# Learning About Creativity from an Artificial Artist

**David Norton**
Brigham Young University
Computer Science Department
ghotikun@hotmail.com

## ABSTRACT
Creativity is a nebulous, albeit vital part of every society. We outline an original computer system designed to produce artefacts that are perceived as creative, to do so through arguably creative processes, and to convey intention to an audience. We show how the computer system, called DARCI (Digital ARtist Communicating Intention), can augment the study of human creativity while advancing the field of artificial intelligence.

### Author Keywords
Artificial creativity, neuroevolution, intention, visual arts

### ACM Classification Keywords
I.2.0 Artificial Intelligence: General—*Cognitive simulation*; J.5 Computer Applications: Arts and Humanities—*Fine arts*

## INTRODUCTION
Despite many philosophical debates and academic treatises, there is still no concrete definition of creativity. Yet society continually attributes creativity to individuals and attempts to foster it. Some feel that creativity, by its very nature, cannot be reproduced in a computer program. Why is this the case? What are the attributes of creativity that cannot be replicated in a program? We don't attempt to dispute the claims that creativity cannot be simulated, but we do argue that by trying to do so, we can learn much about what it means to be creative while advancing the capabilities of artificial intelligence.

We outline an original computer system designed to produce images that are not only perceived by humans as creative products, but that are also produced through arguably creative processes. Additionally, the system will be able to communicate the intent behind its works, thus demonstrating a level of awareness that we feel is an important property of creativity. We hope that this system will be useable in a variety of research related to both human and artificial creativity. Since all good virtual artists need a catchy name, we call ours DARCI (Digital ARtist Communicating Intention).

There are two schools of thought regarding the attribution of creativity. One suggests that the artefact stands alone and should be judged independently of its creator. The other, that the thought and process involved in producing the artefact are paramount in attributing creativity [2]. If not taken to extremes, these schools of thought are actually complementary and should both be taken into consideration when deciding who and what is creative. We will consider two definitions of creativity, spanning both sides of the debate, that we will use to analyze the creativity of DARCI.

Graeme Ritchie presents a list of 18 criteria which could be used to attribute creativity to a series of artefacts, independent of the creator [6]. The list centers around three attributes: novelty, quality, and typicality. We will use these criteria to judge the creativity of the artefacts DARCI produces, and will do so entirely apart from the process used to make the artefacts.

Simon Colton, on the other hand, describes three conditions which must exist for a system to be considered creative. It must be described as skillful, appreciative, and imaginative [3]. In this paper, we will show that DARCI can indeed be described in these terms.

While we may be able to convince a readership that our system is skillful, appreciative, and imaginative, if the audience for DARCI's work isn't convinced of this, we will have ultimately failed. Because of this, we are also concerned with the ability of DARCI to convey the process behind her works, to the audience. The Painting Fool does this by allowing the audience to watch it as it paints [1]. DARCI will eventually do this by communicating intention, both implicitly and explicitly, to the audience.

## DARCI
Before analyzing the creativity of DARCI, we must describe the processes involved in the production of an artefact. DARCI is designed to be modular and expandable. The parts described in this paper represent a core system that will be expanded upon over time.

To begin with, DARCI will simply render photographs, the subject, using Non-photorealistic Rendering (NPR) filters selected according to both the aesthetic qualities of the subject, and constraints imposed by an accompanying description. The description will consist of a series of keywords describing the desired outcome of the image. For example, "Bright, Majestic" might accompany a photograph of a lion. Alternatively, "Dark, Scary" may accompany the same image. These will yield different artefacts. Once a lexicon of keywords and their relationships to image properties have been developed, DARCI will be able to generate a description on her own, when presented with only a subject.

Eventually, we hope that DARCI will be able to parse complex descriptions, choose subjects from a source such as the web, and even create her own descriptions and images using a cognitive model and components of the core system soon to be described. However, these steps are not within the scope of this paper.

Figure 1 shows a graph of how the components that make up DARCI are interconnected, starting with the subject and description inputs. The numbers in the figure refer to important steps that are described below. There are three pipelines: the solid black lines represent the generative pipeline, the pipeline concerned with creating an artefact (steps 1-4); the dashed blue lines represent the self-evaluation pipeline, the process concerned with what DARCI thinks of her own work (steps 5-7); finally the dotted red lines show the external-evaluation pipeline, the pipeline concerned with how DARCI responds to external criticism (steps 8-10).

### Generative Pipeline

**1:** The *image analyzer* acts as DARCI's visual perception. It performs a series of image processing analyses that may include building color histograms, taking brightness/contrast measures, and performing edge detection. The results of these measures are encoded in a data structure called the *image analysis*. The image analyzer is designed to be modular so that new analysis tools can be added at anytime, essentially improving DARCI's visual awareness. We will initialize this component with a variety of available tools and experiment to find a useful set.

**2:** The *language analyzer* acts as DARCI's conceptual awareness as it provides DARCI with an elementary understanding of language. It performs a series of text analyses in an attempt to glean the basic meaning of descriptions provided to DARCI. These results are encoded in a data structure called the *language analysis*. At first, this will consist primarily of the keywords that make up the description. However, as with the image analyzer this component is modular and in the future may yield more sophisticated output.

**3:** The *visuo-linguistic associator* (VLA) forms associations between attributes of the image analysis and language analysis. Since the language analysis will at first be focused on providing keywords in its data string, this component will be primarily forming relationships between image attributes and keywords (later more unusual relationships could be explored). This is a two way association as the VLA will repeatedly act as a translator from image analysis to language analysis and vice versa. For example, DARCI may be presented with a subject but not a description. She will then need to generate a language analysis to fill in for the lack of description.

Once associations have been made (or information gaps have been filled), the VLA combines the analyses into the input for a neural network. Because the lexicon of keywords may grow very large, a more concise means of indicating lexical content is needed. One potential way to achieved this would be to group keywords into categories and then output relevant categories instead of keywords themselves.

**4:** The *renderer* decides how to render the image based on the qualities of the image and the provided description. The output of the VLA is used as input to a population of neural networks. The neural networks' outputs correspond to the priority of, and parameter settings for NPR filters contained within a modular database. Each of the resulting filter configurations will be applied to the original image, creating a series of artefacts.

NeuroEvolution of Augmenting Topologies (NEAT) [7] will be used to evolve the networks. This system was chosen because of its availability and support, robustness in online scenarios, and its ability to change the topology of the network. The latter is particularly useful, as the modular nature of many components in DARCI will require occasional changes to the input and output arrangement of the neural networks.

The fitness function of the algorithm is two-fold. First, DARCI self-evaluates each artefact. Second, once a particular network surpasses a fitness threshold, the resulting artefact will be output for external-evaluation—human evaluation.

### Self-Evaluation Pipeline

**5:** The artefact returns to the image analyzer where it is processed in the same way as the original subject image. The output, called the artefact analysis, proceeds to the VLA through the self-evaluation pipeline.

**6:** The VLA generates a language analysis to match the artefact analysis. This analysis is then compared to the original language analysis with a similarity metric. The VLA then generates a fitness value corresponding in part to the similarity of the two language analyses. In other words, DARCI positively evaluates artefacts that match their description. Other factors, such as relative originality and aesthetics, may also be taken into consideration. This fitness value is passed back to the renderer along with the series of keywords generated in the language analysis to act as a caption for the artefact (note that this is not the same language data passed to the renderer in the generative pipeline).

**7:** After being assigned its fitness value, each network that passes a fitness threshold has its artefact output to the external-evaluation pipeline with its caption affixed (otherwise, the caption is discarded). These chosen networks are then frozen in the evolutionary process until results from the external-evaluation pipeline return.

### External-Evaluation Pipeline

**8:** An evaluation for each of DARCI's "valued" artefacts is performed by humans who judge the quality of an artefact, how relevant the caption is, and also provide their own brief

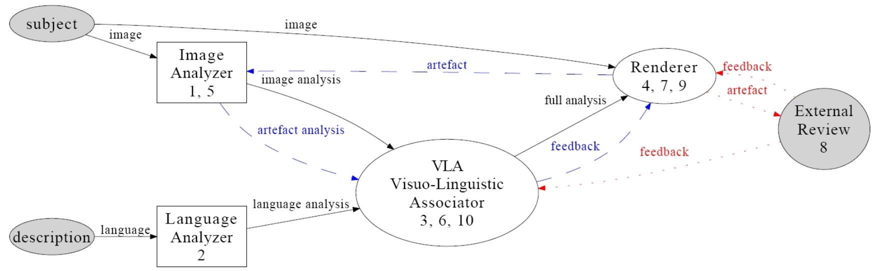


**Figure 1. Model of DARCI (Digital ARtist Communication Intention). Solid black lines show the generative pipeline, the pipeline concerned with creating an artefact (steps 1-4). Dashed blue lines show the self-evaluation pipeline, the process concerned with what DARCI thinks of her own work (steps 5-7). Dotted red lines show the external-evaluation pipeline, the pipeline concerned with how DARCI responds to external criticism (steps 8-10).**

keyword descriptions. After sufficient evaluation, the results are compiled and returned to DARCI. Results pertaining to the artefact as a whole are returned as a fitness function to the renderer; results pertaining to the caption, as well as the list of new keywords, are returned to the VLA.

**9:** The external fitness function is applied to the associated network. Positive feedback will increase the fitness of the network, while negative feedback will lower it.

**10:** Feedback concerning the relevancy of the caption may be used to modify existing associations in the VLA. Additionally, keywords provided by external-evaluators will help build associations.

## ANALYSIS OF CREATIVITY

Earlier we briefly mentioned two complementary definitions of creativity. Ritchie's attributing creativity to the artefact, and Colton's attributing creativity primarily to the process. Unfortunately, we can't evaluate any artefacts until DARCI produces them. So the discussion of Ritchie's 18 criteria will have to remain a discussion of how we intend to use the criteria in the future. However, whether or not our system possesses Colton's three behaviors can largely be determined by looking at DARCI's design. After exploring these two definitions of creativity, we will explain how DARCI will be able to communicate intent.

### Creative Artefacts

To meet our goal, DARCI needs to produce images that are considered creative. According to Ritchie, this means possessing novelty, quality, and typicality in such a way that satisfy at least some of 18 criteria. DARCI's work will periodically be evaluated by humans, but this is only a measure of quality and is meant for DARCI's own development. When assessing the creativity of DARCI's images, we will perform additional tests to obtain quality and typicality measures. Ritchie describes typicality as the extent to which an artefact belongs to the class of artefacts it is intended as. In other words, if an artefact is supposed to be a song, it should be recognized as a song whether or not it is novel or of high quality. The novelty of DARCI's artefacts will be determined as part of the 18 criteria.

### Creative Processes

The first quality Colton describes is skill. Through examples found throughout Colton's work we learn that skill refers to the system's ability to perform a task. The more challenging the task, the more skillful the program. DARCI is designed to produce renderings of images and so can be thought of as skillful. How skillful she is will be determined by the complexity and variety of filters made available to her.

The second necessary behavior Colton explains is appreciation. This means that the program must attribute some value to its work. DARCI evaluates her own work via the self-evaluation pipeline outlined in steps 5-7. DARCI's ability to ascribe value to her own work is more interesting than it may first appear. DARCI makes evaluations by comparing how visually similar an artefact is to its description. But, because the description is not required to relate to the subject image, unpredictable relationships may form in the VLA. This will result in an equally unpredictable, though not random, value metric for DARCI. Thus, over time DARCI will develop her own aesthetic, a trait Colton describes as an ultimate goal for automated painters. Other sources have even considered this ability to evolve an evaluation metric, to be critical in attributing creativity to a system [4]. While it is true that we provide the initial descriptions seeding DARCI's associations, in the future, DARCI will create her own descriptions making this process even more interesting.

In addition to evaluating herself, DARCI uses feedback from humans. This may appear an usurpation of the self-evaluation process; however, we feel that it is imperative that an agent striving for creativity be placed in a social environment where its value system can mature in a culturally meaningful way. Kyle Jennings makes a very good case for this, and proposes a formal method for building a model of social influence within an agent [5]. Eventually, we will investigate using this model with DARCI. As it stands, our method of assimilating external feedback, outlined in steps 9 and 10, will allow DARCI to develop more complex relationships in the VLA than she would otherwise.

Imagination is the third quality Colton describes. This means that the system is able to come up with meaningful and original artefacts rather than just regurgitating learned data. By combining two domains of input, an image and a description, in order to create a single output, an artistic image, DARCI is likely to produce original artefacts that convey some degree of meaning from the two domains. The originality of DARCI's artefacts can be further enhanced by implementing a more sophisticated fitness calculation than the one presented in step 6. We may eventually even consider implementing a fitness evaluation based on Ritchie's criteria.

### Communication of Intention

The communication of intent is a feature of DARCI that will be more fully developed once the core system outlined in this paper is implemented. That being said, there is communication of intent to a small degree in the current design.

The description constraint associated with a subject can be thought of as the intent of DARCI's teacher. Going back to a previous example, if we were to present the picture of a lion with the description "dark, scary", then the intent of the teacher would be for DARCI to render a lion in a dark and scary manner. This would not be DARCI's intention since it was provided to her as input.

As soon as DARCI has developed a robust series of relationships between image qualities and keywords, we will begin presenting DARCI with only an image. At this point it will be DARCI providing the description and thus the intent, not the teacher. It could be argued that we will have trained DARCI to have the intentions she does. But, as the system performs its own self-evaluations and takes into consideration external-evaluations, the relationships DARCI will have formed will become unpredictable and unique.

Concerning the communication of intention, by rendering an image according the constraints of the description, DARCI will be implicitly communicating her (or her teacher's) intent. By literally affixing the description to an artefact, DARCI will be able to explicitly communicate intent. When the descriptions become interesting enough (not just lists of key words), we will proceed to this step.

In future upgrades to DARCI's design, we hope that she will be able to create her own unique descriptions and subjects using a cognitive model and the VLA. At that point, both the subject and the description would be wholly DARCI's, along with the intention they convey. The architecture of DARCI is designed to be extensible, and so will accommodate these improvements.

## CONCLUSIONS

We have outlined the design of a computer system, called DARCI, that we have shown can be considered creative according to one definition of creativity [3]. That is, the design can be shown to represent a skillful, appreciative, and imaginative agent. We have also presented another attribute of our own, the ability to communicate intention, that we argue to be important and within the capability of DARCI. Furthermore we have referenced (the details are far to extensive to include in this paper) a method for empirically evaluating the creativity of the artefacts that DARCI will produce [6].

Once DARCI is implemented, we believe that she may be a useful tool in the study of human creativity. By analyzing deficits in the artefacts she creates, we may be able to learn why the attributes she possesses are insufficient for producing truly creative works. We may also learn more about the types of learning that can foster or inhibit creativity. Should we train the VLA offline until enough language-image relationships are formed, or will we get more creative results if we let it develop from scratch while part of the complete system? Because DARCI is modular in design, we will be able to experiment by adding or removing features from her architecture. For example, what would be the effect of suddenly adding or removing an analysis tool to the image analyzer? Even though DARCI is an artificial system, being able to answer questions like these could provide insights into human creativity.

Finally, the development of DARCI may also aid in artificial intelligence research. The goal of artificial intelligence is to emulate human intelligence to solve a variety of problems. Being creative is an attribute that allows humans to be the problem solvers we are. Any progress made in simulating this creativity will further the study of AI.

## REFERENCES


1. The painting fool. http://thepaintingfool.com/.

2. M. A. Boden. *The Creative Mind: Myths and Mechanisms (second edition)*. Routledge, 2004.

3. S. Colton. Creativity versus the perception of creativity in computational systems. *Creative Intelligent Systems: Papers from the AAAI Spring Symposium*, pages 14–20, 2008.

4. M. Elton. Towards artificial creativity. *Proceedings of LUTCHI Symposium on Creativity and Cognition*, 1993.

5. K. Jennings. Developing creativity: Artificial barriers in artificial intelligence. *Proceedings of the International Joint Conference on Computational Creativity*, pages 1–10, 2008.

6. G. Ritchie. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17:67–99, 2007.

7. K. O. Stanely. *Efficient Evolution of Neural Networks through Complexification*. PhD thesis, The University of Texas at Austin, 2004.