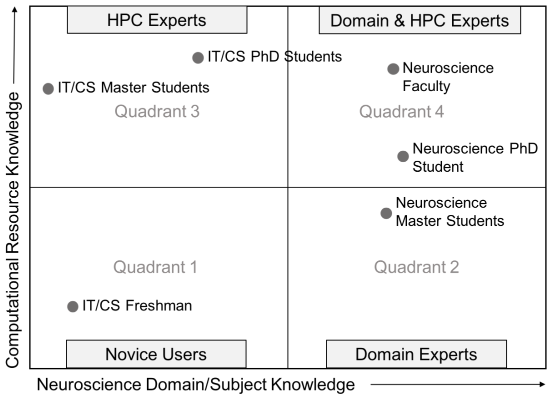

Fig. 2. Classifying users into quadrants based on their proficiency in domain and computational resource knowledge

Algorithm 2 Mamdani inference for recommending resources

```
/*fuzzification*/
resource=fuzzify(usercategory);
/*rule base*/
if usercategory == HPC Novice and neuroscience Novice
resource'= guided user interface
else if usercategory == HPC novice and neuroscience expert
resource'= basic user interface and cloud templates
else if usercategory == HPC expert and neuroscience novice
resource'= Jupyter Notebooks and publications
else if usercategory == HPC expert and neuroscience expert
resource'= CIPRES and cloud templates
/*defuzzification*/
recommender=defuzzify(resource)
return recommender
```

to recommend user resources which corresponds to their proficiency. Fuzzy inference is defined as the process of mapping the given input variables to an output space through a fuzzy logic mechanism which consists of multiple *If-Then rules*, membership functions and fuzzy logical operations such as AND, OR, and so on. The motivation for using this technique is because of its popularity in effectively yielding precise outputs and approximates to the human reasoning and our linguistic nature. We used the Mamdani approach [19] as it gives more straight-forward and convincing outputs in comparison to others.

There are variations in Mamdani Inference system such as multi-input multi-output, two-input single-output etc. Two input - two output is well suited for our case as we have user proficiencies in (1) HPC and (2) neuroscience domain as input and the output being simulation tools and external resources. Simulation tools include a guided user interface, Jupyter Notebooks, and CIPRES, whereas external resources include relevant publications and CI/cloud templates.

The four steps of Mamdani fuzzy inference system as shown in Algorithm 2 is described below:

Step 1 - Fuzzification: In the first step, we transform the numbers representing user proficiency into equivalent membership values of a fuzzy set using different membership functions. At the input, a Gaussian function is chosen because of their convenience for accurate representation and faster optimization of small rule-based data sets. The triangular membership function is selected for output variables as it is simpler and unity partition condition is easily satisfied. It shows that in comparison to other functions, the sum of membership grades for each value x amounts to 1.

Step 2 - Rule-based: The next step applies Implication method where we define the If-Then rules. The rules in the fuzzy set have an antecedent (if) and consequent (then) parts. Antecedent part gives single numerical value and consequent part reshapes the fuzzy set according to the result associated with the antecedent. Following are the few rules we have defined in our inference system:

- Rule 1: *If* user is Novice in HPC and neuroscience,
  Then simulation tool is guided user interface.
- Rule 2: *If* user is Novice in HPC and Expert in neuroscience,
  Then simulation tool is basic user interface and external resource is Cloud Template.
- Rule 3: *If* user is Expert in HPC and Novice in neuroscience,
  Then simulation tool is Jupyter Notebook and external resource is publications.
- Rule 4: *If* user is Expert in HPC and Expert in neuroscience,
  Then simulation tool is CIPRES and external resources are Cloud Template and publications.

Step 3 - Aggregation: The If-Then rules are aggregated to form a single fuzzy set using functions such as max, sum, among others. In our case, we use the max function on four truncated fuzzy sets obtained from the four rules of the previous step, to generate a new fuzzy set which represents the output variables to be supplied to the Defuzzification step.

Step 4 - Defuzzification: In this final step, the fuzzy set produces a single scalar value to quantify the simulation tools and external resources. We have used the popular Center of Gravity method for defuzzification as shown in (2).

$$Z_{COA} = \int_{z} \mu_A(z).zdz / \int_{z} \mu_A(z) \tag{2}$$

where, z is output variable and $\mu_{\mathrm{A}}(z)$ is membership function. Based on the results obtained from the Mamdani fuzzy interference, our conversational recommender will pass on the control to invoke the relevant recommender systems.

C. Notebook, Cloud Solution and Publication Recommender

As discussed earlier, the Mamdani inference system determines the type of simulation and external resources and these are obtained by making a REST API call to a different recommender system. Following gives brief description of each of the recommender used by our conversational recommender and in-depth details of these recommendation systems is out of the scope of this paper.

Notebook Recommender: Our Notebook Recommender algorithm accepts a list of keywords as input and crawls through all the notebooks in the repository to extract the visible text from

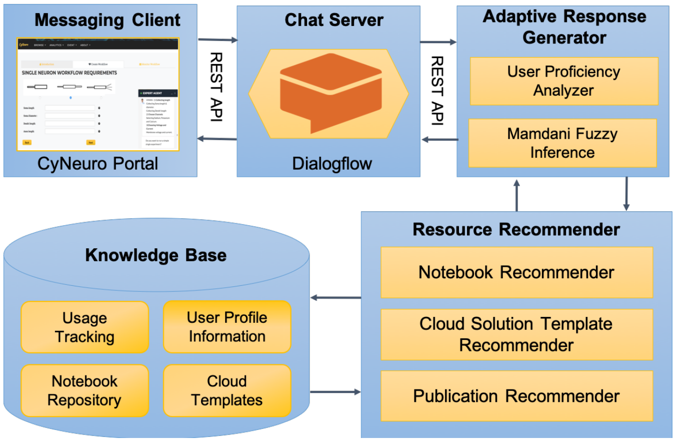

Fig. 3. The architecture of the CyNeuro science gateway and chatbot integration for neuroscience researchers and educators.

it. It uses the Whoosh Python module to index the documents and then queries the index for the input parameters. It uses standard TF_IDF (Term Frequency and Inverse Document Frequency) statistics to score and sort the search results as it accurately gauges the importance of a keyword in a notebook document. The recommender returns the notebooks with the highest TF_IDF scores for the given search string.

External Resources: For HPC and neuroscience expert users, we suggest cloud infrastructure templates, which users can customize as per their simulations and workflow requirements as presented in our prior work [11]. Similar to this, we also have a publication recommender which suggests most relevant publications that are related to a user's simulation/workflow. By recommending publications, a user can find more about researchers who are working on similar subjects with the possibility for future collaborations.

IV. Conversational Recommender Evaluation

In this section, we present the architectural details of Vidura and showing its integration on the CyNeuro Science Gateway [5], and present an evaluation of our conversational recommender for the neuron single cell simulation use case [5].

Vidura Architecture and Implementation: The architectural overview for CyNeuro Science Gateway integrated with our conversational recommender system is shown in Fig. 3. CyNeuro portal is built using Apache, MySQL, and PHP Laravel framework. Front-end user interface is built using AngularJs JavaScript framework. Vidura chatbot widget on the CyNeuro portal is developed using JQuery. Google Dialogflow conversational API is used to combine machine learning, natural language processing, and integrated dialog tools to create conversational flows between our chatbot on CyNeuro portal and the users. We followed the prescribed setup process to create 'intents' for the user dialogue around 'entities' such as templates, notebooks, among others as explained in the previous section. Dialogflow allows the use of external web services to generate custom responses by using the 'fulfillment' feature. It allows Dialogflow chat server to interact with our fuzzy based conversational recommender to recommend resources to the users.

Evaluation of User Quadrant Questionnaire: Reliability of a questionnaire refers to how cohesive each set of assessment questions cluster together, with a high internal consistency, to measure the respective concept of the neuroscience knowledge scale and HPC knowledge scale. We have calculated the Cronbach's alpha for our user quadrant questionnaire in the SPSS (Statistical Package for Social Sciences) software [20], using the reliability test procedure. According to [12], this index value should be above 0.70 to be deemed reliable (i.e, have high internal consistency). We have piloted these questions with responses from a small set of 14 users with varying backgrounds and expertise levels. We found that the HPC oriented questions achieved a cronbach's alpha of 0.749, and the neuroscience domain knowledge questions achieved a cronbach's alpha of 0.919. Both scores are greater than the minimum required value of 0.7, which validates that our user quadrant questionnaire is reliable and consistent.

Evaluation of Custom Response Generated by Vidura: To evaluate user-specific response generator, we have supplied Vidura with sample queries and compared its responses with annotations from human domain experts. They have attested the quality of auto-generated response and its variation with user proficiency. Table 1 shows the responses generated for a sample question "Help me with Neuron Simulation", for different users based on their proficiency. Quadrant 1 user receives more textual information about domain subject matter of the simulation along with basics operational details for the tools to support their poor knowledge about HPC/CI resources and neuroscience domain. In addition, the chatbot prompts them to use the guided user interface with significant chatbot support, as it provides extra help at each simulation stage. For the same question, Quadrant 2 user (neuroscience domain expert with low HPC knowledge), is given less neuroscience related information and more HPC related information to compensate for the knowledge gap. Because of their low HPC/CI knowledge, they will continue to receive a minimalistic guided user interface for the simulation job. They will also receive publications recommendations related to their simulation job.

Quadrant 3 user (HPC experts with low neuroscience domain knowledge) receives Jupyter Notebook as the (neuron) simulation tool recommendation. Rather than using the basic graphical user interface which will slow down their productivity, using notebook should be straight-forward and much faster because of their prior HPC and programming skill set. In addition, cloud templates are given to them to explore alternative options to run their workflows on public cloud services. The HPC and neuroscience expert user belonging to quadrant 4 requires minimalistic guidance. Their main focus is to successfully accomplish complex experiments in a productive manner and hence, continuous chatbot interruptions may appear intrusive for them. Hence, they are redirected to the CIPRES gateway for simulation queries, where they can submit their batch jobs. These users also get extensive publication recommendations related to their research problems.

TABLE I
RESPONSE GENERATED BY VIDURA FOR THE SAMPLE USER QUERY "HELP ME WITH NEURON SIMULATION" OR "RECOMMEND SIMULATION TOOLS FOR NEURON SINGLE CELL SIMULATION"

User Proficiency	**Vidura Response**		
	Text	**Simulation Tool**	**External Resource**
Novice User (Quadrant I)	Neuron is a single cell of our brain. We use NEURON software for simulation. Please use simulation link to get started with your simulation.	Guided User Interface on CyNeuro Portal	Cloud Templates: None Publications:
Domain Expert (Quadrant II)	Use the simulation tool provided to submit your job to super computer. You will be notified when job is completed.	Simple Job Submission UI on CyNeuro Portal	Cloud Templates:None Publications: "Software Automation for Biologically Realistic Neuro Big Data Simulations"
HPC Experts (Quadrant III)	Use the Juputer Notebook which walks you through computational modeling of single neuro cell	Jupyter Notebook on MU Data Center	Cloud Templates:CyVerse, NSG, AWS Publications:None
Domain & HPC Experts (Quadrant IV)	Use CIPRES workflow manager for all of your job submission to HPC system	CIPRES Workflow manager on MU Data Center	Cloud Templates: Publications:"Software Automation for Biologically Realistic Neuro Big Data Simulations"

V. CONCLUSION

In this paper, we have proposed a conversational recommender that is used to provide a guided user interface and a chatbot functionality for neuroscience researchers and educators/students. The intuitionistic fuzzy logic based conversational recommender design and integration with a CyNeuro science gateway (built upon CIPRES) was described to support access of HPC resources for exemplar neuroscience research use cases. One of the main contributions of this work is a methodology for personalization that enables a chatbot viz., Vidura to help users through customized recommendations. This strategy is likely to further promote meaningful use, encourage continuing adoption, and sustain long-term diffusion of the CyNeuro science gateway in the neuroscience community in overcoming the challenges of on-going expert service support. Arguably, the same methodology can be applied to other science gateways across various scientific domains such as bioinformatics or even manufacturing involving data-intensive or computation-intensive simulations.

Our future work is to extend our results into a platform within the neuroscience community for researchers/educators to define new/diverse scientific workflows using local and remote HPC resources and foster collaborative efforts through their knowledge sharing via CI templates and Jupyter Notebooks.

REFERENCES

[1] NSF (2016) Understanding the BRAIN.

[2] NSF (2015). iNeuro Conference Report, https://mdcune.psych.ucla.edu/modules/ineuro/reports/iNeuro_ConferenceReport_20150804.pdf

[3] P. J. Guo, "Software Tools to Facilitate Research Programming", *PhD Thesis, Stanford University*, 2012.

[4] A. C. Rule, "Design and Use of Computational Notebooks", *PhD Thesis, University of California San Diego*, 2018.

[5] P. Calyam, & S. S. Nair, "Experiences from a Multi-disciplinary Course Sequence Development on Cyber and Software Automation in neuroscience", *13th Gateway Computing Environments Conference (Gateways)*, 2018.

[6] A. Schrock, K. Kee, "Communicative capacity building strategies in cyberinfrastructure", *National Communication Association Annual Conference*, 2018.

[7] A. Zonca, R. Sinkovits, "Deploying Jupyter Notebooks at scale on XSEDE resources for Science Gateways and workshops", *Proc. of the Practice and Experience in Advanced Research Computing (PEARC)*, 2018.

[8] M. Miller, W. Pfeiffer, T. Schwartz, "The CIPRES science gateway: a community resource for phylogenetic analyses", *Proc. of the ACM TeraGrid Conference*, 2011.

[9] R. John, N. Potti, J. Patel, "Ava: From Data to Insights Through Conversation", *Proc. of the Biennial Conference on Innovative Data Systems Research*, 2017.

[10] E. Fast, B. Chen, J. Mendelsohn, J. Bassen, M. Bernstein, "Iris: A conversational Agent for Complex Tasks", *Proc. of ACM CHI Conference on Human Factors in Computing Systems*, 2018.

[11] R. Antequera, P. Calyam, A. Chandrashekara, S. Malhotra, "Recommending Resources to Cloud Applications based on Custom Templates Composition", *Proc. of the ACM Computing Frontiers Conference*, pp. 136-145, 2017.

[12] R. F. Devellis, "Scale development: Theory and applications", *Sage publications*, 2016.

[13] Zhongqi Lu, Sinno Jialin Pan, Yong Li, Jie Jiang, Qiang Yang "Collaborative evolution for user profiling in recommender systems", *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI*, 2016.

[14] Manan Jadhav, Shagufta Rizwan, Aditi Nehete "User Profiling based Adaptive Test Generation and Assessment in E Learning System", *3rd IEEE International Advance Computing Conference (IACC)*, 2013.

[15] Mukta Goyal, Divakar Yadav, Alka Choubey "Intuitionistic Fuzzy Approach for Adaptive Presentation in an E-learning Environment", *Proc. of IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2015.

[16] H. Io, C. Lee "Chatbots and Conversational Agents: A Bibliometric Analysis", *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2017.

[17] S. Bieliauskas, A. Schreiber, "A Conversational User Interface for Software Visualization", *IEEE Working Conference on Software Visualization*, 2017.

[18] Y. Lin, W. Berlin, "Fuzzy mode and its applications in survey research", *Proceedings of the 10th WSEAS International Conference on APPLIED MATHEMATICS*, 2006.

[19] W. Sari, O. Wahyunggoro, S. FauziatiA, "Comparative study on fuzzy Mamdani-Sugeno-Tsukamoto for the childhood tuberculosis diagnosis", *American Institute of Physics*, 2016.

[20] A. Agresti, B. Finlay, "Statistical methods for the social sciences: With SPSS from A to Z", *A brief step-by-step manual: Boston: Allyn & Bacon*, 2009.