

# Mappa Mundi: An Interactive Artistic Mind Map Generator with Artificial Imagination

Ruixue Liu<sup>1</sup>, Baoyang Chen<sup>2</sup>, Meng Chen<sup>1</sup>, Youzheng Wu<sup>1</sup>, Zhijie Qiu<sup>2</sup> and Xiaodong He<sup>1</sup>

<sup>1</sup>JD AI, Beijing, China

<sup>2</sup>Central Academy of Fine Arts, Beijing, China

{liuruixue, chenmeng20, wuyouzheng1, xiaodong.he}@jd.com, {chenbaoyang, qiuzhijie}@cafa.edu.cn

## Abstract

We present a novel real-time, collaborative, and interactive AI painting system, Mappa Mundi, for artistic Mind Map creation. The system consists of a voice-based input interface, an automatic topic expansion module, and an image projection module. The key innovation is to inject Artificial Imagination into painting creation by considering lexical and phonological similarities of language, learning and inheriting artist's original painting style, and applying the principles of Dadaism and impossibility of improvisation. Our system indicates that AI and artist can collaborate seamlessly to create imaginative artistic painting and Mappa Mundi has been applied in art exhibition in UCCA, Beijing<sup>1</sup>.

## 1 Introduction

Mind Map is a diagram that connects relevant words, terms or events around a keyword or an idea and it's perceived as an artistic possibility representation of knowledge. Through constant interpretation and abstraction, the artistic creation of Mind Map is a process of subject decision making and information expansion. In this paper, we introduce Mappa Mundi, an interactive Mind Map generator featured with artificial imagination.

The creation of Mind Map requires the interaction between AI and artist on the decision making of idea expansion. While previous AI image generators, such as pix2pix [Isola *et al.*, 2017] framework, depends only on the model performance and less on taking artist's interaction into account. Besides, models constructed with neural approaches barely extract features from big data, whereas features, such as imagination and creativity, are far to reach. As Mind Map represents the universal mind of human thinking, the Kantian thinking of free play of imagination, which is the foundation of modern arts [Kant, 1987], plays an important role in expanding artist's thought and connecting the flourishing information. Different from the traditional framework, Mappa Mundi offers opportunities to enable AI and artist together during art making.

Mappa Mundi enables voice input from human and extracts the keywords for vivid expansion on myriad levels of

domains, which explores the impossibility of improvisation variation arising from standardisation. To improve the imaginative feature of word expansion results, Mappa Mundi employs not only semantic similarity and linguistic features but also establishes a knowledge graph to imitate artist's idea and style. Besides, Mappa Mundi breaks conventions and explores more possibilities for word's connections partially stem from the Dadaism art movement. Through interaction sparked in our system, the free play of human imagination and machine intelligence is consolidated, and eventually constitutes a contemporary Mappa Mundi, which cultivates a more and more electrified arena for endless imagination of both human and machine.

## 2 Related Work and Challenges

Nowadays, AI demonstrates stronger potential for art creation. Many researches have been conducted to involve AI into poem generation [Zhang and Lapata, 2014; Cheng *et al.*, 2018], creation of classical or pop music [Manzelli *et al.*, 2018; Hadjeres *et al.*, 2017] and automatic images generation [van den Oord *et al.*, 2016; Yan *et al.*, 2016; Xu *et al.*, 2018]. Whereas there are few researches exploring the possibility of artificial imagination for artwork creation.

To generate an imaginative and creative Mind Map, the key problem is automatic topic expansion. There are several challenges: first of all, language is an informative system. Thus many linguistic features should be considered in words and relation extension. Second, as an artwork, it should reflect artist's mind and individual experience. Third, for Mind Map creation, artists usually pay less attention on defining accurate relation between concepts, instead, they are defined with imagination. So, to build artistic Mind Map, the system should find a way to break the convention rules for the connection between literal representations.

## 3 System Architecture

The overall architecture of Mappa Mundi is illustrated in Figure 1. It consists of three modules in total. The first one is a voice-based input interface, which translates user's voice input into text and extracts keyword from it. The second is topic expansion module, which expands the keyword into several word candidates that share various connections with it. The third module is image projection, which takes the keyword

<sup>1</sup><http://ucca.org.cn/en/exhibition/qiu-zhijie-mappa-mundi/>

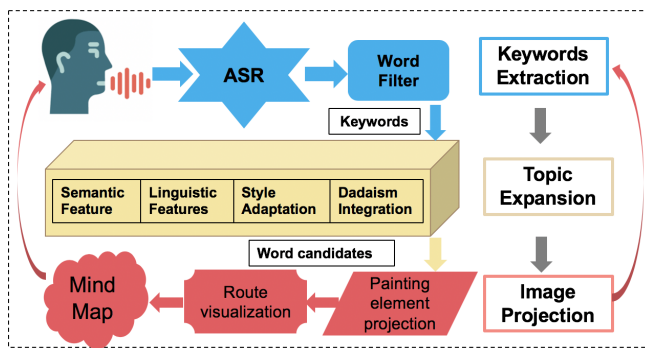


Figure 1: System architecture of Mappa Mundi

and its candidates as input and projects the concepts and relations into a Mind Map image. As an interactive system, Mappa Mundi presents artist’s speech as Mind Map image and this imaginative picture in return provokes artist’s aspiration for further thoughts expansion.

### 3.1 Voice-based Input Interface

Automatic Speech Recognition (ASR) engine is designed for interaction. Both English and Chinese languages are supported. We also performed lexicon and language model adaptation to enhance the recognition of words in art domain. After that, we developed a keyword extraction function to extract meaningful words from the converted text for further expansion. We adopted open-source software jieba<sup>2</sup> for Chinese and Stanford parser [Toutanova and Manning, 2000] for English POS tagging. Only informative words such as noun, verb and adjective words are kept and TFIDF weights are calculated for further filtering. The output of this module is the keywords.

### 3.2 Topic Expansion

As imagination is the soul for artistic Mind Map, Mappa Mundi employs several features to increase information variety [Liu *et al.*, 2019] during topic expansion. It firstly uses word embeddings to find candidates based on semantic similarity [Mikolov *et al.*, 2013; Pennington *et al.*, 2014; Peters *et al.*, 2018]. To enrich linguistic information of expansions, it also takes the morphological and phonological features into account. Words sharing similar characters or morphemes or phonetic syllables are selected as candidates. Moreover, we construct a knowledge graph (KG) by extracting concepts and relations from artist’s original Mind Map masterpieces. Mappa Mundi will adopt those concepts and relations once a topic word is covered by the KG. By doing so, it can better imitate artists’ mind and thoughts. Further, Mappa Mundi breaks the restriction of domains and conventions and explores more possibilities for words’ connections by creating rules following the famous Dadaism principle in art movement [Elger, 2004]. Finally, with all above mentioned methods, Mappa Mundi can create a branch of creative and informative word candidates for a given keyword.

<sup>2</sup><https://github.com/fxsjy/jieba>



Figure 2: Screen-shot of a generated Mind Map image with voice input of *Beijing Bids for 2022 Winter Olympics*

### 3.3 Image Projection

To visualize the abstract concepts, Mappa Mundi includes around 3000 painting elements which are extracted from the representative Mind Map paintings of a famous artist<sup>3</sup>. These elements are sharing traditional Chinese Shan-shui (Mountain-river) painting style and are further analyzed into 5 typical Shan-shui painting categories including *architecture, mountain, river, grassland and lake*. Mappa Mundi will classify each word into one of the five categories above and then connect it with the elements in the corresponding domain. An additional type of painting element *Route* is also involved. The distance between the concept and its candidate is determined by their similarity score and then visualized as *Route* in image. The more similar these two words are, the closer distance will be.

Our Mappa Mundi is not a one-shot completion, it is an interaction of live stimulations, which contains constant vocal inputs via ASR and outputs of generated words and their relations. It is a two-way street. The generated information, after being presented in our system, becomes the inspiration for artist’s next vocal input. This artwork reflects both the development of artist’s thinking and the AI-enabled imagination.

## 4 Conclusion

In this paper, we propose a Mind Map generator, Mappa Mundi, that can inject Artificial Imagination together with human interaction into an artwork. The way of enriching artificial imagination includes considering semantic similarity, integrating informative linguistic features, inheriting artist’s mind, and inserting Dadaism and impossibility of improvisation principles. Mappa Mundi is presented to understand how AI can aid artistic practice. Its development also indicates AI has more implications in art than a way of artistic creation, and it enriches the possibility of artistic practice.

## Acknowledgments

We want to thank Yan Dai, Chuyan Xu, Haozhen Feng and Xiaoyu Guo for their work and support to build this system.

<sup>3</sup>[https://en.wikipedia.org/wiki/Qiu\\_Zhijie](https://en.wikipedia.org/wiki/Qiu_Zhijie)

## References

- [Cheng *et al.*, 2018] Wen-Feng Cheng, Chao-Chung Wu, Ruihua Song, Jianlong Fu, Xing Xie, and Jian-Yun Nie. Image inspired poetry generation in xiaoice. *arXiv preprint arXiv:1808.03090*, 2018.
- [Elger, 2004] Dietmar Elger. *Dadaism*. Taschen, 2004.
- [Hadjeres *et al.*, 2017] Gaëtan Hadjeres, François Pachet, and Frank Nielsen. Deepbach: a steerable model for bach chorales generation. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1362–1371. JMLR. org, 2017.
- [Isola *et al.*, 2017] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *CVPR*, 2017.
- [Kant, 1987] Immanuel Kant. *Critique of judgment*. Hackett Publishing, 1987.
- [Liu *et al.*, 2019] Ruixue Liu, Baoyang Chen, Xiaoyu Guo, Yan Dai, Meng Chen, Zhijie Qiu, and Xiaodong He. From knowledge map to mind map: Artificial imagination. *Proceedings of the Second International Conference on Multimedia Information Processing and Retrieval*, pages 496–501, 2019.
- [Manzelli *et al.*, 2018] Rachel Manzelli, Vijay Thakkar, Ali Siahkamari, and Brian Kulis. Conditioning deep generative raw audio models for structured automatic music. *arXiv preprint arXiv:1806.09905*, 2018.
- [Mikolov *et al.*, 2013] Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 746–751, 2013.
- [Pennington *et al.*, 2014] Jeffrey Pennington, Richard Socher, and Christopher Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014.
- [Peters *et al.*, 2018] Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*, 2018.
- [Toutanova and Manning, 2000] Kristina Toutanova and Christopher D Manning. Enriching the knowledge sources used in a maximum entropy part-of-speech tagger. In *Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics-Volume 13*, pages 63–70. Association for Computational Linguistics, 2000.
- [van den Oord *et al.*, 2016] Aaron van den Oord, Nal Kalchbrenner, Lasse Espeholt, Oriol Vinyals, Alex Graves, et al. Conditional image generation with pixelcnn decoders. In *Advances in Neural Information Processing Systems*, pages 4790–4798, 2016.
- [Xu *et al.*, 2018] Tao Xu, Pengchuan Zhang, Qiuyuan Huang, Han Zhang, Zhe Gan, Xiaolei Huang, and Xiaodong He. Attngan: Fine-grained text to image generation with attentional generative adversarial networks. In *2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2018.
- [Yan *et al.*, 2016] Xinchun Yan, Jimei Yang, Kihyuk Sohn, and Honglak Lee. Attribute2image: Conditional image generation from visual attributes. In *European Conference on Computer Vision*, pages 776–791. Springer, 2016.
- [Zhang and Lapata, 2014] Xingxing Zhang and Mirella Lapata. Chinese poetry generation with recurrent neural networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 670–680, 2014.