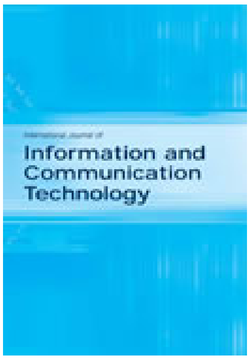
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Application of knowledge graph-enhanced generative diffusion model for brand visual generation

Fengfeng Shi

DOI: 10.1504/IJICT.2025.10075270

Article History:

Received:	26 August 2025
Last revised:	12 November 2025
Accepted:	13 November 2025
Published online:	12 January 2026

Application of knowledge graph-enhanced generative diffusion model for brand visual generation

Fengfeng Shi

College of Cultural Creativity and Tourism,
Yuncheng Vocational and Technical University,
Yuncheng 044000, China
Email: shiff136220@126.com

Abstract: As brand visual content becomes more important in marketing, it has gotten harder to make images that fit the brand's identity. This research presents a generative diffusion approach utilising knowledge graph improvement for brand visual generation to tackle this issue. By putting the brand knowledge graph into the generation process, the system gives semantic direction for making images. The first step in the plan is to construct a knowledge graph that shows the brand's main qualities. Then, a generative diffusion model based on the knowledge graph is made and tested to see how well it works for brand visual production. The model enhances the inception score (IS) by 21.3% and diminishes the Fréchet inception distance (FID) by 19.5% in comparison to the conventional generative model. The model makes pictures that are consistent with the brand and seem beautiful, with good brand customisation and innovation.

Keywords: knowledge graph; generative diffusion model; brand visual generation; brand consistency; visual quality.

Reference to this paper should be made as follows: Shi, F. (2025) 'Application of knowledge graph-enhanced generative diffusion model for brand visual generation', *Int. J. Information and Communication Technology*, Vol. 26, No. 51, pp.50–69.

Biographical notes: Fengfeng Shi received her Master's degree in the North University of China in 2018. She is currently a Lecturer at the Yuncheng Vocational and Technical University. Her research interests are art design and knowledge graph.

1 Introduction

Brand visual elements, such logos, advertising imagery, product packaging, and so on, show the brand's basic values and cultural ideas. They can also show how the brand is different from others. When it comes to traditional brand visual design, the designer's experience and gut feeling are usually all they have to go on. The designer's personal style, market demand, brand positioning, and other considerations might also limit their creativity. These aspects are vital in the design process, but there is also some subjectivity, and the results do not always satisfy the needs of brand communication. As computer vision technology has quickly advanced, data-driven brand vision generation has become a major area of study (Quan et al., 2023).

The difficulty of creating brand visuals is not only in how good the images look, but also in how to make them fit well with the company's unique traits, market position, customer preferences, and other aspects. Knowledge graph has become a significant area of research in brand visual generation in recent years to solve this challenge. Knowledge graph is a systematic technique to show knowledge and information in a graphical style and show the many levels and dimensions of relationships between objects. In the brand visual generation task, the knowledge graph can give the generation model a lot of background and semantic information. This helps the model create the semantic guidance for the image, so that the image not only looks good but also accurately shows the brand's traits. Knowledge graphs can give more accurate guidance for the generation process by turning information about the brand's core values, consumer preferences, market trends, and other factors into a graph and combining it with the generation model. This makes sure that the images that are generated fit the brand's needs and position perfectly.

Knowledge graphs have a lot of uses in many fields, but they still have certain problems when it comes to making brand vision. First, brand-related knowledge graphs often have information that is broken up and not updated quickly enough, which makes them less useful as guides in the generating process. Second, it is still a technological problem to figure out how to turn the abstract ideas and complicated connections in a knowledge graph into useful information that the generative model can understand (Zhong et al., 2023). At present, the building of numerous knowledge graphs mostly emphasises the aggregation of static information, resulting in a deficiency of dynamic updates and tailored adaptations of brand information, which therefore impairs their flexibility in practical applications. So, figuring out how to make and use brand knowledge graphs and integrate them with generative diffusion models has become a big difficulty in the field of brand visual generation.

This research provides a generative diffusion model based on knowledge graph enhancement to address the aforementioned issues, with the objective of enhancing the quality and relevance of brand visual generation. The main new idea in this paper is to use a knowledge graph with a generative diffusion model. The knowledge graph gives semantic guidance to the generation process, so the images that are made not only look great, but they also accurately show the brand's traits, core values, and the needs of the customers (Dey, 2024). This work introduces an innovative generative diffusion model architecture that incorporates brand knowledge information during the generation process. With the semantic guidance of the knowledge graph, the resulting images can better align with the specific requirements of the brand. This study presents a novel technical method in the domain of brand visual generation, aiming to address the shortcomings of current technologies and enhance the relevance and novelty of the created images.

The study aims of this paper concentrate on three principal aspects. First, to create a knowledge graph with a lot of brand information that can help the generation process by giving it more accurate semantic guidance. Second, create and use a generative diffusion model that builds on knowledge graph enhancement to make the most of the knowledge graph's role in the generation process and improve the quality and brand consistency of the images that are generated. Lastly, experimental validation is performed to assess the model's efficacy and superiority in brand visual generation tasks through comparative experiments, and to confirm the practical applicability of the proposed method in this paper.

The contribution of this thesis is in the introduction of a novel approach for brand visual production, as well as the provision of an innovative technical concept through the integration of a knowledge graph and a generative diffusion model. The process not only makes the photographs better, but it also makes the connection between the images and the brand stronger, which makes brand visual generation work better. This work also goes into great depth on how to build a dataset, pre-process it, and train a model, which will be useful for future research.

This chapter outlines the study context, rationale, and constraints of current methodologies, while delineating the research aims and contributions of this paper. Next, Section 2 will give an overview of related work, focusing on the current state of research on knowledge graphs, generative diffusion models, and brand visual generation. It will also look at the pros and cons of the current methods to set the stage for future research. Section 3 will go into great depth on how to design and use the generative diffusion model that this study suggests based on improving knowledge graphs. Section 4 will present the experimental data and offer a comprehensive analysis of the model's performance. Section 5 will finally sum up the main points of this paper, talk about the research's limitations, and look ahead to where future research might go.

In general, the research in this paper is very important for the field of brand visual generation, both in terms of theory and practice. This paper integrates knowledge graphs with generative diffusion models, offering a novel technical approach for brand visual generation and robust technical support for brand marketing, advertisement design, and product packaging. The research findings of this work offer innovative concepts for the advancement of brand visual generation and furnish technological assurance for the implementation of associated applications.

2 Relevant work

2.1 Knowledge graph

A knowledge graph is a structured network that uses nodes and edges to show how things, concepts, and their relationships in knowledge are related to each other. The main idea behind knowledge graphs is to help computers understand, store, and think about complicated information in the human world by organising and representing it in a logical way. Since Google formally proposed the concept of knowledge graph in 2012, knowledge graph, as an important research topic in the field of AI, has received widespread attention and achieved important applications in several fields (Peng et al., 2023). Knowledge graphs can not only provide rich semantic information, but also provide strong support for various intelligent tasks, such as natural language processing (NLP), recommender systems, semantic search, etc.

The basic building blocks of a knowledge graph include entity, relation and attribute. Entity is the basic building block in a knowledge graph and represents a specific thing, concept or event. In a cinema knowledge graph, the director, the actor, and the film are all entities. On the other hand, relationships show how entities are connected or interact with each other, as a director directing a movie or an actor playing in one. Attributes are things that describe or define the entities, such the year the movie came out, the director's nationality, and so on. By connecting these entities, relationships and attributes, the knowledge graph can form a comprehensive knowledge network that provides complex knowledge reasoning and speculation capabilities for intelligent systems (Chen et al., 2022).

The creation of a knowledge graph usually involves multiple steps. In the data collection phase, it is necessary to obtain information from a variety of sources, which can come from text, databases, web pages, social media, and so on. Through information extraction techniques, entities, relationships and attributes are extracted from the massive unstructured or semi-structured data, and a preliminary knowledge graph is constructed based on this information. Then the data collected from different sources need to be integrated to ensure the accuracy of the knowledge graph. The extracted entities and relationships are then transformed into a format that can be understood by computers (Nasar et al., 2021). Finally, the knowledge inference phase deduces the implied information through in-depth analyses of the knowledge graph to provide richer knowledge services to users.

Knowledge graphs are also very crucial for recommender systems. Collaborative filtering or content filtering are used to make recommendations in classic recommender systems; however, these algorithms typically miss the deeper semantic information that explains why users do what they do. On the other hand, recommender systems can better comprehend what users like and need by using knowledge graphs. This lets them give customised recommendations based on what users already know, what they like, and other related information. For instance, in a movie recommendation system, the knowledge graph can not only keep track of what movies the user has watched, but it can also use information like the director, actors, and genre of the movie to provide better and more varied recommendations.

Many fields employ knowledge graphs. First, knowledge graphs can aid NLP tasks including text understanding, information extraction, named entity recognition, and sentiment analysis with semantic assistance. Second, knowledge graphs improve search engine semantic understanding and accuracy. Smart question-and-answer systems are another major use. Q&A systems can understand entities and relationships in a question and reason from knowledge graphs to obtain accurate responses.

Despite the remarkable applications of knowledge graphs in several fields, their development still faces some challenges. First, the construction of knowledge graphs requires a large amount of data support, and there are great challenges especially in cross-domain and multi-modal data fusion. Knowledge graphs from different domains may have semantic differences, requiring unified modelling of knowledge from different domains, which is often very complicated in practical applications (Ji et al., 2021). Second, it is also a technical challenge to effectively extract valuable information from massive data and transform it into entities and relationships in a knowledge graph. Currently, the construction of most knowledge graphs relies on manual annotation and rule-based reasoning, lacking automation and intelligence. With the development of big data and AI technology, how to use automation and deep learning (DL) technology to improve the efficiency and accuracy of knowledge graph construction has become an urgent problem. In addition, updating and maintaining the knowledge graph is also an important issue.

Nevertheless, the great potential of knowledge graph in intelligent applications still makes it a research hotspot in the field of AI. With the continuous progress of technology, especially driven by NLP, machine learning (ML) and graph databases, the application of knowledge graphs will become more and more widespread and play an increasingly important role in the future intelligent society.

2.2 Generative diffusion model

There are usually two basic parts to generative diffusion models: the forward process and the inverse process. The forward process is a way to add noise to data samples that does not change over time. It normally works by slowly adding noise to the samples until they are just noise. The model adds noise to the samples at each step based on their current condition. This makes the samples slowly move away from the original data distribution. The opposite process, on the other hand, is a process of learning denoising, in which the model slowly turns pure noise samples back into true images or data. To make the reverse process work, you need to train a denoising network so that the model can get rid of noise at each step and slowly bring back the true samples.

The diffusion model is better since it makes better generations. Diffusion models are more stable throughout training than classic generative adversarial networks (GANs) and are less likely to run into the mode collapse problem (Cobbinah et al., 2025). Because the generator and discriminator are playing a game in GAN, the training process could lead to the generator making samples that are too similar and do not have enough variety. Diffusion models, on the other hand, use a step-by-step denoising procedure to get around this obstacle, which lets them create more varied and high-quality samples.

The diffusion model is also quite easy to adjust when it comes to making images. Researchers can change the style, quality, or details of the findings by changing how the noise is removed at each step of the inverse process. This is because the creation process is done by step-by-step denoising. By properly training the diffusion model, it is possible to control some features, including the style, colour, or topic matter of the images it makes. Because they can be controlled, diffusion models are quite popular for making images, especially in the domains of art and design.

Despite its significant advantages, the diffusion model has some challenges. Firstly, the process of generating samples is more time-consuming than other generative models because the diffusion process requires noise addition and removal through multiple iterations. This means that diffusion models normally need more computer power and time to create high-quality samples. In actual life, this difficulty could make it hard to apply diffusion models, especially for jobs that need to be done quickly. Also, diffusion models normally take longer to train, especially when the datasets are complicated. The training procedure may take longer and need a lot of computing power.

Second, diffusion models could have trouble working with big datasets. Diffusion models typically necessitate substantial training data to comprehend the generation process, particularly in high-resolution image generating applications where extensive training data and computer resources are essential. Getting a lot of high-quality training data can be hard in some fields, especially when the data is specific to that field (Dou et al., 2023). For instance, in medical picture creation, the expense of data collecting and labelling is significant, making the effective training of diffusion models with limited data an urgent challenge.

Additionally, diffusion models exhibit great generating quality and stability; yet their generative process lacks interpretability. The generation process of diffusion models is more complicated than that of some classical generative models, and it does not have a simple way to explain it. This means that researchers have a hard time figuring out exactly what the model does at each step of the creation process. Generative diffusion models could not be useful in some situations where a high level of interpretability is needed, such medical diagnosis and financial forecasts, because they are hard to understand.

Even though diffusion models have certain problems and limitations, their benefits in terms of generation quality, stability, controllability, and diversity have made them an important area of research in DL in recent years. As processing power increases and algorithms are constantly improved, diffusion models are likely to make bigger advances in future study and be employed in many areas. The diffusion model has a lot of potential uses, such as picture production, NLP, image restoration, super-resolution reconstruction, and more. As model training methods get better and computers get faster, diffusion models will be used in more domains, and their generation effects will get better. They will become an important aspect of AI generation models.

2.3 Brand vision generation

Brand visual generation is a key part of brand communication. Its goal is to use visual design to show off the brand's uniqueness, core values, and cultural meanings. As technology has advanced, brand visual production has moved from traditional manual design to automated design that uses computer graphics, ML and DL. Generative models, especially GAN and generative diffusion models, have given brand visual generation new options in the last few years. This has made building a brand image smarter, more personalised, and more efficient.

In the past, designers relied a lot on their gut feelings and experience when making brand visuals. Designers do visual designs to make things like logos, ads, packaging, and other visual content that fits with the brand's image and is appealing to the target demographic and market needs (Trehan and Kalro, 2024). This technique relies heavily on the designer's unique skills and originality. However, conventional design frequently lacks universality and a high level of innovation because brand visual development is subjective. Also, manual design takes a long time and costs a lot of money, especially when you need to develop and update a lot of brand visual aspects at once. Traditional approaches are less efficient.

Digital design slowly took the place of manual design as computer visuals got better. The late 1980s saw the rise of graphic computers and the release of professional graphic design software. This took brand visual generation into the digital age. During this time, computer-aided design technologies made the design process faster, but brand visual generation still depended on designers' ingenuity and hard work. Graphic design tools like Adobe Illustrator and CorelDRAW made it possible for designers to do graphic design and editing on computers, which made design far more accurate and flexible (Ametordzi and Olalere, 2024). But these tools still cannot replace the human designer, and they cannot automate smart design either.

The 21st century has seen fast progress in DL and AI, which has changed the way brands create visions. Deep neural networks (DNN), convolutional neural networks (CNN), and other technologies have made it possible for computers to learn and grasp the structure, style, and features of an image on their own (Khan et al., 2020). As a type of generative model, GAN lets computers make realistic visuals through adversarial training. It has slowly become a significant tool for creating brand vision. Ian Goodfellow and others came up with the GAN model. Ian Goodfellow et al. came up with the GAN model in 2014. It works by having a generator and a discriminator play a game, and the generated image can realistically replicate the distribution of the training data.

On the basis of GAN, generative diffusion model, as a new type of generative model, has gradually become a research hotspot in brand visual generation. Diffusion models are frequently utilised in many sectors, such as art, product design, and other areas, to make high-quality images. The generative diffusion model may make the brand visual content that it creates more detailed, realistic, and in accordance with the brand's style and positioning by slowly changing the noise reduction method.

Brand visual generation has also used other methods besides GAN and diffusion models. For instance, generative models based on variational auto-encoders (VAEs) are used to make and rebuild brand images. By learning the distribution of latent space, VAE can make images that are both diverse and powerful. VAE is used in brand vision generation for activities like automated design and tailored recommendations. For instance, models based on VAE can make visual designs that follow brand rules depending on user-provided brand information.

Brand visual generation research encompasses the integration of multimodal learning and knowledge graphs, alongside simply generative models. Multimodal learning is when you learn by using the connections between different types of information, such text, images, videos, and so on. Multimodal learning may blend a brand's textual information (such its name, position, cultural history, etc.) with image data to make visual material that better fits the brand's personality. For instance, a DL model can use semantic information in a brand's text to combine a text description of the brand with an image generating assignment to make a visual design that fits the brand image. Multimodal learning not only improves the semantic information in the images it creates, but it also gives higher-level direction for creating brand visuals.

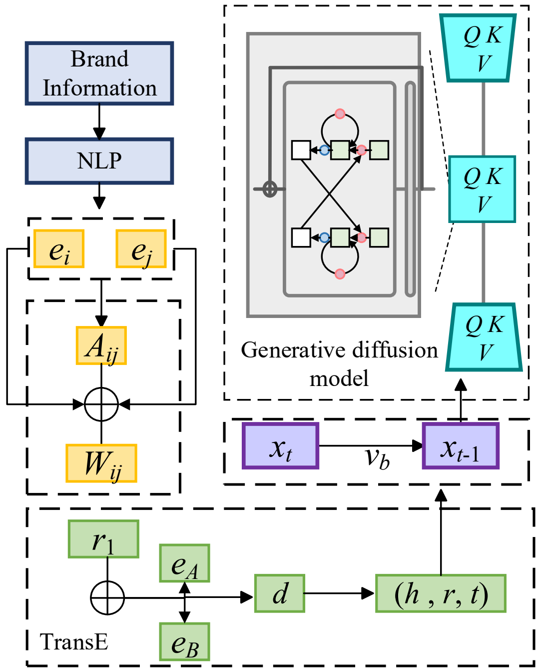
Brand visual generation also uses knowledge graphs. Knowledge graph is a systematic way to describe knowledge. It can give the generative model a lot of semantic information, which helps the generative system comprehend the brand's traits, values, cultural background, and other things. By using a knowledge graph to characterise the brand's different traits and information and combining it with the generative model, the system may make visual material that fits the brand image better. The knowledge graph and generative model work together to make brand visual generation smarter and more accurate, and they also make it easier to create customised and unique designs.

Brand visual generation is also being used in more situations. Brand visual generation has gradually moved into areas like virtual reality (VR), game development, and e-commerce, in addition to more typical design duties like brand logos, advertising ideas, and product packaging. In summary, the research on brand visual generation has experienced a shift from traditional design methods to AI technology, and the introduction of generative modelling has provided a more intelligent and efficient solution for brand visual design.

3 Design of knowledge graph-enhanced generative diffusion model

The generative diffusion model based on knowledge graph enhancement proposed in this paper aims to improve the quality and consistency of brand visual content by combining the semantic information of the brand with the image generation capability of the generative model. The design of the model consists of five key components covering from the construction and embedding of the knowledge graph to the fusion and optimisation during the generation process, sees Figure 1.

Figure 1 Design of generative diffusion model enhanced by knowledge graphs (see online version for colours)



3.1 *Knowledge graph construction*

A knowledge graph is a structured semantic representation that uses graphics to show entities, properties, and relationships between entities. This gives the creation process a lot of semantic information. For instance, the name ‘Apple’ might be seen as an entity, and its traits might be ‘innovation’, ‘technology’, ‘high-end market positioning’, ‘positioning’ and so forth. The knowledge graph will be made up of these entities and their properties. These entities and their characteristics will comprise the fundamental components of the knowledge graph. After that, relationship extraction methods are employed to find the connections between the entities. The connection between ‘Apple’ and ‘innovation’ is ‘brand characteristics’, whereas the connection between ‘Apple’ and ‘innovation’ is ‘brand features’. For instance, ‘brand identity’ is the link between ‘Apple’ and ‘innovation’, while ‘market positioning’ is the link between ‘Apple’ and ‘premium market positioning’. The construction of a knowledge graph provides semantic guidance for the generative diffusion model in the task of brand visual generation. This means that the images that are created not only have great visual effects, but they also accurately show the brand’s core values, unique features, and market position. So, building a knowledge graph is quite important for this paradigm.

To build a brand knowledge graph, you first need to get essential information about the brand from a number of different data sources. This information contains things like the brand name, brand values, brand culture, market positioning, product categories, target consumer groups, and so on. These kinds of information are often found in unstructured form in many texts, like the brand’s official website, ads, social media posts, customer reviews, and more. To turn this data into a structured knowledge graph, we need to use NLP methods, specifically information extraction methods, to work with and analyse text data.

The initial step in information extraction is recognising entities, which means finding terms in the text such brand name, product kind, market positioning, etc. The next stage is to build the graph's nodes and edges after getting the entities and relationships. The relationships between entities are the edges of the graph, and each entity is a node. Each edge of the graph will have a unique property that tells you what kind of relationship there is between the entities (Fan et al., 2024). For instance, the connection between the brand 'Apple' and 'innovation' can be shown as an edge with the property 'brand characteristics'. In this method, different kinds of information about a brand may be put into a graph, which can then be used by later generative diffusion models.

We also need to improve the graph so that the brand knowledge graph we made works well in the generating process. The optimisation process involves getting rid of unnecessary nodes and edges, making the network more hierarchical, and improving the semantic links between entities and connections. Some common ways to optimise are filtering entities by weight, optimising relationships by semantic similarity, and so on. These optimisation techniques can guarantee that the graph's information is not only complete but also semantically consistent.

Mathematical models are often used to show how strong the link is between entities and relationships when building a knowledge graph. The neighbourhood set of a graph is one way to show how things are related to each other. If we have a set of brand entities E, we may show the graph's neighbour set as:

$$A_{ij} = \begin{cases} 1 & \text{if there is a relationship between } e_i \text{ and } e_j \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where A_{ij} shows if there is a link between node e_i and node e_j. This set helps us figure out how brand entities are related to each other throughout the construction process and helps us train the model later.

In addition, each node and edge of the knowledge graph may also be assigned a different weight value indicating the relative importance of that node or edge. Specifically, a weight set W may be used to represent the weight of each edge, defined as:

$$W_{ij} = f\left(e_i, e_j\right) \tag{2}$$

where $f(e_i, e_j)$ is a weight function that finds the weights between entities e_i and e_j based on how important they are. The knowledge network gives the generative model semantic information and also changes the effect of different brand aspects on the picture generating process by using weights.

After the knowledge graph is built and improved, it can be combined with the generative diffusion model to be a major conditional input in the model creation process. The knowledge graph's nodes and edges, notably the brand's fundamental information and market position, will give the generative diffusion model and image generation a lot of semantic contexts to work with (Senior et al., 2025). The main purpose of this procedure is to use the knowledge graph to make the generated images look more like the brand and show how distinct and customised the brand is.

To sum up, building a knowledge graph is not only the first step in making a generative diffusion model, but it also gives the model better semantic guidance. By creating a complete and optimised brand knowledge graph, we can give high-quality brand information to help the next generation process, which will make the images that are made better quality and more consistent with the brand.

3.2 Knowledge graph embedding and coding

After constructing a brand knowledge graph, the next step is to transform its information into variable representations usable by the generative diffusion model, a process known as knowledge graph embedding and encoding. In this paper, translation-based embedding (TransE) technique is used for embedding and encoding to ensure that the semantic information of the brand knowledge graph can be efficiently transferred to the generative model.

The core idea of the TransE model is to map the entities and relationships in the graph to the same variable space, such that the relationships in the graph can be represented by the translation of the variables (Guo and Zhao, 2022). The TransE model expects to be realised by variable manipulation:

$$e_B = e_A + r_1 \tag{3}$$

where e_A and e_B denote the variable representations of brand A and brand culture B, respectively, and r_1 is the relationship variable. In practice, the TransE model learns the variable representations of entities and relationships by optimising the following objective function:

$$L = \sum_{(h,r,t)\in S} [\gamma + d(e_h + r_r - e_t)]_+ \tag{4}$$

where S is the set of all triples (h, r, t), h and t are the head and tail entities, r is the relationship between them, $d(\cdot)$ is the Euclidean distance metric, γ is a constant margin, and $[\cdot]_+$ is the maxima operation, which makes sure that the distances of the negative samples are greater than the distances of the positive samples.

The TransE model learns several ways to represent each item and relationship by minimising this loss function. This lets relationships between entities be shown by simple variable addition operations. The generative diffusion model will use these variable representations as inputs to help the model semantically guide the brand message while it is being generated.

In brand visual generation, the knowledge graph's inherent information gives the generative model exact semantic direction. The generative diffusion model can use the information about brand culture and market positioning in the knowledge graph embedding to come up with advertising ideas, product packaging, or brand logos that are very similar to the brand's characteristics (Yenduri et al., 2024). This semantic assistance derived from knowledge graph embedding can markedly enhance the brand relevance and customisation of the generated images.

In short, knowledge graph embedding and encoding methods are very important for making brand visuals. We can successfully incorporate semantic information from the brand knowledge graph into the low-dimensional variable space and transfer this information to the generative diffusion model by using the TransE model. This gives the generation process substantial semantic guidance. In the end, this method not only makes the photos better, but it also makes sure that they fit in well with the brand's core values and culture.

3.3 Design of generative diffusion model

In the implementation of the diffusion model, the denoising operation can be accomplished by training a conditionalised denoising network, where the model is updated at each step based on the noisy image at the current time step as well as the conditional inputs, (i.e., the embedded information in the brand knowledge graph), to generate images that are more in line with the brand's requirements. Assuming that the time step in the generation process is t, the diffusion process can be described as:

$$x_t = x_{t-1} + \alpha_t \left(x_0 - x_{t-1}\right) \tag{5}$$

where x_t is the image at time step t, α_t is the diffusion coefficient at each step, which controls the rate of denoising, and x_0 is the final generated image. The model may slowly make a visually realistic image from the original noisy image that matches the brand features by continuously denoising and adjusting it.

To ensure that the semantic information in the brand knowledge graph successfully guides the creation process, we use the embedded variables in the knowledge graph as conditional inputs to the denoising process at every step. The embedded variable v_b of the brand knowledge graph is employed as conditional information that affects the picture generation process at each step of the model (Sun et al., 2024). In each denoising stage, the brand information in the embedded variables will work with the picture content to make sure that the resulting image not only looks good, but also accurately shows the brand's essential values and the demands of the consumer. For instance, when making an advertising poster, the knowledge graph's information about the brand's personality, market position, and other things will affect the colour tone, composition, and elements of the image. This way, the final advertising poster will fit with the brand's overall image.

The conditional input v_b is added at each denoising step in the diffusion process. The denoising network's goal is not only to make the noisy image as close to the final target image as possible, but also to find the best way to improve the brand consistency of the generated image using the conditional input v_b. The following equation can be used to show this process:

$$L_{\text{total}} = L_{\text{noise}} + \lambda \cdot L_{\text{brand}}\left(v_b\right) \tag{6}$$

where L_{noise} is a standard denoising loss that shows how different the generated image is from the real image, $L_{\text{brand}}(v_b)$ is a brand consistency loss that shows how consistent the generated image is with the brand knowledge graph's embedded information, λ is an important hyperparameter that balances the noise loss and the brand consistency loss.

In summary, the generative diffusion model designed in this paper combines the powerful image generation capability of the generative model with brand-specific semantic guidance by introducing information from the brand knowledge graph.

3.4 Integration of knowledge graph information with diffusion process

In the generative diffusion model, the introduction of knowledge graph information provides important semantic guidance for the image generation process, especially in the brand visual generation task, which can help the generated images better reflect the core features of the brand, its market positioning and consumer needs.

Specifically, the brand information in the knowledge graph is fused with the current noisy image at each stage of generation. Assuming that at step t, the generated noisy image is x_t and the embedding variable of the brand knowledge graph is v_b, the updating process of the diffusion model can be expressed as:

$$x_t = x_{t-1} + \alpha_t \left(x_0 - x_{t-1} \right) + \beta_t v_b \tag{7}$$

where α_t controls the rate of noise removal and β_t is a hyperparameter that regulates the impact of the embedded information of the brand knowledge graph on the denoising process.

The model not only reduces the distance between the noisy image and the final target image at each phase of generation, but it also increases the match between the created image and the brand attributes (Tian, 2020). If $f(x_0)$ represents the visual attributes of the created image and v_b is the brand embedded variable, then the brand consistency loss can be written as:

$$L_{\text{brand}} = \left\| f\left(x_0 \right) - v_b \right\|_2^2 \tag{8}$$

The model makes sure that the images it makes are compatible between the visual content and the brand semantics by minimising this loss function. Also, during the denoising step of the diffusion model, the brand embedding variable v_b is updated by back propagation, which makes it much easier for brand information to guide the generation process.

The generative diffusion model can effectively control the brand consistency of the images it creates by combining information from the knowledge graph. This avoids the problems with traditional methods, which only make sure that the images fit the visual distribution and not the brand semantics. At each step, the denoising update considers the brand information. This makes the images that are made more in line with the brand's positioning and values in terms of look and meaning. Not only does combining the semantic information of the brand knowledge graph make the image better, but it also makes the brand features easier to recognise, which means the image can better meet the brand's visual marketing needs.

The fusion approach of knowledge graph information and diffusion process suggested in this paper successfully enhances the brand consistency of the generated images by integrating brand semantic information at each stage of the denoising process. The created photos not only look great, but they also accurately show the brand's main attributes, making them a new method to create brand visuals.

3.5 Adaptive adjustment and learning mechanism

In the generative diffusion model, achieving brand consistency depends not just on the knowledge graph's guidance but also on making changes that fit the needs of each brand. Different brands need different things when it comes to how they look. Some firms may want to be seen as high-end, while others may want to be seen as innovative or emotionally connected. So, figuring out how to change the way images are made based on the brand's fundamental values and market demand is the key to making the brand more consistent and the images better (Hofmann et al., 2021). To accomplish this, this study suggests an adaptive adjustment and learning mechanism that may change the model's generation approach on the fly based on the brand's traits at each stage of generation. This will lead to high-quality and highly personalised brand visual production.

The main idea behind this adaptive adjustment method is to use a feedback loop to automatically change the strategy of the generating process based on how the created images deviate from the brand goals. To start, we designate the feature variable of the brand target image y_b. This variable stands for the brand's basic traits, like its cultural values and how it positions itself in the market. After that, the function $f(x_0)$ gets the features of the created image and compares them to the brand target image y_b. To reduce the disparity between the generated picture and the brand target, we implement a dynamic adjustment system that relies on the loss function.

The following equation can be used to show how the generating process can adapt:

$$L_{\text{adapt}} = \left\| f\left(x_0\right) - y_b \right\|_2^2 + \lambda \cdot L_{\text{brand}}\left(v_b\right) \tag{9}$$

where L_{adapt} is the adaptive adjustment loss; $L_{\text{brand}}(v_b)$ is the brand consistency loss; and λ is an equilibrium hyperparameter that moderates the effect between the two.

The learning process is another important feature of the adaptive adjustment mechanism. To make the generation effect even better, the model has to change the hyperparameters and strategies in the generation process based on what it learns from feedback. Each brand's desired image has its own distinctive qualities that may need varied denoising rates, weight changes, and so on. The model changes the diffusion coefficient α_t and the influence weight β_t of the brand embedded variables in the generation stage by looking at the variations in features between the created image and the brand's target image. A reinforcement learning technique that optimises the parameter settings in the generation process depending on how similar the generated image is to the target image each time is used to change these hyperparameters. Assuming that the diffusion coefficient and brand embedding weights are α_t and β_t, respectively, at step t, the model can use the reinforcement learning process to change these two values depending on loss feedback:

$$\alpha_t^{(new)} = \alpha_t - \eta \frac{\partial L_{\text{adapt}}}{\partial \alpha_t} \tag{10}$$

$$\beta_t^{(new)} = \beta_t - \eta \frac{\partial L_{\text{adapt}}}{\partial \beta_t} \tag{11}$$

where η is the learning rate, which regulates how big the adjustment steps are, and the partial derivatives of the loss function L_{adapt} with respect to α_t and β_t shows how much each parameter affects the loss, which shows how different the brand attributes are from the generated image. This technique allows the brand to change each hyperparameter in the generation process in real-time to better fulfil the purpose of brand visual generation.

The generative diffusion model can change the way it generates images based on the brand's needs thanks to this adaptive adjustment and learning mechanism. This makes sure that the final image looks good and is consistent with the brand. The model constantly improves the generation strategy during training by using feedback information. This makes the generated images more and more in line with the brand's key principles, and it also makes it easier to tell the difference between other brands and make them unique.

In general, the adaptive adjustment and learning mechanism described in this research makes the generating process smarter and more personalised by adding brand feedback information and dynamic adjustment tactics. The generative diffusion model not only creates images that are visually appealing, but it also precisely shows how distinctive each brand is. This makes the visual output of each brand very relevant and new.

4 Experimental results and analyses

4.1 Dataset and pre-processing

For this study, we used the Open Images Dataset as our main source of data. Google gave us the dataset, which has more than 9 million labelled photos in many categories, such as items, brand logos, scenery, and more. This dataset may be used to create a lot of visual material for brand visual generating activities. The dataset is great for training generative diffusion models since it has a lot of brand logo pictures and also lets brand characteristics and knowledge graphs work together semantically.

The Open Photographs Dataset has photographs and a lot of labels. Each image has information like object categories, bounding boxes, brand logos, and more. We may use this information to find the brand's visual features and add them to the brand's knowledge graph for the picture generating assignment. Table 1 shows the essential details about the dataset.

Table 1 Information about open images dataset

Attribute	*Description*
Number of images	Over 9 million images
Number of brand categories	Over 500 categories, including brand logos and trademarks
Label types	Object categories, brand labels, bounding boxes, scene labels, etc.
Image resolution	Images of varying resolutions, high enough for generating high-quality images
Annotation information	Each image includes multiple labels such as categories, bounding box coordinates, brand information
Download method	Available via Google's API or direct dataset download, supporting category-based selection
Application domains	Object detection, image classification, brand recognition, image generation, etc.

Data preparation is an important step to make sure that the data can be used correctly in the brand vision generating model. We employed a number of methods to modify the data, such as rotation, flipping, colour dithering, and scaling, to make the photos more diverse, make the model more robust, and lower the danger of overfitting. We then used the bounding box information in the dataset to trim the brand logo areas in the photos to make them even more accurate. This manner, we make sure that the photos we create focus on the brand's most important visual characteristics.

We took the tags that were related to the brand in the image and linked them to the brand's semantic information while integrating them with the knowledge graph. The attributes of each image include embedded variables from the brand knowledge graph, such as the brand's core values and market positioning. This can help the generative diffusion model training process by giving more exact semantic guidance for the image production process.

Finally, the dataset is split into three parts: training, validation, and test. This makes sure that the training and validation processes are entirely representational of the data. The dataset is split up so that 80% goes to the training set, 10% goes to the validation set, and 10% goes to the test set (D'Souza et al., 2020). Each subset has a fairly even mix of brand categories.

The dataset is now ready for model training and experimental validation after these preparatory processes. Each training sample has photos and brand labels, and it also uses the brand's knowledge graph to make sure that the images it makes are not only pretty but also show the brand's essential qualities accurately.

4.2 Visual quality comparison experiment

To assess the efficacy of a generative diffusion model predicated on knowledge graph enhancement in a brand visual creation assignment, we devised an experiment for comparing brand consistency and visual quality. The project seeks to ascertain whether the integration of brand knowledge graph data may significantly improve the visual quality of the generated images, in comparison to the conventional generative diffusion model. We specifically evaluate the generative diffusion model based on knowledge graphs against the traditional generative diffusion model to see which one makes better photos.

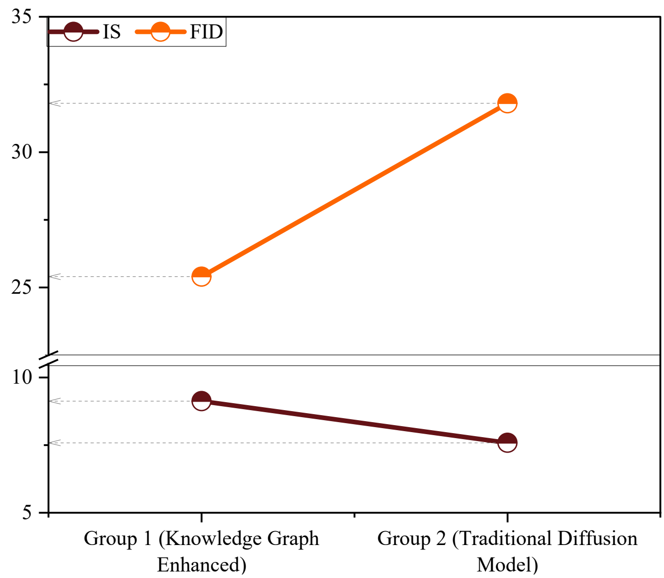
There are two groups in the experiment: experimental group 1 employs the knowledge graph-enhanced generative diffusion model for brand visual generation, integrating the brand's knowledge graph data into the generation process to ensure that the produced images accurately represent the brand's core attributes, including brand colours, logo morphology, and market positioning. In contrast, experimental group 2 utilises the conventional generative diffusion model for brand visual generation. This experimental group solely uses picture data to create things; it does not use the brand's knowledge graph information.

To check how good the generated images are, we utilise two standard metrics: inception score (IS) and Fréchet inception distance (FID) (Buzuti and Thomaz, 2023). The experimental results indicate that the generative diffusion model, enhanced by knowledge graphs, exhibits superior visual quality relative to the traditional model. Figure 2 shows the exact results.

The experimental results yield several conclusions. First, the IS findings show that experimental group 1 (knowledge graph enhanced model) got a score of 9.12, which is much higher than the score of experimental group 2 (conventional diffusion model), which was 7.58. This shows that the generative diffusion model works better for image variety and quality when the brand knowledge graph is added. The greater IS value shows that the knowledge graph-based model can make images that are richer and have better visual quality, which makes the images look more real and trustworthy.

Secondly, from the results of FID, the FID value of experimental group 1 is 25.4, while the FID value of experimental group 2 is 31.8. The FID value indicates that the images generated by the generative diffusion model based on knowledge graph enhancement have a higher similarity with the real images.

Figure 2 Visual quality comparison experiment results (see online version for colours)



Comprehensively analysing the above two indicators, the IS and FID of experimental group 1 are better than that of experimental group 2, indicating that the embedding of the brand knowledge graph effectively enhances the quality of the images, which makes the generated images more in line with the visual requirements of the brand, and can reflect the core features of the brand in the generation process.

Overall, the experimental results confirm the efficacy and benefits of the generative diffusion model predicated on knowledge graph enhancement in the brand vision generation task, particularly in enhancing image quality and the congruence between the generated images and the authentic brand vision.

4.3 Evaluation of brand customisation generation effect

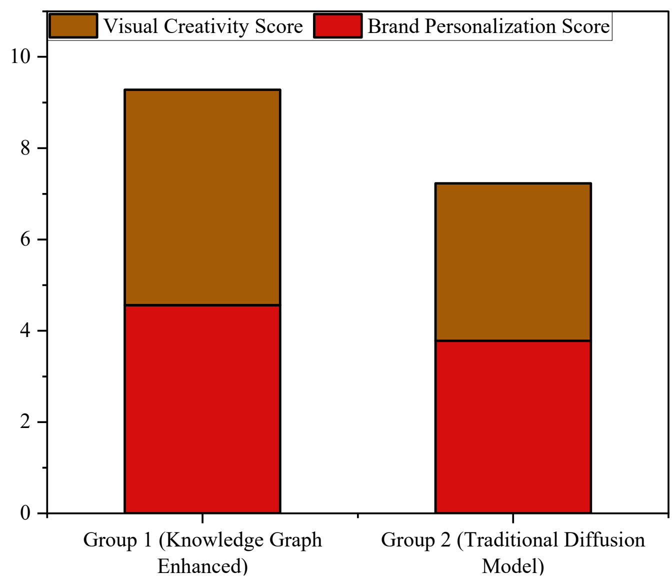
We conducted an experiment to see how well the generative diffusion model with a knowledge graph works for brand visual generation and to see how well it can customise brands. The experiment seeks to ascertain if the model can be tailored to the distinct attributes of the brand during the creation of specific brand visual content, so ensuring that the produced images accurately reflect the company's identity and market positioning.

There are two groups in the experiment: experimental group 1 employs a generative diffusion model augmented with a knowledge graph to produce brand-specific visual content. In this process, the model utilises brand attributes (such as logos, colours, market positioning, etc.) from the brand's knowledge graph to guide the generation, ensuring that the resulting images fulfil the brand's personalisation requirements. Conversely, experimental group 2 utilises a conventional generative diffusion model to create visual content for the same brand. The creation process in experimental group 2 is not based on a brand knowledge graph like it is in experimental group 1; instead, it is based only on image data.

Two evaluation indicators are utilised to measure how well the generated graphics personalise the brand. The first is brand personalisation performance, which uses expert ratings or customer surveys to see if the images show off the brand's personality and originality (Cavdar Aksoy et al., 2021). This type of evaluation uses a survey to get users' opinions on the style of the images and the brand's viewpoint. The second is visual creativity and innovation, which is scored by experts to assess whether the generated images are innovative while conforming to the brand's style (Pagani and Wind, 2025).

Figure 3 displays the outcomes of the experiment. It indicates that the generative diffusion model based on knowledge graph enhancement does a better job of brand customisation and visual innovation than the classic generative model.

Figure 3 The effect of brand customisation (see online version for colours)



The analysis of the experimental results indicates that, in terms of brand personalisation performance, the generative diffusion model enhanced by the knowledge graph (experimental group 1) achieved a score of 4.56, significantly surpassing the traditional generative diffusion model (experimental group 2), which scored 3.78. This suggests that the generative model combined with the brand knowledge graph can more accurately capture the core features of the brand and effectively reflect the personality and market positioning of the brand in the generation process. In contrast, the images generated by the traditional generative model fail to effectively present the personality of the brand, showing differences and deficiencies in brand style.

Second, from the perspective of visual creativity and innovation, the score of experimental group 1 is 4.72, which is higher than that of experimental group 2, which is

3.45. This indicates that the generative diffusion model based on the knowledge graph not only excels in brand style consistency but also innovates on the basis of brand style. The traditional generative model, on the other hand, generates more conventional and conservative visual content that lacks sufficient creativity, resulting in generated images that are deficient in brand differentiation and visual innovation.

In summary, the generative diffusion model based on knowledge graph enhancement outperforms the traditional generative diffusion model in terms of brand personalisation performance and visual creativity and innovation. Through this experiment, we can conclude that the knowledge graph-enhanced generative diffusion model can not only better convey the core value of the brand but also bring more creativity and flexibility to the generation of the brand's visual content, which can satisfy the modern brand's demand for visual content innovation and differentiation.

5 Conclusions

In this paper, we propose a generative diffusion model based on knowledge graph enhancement and apply it to the brand visual generation task. Validated by two experiments, the results show that the knowledge graph-based generative diffusion model outperforms the traditional generative diffusion model in terms of brand consistency, visual quality and brand personalisation. The testing results show that the method works well and might be used in the field of brand visual generation.

The model developed in this paper has commendable experimental outcomes in various aspects; nonetheless, it possesses certain limitations. The model's dependence on knowledge graphs necessitates the precision and completeness of brand information. Second, this study's dataset is mostly from the Open Images Dataset, which does not have enough brand visual data and does not cover all brand types. Furthermore, while the model excels in customised generation, it requires additional optimisation and enhancement in practical applications to increase the diversity and innovation of the generated images.

Future research can improve the generative diffusion model in a number of ways by enhancing the knowledge graph. First, for the building of the knowledge graph, more detailed brand information and market data can be incorporated in the future to make brand semantics even more accurate. Second, the existing model does a good job of keeping the brand consistent, but the inventiveness and variety of the visuals it makes might still be improved. In the future, there may be additional ways to customise the generated images so that they exhibit more innovation and variations while yet keeping the brand design the same. Finally, as multimodal technology improves, merging generative diffusion models with video generation, VR, and other technologies can open up new ways for brand visual production to be used. In the future, we could be able to learn more about how to make material that fits the brand's stance in changing situations or improve the brand experience through engagement in virtual spaces.

Declarations

All authors declare that they have no conflicts of interest.

References

Ametordzi, S. and Olalere, F. (2024) 'The effect of digital technology on graphic design ideation output', *IDA: International Design and Art Journal*, Vol. 6, No. 1, pp.144–156.

Buzuti, L.F. and Thomaz, C.E. (2023) 'Fréchet AutoEncoder distance: a new approach for evaluation of generative adversarial networks', *Computer Vision and Image Understanding*, Vol. 235, p.103768.

Cavdar Aksoy, N., Tumer Kabadayi, E., Yilmaz, C. and Kocak Alan, A. (2021) 'A typology of personalisation practices in marketing in the digital age', *Journal of Marketing Management*, Vol. 37, Nos. 11–12, pp.1091–1122.

Chen, Y., Li, H., Li, H., Liu, W., Wu, Y., Huang, Q. and Wan, S. (2022) 'An overview of knowledge graph reasoning: key technologies and applications', *Journal of Sensor and Actuator Networks*, Vol. 11, No. 4, p.78.

Cobbinah, M., Nunoo-Mensah, H., Ebenezer Adjei, P., Adoma Acheampong, F., Acquah, I., Tutu Tchao, E., Selasi Agbemenu, A., John Kponyo, J. and Abaidoo, E. (2025) 'Diversity in stable GANs: a systematic review of mode collapse mitigation strategies', *Engineering Reports*, Vol. 7, No. 6, p.e70209.

D'Souza, R.N., Huang, P-Y. and Yeh, F-C. (2020) 'Structural analysis and optimization of convolutional neural networks with a small sample size', *Scientific Reports*, Vol. 10, No. 1, p.834.

Dey, L. (2024) 'Knowledge graph-driven data processing for business intelligence', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 14, No. 3, p.e1529.

Dou, B., Zhu, Z., Merkurjev, E., Ke, L., Chen, L., Jiang, J., Zhu, Y., Liu, J., Zhang, B. and Wei, G-W. (2023) 'Machine learning methods for small data challenges in molecular science', *Chemical Reviews*, Vol. 123, No. 13, pp.8736–8780.

Fan, W., Lu, P., Pang, K., Jin, R. and Yu, W. (2024) 'Linking entities across relations and graphs', *ACM Transactions on Database Systems*, Vol. 49, No. 1, pp.1–50.

Guo, X. and Zhao, L. (2022) 'A systematic survey on deep generative models for graph generation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, No. 5, pp.5370–5390.

Hofmann, J., Schnittka, O., Johnen, M. and Kottemann, P. (2021) 'Talent or popularity: what drives market value and brand image for human brands?', *Journal of Business Research*, Vol. 124, pp.748–758.

Ji, S., Pan, S., Cambria, E., Marttinen, P. and Yu, P.S. (2021) 'A survey on knowledge graphs: representation, acquisition, and applications', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 2, pp.494–514.

Khan, A., Sohail, A., Zahoora, U. and Qureshi, A.S. (2020) 'A survey of the recent architectures of deep convolutional neural networks', *Artificial Intelligence Review*, Vol. 53, No. 8, pp.5455–5516.

Nasar, Z., Jaffry, S.W. and Malik, M.K. (2021) 'Named entity recognition and relation extraction: state-of-the-art', *ACM Computing Surveys (CSUR)*, Vol. 54, No. 1, pp.1–39.

Pagani, M. and Wind, Y. (2025) 'Unlocking marketing creativity using artificial intelligence', *Journal of Interactive Marketing*, Vol. 60, No. 1, pp.1–24.

Peng, C., Xia, F., Naseriparsa, M. and Osborne, F. (2023) 'Knowledge graphs: opportunities and challenges', *Artificial Intelligence Review*, Vol. 56, No. 11, pp.13071–13102.

Quan, H., Li, S., Zeng, C., Wei, H. and Hu, J. (2023) 'Big data and AI-driven product design: a survey', *Applied Sciences*, Vol. 13, No. 16, p.9433.

Senior, H., Slabaugh, G., Yuan, S. and Rossi, L. (2025) 'Graph neural networks in vision-language image understanding: a survey', *The Visual Computer*, Vol. 41, No. 1, pp.491–516.

Sun, H., Xu, H., Wu, J., Sun, S. and Wang, S. (2024) 'Let pictures speak: hotel selection-recommendation method with cognitive image attribute-enhanced knowledge graphs', *International Journal of Contemporary Hospitality Management*, Vol. 36, No. 12, pp.4296–4318.

Tian, Z. (2020) 'Dynamic visual communication image framing of graphic design in a virtual reality environment', *IEEE Access*, Vol. 8, pp.211091–211103.

Trehan, S. and Kalro, A.D. (2024) 'A comprehensive framework of visual elements in brand logos: content analysis of top 500 global brands', *Journal of Product & Brand Management*, Vol. 33, No. 8, pp.1055–1072.

Yenduri, G., Ramalingam, M., Selvi, G.C., Supriya, Y., Srivastava, G., Maddikunta, P.K.R., Raj, G.D., Jhaveri, R.H., Prabadevi, B. and Wang, W. (2024) 'GPT (generative pre-trained transformer) – a comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions', *IEEE Access*, Vol. 12, pp.54608–54649.

Zhong, L., Wu, J., Li, Q., Peng, H. and Wu, X. (2023) 'A comprehensive survey on automatic knowledge graph construction', *ACM Computing Surveys*, Vol. 56, No. 4, pp.1–62.