Understanding Uncertainty in Atomistic Simulations

Aik Jun Tan*
Massachusetts Institute of Technology

Franklin Morgan†
Massachusetts Institute of Technology

Daniel Schwalbe-Koda‡
Massachusetts Institute of Technology

Abstract

Academic work is often communicated through peer-reviewed research papers, conference presentations and occasional media coverage. While these formats allow for a full or partial understanding of the results, they are often limited in availability or require considerable effort to be read by an expert outside of the field. In this work, we use an interactive visualization to summarize the main insights from a research paper. As an example, we focus on a recent paper addressing uncertainty quantification in materials simulation. To address the known difficulties of representing uncertainty in high-dimensional spaces, we use a scrollytelling narrative to gradually build the concepts of the paper. In this framework, dynamic images allow the viewer to interact with data not readily available in the article, thus improving the reading experience. We qualitatively validate our approach by showing the visualizations to experts and non-experts in the field. Viewers exposed to the simplified approach report a higher curiosity towards the research results regardless of their level of expertise. Moreover, expert viewers report that this interactive article could greatly help their understanding of new research articles. We believe that interactive visualizations can increase the exposure of academic articles and communicate their results more efficiently both to experts and non-experts.

Index Terms: Human-centered computing—Visualization—Visualization design and evaluation methods; Computing methodologies—Machine learning—Machine learning algorithms—Ensemble methods

1 Introduction

Communication of academic research is traditionally performed through peer-reviewed articles summarizing the scientific findings of an investigation. Formal records of scholarly discoveries hark back to ancient Greece [23], where the concept of Academia was founded. However, written reports of oral debates or scientific inquiry had to be hand-copied throughout generations to be preserved. It was only centuries later when the invention of the printing press (circa 1440 AD) enabled the mass dissemination of books and articles. The creation of scientific journals is best illustrated by the establishment of the Philosophical Transactions of the Royal Society in 1665 AD, which centralized the review and written discussion of science in the United Kingdom for centuries. This slow-paced communication with editorial offices and peers was only disrupted a few decades ago with the widespread usage of internet. While this technology enabled most works, including science, to advance at unprecedented speed, it also imposed a taxing demand on workers' time and attention.

Nowadays, fields such as machine learning (ML) are among the most fast-paced areas of research worldwide, with more than 100 new papers uploaded daily to the arXiv preprint server under this category. As such, staying informed about the literature has become an overwhelming task for researchers in the area. Academic papers continue to demand the same careful perusal they did decades ago. While their format is adequate for the rigorous descriptions reported therein, it might prevent researchers slightly outside of the field from interpreting the main insights of the paper without time-consuming deliberation.

In this work, we propose summarizing academic papers with interactive narratives to simplify their understanding for scientific audiences. As an example, we translate this concept to an article quantifying uncertainty in ML models for physical sciences. While uncertainty-aware data visualizations can expose the limitation of models for physical research, designing static graphs that can demonstrate data uncertainty is not straightforward. Our animated story walks the viewer through the method and gradually builds the complexity of explanations, thus providing an overview of the research article without relying on equations or excessive jargon. When exposed to the content, viewers outside and inside the field of research demonstrated curiosity towards the article. While viewers outside of the field engaged with the introductory content, viewers in the field understood the concepts with less self-reported effort. We believe using animated stories to simplify publicizing academic research can raise awareness about different fields of research while avoiding "siloing" the information to domain experts.

2 Related Work

Interactive academic research: analysis on the effective use of interactions and animations [10, 25]. The use narrative visualization to affect end-user interpretation [12].

Visualizing uncertainty: the use of gradient and violin plots [5], hypothetical outcome plots [13], quantile dotplots [6, 16] and others have been explored as possible ways to visualize uncertainty. The challenges associated to communicating uncertainty are often associated to these representations, and may lead to different interpretations of the final data [11, 16].

Uncertainty quantification in neural networks: Bayesian NNs [2], Monte Carlo dropout [7] or ensemble methods [4, 20] are typical ways to estimate uncertainty of ML models by generating a distribution of predictions for a given input. In particular, NN ensembles [9, 17, 28] have been widely successful in generating distributions of predictions with increased reliability and scalability.

Uncertainty quantification in atomistic simulations: uncertainty quantification applied to atomistic simulations is a recent field [18, 19, 24, 27]. ML potentials from Gaussian processes display native error quantification and have been used for active learning [8, 15, 26]. NN ensembles were used to control simulations [1, 3], inform active learning loops [14, 21], or sample new geometries [22].

3 Methods

The main goal of the project was to propose a visualization technique that summarized an academic paper for researchers inside or slightly outside of the field explored in the paper. Differently from mass media articles — which focus on a broader public — or academic posters, talks or papers — which have limited accessibility or are time-consuming — an interactive summary of the paper has enough complexity to serve as a stepping stone to the main research findings (Fig. 1). The interactive features also enable viewers to explore

*e-mail: atan10@mit.edu
†e-mail: efmorgan@mit.edu
‡e-mail: dskoda@mit.edu

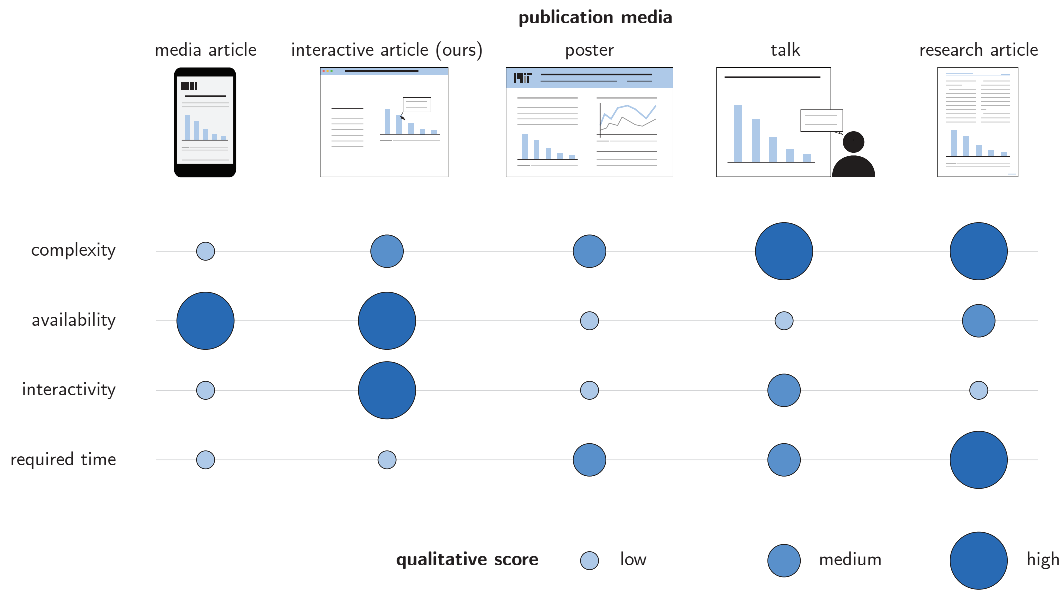

Figure 1: Characteristics of different publication types according to their media.

sideways of the results that are not included in the main paper for the sake of conciseness.

To conform to the linearity of most papers, we selected a scrollytelling visualization. The technique allows the research paper to be simplified while also increasing in complexity and engagement. These narratives also enable the structure of papers to be broken down into small pieces that build in complexity. Overall, the scrollytelling tool is comprised of four primary components: (1) the scroller, (2) the descriptions, (3) the visuals, and (4) the navigator. Fig. 2 exemplifies the components of the tool as implemented in this work.

The scroller is the key engine behind the webpage, allowing users to control the pace of the read. In building our scroller, we set triggers at select page intervals to ensure that the right mix of descriptors and visuals appear at the right time. This is extremely important for emphasis and flow; specific ideas should be focused on only after pre-requisite knowledge has been imparted. Users can control how fast the ideas flow by using the scroller.

The descriptions are the paragraphs which provide contextual information on the visuals being shown. We positioned them on the left side of the page window. They are initially set to zero opacity and are designed to fade into vision along with the visuals as the section is scrolled into view. They will fade away as the page is further scrolled. In producing the description, we made a conscious decision to include only the minimal text required to describe a visual. This is done to avoid cluttering and to generate focus.

The visuals are the images and graphs that condense all the findings of the research. They are rich in information and are complementary to the textual description. Unlike the descriptive text, whose position and opacity change with scrolling, in most times the opacity of the visuals changes with scrolling. This transition helps the viewer to understand how concepts are built together to form the main paper. To serve as a stepping stone for the manuscript, images from the paper fade into vision as user scrolls through the page. The positions of the visuals are mostly fixed and do not change with scrolling to aid the viewer perceive the gradual changes that are made as the sections pass by. We also included some interactivity with the visuals to allow users to quickly explore a few points within a complex manifold and see the impact of their proposed changes.

The navigator is a shortcut function on the top left side of the page which allows users to quickly navigate to a section without being constrained by the flow of the scroller. This tool helps users to revise and reinforce concepts quickly by jumping between different sections instantaneously.

4 Results

To verify the usage of an interactive narrative to summarize a research paper, we have gathered the feedback of five viewers outside of this research domain through the 6.859 class feedback, and three expert viewers from our research groups. All viewers were asked questions on whether the visualization was effective in terms of content, engagement and interactivity. Expert reviewers were asked additional questions on whether the webpage has improved their knowledge of the field or increases the likelihood that the paper will be read later.

Viewers were overall positive about the visualization. Both experts and non-experts rated the interactive article effective in conveying information at a lower complexity than a research article, with expert reviewers reporting that the webpage would be particularly useful if the topic had been unknown to them. Expert viewers also reported a higher likelihood of reading the paper after engaging with the interactive webpage.

The main source of disagreement between experts and non-experts is the amount of content shown in the webpage. Non-experts felt that the webpage tried to convey an excessive amount of information, while expert viewers suggested more content could be added to cover other topics, such as adversarial attacks in image classification or uncertainty quantification for molecular structures. The amount of interactivity is also deemed as “just right” for expert viewers. On the other hand, non-experts suggested more interactivity to the visualization, including more tooltips, hovering functionalities, and other complex tasks. Non-experts were also more likely to recommend new figures, whereas expert viewers liked the fact that figures from the manuscript itself were deconstructed and explored in detail.

Uncertainty visualization, one of the main topics of the article, was reported by the viewers as properly treated by the interactive article. The explanation of the two-pane prediction-uncertainty visualization helped convey the meaning of uncertainty quantification

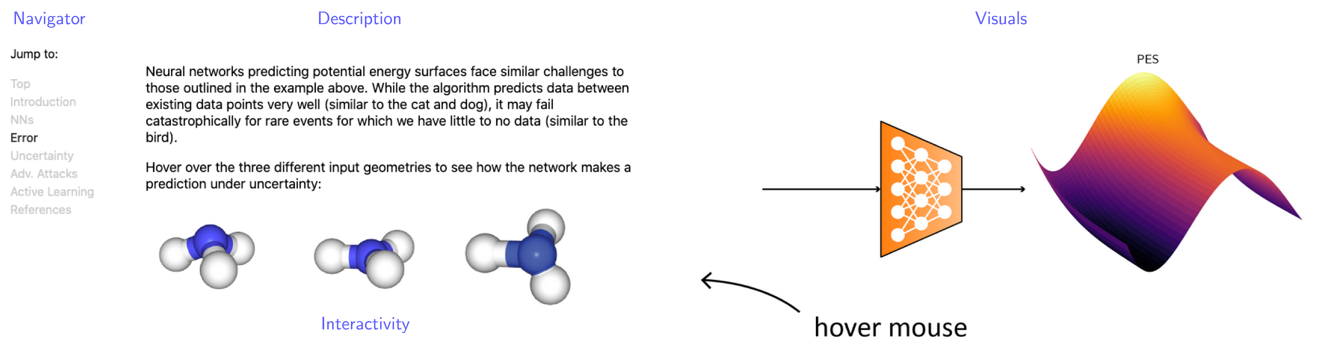

Figure 2: Screenshot of the scrollytelling narrative implemented in this work. The webpage is comprised by the scroller (engine, not shown), a navigator, the description, and the visuals.

for two variables. Moreover, the expert reviewers reported that the interactivity of the active learning loop, which compares the accuracy of neural networks across different generations, greatly helps understanding the main takeaways of the article.

5 Discussion

The qualitative analysis of the results indicates that the interactive article may effectively summarize a research paper both for non-experts and experts. An expert audience seems to benefit from the complexity build-up, while non-expert viewers may require more visual content to engage with the discussion. While the amount of text in the interactivity may vary with the topic, several popular scrollytelling articles provide more content than opinion pieces or mass media articles. Interestingly, the interactivity and animation are also used as ways to display the contents in a memorable way instead of being the central part of the narrative. Nevertheless, converting a research paper into an interactive visualization may be a reasonable approach to aid all viewers, even at low levels of interactivity.

While the main messages of the paper were understood by the viewers, both experts and non-experts suggested visualizations that could greatly improve the appeal of the webpage. As an example, a map between molecular structures and the potential energy surface was recommended as being an interesting piece of information. Although we believe this could be a great enhancement to the project, we also recognize the amount of effort to make this possible. Interatomic interactions are expressed by many-body electronic effects, which can either be calculated through quantum chemistry, or parametrized using force fields. The former is more accurate, although significantly costly even for small systems. Dedicated packages are used in supercomputer centers to perform these calculations over multiple processing units, which renders an analogous, web-based implementation impractical. Using parametrized force field calculations could be implemented, but would require programming the calculation of neighbor lists and interatomic parameters on-the-fly in this web application. Despite the feasibility of this last approach, it contradicts the purpose of the article, which is using a neural network to parametrize these interactions. Embedding the adequate architectures and weights into a web application is an interesting direction of work, but not viable at the moment due to the complexity of the autodifferentiation software used to implement most of the evaluation and uncertainty quantification. Moreover, hundreds of megabytes are required to store the neural network weights, hindering their use in browsers.

On the other hand, mapping a potential energy surface to an atomic structure is only possible under simplified assumptions. The degeneracy of molecular energies implies that structure-energy maps are not invertible unless limited to certain didactic constraints. We have experimented with replotting the article figures on the web visualization, but the amount of data used when drawing the contour surfaces made the visualizations unresponsive to user commands without latency. Simplifying the representation, such as plotting a one-dimensional well, could make this example feasible, but may compromise the overall understanding of a high-dimensional optimization that is performed through adversarial attacks.

6 Future Work

Summarizing academic work in this interactive manner requires a combination of expert knowledge and visualization techniques. Future work could focus on producing software packages that abstract away the process of creating a storytelling narrative, similarly to what is performed in slideshow applications. Alternatively, creating a well-documented library that builds the website, e.g. with Ruby, could be enough to lower the barrier for scientists to translate their research findings without having to deal directly with d3.js and HTML. While the library of interactions would be limited, it could help academic researchers to publish informal summaries about their own research findings and achieve a broader audience.

The interactivity in this work could be refined by providing the viewers with more options to explore the data from the main research article. Instead of being constrained to a summary, this interactive research article could also be a supporting information, guiding experts on how to reproduce their research or other interesting insights that were not highlighted in the main paper.

Particularly in the case of atomistic simulations, visualizing and manipulating 3-D geometries could be extremely engaging for viewers, despite the challenge posed by programming such tool and the limited application. We believe such interactivity would not offer valuable insights to the viewer from an academic perspective.

Importantly, the framework shown in this work could be extended to any other research article. To support the informal discussion in this report, a full study on the availability of research media shown in Fig. 1 could be proposed, along with a study on the demands of academic reading. We believe understanding science communication could enable a fast-pace evolution of science without increase of burden for researchers.

Acknowledgments

The authors thank the anonymous reviewers of the interactive visualization for the valuable feedback.

References

[1] J. Behler. Representing potential energy surfaces by high-dimensional neural network potentials. *Journal of Physics: Condensed Matter*, 26(18):183001, 5 2014.

[2] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra. Weight Uncertainty in Neural Network. In F. Bach and D. Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 1613–1622, Lille, France, 2015. PMLR.

[3] L. Chen, I. Sukuba, M. Probst, A. Kaiser, I. Sukuba, and M. Probst. Iterative training set refinement enables reactive molecular dynamics: Via machine learned forces. *RSC Advances*, 10(8):4293–4299, 1 2020.

[4] R. T. Clemen. Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5(4):559–583, 1989.

[5] M. Correll and M. Gleicher. Error bars considered harmful: Exploring alternate encodings for mean and error. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2142–2151, 2014.

[6] M. Fernandes, L. Walls, S. Munson, J. Hullman, and M. Kay. Uncertainty displays using quantile dotplots or cdfs improve transit decision-making. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2018.

[7] Y. Gal and Z. Ghahramani. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In M. F. Balcan and K. Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1050–1059, New York, New York, USA, 2016. PMLR.

[8] J. A. Garrido Torres, P. C. Jennings, M. H. Hansen, J. R. Boes, and T. Bligaard. Low-Scaling Algorithm for Nudged Elastic Band Calculations Using a Surrogate Machine Learning Model. *Physical Review Letters*, 122(15):156001, 4 2019.

[9] L. K. Hansen and P. Salamon. Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):993–1001, 1990.

[10] J. Heer and G. G. Robertson. Animated transitions in statistical data graphics. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1240–1247, 2007.

[11] J. Hullman. Why authors don't visualize uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):130–139, 2019.

[12] J. Hullman and N. Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2231–2240, 2011.

[13] J. Hullman, P. Resnick, and E. Adar. Hypothetical outcome plots outperform error bars and violin plots for inferences about reliability of variable ordering. *PloS one*, 10(11):e0142444, 2015.

[14] G. Imbalzano, Y. Zhuang, V. Kapil, K. Rossi, E. A. Engel, F. Grasselli, and M. Ceriotti. Uncertainty estimation for molecular dynamics and sampling. *The Journal of Chemical Physics*, 154(7):74102, 2 2021.

[15] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, and M. Bokdam. Phase Transitions of Hybrid Perovskites Simulated by Machine-Learning Force Fields Trained on the Fly with Bayesian Inference. *Physical Review Letters*, 122(22):225701, 6 2019.

[16] M. Kay, T. Kola, J. R. Hullman, and S. A. Munson. When (ish) is my bus? user-centered visualizations of uncertainty in everyday, mobile predictive systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5092–5103, 2016.

[17] B. Lakshminarayanan, A. Pritzel, and C. Blundell. Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30, pages 6402–6413. Curran Associates, Inc., 2017.

[18] F. Musil, M. J. Willatt, M. A. Langovoy, and M. Ceriotti. Fast and Accurate Uncertainty Estimation in Chemical Machine Learning. *Journal of Chemical Theory and Computation*, 15(2):906–915, 2019.

[19] A. A. Peterson, R. Christensen, and A. Khorshidi. Addressing uncertainty in atomistic machine learning. *Physical Chemistry Chemical Physics*, 19(18):10978–10985, 2017.

[20] D. N. Politis and J. P. Romano. Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions. *Ann. Statist.*, 22(4):2031–2050, 1994.

[21] C. Schran, K. Brezina, and O. Marsalek. Committee neural network potentials control generalization errors and enable active learning. *Journal of Chemical Physics*, 153(10):104105, 9 2020.

[22] D. Schwalbe-Koda, A. R. Tan, and R. Gómez-Bombarelli. Differentiable sampling of molecular geometries with uncertainty-based adversarial attacks. *arXiv:2101.11588*, 1 2021.

[23] R. Spier. The history of the peer-review process. *TRENDS in Biotechnology*, 20(8):357–358, 2002.

[24] K. Tran, W. Neiswanger, J. Yoon, Q. Zhang, E. Xing, and Z. W. Ulissi. Methods for comparing uncertainty quantifications for material property predictions. *Machine Learning: Science and Technology*, 1(2):025006, 5 2020.

[25] B. Tversky, J. B. Morrison, and M. Betrancourt. Animation:can it facilitate. *International Journal of Human-Computer Studies*, 57:247–262, 2002.

[26] J. Vandermause, S. B. Torrisi, S. Batzner, Y. Xie, L. Sun, A. M. Kolpak, and B. Kozinsky. On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. *npj Computational Materials*, 6(1):20, 2020.

[27] S. Venturi, R. L. Jaffe, and M. Panesi. Bayesian Machine Learning Approach to the Quantification of Uncertainties on Ab Initio Potential Energy Surfaces. *Journal of Physical Chemistry A*, 124(25):5129–5146, 6 2020.

[28] Y. Zhao, J. Gao, and X. Yang. A survey of neural network ensembles. In *2005 International Conference on Neural Networks and Brain*, volume 1, pages 438–442, 2005.