Artist in the Cloud: Towards an Autonomous Artificial Artist

Gene Kogan
gene@genekogan.com

Abraham https://abraham.ai

Objective

We propose to build Abraham, an agent which autonomously creates unique and original art.

The agent should satisfy the following criteria.

- **Autonomy**: Act independently of its authors.
- **Originality**: Exhibit novel creativity not derivative of its authors.
- **Uniqueness**: Artworks can not be reproduced by external means.

Strategy

We propose to frame the "agent" as an open decentralized network of collaborators who design, train, and curate a generative art program. The agent's output is a non-linear function of the collective actions of its contributors. We claim this constitutes autonomy. To support this claim, we make an analogy to the "collective intelligence," "superorganism," or "hive mind" metaphors in which a separate intelligence apparently emerges out of a complex system of interacting smaller agents.

Figure 1: examples of emergence or collective intelligence in nature (from left to right: coral reef, termite mound, bee hive, flock of birds)

By crowd-sourcing the construction and training of the generative art program, the agent's artistic palette emerges unpredictably from the decisions of its many contributors, and we claim this facilitates its originality. We compare this to a human artist learning an original style by amalgamating a unique set of influences.

To enforce decentralization, we seek to apply techniques from secure and privacy-preserving machine learning to make the generative model's parameters secret and its training data irreproducible, so that no individual party can copy or retrain the model. This ensures the uniqueness of the artworks, as no external process can recreate them.

Methodology

Deep generative models

We use one or more deep generative models as the basis for the generative art program. Generative models are favorable for decentralization, as we can crowd-source their training data. In contrast, a procedural generative art program is hand-engineered with complex heuristics, making it difficult for multiple parties to collaborate over.

Secure multi-party computation (MPC)

MPC protocols require multiple parties to jointly compute a function. For a neural network, each operation is split into several sub-operations, each of which is meaningless alone, but recombine to the correct output. This allows the parties to cooperate over training and running a model, without revealing the learned weights to any individual.

A variant is model splitting, a scheme where instead of jointly computing every operation, the model is split into pieces, which are distributed intact across parties. For example, it can be split simply by layers, which cuts down on communication costs, but at some cost to privacy.

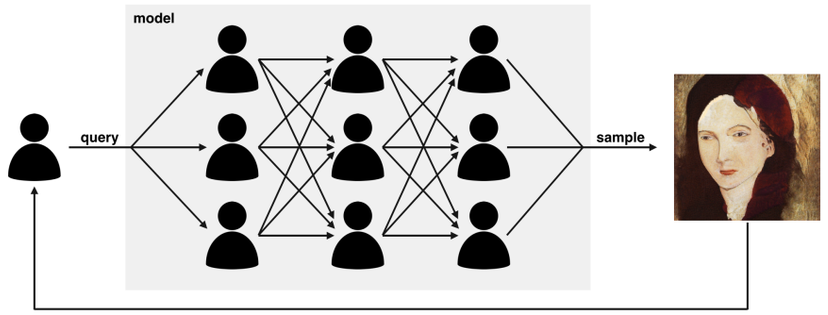

Figure 2: A neural network is split across a peer-to-peer network and held as a shared secret.

MPC enables a collective to securely hold the weights of a neural network as a shared secret. If training data is also private, the learned model is irreproducible, and no party can perform a forward pass alone or download the full model. This satisfies the uniqueness constraint.

Federated learning

A copy of the model is sent to individual clients who compute gradients on their own private data, and send them to a third party who securely aggregates them to update the model. With proper differential privacy, the original training data can never be reconstructed.

Motivations

The Abraham project is motivated by the opportunity to create a novel type of generative art which is based on collective intelligence and demonstrates intrinsic agency beyond that of previous AI art programs.

Another motivation is to serve as an educational vehicle and testing ground for experimental and controversial technologies from AI, distributed computing, and cryptography.

Inspirations and foundations

The idea of an autonomous artificial artist builds upon work from multiple fields, and is inspired by numerous prior and contemporary projects with similar goals.

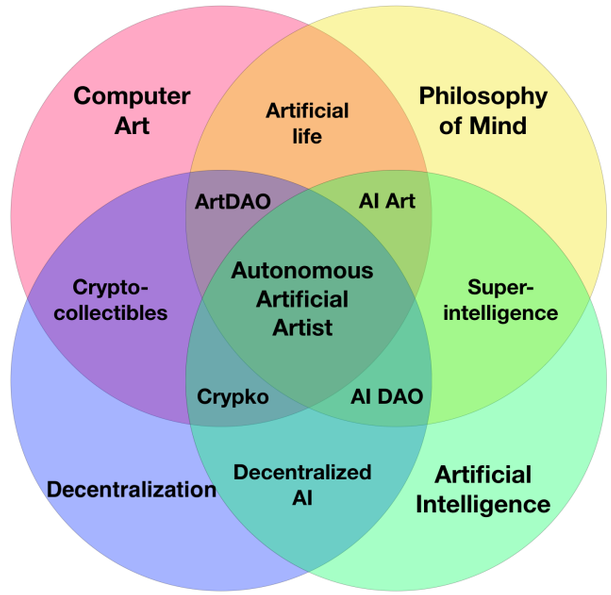

Figure 3: Intellectual foundations of autonomous artificial artists.

Experimental results

None yet :(
We are developing a small mobile demo in which a group of people train a federated variational autoencoder on images from their mobile cameras.

Future work

- Research each of the technical components.
- Make initial prototypes which meet limited versions of the full criteria towards an eventual proof-of-concept.
- Compile educational resources.