

Interpolating GANs to Scaffold Autotelic Creativity

Ziv Epstein^a, Océane Boulais^a, Skylar Gordon^a and Matt Groh^a

^aMIT Media Lab, 75 Amherst Street, Cambridge MA 02139

Abstract

The latent space modeled by generative adversarial networks (GANs) represents a large possibility space. By interpolating categories generated by GANs, it is possible to create novel hybrid images. We present “Meet the Ganimals,” a casual creator built on interpolations of BigGAN that can generate novel, hybrid animals called ganimals by efficiently searching this possibility space. Like traditional casual creators, the system supports a simple creative flow that encourages rapid exploration of the possibility space. Users can discover new ganimals, create their own, and share their reactions to aesthetic, emotional, and morphological characteristics of the ganimals. As users provide input to the system, the system adapts and changes the distribution of categories upon which ganimals are generated. As one of the first GAN-based casual creators, Meet the Ganimals is an example how casual creators can leverage human curation and citizen science to discover novel artifacts within a large possibility space.

Keywords

GANs, computational creativity,
citizen science, artificial intelligence

1. Introduction

Generative adversarial networks (GANs) [1] are a subclass of generative models that enable anyone to generate photo-realistic images with a single click of a button [2]. Some have heralded this innovation as the end of design, whereby machine intelligence will replace human creation [3]. However, a larger contingent considers these new generative technologies as yet another tool in an artist’s toolkit, which offers new expression with its novel affordances [4]. Since the application of generative models does not require formal artistic training or technical expertise, these models can serve as scaffolding to generate large possibility spaces that can be embedded in casual creator systems [5]. Recent platforms such as RunwayML, GANBreeder, GANPaint, and DeepAngel have already started to use the new medium of GANs for casual creation.

The key challenge for GAN-based casual creators is designing systems that “support a state of creative flow” - whereby users and the generative models can co-create new artifacts in a collaborative, coordinated and organic dialogue, towards the idea of *mixed-initiative co-creativity*. [6, 7].

We introduce one such casual creator system, Meet the Ganimals, that allows users to selectively create new artificial hybrid species by interpolating between categories modeled by

Joint Proceedings of the ICCV 2020 Workshops (ICCV-Ws 2020), September 7-11 2020, Coimbra (PT) / Online

✉ zive@mit.edu (Z. Epstein)



© 2020 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

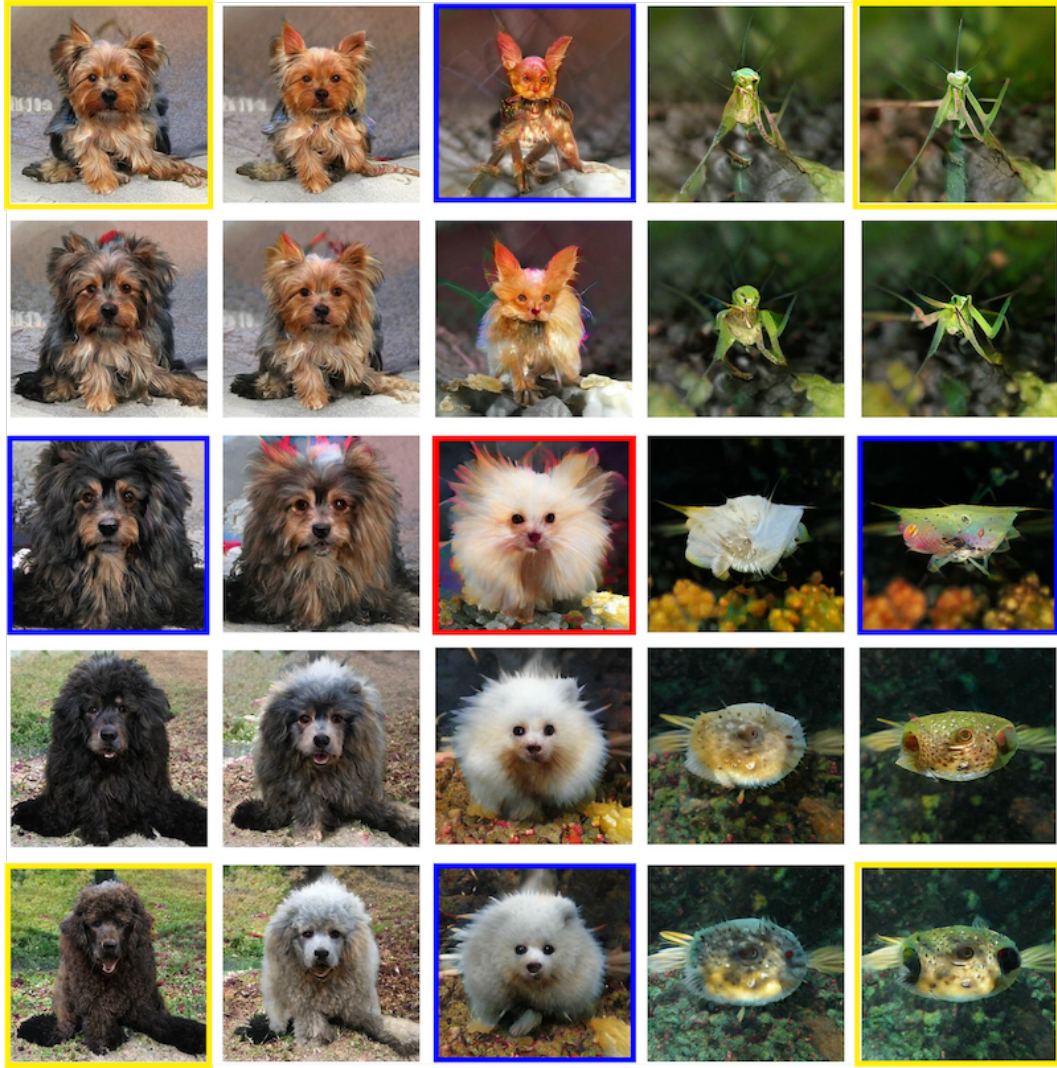


Figure 1: Schematic for interpolation. The yellow images in the four corners represent the praying mantis, a boston terrier, a pufferfish and poodle categories respectively, which we call generation zero (G_0) ganimals, the images in the four outer mid-points in blue are hybrids of two G_0 ganimals which we call G_1 ganimals, and the center image is a G_2 ganimal, which is a combination of all four G_0 ganimals.

BigGAN [8]. Trained on images with 1000 categorical labels, BigGAN embeds each category in a high-dimensional latent space. This space can be smoothly traversed such that images of mixed categories can be synthesized via interpolating the categories. Figure 1 presents examples of images generated from single and mixed categories. The original BigGAN model was trained on 1000 categories, and we restrict BigGAN to the 396 animal categories. The goal of constraining the model to animal categories is to focus the experience on discovering and breeding hybrid animals — what we call “ganimals.”

Unlike professional creative systems which provide users with many precise tools to craft

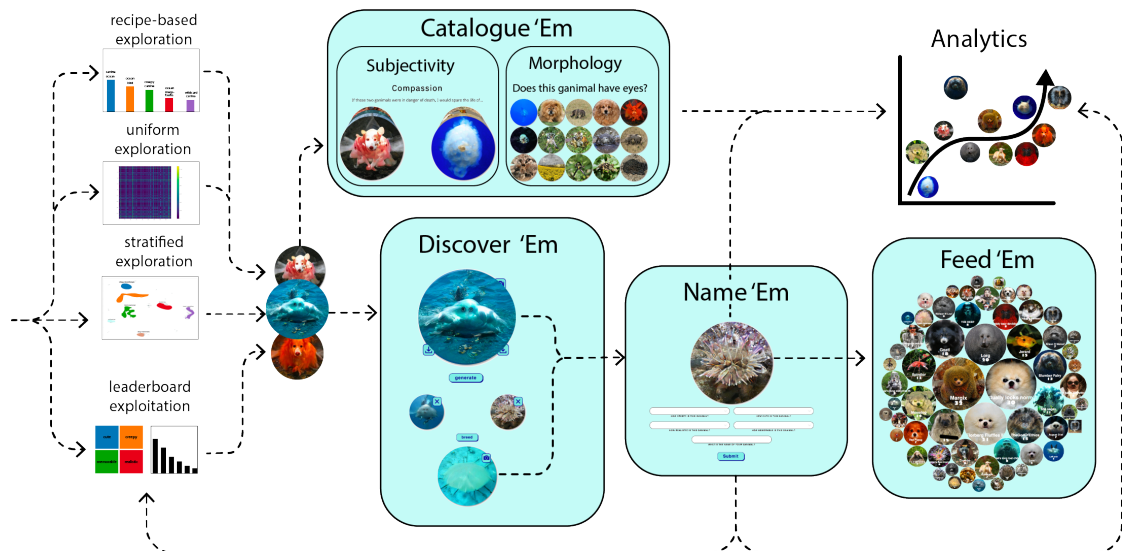


Figure 2: System map of the Meet the Ganimals Platform

artifacts directly, Meet the Ganimals is a simple interface designed to promote the exploration of a vast possibility space. The possibility space includes three generations of ganimals; 396 G_0 ganimals that correspond directly to the ImageNet animal categories, 78,210 G_1 ganimals from hybrid pairs of G_0 ganimals, and 3,058,362,945 G_2 ganimals come from hybrid quadruples of G_0 ganimals (see Figure 1). The possibility space is even larger when accounting for variations that come from truncation inputs and random seeds.

With such a large possibility space, it is difficult to find the near-superlative ganimals e.g. the cutest, creepiest, most memorable ganimals. Meet the Ganimals confronts this seemingly intractable search problem with two innovations. First, Meet the Ganimals simplifies exploration into a two-part creation interface: Users generate a large number of G_1 ganimals to find the ones they like best, and then they breed the chosen G_1 ganimals into G_2 ganimals. Second, instead of randomly combining categories to generate G_1 ganimals in this first stage, the system instead balances exploring new permutations, and exploiting previously popular permutations, as indicated from crowd signals from other users. These innovations build on earlier interfaces where user-generated landmarks serve as navigation elements in a large parametrically-defined possibility space [9, 10].

2. Related Work

Several recently developed platforms explore how GANs serve as scaffolding for autotelic creativity. RunwayML provides a simple interface such that non-technical artists can use state-of-the-art neural network models (e.g. style transfer and super resolution) in their work. GAN-Paint is a scene drawing tool that allows users to add or remove trees, grass, and other natural features with a simple click [11]. DeepAngel is an online tool that removes objects from images users upload via masking and generative inpainting [12]. While such functionality exists

in Photoshop (e.g. content-aware fill), DeepAngel provides a one-click interface, sidestepping the technical skills required for photo-editing.

Other related platforms have explored creating collaborative media via crowd signals and genetic algorithms. The R/Place experiment on Reddit allowed users to collectively recolor pixels of a dynamically changing image [13]. Electric Sheep generates procedural animations using crowd signals in an evolutionary algorithm [14]. PicBreeder demonstrated how pictures could be evolved collaboratively to rapidly explore a possibility space and proliferate fascinating discoveries while promoting individual exploration [15, 16].

ArtBreeder is an example of a casual creator built on interpolating GANs [17]. From the perspective of a user, ArtBreeder is a platform for creating new artworks by blending existing images.

Meet the Ganimals builds on these projects to combine GAN scaffolding with collective feedback while focusing on the domain of hybrid animals.

3. System Overview

Meet the Ganimals is designed with modern UI/UX paradigms built on BigGAN to serve as a mixed-initiative co-creation tool whereby users are both creators and consumers of the possibility space. From April 20th to May 20th, 51,110 ganimals were generated, and 10,587 ganimals were bred by 4,392 users. The system map for the platform is shown in Figure 2.

3.1. Random Stimulus for Exploring Possibility Space

In *The Book of Imaginary Beings*, Jorge Luis Borges wrote about his compilation of created creatures, saying that “the book ... is not meant to be read straight through; rather, we should like the reader to dip into these pages at random, just as one plays with the shifting patterns of a kaleidoscope” [18]. Echoing Borges, this system leverages the idea of the random stimulus principle of lateral thinking to offer a stochastic exploration of the possibility space [19]. In the “Discover ‘Em” page, users are shown ganimals randomly generated using a bandit algorithm that balances exploration of an unseen possibility space with the popularity of the discovered space. In particular, G_1 ganimals are generated and presented to users according one of four selection procedures: (1) 30% of the time that ganimal is generated using a carefully feature engineered stochastic process that samples pairs that we as designers found to be compelling (“recipe-based exploration”), (2) 30% of the time by random uniformly sampling two animal categories to breed (“uniform exploration”), (3) 30% of the time by randomly sampling two animal categories to breed stratified by species (“stratified exploration”), and (4) 10% of the time by sampling a ganimal from the top rated ganimals, proportional to its order in the leaderboard for a random one of the following characteristics: cute, creepy, realistic, or memorable (“leaderboard exploitation”).

For the recipe-based exploration, we carefully curated an ad-hoc generative process that we found created high-quality ganimals. In particular, we defined five cores - sets of conceptually similar ImageNet categories that are well-suited for blending for aquatic, canine, bird, megafauna, and wildcard categories. We then randomly blend these cores in order to create diversity in the resulting ganimals.

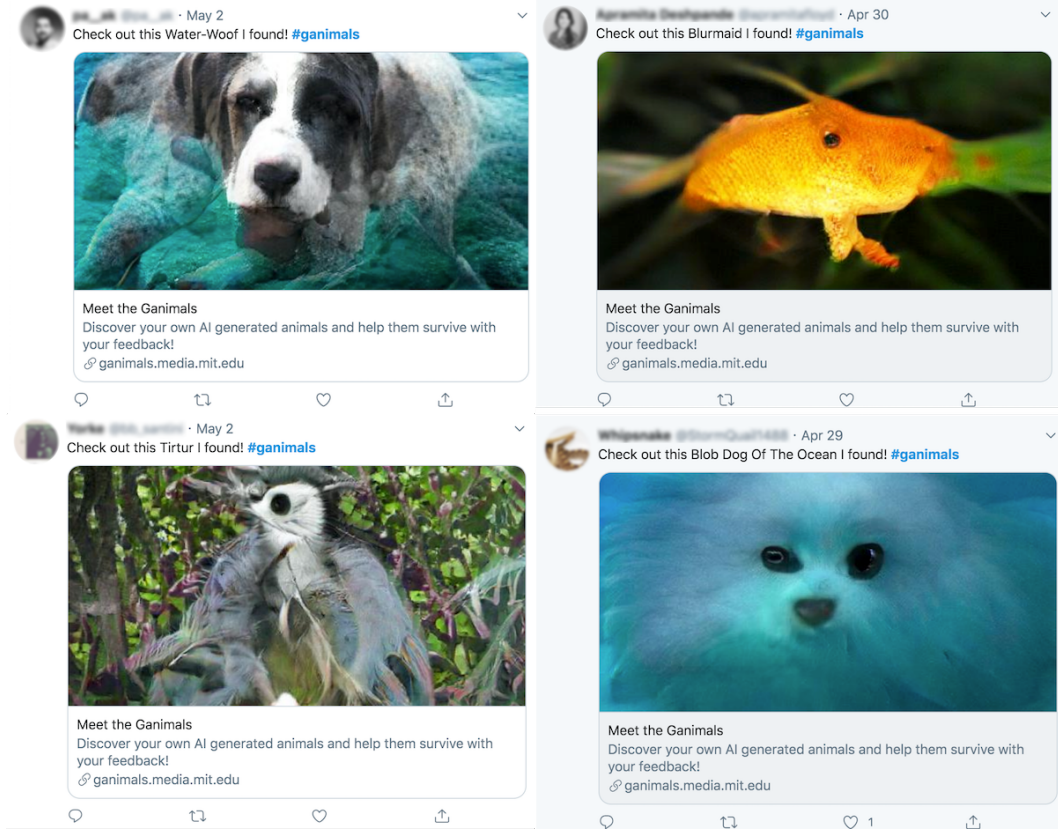


Figure 3: Selected screenshots of tweets from (anonymized) users sharing the ganimals they have discovered and named.

For stratified exploration, we uniformly sample a pair of animal species and sample an ImageNet category that corresponds to that species. For the majority of categories, there is a one-to-one correspondence between ImageNet categories and species. However, there are 118 categories of dogs (*Canis lupus familiaris*). The stratified exploration downsamples the frequency of dogs relative to the frequency in which dogs appear in ImageNet categories to promote diversity in the kinds of ganimals created. Users can curate G_1 ganimals and blend G_1 ganimals to create their own G_2 ganimal, which can be named and given its own unique hyperlink. This process supports the creative flow that allows users to efficiently explore the system’s possibility space, view a diverse array of combinations, and add their own creativity to the that of the system to create their own artifacts.

Users feel a sense of pride and ownership over the ganimals they create, which they have shown by sharing discoveries on social media (see Figure 3).

3.2. Towards a Citizen Science of GAN Subjectivity

From identifying exoplanets [20] and Christmas birds [21] to detecting changes in climate [22] and coral reef coverage [23], citizen science projects have been core part of engaging the public

in global scientific operations. Ultimately, participants engage in these projects because it is of individual interest to participate in a public (and therefore social) network [24], and because they are aesthetically pleasing and easy-to-use [25].

Meet the Ganimals has no metric for success, efficiency, or productivity, and there is no way for a user to demonstrate technical artistic or design skills. Instead, users are motivated by naming privileges for unseen ganimals, and general curiosity. As such, Meet the Ganimals is well suited for casual citizen science.

In the “Catalogue ’Em” page, users have the opportunity to take on the role of citizen scientists (i.e. “casual” scientists) to help answer scientific questions about the ganimals: Do ganimals with canine morphological features look cuter than the rest [26]? Does curation decrease as the underlying animal categories diverge in evolutionary time [27]? Do descendents of charismatic megafauna emerge as the most popular [28]?

To explore such questions, users can recount their subjective and emotional perspectives of the ganimals, as well as annotate their morphological features. For morphology features, users can annotate whether or not the ganimals have a head, eyes, a mouth, a nose, legs, hair, scales, feathers, live underwater, or are bigger than a house cat. For subjective perspectives, users can annotate how much compassion and empathy they feel towards the ganimals, as well as how cute, memorable, realistic and creepy they are.

This process allows for a deeper understanding of how animal morphologies relate to subjective perception. In particular, we can correlate subjective evaluations of ganimals with their other characteristics - such as crowd annotated morphology features, or the number of “dog” categories present within that ganimal. As a preliminary analysis, we find that ganimals contain at least one dog are statistically significantly cuter than those that do not, and that ganimals that contain at least one insect are statistically significantly less cute than those that do not (see Figure 4). Such knowledge can inform the design of future generative algorithms that use crowdsourced labels to surface maximally compelling artifacts.

When users take on the role of an explorer in the realm of ganimals, they not only become a creator, but also take on a role as a participant in the scheme of a broader investigation. By departing from the traditional role of a data-driven “scientist” who seeks a conclusion to a hypothesis, Meet the Ganimals empowers the exploration-driven creator: casual citizen scientists do not participate in the context of accomplishing a specific task or aim, and largely are not driven by a particular hypothesis. This unique dynamic thereby can leverage discovery to collect in-the-field results in a host of different ecosystems.

In addition to discovery, this framework for casual citizen science also raises questions of authorship and attribution. For in contexts where many people are collectively contributing to a AI system, particularly in the creative domain, platform designers must take extra care to consider how credit is distributed [29, 30].

4. Random World Assignment to Explore Local Ecologies

In line with the Music Lab experiment [31], a final ingredient for the Meet the Ganimals system is the random assignment of each participant to a “world,” with its own local ecology that involves independently from those of other worlds. Each world is initialized with a fixed “seed

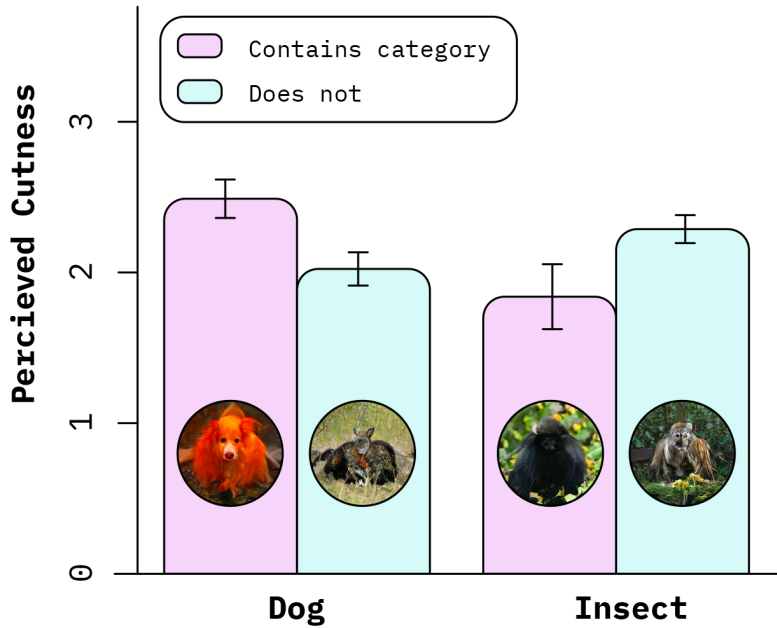


Figure 4: Perceived cuteness rating of ganimals with or without specific categories. Images are exemplary ganimals from that set.

set” of 100 ganimals, randomly selected using the random stimulus approach discussed above. Then users assigned to a particular world interact with this seed set plus the ganimals discovered and bred by users assigned to that world. In the “Feed Em” page, users can feed the ganimals they like the best. Well-fed ganimals are promoted and remain in this view, while unfed ganimals disappear. In addition, the fourth selection procedure from the bandit algorithm (“leaderboard exploitation”) only pulls ganimals from the world corresponding to that user.

The design of the “Feed Em” page can differ across worlds, allowing cross-world comparisons to serve as an A/B test for how different UX/UI patterns affect emergent ecologies. For example, one might ask how the layout of the “Feed Em” page (e.g. a linear feed-like view versus a more spatial ecological view) changes the resulting diversity of the ganimals in that world.¹ The random assignment of users to different worlds that evolve independently provides a virtual laboratory to compare behavior and curation across worlds, and causally assess the impact of design interventions.

5. Discussion

GAN architectures force two computational agents, the generator and discriminator, to compete against each other with the goal of creating a statistical model resembling the training

¹A deep dive into the experimental design and measurement approach of such a research question is beyond the scope of this paper, which focuses on the overview of the casual creator itself. However interested readers can learn more by reading the pre-analysis plan here: <https://asppredicted.org/65nv7.pdf>

data. Casual creators built on GAN architectures introduce a third agent, a casual human collaborator, into the loop to explore the most intriguing parts of latent space.

With a simple interface for creating and curating images of hybrid AI generated animals, users are motivated to engage in computational creativity for no other reasons than their own curiosity and the chance to name their creations and discoveries. The autotelic motivation drives the interactions within the casual creator and as a result, the system provides insights into what intrigues people [6].

In many creative endeavors, the production and consumption of artifacts are separated, which can lead to undifferentiated production and passive consumption. Human-in-the-loop casual creators built upon GANs are a new medium that blends production and consumption of media into a singular creative process. While Meet the Ganimals focuses on generating images of hybrid animals in particular, it is but one of a growing number of casual creators built on interpolating the GAN latent space of other cultural artifacts. Beyond animals, artifacts as varied as facial expressions, architectural landmarks and fashion are emerging domains where GANs could serve as scaffolding for casual creators [32], paving the way for new forms of human-AI collaboration. With a well-constrained casual creator, the frontiers of the GAN latent space are within reach.

Acknowledgments

We would like to thank Abhimanyu Dubey, Aurélien Miralles and the rest of the Ganimals team. We would also like to thank the anonymous reviewers for helpful feedback.

References

- [1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [2] T. Karras, S. Laine, T. Aila, A style-based generator architecture for generative adversarial networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4401–4410.
- [3] Nichols, Gabriel, These Stunning A.I. Tools Are About to Change the Art World, <https://slate.com/technology/2017/12/a-i-neural-photo-and-image-style-transfer-will-change-the-art-world.html>, 2017. Online; accessed 19 May 2020.
- [4] A. Hertzmann, Can computers create art?, in: *Arts*, volume 7, Multidisciplinary Digital Publishing Institute, 2018, p. 18.
- [5] K. Compton, M. Mateas, Casual creators., in: *ICCC*, 2015, pp. 228–235.
- [6] G. N. Yannakakis, A. Liapis, C. Alexopoulos, Mixed-initiative co-creativity mixed-initiative co-creativity (2014).
- [7] D. Acharya, N. Wardrip-Fruin, Building worlds together: understanding collaborative co-creation of game worlds, in: *Proceedings of the 14th International Conference on the Foundations of Digital Games*, 2019, pp. 1–5.

- [8] A. Brock, J. Donahue, K. Simonyan, Large scale gan training for high fidelity natural image synthesis, arXiv preprint arXiv:1809.11096 (2018).
- [9] J. Talton, D. Gibson, P. Hanrahan, V. Koltun, Collaborative mapping of a parametric design space, Technical Report, Citeseer, 2008.
- [10] J. Harding, C. Brandt-Olsen, Biomorpher: Interactive evolution for parametric design, *International Journal of Architectural Computing* 16 (2018) 144–163.
- [11] D. Bau, J.-Y. Zhu, H. Strobelt, B. Zhou, J. B. Tenenbaum, W. T. Freeman, A. Torralba, Gan dissection: Visualizing and understanding generative adversarial networks, arXiv preprint arXiv:1811.10597 (2018).
- [12] M. Groh, Z. Epstein, N. Obradovich, M. Cebrian, I. Rahwan, Human detection of machine manipulated media, arXiv preprint arXiv:1907.05276 (2019).
- [13] J. Rappaz, M. Catasta, R. West, K. Aberer, Latent structure in collaboration: the case of reddit r/place, in: Twelfth International AAAI Conference on Web and Social Media, 2018.
- [14] S. Draves, The electric sheep screen-saver: A case study in aesthetic evolution, in: *Workshops on Applications of Evolutionary Computation*, Springer, 2005, pp. 458–467.
- [15] J. Secretan, N. Beato, D. B. D Ambrosio, A. Rodriguez, A. Campbell, K. O. Stanley, Picbreeder: evolving pictures collaboratively online, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2008, pp. 1759–1768.
- [16] J. Secretan, N. Beato, D. B. D’Ambrosio, A. Rodriguez, A. Campbell, J. T. Folsom-Kovarik, K. O. Stanley, Picbreeder: A case study in collaborative evolutionary exploration of design space, *Evolutionary computation* 19 (2011) 373–403.
- [17] J. Simon, Ganbreeder, Ganbreeder. Accessed March 29 (2019).
- [18] J. L. Borges, *The book of imaginary beings*, Random House, 1957.
- [19] M. Beaney, *Imagination and creativity*, volume 4, Open University Worldwide Ltd, 2005.
- [20] J. K. Zink, K. K. Hardegree-Ullman, J. L. Christiansen, I. J. Crossfield, E. A. Petigura, C. J. Lintott, J. H. Livingston, D. R. Ciardi, G. Barentsen, C. D. Dressing, et al., Catalog of new k2 exoplanet candidates from citizen scientists, *Research Notes of the AAS* 3 (2019) 43.
- [21] The National Audobon Society, The audobon christmas bird count project, <https://www.audubon.org/birdst>, 2019. Online; accessed 16 July 2020.
- [22] SciStarter, Climateprediction.net, <https://scistarter.org/climatepredictionnet>, 2019. Online; accessed 16 July 2020.
- [23] V. Raoult, P. A. David, S. F. Dupont, C. P. Mathewson, S. J. O’Neill, N. N. Powell, J. E. Williamson, Gopros™ as an underwater photogrammetry tool for citizen science, *PeerJ* 4 (2016) e1960.
- [24] R. Lukyanenko, A. Wiggins, H. K. Rosser, Citizen science: An information quality research frontier, *Information Systems Frontiers* (2019) 1–23.
- [25] R. Bonney, C. B. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. V. Rosenberg, J. Shirk, Citizen science: a developing tool for expanding science knowledge and scientific literacy, *BioScience* 59 (2009) 977–984.
- [26] J. Kaminski, B. M. Waller, R. Diogo, A. Hartstone-Rose, A. M. Burrows, Evolution of facial muscle anatomy in dogs, *Proceedings of the National Academy of Sciences* 116 (2019) 14677–14681.
- [27] A. Miralles, M. Raymond, G. Lecointre, Empathy and compassion toward other species decrease with evolutionary divergence time, *Scientific Reports* 9 (2019) 1–8.

- [28] N. J. Bennett, R. Roth, S. C. Klain, K. Chan, P. Christie, D. A. Clark, G. Cullman, D. Curran, T. J. Durbin, G. Epstein, et al., Conservation social science: Understanding and integrating human dimensions to improve conservation, *Biological Conservation* 205 (2017) 93–108.
- [29] Z. Epstein, S. Levine, D. G. Rand, I. Rahwan, Who gets credit for ai-generated art?, *Iscience* 23 (2020) 101515.
- [30] J. K. Eshraghian, Human ownership of artificial creativity, *Nature Machine Intelligence* (2020) 1–4.
- [31] M. J. Salganik, P. S. Dodds, D. J. Watts, Experimental study of inequality and unpredictability in an artificial cultural market, *Science* 311 (2006) 854–856.
- [32] J. Zhu, Y. Shen, D. Zhao, B. Zhou, In-domain gan inversion for real image editing, *arXiv preprint arXiv:2004.00049* (2020).