

# From Pixels to Generalization: Ensuring Information Security and Model Performance with Design Principles for Synthetic Image Data in Deep Learning

Martin Böhmer<sup>1</sup>

<sup>1</sup> Martin Luther University Halle-Wittenberg, Universitätsring 3, 06108 Halle (Saale), Germany

## Abstract

This paper explores the ethical and effective utilization of synthetic image data in computer vision deep learning. It addresses the challenges of acquiring real-world training data and proposes design principles for selecting, generating, and integrating synthetic images. These principles cover aspects such as ethical compliance, privacy protection, scene diversity, and complexity management. By adopting a design science research approach and using a multi-method research design, the study provides actionable guidance for researchers and practitioners, as these design principles ensure responsible use of synthetic image data while improving model performance and privacy protection. The paper contributes to design knowledge in the general IS, deep learning, and IS ethics field, highlighting the theoretical and practical relevance of the proposed principles. The reusability of the design principles promotes the efficient use of synthetic image data in computer vision and has been positively evaluated.

## Keywords

Design Principles, Deep Learning, Synthetic Image Data, Information Security, AI Ethics

## 1. Introduction

Since computer vision deep learning models often consist of millions or even billions of parameters, they rely on large amounts of training data to achieve high performance and generalization [1]. However, acquiring real-world training data for artificial intelligence (AI) applications can be costly, error-prone, limited, or imbalanced [2, 3, 4]. Synthetic image data (e.g. in the form of video game engine generated scenes) has emerged as a promising alternative, offering scalability, precision, and potentially more robust and accurate models [5, 2, 6]. Nonetheless, guiding design knowledge on how to utilize synthetically generated image data in deep learning remains scarce. Moreover, the synthetic illustration of humans, including their separate or related characteristics such as body parts, raises ethical considerations regarding privacy, consent, and the potential for misrepresentation or discrimination. In addition, the use-case context of synthetic image data often revolves around human-related domains [7, 8, 9], such as medicine or surveillance. These domains inherently involve sensitive information and human interactions, making it crucial to design technologies that align with ethical standards and user values. Given the nascent state of the synthetic imagery domain, this paper therefore defines the following guiding research question:

**RQ:** *How to ethically, effectively, and robustly utilize synthetic image data in computer vision deep learning environments?*

---

CIISR 2023: 3rd International Workshop on Current Information Security and Compliance Issues in Information Systems Research, co-located with the 18th International Conference on Wirtschaftsinformatik (WI 2023), September 18, 2023, Paderborn, Germany

✉ martin.boehmer@wiwi.uni-halle.com (M. Böhmer)

ORCID 0000-0002-6249-3184 (M. Böhmer)



© 2023 Copyright for this paper by its authors.  
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



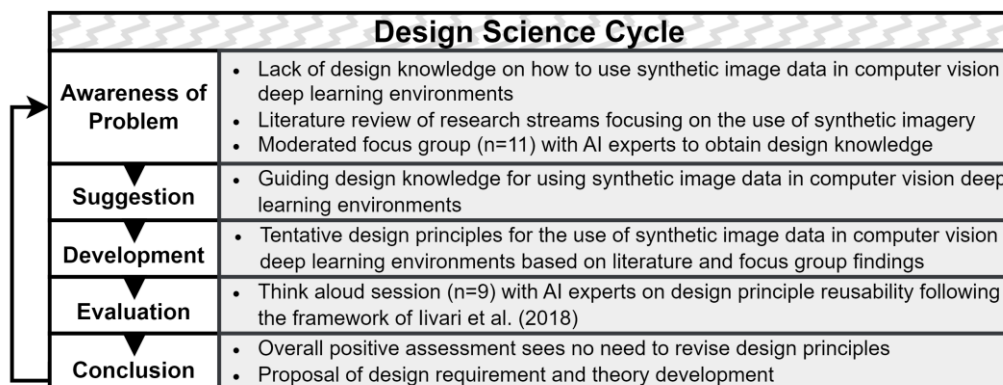
CEUR Workshop Proceedings (CEUR-WS.org)

To answer this question and to contribute prescriptive knowledge, the design science research paradigm [10] and design science process model [11] are adopted, with a focus on the *value sensitive design* theory [12] as the guiding theoretical lens. The research approach employs a multi-method research design, combining qualitative methods such as moderated focus groups and think aloud sessions to ensure the validity and comprehensiveness of the study design. Therefore, the research aims to fill the aforementioned research gap by deriving design principles based on kernel theories, the literature, and practical insights. The design principles address key aspects such as ethical compliance, privacy protection, data governance, scene diversity, controlled composition, complexity management, and data augmentation, providing actionable guidance for researchers and practitioners in the selection, generation, and integration of synthetic image data for training deep learning models. By adopting these design principles, practitioners can ensure ethical and responsible use of synthetic image data while enhancing model performance, privacy protection, and generalization. The reusability of these principles in similar contexts contributes to their wider application and adoption, addressing the current lack of design knowledge and promoting efficient utilization of synthetic image data in computer vision deep learning environments.

The following sections present the research design, theoretical and practical foundations, design principles, and the evaluation of the proposed principles. The conclusion highlights the contributions of this study to the design knowledge in the field of utilizing synthetic image data in computer vision and identifies potential areas for future research.

## 2. Research Design

In order to contribute prescriptive rather than descriptive knowledge [13], the design science research paradigm [13] was used for the purpose of this study. In addition, value sensitive design theory [12] was used as the guiding theoretical lens (see Section 3.1), which served as the theoretical framework for the methodological techniques and design of the approach undertaken in this paper. Therefore, a multi-method research approach [14] consisting of different qualitative methods was chosen to address the shortcomings of single methods and to ensure the validity of the design. The methods used for this study are based on value sensitive design [12] and include a moderated focus group for data collection and a think aloud session for evaluating the design principles.



**Figure 1:** Design science research approach

Since the design science research paradigm can be operationalized through various methodological approaches [15], this study utilized the framework proposed by Vaishnavi and Kuechler [11] due to its explicit focus on theoretically grounded design principles. As shown in Figure 1, the approach includes the five steps: awareness of problem, suggestion, development, evaluation, and conclusion. Furthermore, and particularly with respect to the ethical and information security scope of this paper, the ethical design science framework proposed by Durani et al. [16] was used to derive ethical and, in addition, negative design principles (discussed

in detail in Section 4), addressing the disruptive nature of recent advancements in technology and especially deep learning. In this paper, we delve into the intricacies of design principles, because they form the foundation of design knowledge and play a pivotal role in solving the problem at hand [17], which in this case is the theoretical void of guiding design knowledge on how to use synthetic image data in deep learning.

### 3. Theoretical and Practical Foundations

Given that state-of-the-art deep learning models for computer vision comprise millions, if not billions, of parameters, the training process for these models necessitates an immense quantity of training data, which is frequently absent or imbalanced [1]. Therefore, the impulse for the underlying research stems from a genuine real-world circumstance, namely, the provision of scalable, precise, and ethical computer vision deep learning models and their respective training data. The highlighted problem originates from several prior studies that found synthetically trained computer vision models to be more robust, accurate, and less error-prone [5, 2, 6]. In addition, synthetic image data achieves photorealism and can be generated and scaled infinitely, making it a genuine alternative to conventional real imagery approaches [3]. Therefore, a thorough analysis of the scientific literature was conducted in the scope of this study to identify relevant research streams and kernel theories. In addition, the theoretical foundations are further supported by the results of a moderated focus group, which serve as practical foundations for the development of the design principles.

#### 3.1. Kernel Theory

To ensure scientific rigor and stringency, design science research endeavors can use kernel theories to derive design principles. Broadly speaking, kernel theory functions as a form of justificatory knowledge within the realm of design knowledge development, as indicated by the work of Gregor and Hevner [10], such as in the form of design principles [18]. Henceforth, this study adopts *the analyze with lens*-mechanism proposed by Möller et al. [19], drawing upon the theoretical foundations of employing kernel theories as a means of analysis. The use of a theoretical lens allows researchers to derive concepts indirectly, guiding the analysis or framing of data within the conceptual borders of a specific theory. This approach aligns with the perspective of Niederman and March [20] on the theoretical lens, which emphasizes its role in aiding the theorization process, leading to the formulation of design principles or meta-requirements based on a data foundation. Thus, by adopting the *analyze with lens*-mechanism proposed by Möller et al. [19], this study aims to analyze the data through a theoretical lens, allowing for a more robust and informed exploration of the underlying concepts and patterns. As the most appropriate kernel theory for the scope of this study, the *value sensitive design* theory [12] was chosen as it epitomizes a theoretically grounded approach that considers human values in a principled and comprehensive manner throughout the design process, which aligns perfectly with the research goal of developing technology that respects and incorporates user values while ensuring ethically responsible and user-centered design decisions. In the specific use-case of synthetic image data utilization for deep learning tasks, this especially connects to the synthetic illustration of humans (including separate or related characteristics, e.g. body parts), the use-case context in which the synthetically generated image data is used (often human-related, e.g. medicine or surveillance), and the potential ethical implications that arise from the creation and utilization of synthetic images. The theoretical lens of *value sensitive design* [12] thus helps not only to derive design principles, but also to establish a robust elicitation and evaluation (e.g., think aloud sessions) of these principles. Moreover, the theory recognizes that technology is not neutral as it can influence behavior, perception, and societal structures, and thus should be designed in a way that reflects positive human values and respects ethical principles, promoting a holistic approach to technology design that goes beyond functionality in order to consider the broader impact on individuals, society, and the environment.

### 3.2. Theoretical Foundations

To establish the objectives for addressing the aforementioned problem, a comprehensive analysis of the scientific literature focused on the research area of synthetic image data generation in deep learning, in line with our DSR methodology, was conducted. As aforementioned, training deep learning models in computer vision requires large amounts of data to achieve a fairly high degree of generalization, which is often costly, missing, or unbalanced [1]. Several studies have highlighted the effectiveness of synthetic data for training deep learning models in various computer vision tasks. Lee et al. [21] and Krump et al. [6] utilized synthetic datasets for deep learning-based object detection, specifically in underwater sonar imaging and vehicle detection on UAV platforms, respectively. Body et al. [22] and Condrea et al. [23] demonstrated the value of artificial augmented textual data and purely synthetic training data, respectively, for sentiment analysis models and vital signs detection in videos. Similarly, Liu et al. [8] and Zaki et al. [24] employed synthetic data for pose estimation and semantic object/scene categorization. Hence, these studies collectively highlight the effectiveness of synthetic data in various computer vision domains.

Additionally, domain adaptation, which is a technique that involves adapting a model trained on one domain (synthetic) of data to perform well on a different but related domain (real), and transfer learning have been extensively explored in the context of synthetic data. Lahiri et al. [25], Venkateswara et al. [26], and Kuhnke and Ostermann [7] focused on unsupervised domain adaptation for synthetic data, learning transferable feature representations, and domain adaptation for pose estimation, respectively. Seib et al. [3] conducted a comprehensive review of current approaches that combine real and synthetic data to enhance neural network training, supporting the argument for a combination of training data and data augmentation. Aranjuelo et al. [27] discussed key strategies for synthetic data generation in people detection from omnidirectional cameras, emphasizing the effective use of both real and synthetic data. Valtchev and Wu [28] demonstrated the utility of domain randomization for neural network classification, showcasing the effectiveness of synthetic data in training robust models. These studies provide insights into the adaptation of synthetic data to real-world scenarios.

Moreover, the combination of synthetic and real training data has been investigated by several researchers. Wan et al. [29], Bird et al. [5], and Abu Alhaija et al. [30] utilized mixed datasets, comprising both synthetic and real data, for document layout analysis, scene classification, and object detection in augmented reality, respectively. Thereby, these studies highlight the benefits of leveraging both synthetic and real data for training computer vision models.

Furthermore, the use of synthetic data generation techniques and simulators has been explored. Müller et al. [31] introduced a photorealistic simulator for generating synthetic data for computer vision applications, whereas Zhang et al. [4] proposed a stacked multichannel autoencoder framework for efficient learning from synthetic data. Valerio Giuffrida et al. [32] generated synthetic training data for the detection of synthetic *Arabidopsis* plants using generative adversarial networks. Scheck et al. [33] introduced a synthetic dataset that serves as a valuable resource for training and evaluating deep learning models. These works provide insights into the generation and utilization of synthetic data for training deep learning models.

Despite the considerable research on utilizing synthetic image data for computer vision deep learning models, there remains a notable research gap in terms of a comprehensive framework or guidelines that provide design knowledge to effectively and systematically utilize synthetic data in this context. While individual studies have demonstrated the benefits and effectiveness of synthetic data in specific tasks, there is a lack of unified principles or guidelines that guide researchers and practitioners in the selection, generation, and integration of synthetic image data for training deep learning models. Hereby, the absence of such design knowledge hinders the widespread adoption and consistent utilization of synthetic data, leading to potential inefficiencies, suboptimal performance, and challenges in real-world deployment.

### 3.3. Practical Foundations

To ensure scientific rigor, and after analyzing the aforementioned research streams and kernel theory, it seemed reasonable to conduct a moderated focus group with AI experts to rigorously derive design knowledge, compare it to the literature findings, and incorporate it into the design principles. Moderated focus groups are especially predestined for extensive qualitative insights into a subject [34] and align with the kernel theory of value sensitive design [12].

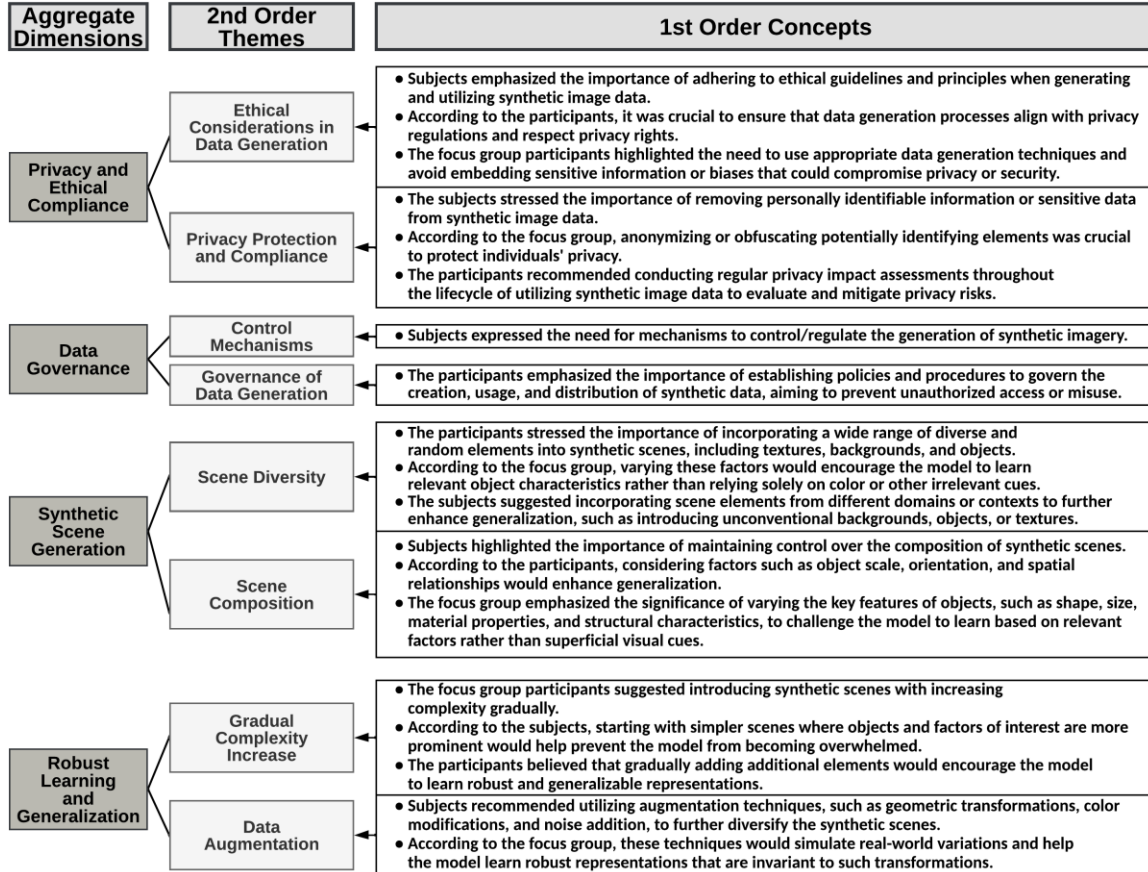


Figure 2: Focus group results

The conducted focus group consisted of n=11 participants, including three AI senior scholars, two AI research associates, two computer vision project leads, and four IS researchers, all with professional experience ranging from 3-17 years. The goal of the moderated focus group was to develop design knowledge (e.g., user-specific requirements, characteristics, process steps), but without incorporating the literature findings to avoid any bias. To ensure qualitative rigor during and after this session, the well-established methodology outlined by Gioia et al. [35] was followed throughout, which involved formulating *first order concepts*, *second order themes*, and *aggregate dimensions (AD)* based on the subjects' expressed statements. The focus group findings are shown in Figure 2. The focus group findings revealed key insights regarding the utilization of synthetic image data in computer vision deep learning settings. Participants emphasized the importance of adhering to ethical guidelines and privacy regulations throughout the data generation process. They also stressed the need to remove personally identifiable information and conduct privacy impact assessments regularly to mitigate privacy risks - resulting in **AD1** (Privacy and Ethical Compliance). The subjects also highlighted the significance of implementing mechanisms for generation control and data governance to prevent unauthorized access or misuse which was epitomized by **AD2** (Data Governance). Additionally, the focus group emphasized the need for synthetic scene diversity, recommending the incorporation of various elements and cross-

domain scene randomization to enhance generalization. While promoting scene diversity, they emphasized the importance of maintaining control over scene composition to ensure proper representation of intended features and factors of interest. This resulted in **AD3** (Synthetic Scene Generation). The participants further suggested gradually increasing the complexity of synthetic scenes to prevent overfitting and promote robust learning and generalization. They also recommended data augmentation techniques, such as geometric transformations and color modifications, to diversify synthetic scenes and enable the learning of robust representations invariant to real-world variations – illustrated by **AD4** (Robust Learning and Generalization).

Overall, these focus group findings provided valuable insights for the proposed design principles, which aim to guide the ethical and effective utilization of synthetic image data in computer vision research and applications.

#### 4. Design Principles for Using Synthetic Image Data

As aforementioned, the performed design cycle was dedicated to creating design knowledge and developing theoretically sound design principles as the main artifact. As design principles embody a general design solution for a class of problems [17], they are of prescriptive and universal nature, specifying how a solution should be designed to achieve the desired objective [36]. In this context, the DPs were derived from a supportive approach and the conceptual schema of Gregor et al. [18], whose a priori specification suggests prescriptive wording [36], thereby allowing us to formulate accessible, precise, and expressive design knowledge, as elucidated by the framework [18]. In addition to utilizing the *anatomy of a design principle* [18] and the kernel theory of *value sensitive design* [12], the development and wording of the design principles were guided by the ethical design science research framework proposed by Durani et al. [16], resulting in more prescriptive guidance for leveraging the positive impact of the artifact and minimizing its adverse effects.

The design principles were rigorously derived from the literature and the aggregate dimensions from the qualitatively analyzed focus group results. As shown in Figure 3, seven specific design principles were developed to address the identified problem of lacking design knowledge on how to utilize synthetic image data in computer vision deep learning environments.

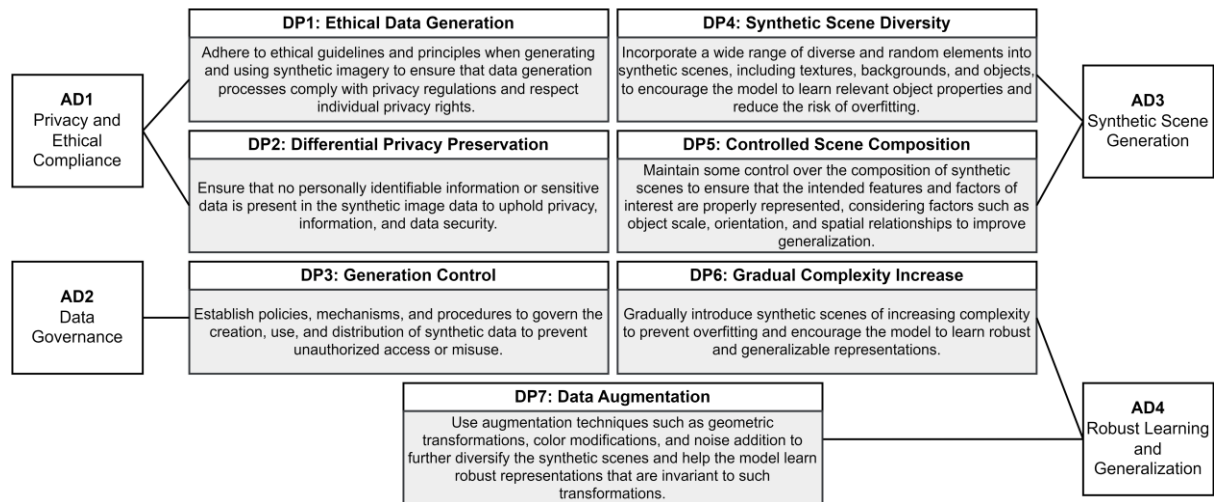
**DP1** draws from *AD1* and states that ethical guidelines and principles should be followed when generating and utilizing synthetic image data. Incorporating value sensitive design [12], it is important to align data generation processes with privacy regulations and to show respect for individual privacy rights. Therefore, care should be taken to employ suitable data generation techniques (e.g. via *Unity3D*) and to refrain from incorporating sensitive information or biases that could potentially compromise the privacy or security of individuals.

**DP2** also builds on *AD1* and addresses the need for the synthetic image data to contain no personally identifiable information (PII) or sensitive data. It's necessary to anonymize or obfuscate any elements that could potentially reveal an individual's identity. Throughout the process of using synthetic imagery, regular privacy impact assessments should be conducted to assess the privacy risks associated with the generation, storage, transmission, and use of the data. Appropriate measures should then be implemented to mitigate the identified risks and ensure ongoing compliance with privacy regulations, thus aligning with value sensitive design [12]. In this regard, it is recommended that differential privacy mechanisms be incorporated into the generation and use of synthetic image data, where controlled noise or perturbations are introduced during data generation to prevent individual data points from being distinguished with a high degree of certainty [3, 4]. This approach can protect the privacy of individuals even in the presence of external information.

Based on *AD2*, **DP3** states that mechanisms should be implemented to control and regulate the generation of synthetic image data, such as process frameworks, toolkits, virtual environments, or guidelines. Hence, policies and procedures need to be established to govern the creation, usage,

and distribution of synthetic data in order to prevent unauthorized access or misuse, ensuring value sensitive design [12].

**DP4** stems from *AD3* and specifies that a wide range of diverse and random elements should be incorporated into synthetic scenes, including textures, backgrounds, and objects. By varying these factors, the model will be encouraged to learn relevant object characteristics instead of relying on color or other irrelevant cues [33, 3]. To further improve generalization, cross-domain scene randomization should be used, which involves incorporating scene elements from different domains or contexts (e.g., non-healthcare elements in healthcare settings). Introducing unconventional backgrounds, objects, or textures that are not typically associated with the objects of interest can push the model to learn their intrinsic properties, thereby promoting adaptability to real-world scenarios [3, 28].



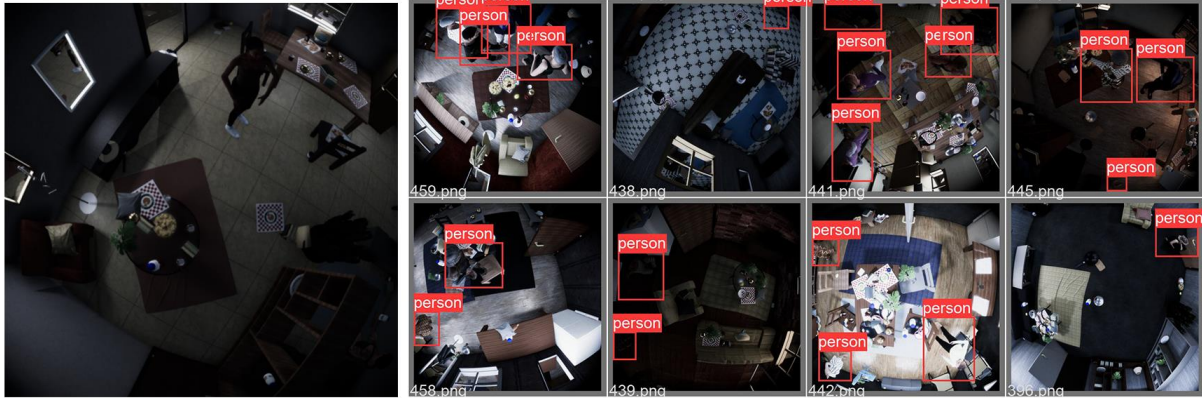
**Figure 3:** Design principles for using synthetic image data in computer vision

Note: DP = design principle; AD = aggregate dimension

**DP5** closely connects to **DP4** and further relates to *AD3*, stating that while aiming to promote scene diversity (and randomness), it is important to maintain a level of control over the composition of synthetic scenes. This ensures that the intended features and factors of interest are properly represented where factors such as object scale, orientation, and spatial relationships should be considered to enhance generalization [6, 33]. Rather than relying solely on changing the appearance of objects, the focus should be on varying their key features, and changing attributes such as shape, size, material properties, and structural characteristics will challenge the model to learn object representations based on these relevant factors rather than superficial visual cues.

**DP6** draws from *AD4* and addresses the gradual introduction of synthetic scenes with increasing complexity. Training the deep learning model should begin with simpler scenes that highlight objects and factors of interest more prominently, and gradually incorporate additional elements to prevent overwhelming the model [1]. This approach helps prevent overfitting and encourages the model to learn robust and generalizable representations, which is highly relevant when working with synthetic rather than real data [5, 3, 29].

**DP7** also stems from *AD4* and states that augmentation techniques, such as geometric transformations, color modifications, and noise addition, should be utilized to enhance the diversity of synthetic scenes [31, 3, 4]. These techniques simulate real-world variations and assist the model in learning robust representations that remain invariant to such transformations, which further mitigates the risk of model overfitting [1].



**Figure 4:** Exemplary synthetic image data instantiation

Figure 4 shows an exemplary instantiation of the design principles in a computer vision deep learning setting for person detection, which is therefore concerned with ethical data generation. By following value sensitive design [12], abstract and non-genuine characters were generated for the scenes, ensuring ethical data generation (DP1) and differential privacy preservation (DP2). Using a video game engine and various mechanisms to generate the synthetic image data (DP3), a wide variety of scenes (DP4) and compositions (DP5) were achieved, significantly and intentionally varying the key features. In addition, several scenes are epitomized by reduced complexity (DP6) to achieve better generalization, where this can be scaled infinitely. Finally, data augmentation techniques (DP7) in the form of geometric transformations (i.e., lens distortion) were used to reduce the risk of overfitting. By incorporating these design principles, the depicted figure ensures ethical data generation practices in person detection, promoting accurate and reliable results while considering privacy, diversity, generalization, and controlled complexity.

## 5. Evaluation

To ensure scientific rigor in the evaluation of this design cycle, the well-established *FEDS* framework proposed by Venable et al. [37] was used. the evaluation phase is highly relevant in design science research [11, 37], it is necessary to select an appropriate strategic process and determine the constructs to be evaluated. Given the small and rather simple design of the main artifact, embodied in a set of design principles that result in low social and technical risk and uncertainty, the evaluation strategy of *quick & simple* [37] was chosen. Thus, the goal was to conduct an evaluation episode to complete the design cycle and to move quickly to a summative evaluation. The evaluation schema employed in this study takes into account the roles of key stakeholders involved the formulation of design principles. This schema allows design science researchers to assess the usability of generated design principles for different user groups. Two critical questions arise from this perspective: first, whether the design principles are understandable and useful to implementers, and second, whether they effectively serve the goals of users who implement the resulting instantiations [18]. Therefore, *evaluation activity 2* [38] was used, which describes an artificial activity, since the artifact has not yet been properly instantiated (note that Figure 4 serves only as an exemplary visualization). This activity aims to validate the principles of form and function, which have been developed during the design cycle [38].



**Table 1**  
**Think aloud results and reusability framework application**

Reusability Category	Verbalized Think Aloud Results
<i>Accessibility</i>	<p>The participants stated that they found the design principles to be highly accessible, emphasizing the clarity and understandability of the language used. Particularly in terms of privacy and data protection (<b>DP1, DP2, DP3</b>), the participants recognized the importance of practitioners being able to comprehend and implement these principles effectively, ensuring ethical and privacy-compliant use of synthetic image data.</p> <p>However, participants noted that while the design principles presented were well constructed, there was a suggestion to consider including explanations for technical terms such as "differential privacy mechanisms". They mentioned that providing brief definitions for such terms could help readers who may not be deeply familiar with the field to better understand the content.</p>
<i>Importance</i>	<p>The participants also highlighted the significant importance of the design principles. They acknowledged that these principles addressed crucial concerns related to privacy, data anonymization, information security, and regulatory compliance (<b>DP1, DP2, DP3</b>). By incorporating these principles into deep learning environments, the participants emphasized the practical relevance and significance of adhering to them, fostering trust and responsible use of synthetic image data.</p>
<i>Novelty and Insightfulness</i>	<p>The participants expressed their appreciation for the design principles, stating that they introduced fresh perspectives to the generation and utilization of synthetic image data.</p> <p>The emphasis on diversity and randomness in scene composition, the incorporation of unconventional elements, and cross-domain randomization were noted as innovative approaches (<b>DP4, DP5</b>). According to the participants, these principles challenged traditional methods and encouraged thinking beyond the conventional, promoting adaptability to real-world scenarios.</p>
<i>Actability and Guidance</i>	<p>The participants commended the actability and appropriate guidance provided by the design principles. They highlighted the clear frameworks, policies, and mechanisms suggested to regulate the generation and usage of synthetic image data (<b>DP3</b>). The participants found the gradual introduction of complexity during training and the utilization of augmentation techniques as practical suggestions aligned with their deep learning workflows (<b>DP6, DP7</b>). The guidance provided struck a balance between providing direction and allowing for creative application of the principles.</p>
<i>Effectiveness</i>	<p>The effectiveness of the design principles was evident to the participants. They recognized the emphasis on preventing privacy risks (<b>DP1, DP2</b>), mitigating overfitting, and enhancing model generalization (<b>DP4, DP5, DP6, DP7</b>). By adhering to these principles, the participants noted that robust deep learning models could be developed, yielding high performance on real-world data.</p> <p>They found the strategies of gradually introducing complexity (<b>DP6</b>) and using augmentation techniques (<b>DP7</b>) to be effective in optimizing performance and ensuring the practical utility of synthetic image data.</p> <p>However, participants expressed that while the concepts of <b>DP6</b> and <b>DP7</b> were intriguing, they suggested that a comparative analysis be included which could contrast the proposed principles with existing methodologies and highlight the unique advantages and improvements.</p>

To ensure the objectives of feasibility, accessibility, completeness, and applicability, it seems reasonable to apply the framework of design principle reusability proposed by Iivari et al. [39]. This framework provides a systematic approach to evaluating the design principles generated during the design cycle and, by assessing the reusability of these principles, researchers can determine their potential for wider application and adoption in similar contexts [39]. For this purpose, and in accordance to the kernel theory of value sensitive design [12], a qualitative think aloud session addressing the reusability of the proposed design principles was conducted. Therefore, the method of *concurrent think-aloud* [40] was employed with n=9 AI experts, where the sample size was decided based on the “10±2 rule” for think aloud sessions [41]. The participants were asked to verbalize their thoughts about the design principles in terms of the reusability categories proposed by Iivari et al. [39]. The experts were provided with a detailed textual description of the design principles, along with visual examples of synthetic image data. Table 1 presents the qualitative think aloud results, including the categories of the reusability framework and the clustered verbalized thoughts of the participants.

Overall, the participants of the think aloud session positively evaluated the design principles, emphasizing their accessibility, importance, novelty and insightfulness, actability with appropriate guidance, and overall effectiveness in enhancing the use of synthetic image data in deep learning environments. Their feedback underscored the value and especially the reusability of these principles in guiding practices and ensuring responsible and efficient utilization of synthetic image data in deep learning. Nonetheless, a few areas for improvement have been identified by the participants. However, a number of potential areas for refinement emerged from their constructive feedback. Participants noted that while the structure of the design principles was commendable, they suggested that clarifications of technical terms such as "Differential Privacy Mechanisms" could be included. It was suggested that providing concise definitions for these terms could serve to help readers or researchers less familiar with the field to better understand the content. In addition, participants expressed the notion that despite the appeal of *DP6 (Gradual Complexity Increase)* and *DP7 (Data Augmentation)*, it might be prudent to introduce a comparative analysis that could compare the proposed principles with existing methodologies, thereby highlighting their particular merits and improvements. These suggestions for refinement, which come from the participants and should be picked up in subsequent design science cycles, are intended to increase the accessibility and effectiveness of the design principles, serve a wider range of readers, and further substantiate their utility.

## 6. Conclusion

The paper proposes general design principles for the use of synthetic image data in computer vision deep learning environments to ensure more ethical, robust, traceable, and effective development and implementation of such models. Consequently, to answer the initially formulated research question of this paper, the results of a completed design science research cycle have been presented. Hereby, the positive evaluation of the design principles substantiates the theoretical and practical relevance of the design principles and researchers can adapt these to develop, utilize, or modify deep learning models based on synthetic image data. By using the DSR paradigm [10], the study moves beyond descriptive knowledge and aims to provide prescriptive knowledge, focusing on the design principles for utilizing synthetic image data in deep learning. This integration of the design science research paradigm contributes to the advancement of design knowledge in the field along with the IS design science knowledge base according to Woo et al. [42]. The paper also contributes theoretically by employing the value sensitive design theory, as proposed by Friedman et al. [12], which enhances the understanding of the ethical implications and user values in the context of synthetic image data utilization. The practical implications of these design principles include improved performance, enhanced privacy protection, and responsible and efficient utilization of synthetic image data in real-world applications, while the reusability of these principles in similar contexts contributes to their

wider application and adoption, promoting responsible and efficient utilization of synthetic image data in computer vision.

Meanwhile, in the context of the positive evaluation episode, the following limitations should be considered: First, design principles and their development are tied to the subjective creativity of the researcher, even after various data collection episodes and literature reviews. However, not all design decisions can or should be derived from behavioral or mathematical theories, as some degree of creativity is essential to developing an innovative design artifact [43, 44], whereas a certain degree of rigor can be implemented such as the utilized methodological approaches of Gregor et al. [18], Möller et al. [19], or Fu et al. [36]. Second, as with any other evaluation, the results describe only one sample, meaning that different results could be expected if a different sample were chosen. Therefore, this particular limitation could be addressed in future research, while the application of the design principles (in various domains such as digital health, etc.) as part of a case study or framing guideline seems highly interesting. It would be presumptuous to assume that the design principles contain all the necessary information that will need to be either refined, adapted, or expanded in future research efforts. Moreover, the highlighted areas for improvement of the design principles based on the reusability framework could be addressed in a subsequent design science cycle and future research.

## References

- [1] L. Alzubaidi, J. Zhang, et al., Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions, *Journal of Big Data* 8 (2021), 1-74.
- [2] S. Hinterstoisser, O. Pauly et al., An annotation saved is an annotation earned: Using fully synthetic training for object detection in 'Proceedings of the IEEE/CVF' (2019).
- [3] V. Seib, B. Lange, S. Wirtz, Mixing Real and Synthetic Data to Enhance Neural Network Training - A Review of Current Approaches, 2020, arXiv:2007.08781.
- [4] X. Zhang, Y. Fu, S. Jiang, X. Xue, Y.G. Jiang, G. Agam, Stacked multichannel autoencoder—an efficient way of learning from synthetic data. *Multimedia Tools and Applications* 77 (2018), 26563-26580.
- [5] J.J. Bird, D.R. Faria, A. Ekárt, P.P. Ayrosa, From simulation to reality: CNN transfer learning for scene classification, in 2020 IEEE 10th International Conference on Intelligent Systems (2020), IEEE, 619-625.
- [6] M. Krump, M. Ruß, P. Stütz, Deep learning algorithms for vehicle detection on UAV platforms: first investigations on the effects of synthetic training, in *Modelling and Simulation for Autonomous Systems: 6th International Conference, MESAS 2020, Palermo, Italy*, 50-70.
- [7] F. Kuhnke, J. Ostermann, Deep head pose estimation using synthetic images and partial adversarial domain adaption for continuous label spaces, in *Proceedings of the IEEE/CVF International Conference on computer vision* (2019), 10164-10173.
- [8] J. Liu, H. Rahmani, N. Akhtar, A. Mian, Learning human pose models from synthesized data for robust RGB-D action recognition. *International Journal of Computer Vision* 127 (2019), 1545-1564.
- [9] C. Taleb, L. Likforman-Sulem, C. Mokbel, Improving deep learning Parkinson's disease detection through data augmentation training, in *Pattern Recognition and Artificial Intelligence: Third Mediterranean Conference, MedPRAI, 2020, Istanbul, Turkey*, 79-93.
- [10] A. Hevner, S. March, J. Park, S. Ram, Design science in information systems re-search, *MIS Quarterly* 28(1) (2004), 75-105.
- [11] V.K. Vaishnavi, W. Kuechler, *Design science research methods and patterns: innovating information and communication technology*, 2015, Crc Press.
- [12] B. Friedman, P.H. Kahn, A. Borning, A. Hultdtgren, *Value sensitive design and information systems. Early engagement and new technologies: Opening up the laboratory*, 2013, 55-95.
- [13] S. Gregor, A.R. Hevner, Positioning and presenting design science research for maximum impact. *MIS Quarterly* (2013), 337-355.

- [14] J. Mingers, J. Brocklesby, *Multimethodology: Towards a framework for mixing methodologies*. *Omega*, 25(5), 1997, 489-509.
- [15] J.R. Venable, J. Pries-Heje, R.L. Baskerville, *Choosing a design science research methodology in ACIS 2017 Proceedings*, 2017, 112.
- [16] K. Durani, A. Eckhardt, T. Kollmer, *Towards ethical design science research in ICIS 2021 Proceedings*, 2021, 3.
- [17] R. Baskerville, J. Pries-Heje, *Explanatory design theory*. *Business & Information Systems Engineering*, 2 (2010), 271-282.
- [18] S. Gregor, L. Chandra Kruse, S. Seidel, *Research perspectives: the anatomy of a design principle*. *Journal of the Association for Information Systems*, 2020.
- [19] F. Möller, T. Schoormann, G. Strobel, M.R.P. Hansen, *Unveiling the Cloak: Kernel Theory Use in Design Science Research in ICIS 2022 Proceedings*, 2022.
- [20] F. Niederman, S. March, *The "theoretical lens" concept: We all know what it means, but do we all know the same thing?* *Communications of the Association for Information Systems*, 44(1), 2019, 1.
- [21] S. Lee, B. Park, A. Kim, *Deep learning based object detection via style-transferred underwater sonar images*. *IFAC-PapersOnLine*, 52(21), 2019, 152-155.
- [22] T. Body, X. Tao, Y. Li, L. Li,, N. Zhong, *Using back-and-forth translation to create artificial augmented textual data for sentiment analysis models*. *Expert Systems with Applications*, 178, 2021, 115033.
- [23] F. Condrea, V.A. Ivan, M. Leordeanu, *In search of life: Learning from synthetic data to detect vital signs in videos*, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, 298-299.
- [24] H.F. Zaki, F. Shafait, A. Mian, *Viewpoint invariant semantic object and scene categorization with RGB-D sensors*. *Autonomous Robots*, 43 (2019), 1005-1022.
- [25] A. Lahiri, A. Agarwalla, P.K. Biswas, *Unsupervised domain adaptation for learning eye gaze from a million synthetic images: An adversarial approach*, in *Proceedings of the 11th Indian Conference on Computer Vision (2018), Graphics and Image Processing*, 1-9.
- [26] H. Venkateswara, S. Chakraborty, S. Panchanathan, *Deep-learning systems for domain adaptation in computer vision: Learning transferable feature representations*. *IEEE Signal Processing Magazine*, 34(6), 2017, 117-129.
- [27] N. Aranjuelo, S. García, E. Loyo, L. Unzueta, O. Otaegui, *Key strategies for synthetic data generation for training intelligent systems based on people detection from omnidirectional cameras*. *Computers & Electrical Engineering*, 92 (2021), 107105.
- [28] S.Z. Valtchev, J. Wu, *Domain randomization for neural network classification*. *Journal of big Data*, 8(1), 2021, 94.
- [29] L. Wan, J. Zhou, B. Zhang, *Data Synthesis for Document Layout Analysis*, in *International Symposium on Emerging Technologies for Education*, 2020, 244-252.
- [30] H. Abu Alhaija, S.K. Mustikovela, L. Mescheder, A. Geiger, C. Rother, *Augmented reality meets computer vision: Efficient data generation for urban driving scenes*. *International Journal of Computer Vision*, 126, 2018, 961-972.
- [31] M. Müller, V. Casser, J. Lahoud, N. Smith, B. Ghanem, *Sim4cv: A photo-realistic simulator for computer vision applications*. *International Journal of Computer Vision*, 126, 2018, 902-919.
- [32] M. Valerio Giuffrida, H. Scharr, S.A. Tsiftaris, *Arigan: Synthetic arabidopsis plants using generative adversarial network*, in *Proceedings of the IEEE international conference on computer vision workshops*, 2017, 2064-2071.
- [33] T. Scheck, R. Seidel, G. Hirtz, *Learning from theodore: A synthetic omnidirectional top-view indoor dataset for deep transfer learning*, in *Proceedings of the IEEE/CVF Winter conference on applications of computer vision (2020)*, 943-952.
- [34] D.L. Morgan, *Qualitative Research Methods: Focus groups as qualitative research (2)*, 1997, Thousand Oaks SAGE Publications, Inc.
- [35] D.A. Gioia, K.G. Corley, A.L. Hamilton, *Seeking qualitative rigor in inductive research: Notes on the Gioia methodology*. *Organizational research methods*, 16(1), 2012, 15-31.

- [36] K.K. Fu, M.C. Yang, K.L. Wood, Design principles: Literature review, analysis, and future directions. *Journal of Mechanical Design*, 138(10), 2016, 101103.
- [37] J. Venable, J. Pries-Heje, R. Baskerville, FEDS: a framework for evaluation in design science research. *European Journal of Information Systems*, 25 (2016), 77-89.
- [38] C. Sonnenberg, J. Vom Brocke, Evaluation patterns for design science research artefacts, in *Practical Aspects of Design Science: European Design Science Symposium (2012)*, Leixlip, Ireland, 71-83.
- [39] J. Iivari, M.R.P. Hansen, A. Haj-Bolouri, A framework for light reusability evaluation of design principles in design science research, in *13th International Conference on Design Science Research and Information Systems and Technology: Designing for a Digital and Globalized World*, 2018.
- [40] M. Van Den Haak, M. De Jong, P. Jan Schellens, Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue. *Behaviour & Information Technology* (2003), 22(5):339-351.
- [41] W. Hwang, G. Salvendy, Number of people required for usability evaluation: the 10±2 rule. *Communications of the ACM* (2010) 53(5):130-133.
- [42] C. Woo, A. Saghafi, A. Rosales, What is a Contribution to IS Design Science Knowledge?, in *Thirty Fifth International Conference on Information Systems*, 2014, Auckland.
- [43] A. Hevner, S. Chatterjee, *Design science research in information systems*, Design research in information systems, 2010, Springer, Boston, 9-22.
- [44] R. Baskerville, M. Kaul, J. Pries-Heje, V.C. Storey, E. Kristiansen, Bounded creativity in design science research, in *ICIS 2016 Proceedings*.