PROCEEDINGS OF SCIENCE

Dynamical modelling of player engagement and skill in the Quantum Moves 2 citizen science game

Paul Saurou,[a] **Arthur Hjorth,**[a] **Miroslav Gajdacz,**[a] **Robert Heck,**[a] **Kobi Gal,**[b] **Jacob Sherson**[a,*] **and Quantum Moves 2 players**

[a] *ScienceAtHome, Center for Hybrid Intelligence,*
Department of Management, BSS, Aarhus University, Denmark

[b] *Ben-Gurion, University of the Negev, Israel and University of Edinburgh, U.K*

E-mail: sherson@mgmt.au.dk

Citizen science games and gamified tasks have shown promise in engaging citizens in scientific work through often digitalized version of real-life scientific tasks while producing useful data. This work has spurred an increasing interest in investigating how to design such interactions to entice players to produce as high quantities of scientifically useful data as possible. Player modelling, the computational analysis of player interactions with and within games, has been applied to better understand which UI features or particular player interactions can best predict player behavior. However, much of this work has relied on modelling approaches that exhibit low degrees of explainability, and that rely on large quantities of player data. In this paper we present Saturable Dynamic Bayesian Network Models (SDBN), a novel player modelling approach that avoids these two issues using a *feature activation network* that quantifies the user's gradually extending familiarity with the interface. This makes our approach usable for more fine-grained design-decisions when building and improving citizen science games, and it makes our approach applicable to games with smaller numbers of players. We apply our novel method to data collected from approximately 2,000 players of the citizen science game Quantum Moves 2. Our study shows that the SDBN approach can reliably model player drop out down to 10 participants and allows for both bulk and individual level insights of how player interactions with and within our game predicts player engagement and the quality of consequent data.

Engaging Citizen Science Conference 2022 (CitSci2022)
25-26 April 2022
Aarhus University, Denmark

*Speaker

1. Quantity and Quality in Citizen Contributions to CitSci Projects

Citizen science has become a popular approach to collecting data or conducting scientific research through the power of the crowd [1]. Despite the growth, however, many citizen science projects struggle to achieve the goal of collecting sufficient data for robust scientific research. Stated plainly, a core challenge is to generate data of sufficient quantity *and* quality.

With increasing levels of digitalization in citizen science projects—chiefly, gamified science tasks—[2], participant contributions can be tracked and quantified. The quantity of contribution per participant in such volunteer based projects tends to have a long-tailed distribution, in which a large group of participants contribute leisurely and a small core of participants contribute significantly more [3–6]. This difference may in part be attributable to varying levels of expertise, either with playing video games in general, with the core scientific content, or both [7, 8] however is not at the moment understood in detail. The probability of disengaging is in commercial applications often referred to as the "churn rate" [9, 10]. Figure 1 demonstrates the churn rate of Quantum Moves 2 [11], a citizen science game analyzed within this work.

A core question in citizen science is, therefore, how do we design games that engage players so they contribute larger quantities of data, while supporting them in making scientific contributions to these projects of sufficient quality?

To answer this question, citizen science scholars have investigated the efficiency of user inter-faces using design experimentation either through natural [12] or deliberate experiments, [13] but ideally such considerations should be based on intricate knowledge of individual player behavior in specific games [14, 15]. As an example, Gal et al. [14] presented a player model-based method to displaying motivational messages to users at the optimal moment to increase their play-time. They used reinforcement learning on data from Galaxy Zoo, one of the largest citizen science platforms. In the video game research literature other sophisticated approaches have been explored. For in-stance, the winning team of the IEEE CIG 2017 Game Data Mining Competition [16] used deep neural networks to predict the remaining lifetime of each player using real world data. In another example, machine learning models predict the churn rate, the expected expenses or the fun of the player to adapt the game experience [17, 18].

However, we see a number of problems with current approaches to understand player engagement in citizen science games. First, most of the models used in participant modeling efforts are neural networks generated using deep or reinforcement learning methods. These methods are efficient and can fit a model with little prior knowledge about the subject [19, 20]. But they are susceptible to overfitting and spurious effects, and because of their lack of *explainability* this is difficult to detect. Additionally, these models are generally very specific. They answer very precise questions and focus on only one parameter that they try to predict, for example, when is the best time to display a motivational message. These models cannot easily be used for other purposes. Finally, these models are highly dependent on the quantity of data available, thus limiting the number of small projects that can use these models.

In this paper, we present a new approach to modelling player interactions in citizen science games, a variant of Dynamic Bayesian Network-models [21], that emphasizes explainability can be broadly used in many different contexts and does not require large amounts of data in order to provide insights into player interactions and engagement. As a case study, we apply it to

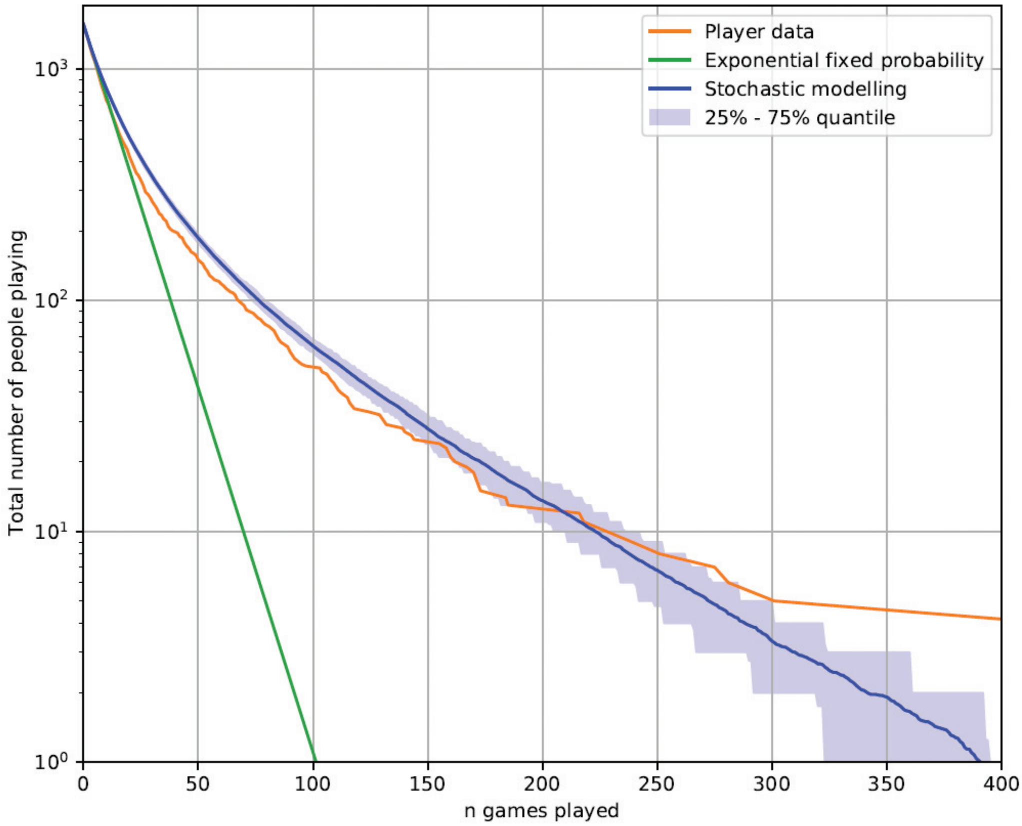

Figure 1: Player churn or total number of active players as a function of the number of games played (n) in the citizen science game Quantum Moves 2. The orange trace represents the player retention data (plotted on a log scale). The green line demonstrates that the player loss is poorly described by an exponential function with a fixed leave probability because there is a significant long-tail of persistent players. The stochastic model result is plotted in blue, where the solid line represents the mean and the shaded area the 25% to 75% quantile.

Quantum Moves 2 (QM2), which along with its predecessor Quantum Moves has been played by more than 250,000 citizen scientists worldwide. The scientific topic of QM2 is quantum physics and the challenge is to provide solutions to transfer atoms—represented by their quantum mechanical wave function—as fast as possible from an initial state to some predefined target state. Scientific results have demonstrated that players search the complex landscape differently and sometimes more efficiently than traditional algorithms [11]. Participants have been studied previously using surveys and interviews [5, 7] and the game design elements of the leaderboard investigated in a 10,000+ participant natural experiment [12].

For this study, we focus on one QM2 level, played by more than 2.000 participants. Compared to the original, QM2 includes a number of optional UI features (such as replay, ghost function and machine optimization of player solutions, see figure 2) designed to assist and enrich the player experience. The purpose of this paper is to show that our so-called SDBN model can be used for this purpose.

Our main research questions are:

1. Does SDBN offer an explainable, generalizable approach to explaining churn?

2. Can the insights offered by SDBN help us better design games to increase the quality and quantity of player-generated data?

To give the reader an understanding of the details of this study, we first introduce our model, then present the data that we collected from QM2 and how we analyzed these data with this model. Finally, we present results, and discuss how these results can help design citizen science games and interfaces that can increase the quality and the quantity of scientific player contributions.

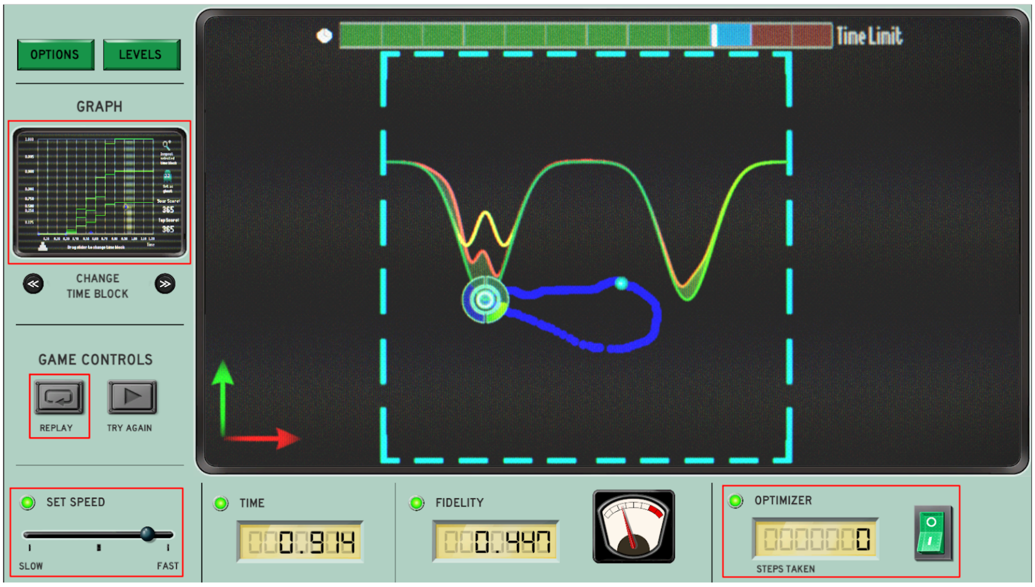

Figure 2: User interface of the game : The player has to move the wave from the right to the left as precisely as possible within the time limit. The red rectangles highlight optional UI features designed to help the players and keep them engaged. Clicking on the top left graph opens a menu with more optional features.

2. Our Model

As explainability is a primary focus for our approach, we base our model on Dynamic Bayesian Network-models, a general tool for behavior modelling [21]. A Bayesian Network (BN) is a probabilistic causal model represented by a directed acyclic graph (no loops). A *Dynamic* BN (DBN) takes into account the evolution of the system through time. The fundamental question that a DBN answers is: to what extent does the occurrence of other events at one point in time predict the occurrence of an event at the next point in time. This can help explain how certain features or particular player achievements in a game leads certain players to play for longer, or to produce better data. Concretely, the nodes of a DBN represent variables of the game system and player, either observable quantities, latent factors or hypotheses at a point in time. Each edge connects two nodes at time t and t+1 and is associated with conditional dependencies between the two nodes.

The crucial, manual step in setting up a DBN is to choose a) the system states that would each be represented by nodes in our graph and b) how to represent time. Fundamentally, we were interested in understanding what makes players play for longer (generating higher quantities of data) and what made players play well (generating data of higher quality). Based on expert knowledge of the game mechanics and observations of prominent player behavior, we therefore chose a set of 15 observables features and no additional latent factors. Eight of these indicate the point in time in which a user uses an optional UI feature for the first time. Four were game milestones features that represent the players' progressing scores in the game. Two were special game techniques that the authors had observed some players deploying. The final state was the point in time at which the player stops playing the game. As time unit, we chose each player's attempt at the level to study the evolution of the player's achievements and the impact that UI features have on it.

In general DBN-features can take on any range of values and can fluctuate non-monontonically in time. This in principle allows for modeling a very rich, granular dynamics but also increases data requirements and challenges explainability. We therefore introduce a novel, simplified variant *saturable* DBN (SDBN) in which features can only take on the values 0 or 1 and once activated (discovered or achieved) will remain at unity. Rather than attempting to model the actual use of the interface elements, this *feature activation network* explores the user's gradually extending familiarity with the interface.

For implementing the SDBN model, we characterise the player state before and after each game trial as an activation vector with Boolean components, $\mathbf{X_t}$ defined at each game trial t . Then we model the probabilities of transitions in the activation vector for the next game trials in the following way:

$$\mathbf{P}(\mathbf{X_{t+1}} \| \mathbf{X_t}) = sig(\mathbf{p} + \mathbf{Q} * \mathbf{X_t})(1 - \mathbf{X_t}) + \mathbf{X_t} \tag{1}$$

where $\mathbf{p}$ is a vector describing the probabilities of discovering each feature "by itself" and $\mathbf{Q}$ a matrix defining the conditional probabilities between features. A sigmoid regularizes the resulting values to represent probabilities $\mathbf{P}(\mathbf{X_{t+1}} \| \mathbf{X_t}) \in [0,1]$. The additional factors containing $\mathbf{X_t}$ ensure the saturation aspect of our model.

To fit the $15x16$ parameters of the model and determine the causal inferences between nodes, we used stochastic gradient descent with the data collected with cross-entropy as the cost function. The stability was ensured using 5-fold cross validation and taking the mean of the folds as the final result.

3. Finding: The DBN can successfully model engagement

As an initial test of our model, we see how well it predicts player churn (i.e., when the players play their last game.) As shown in figure 1, the model was able to explain the churn rate from the initial 2,000 down to less than 10 participants. This suggests that our 15 features have strong explainable power in predicting player churn. In contrast, the best single-parameter exponential decay clearly only explains player dropoff over much less than a decade.

Our discovery and achievement probability network also gives insight into why people keep playing the game. Indeed, figure 3 shows the transition matrix of the model demonstrating both positive and negative correlations. The two circled dependencies represent examples of causal inferences. The bottom one indicates that achieving milestones reduces the chances that the player leaves the game just after, and the top one indicates that achieving a milestone increases the chance of reaching a better milestone just after.

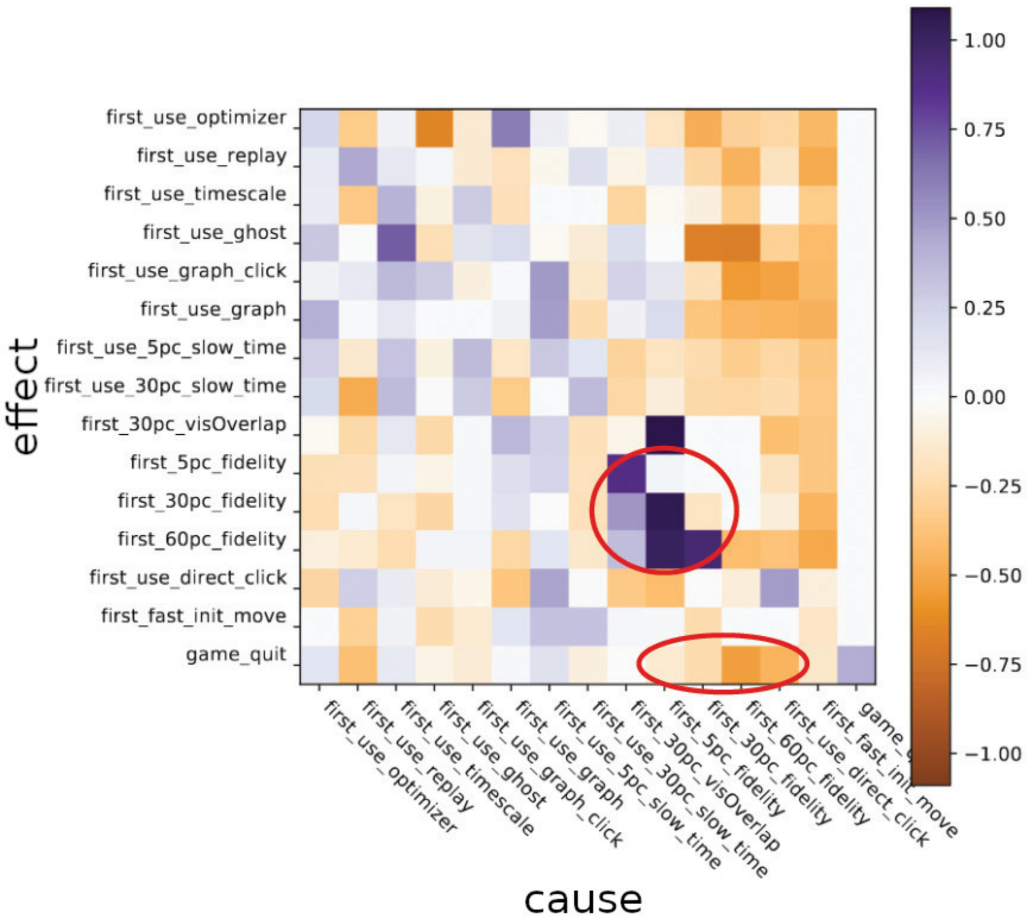

Figure 3: This matrix shows the influence of one feature on the subsequent activation of another one. A blue square represents a strong causal impact and a brown square represent a reverse causal impact.

4. Finding: The DBN exhibits different clusters of volunteers

This general model predicts the mean behavior of players. However it can be extended by making the spontaneous (p) and interaction (Q) parameters player specific. These are regrouped to prevent overfitting in a few (6) fit parameters: global interactions with UI or learning and discovery of special strategies. These parameters are focused on the links between UI and learning the game. This allows us to train an individual model for each player. This method provides us player-specific models documenting each player's specificities.

To better understand the individual SDBN-models, we clustered the players using k-means. Figure 4 highlights the very different progression of players. The participants in cluster 4 almost immediately use the machine optimization feature and manually achieve the first milestone (5%) and quickly advance to the second one (30%), whereas the third (60%) is only achieved after more experimentation. Within cluster 0, progress in feature exploration and game achievements is much slower. Unlike players in cluster 4, they do however seem to discover the "ghost" feature and soon thereafter start achieving game milestones. Surprisingly, when behavior over a longer timescale is investigated a significantly larger fraction of the initially slow learners in cluster 0 achieve the highest

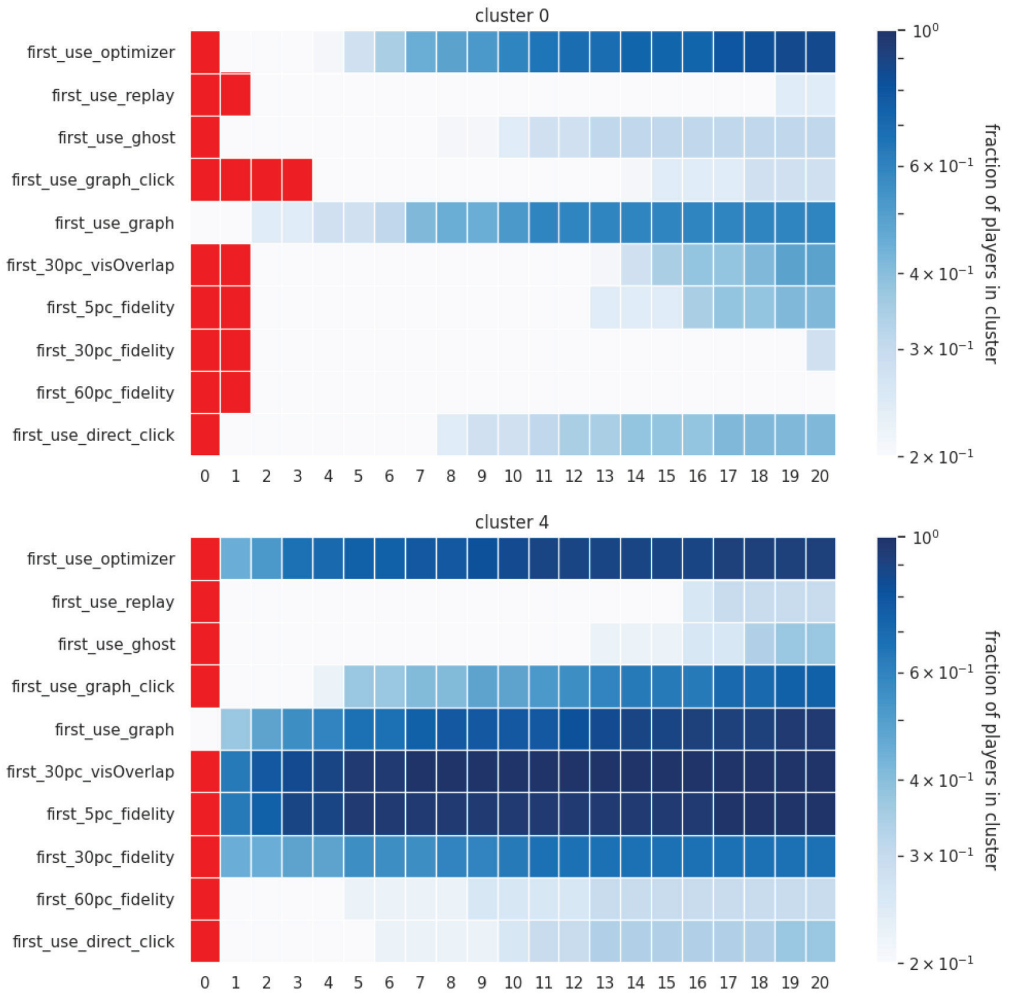

Figure 4: Cluster analysis of player behavior on their 20 first games. Each individual player has been modelled by a SDBN and their behaviors have been clustered using K-Means. The blue scale is logarithmic and features not discovered yet are marked in red.

game milestone compared to cluster 4 players. This granular behavioral insight demonstrates the dramatically different required interface changes or motivational messages required for different player clusters.

5. Conclusion

In this paper, we presented a novel feature activation model, Saturable Dynamic Bayesian Networks. The model addresses shortcomings in current citizen science game-player modeling, specifically relating to explainability and the necessity for large quantities of data. Applying the model to data from approximately 2,000 players of Quantum Moves 2, we show that the model can predict both aggregate level data—churn rate across all players—and explain the degree to which the discovery of particular gameplay features cause players to play longer and provide better solutions.

We hope that this model can help towards better understanding how to make sense of player data, and how to use this knowledge to design citizen science games that produce more and better data.

Acknowledgements

The authors acknowledge support of the Carlsberg and Novo Nordisk Foundations and thank the ScienceAtHome game development and outreach teams as well as the Quantum Moves 2 players around the world.

References

[1] R. Bonney, J.L. Shirk, T.B. Phillips, A. Wiggins, H.L. Ballard, A.J. Miller-Rushing et al., Next Steps for Citizen Science, *Science* **343** (2014) 1436.

[2] J. Rafner, M. Gajdacz, G. Kragh, A. Hjorth, A. Gander, B. Palfi et al., Revisiting Citizen Science Through the Lens of Hybrid Intelligence, *arxiv.2104.14961* (2021) .

[3] O. Nov, O. Arazy and D. Anderson, Dusting for science: motivation and participation of digital citizen science volunteers, in *Proceedings of the 2011 iConference*, (Seattle Washington USA), pp. 68–74, ACM, Feb., 2011, DOI.

[4] A. Eveleigh, C. Jennett, A. Blandford, P. Brohan and A.L. Cox, Designing for dabblers and deterring drop-outs in citizen science, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (Toronto Ontario Canada), pp. 2985–2994, ACM, Apr., 2014, DOI.

[5] A. Lieberoth, M.K. Pedersen, A.C. Marin, T. Planke and J.F. Sherson, Getting Humans to do Quantum Optimization—User Acquisition, Engagement and Early Results from the Citizen Cyberscience Game Quantum Moves, *Journal of Human Computation* **1** (2014) 221.

[6] A. Segal, Y. Gal, E. Kamar, E. Horvitz, A. Bowyer and G. Miller, Intervention strategies for increasing engagement in crowdsourcing: Platform, predictions, and experiments, in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, IJCAI'16, p. 3861–3867, AAAI Press, 2016.

[7] C. Díaz, M. Ponti, P. Haikka, R. Basaiawmoit and J. Sherson, More than data gatherers: exploring player experience in a citizen science game, *Quality and User Experience* **5** (2019) 1.

[8] J.A. Miller and S. Cooper, Barriers to Expertise in Citizen Science Games, in *CHI Conference on Human Factors in Computing Systems*, (New Orleans LA USA), pp. 1–25, ACM, Apr., 2022, DOI.

[9] J. Runge, P. Gao, F. Garcin and B. Faltings, Churn prediction for high-value players in casual social games, *2014 IEEE Conference on Computational Intelligence and Games* (2014) .

[10] N. Lu, H. Lin, J. Lu and G. Zhang, A Customer Churn Prediction Model in Telecom Industry Using Boosting, *IEEE Transactions on Industrial Informatics* (2014) .

[11] J.H.M. Jensen, M. Gajdacz, S.Z. Ahmed, J.H. Czarkowski, C. Weidner, J. Rafner et al., Crowdsourcing human common sense for quantum control, *Physical Review Research* **3** (2021) 013057.

[12] M.K. Pedersen, N.R. Rasmussen, J.F. Sherson and R.V. Basaiawmoit, Leaderboard Effects on Player Performance in a Citizen Science Game, *Proceedings of the 11th European Conference on Game Based Learning* (2017) 531.

[13] N. Prestopnik, K. Crowston and J. Wang, Gamers, citizen scientists, and data: Exploring participant contributions in two games with a purpose, *Computers in Human Behavior* **68** (2017) 254.

[14] A. Segal, K. Gal, E. Kamar, E. Horvitz and G. Miller, Optimizing Interventions via Offline Policy Evaluation: Studies in Citizen Science, *Proceedings of the AAAI Conference on Artificial Intelligence* **32** (2018) .

[15] J.A. Miller, U. Narayan, M. Hantsbarger, S. Cooper and M.S. El-Nasr, Expertise and engagement: re-designing citizen science games with players' minds in mind, in *Proceedings of the 14th International Conference on the Foundations of Digital Games*, (San Luis Obispo California USA), pp. 1–11, ACM, Aug., 2019, DOI.

[16] A. Guitart, P.P. Chen and A. Periáñez, The Winning Solution to the IEEE CIG 2017 Game Data Mining Competition, *Machine Learning and Knowledge Extraction* **1** (2018) 252.

[17] A. Guitart, P.P. Chen, P. Bertens and Periáñez, Forecasting Player Behavioral Data and Simulating in-Game Events, in *Advances in Intelligent Systems and Computing*, pp. 274–293, Springer International Publishing (2017), DOI.

[18] A. Fortin-Cote, C. Chamberland, M. Parent, S. Tremblay, P. Jackson, N. Beaudoin-Gagnon et al., Predicting video game players' fun from physiological and behavioural data, in *Advances in Information and Communication Networks*, pp. 479–495, Springer International Publishing (2019).

[19] Y. Xue, I. Davies, D. Fink, C. Wood and C.P. Gomes, Behavior Identification in Two-Stage Games for Incentivizing Citizen Science Exploration, in *Principles and Practice of Constraint Programming*, M. Rueher, ed., vol. 9892, (Cham), pp. 701–717, Springer International Publishing (2016), DOI.

[20] J. Yu, W.-K. Wong and R.A. Hutchinson, Modeling Experts and Novices in Citizen Science Data for Species Distribution Modeling, in *2010 IEEE International Conference on Data Mining*, pp. 1157–1162, Dec., 2010, DOI.

[21] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 4 ed. (2020).

PoS(CitSci2022)074