

Innovating with Artificial Intelligence: Capturing the Constructive Functional Capabilities of Deep Generative Learning

Peter Hofmann
Project Group Business and
Information Systems Engineering of
the Fraunhofer FIT,
University of Bayreuth
peter.hofmann@fit.fraunhofer.de

Timon Rüchel
FIM Research Center,
University of Augsburg
timon.ruechel@fim-rc.de

Nils Urbach
Project Group Business and
Information Systems Engineering of
the Fraunhofer FIT, Frankfurt
University of Applied Sciences
nils.urbach@fim-rc.de

Abstract

As an emerging species of artificial intelligence, deep generative learning models can generate an unprecedented variety of new outputs. Examples include the creation of music, text-to-image translation, or the imputation of missing data. Similar to other AI models that already evoke significant changes in society and economy, there is a need for structuring the constructive functional capabilities of DGL. To derive and discuss them, we conducted an extensive and structured literature review. Our results reveal a substantial scope of six constructive functional capabilities demonstrating that DGL is not exclusively used to generate unseen outputs. Our paper further guides companies in capturing and evaluating DGL's potential for innovation. Besides, our paper fosters an understanding of DGL and provides a conceptual basis for further research.

1. Introduction

VanRullen and Reddy [64] developed a deep learning algorithm that can translate a brain activity signal capturing the stimulus of looking at a human face image into a reliable estimation of that picture. To take this “mind-reading” [64, p. 7] experiment one step further, they asked their probands only to imagine certain faces and reconstructed pictures of these faces with 84.2% correct image decoding [64]. As another example, Grisoni et al. [26] trained a neural network on known anticancer peptides to create novel amino acid sequences. By testing their activity against cancer cells, half of the generated peptides revealed as being effective without affecting human blood cells [26].

Both examples deploy deep generative learning (DGL) models, a currently emerging type of deep learning models [41]. Especially Generative Adversarial Networks (GANs) [25] and Variational

Autoencoders (VAE) [39, 54] are recent and viable approaches to implementing DGL [33]. Generally, DGL models' capability to capture complex, non-linear patterns of the input data enables the generation of previously unseen but plausible data [41, 50]. Thus, DGL models can create an unprecedented variety of new outputs [41]. For example, DGL models can already change the posture of humans within images [4], translate text into images [76], or create music [63] and might soon create visual, text, and audio data that is tough to distinguish from real-world data [51].

However, research on DGL has so far predominantly focused on technical improvements despite its growing awareness. While Artificial Intelligence (AI), in general, raises high expectations for product and process improvements [53] as well as the innovation of business models [6], the variety of technological breakthroughs in DGL leaves companies with a vague and elusive understanding of DGL's potential for innovation. Although existing research about AI in general already conceptualizes AI's functions [e.g., 31, 52] and its business impacts [e.g., 7, 32, 52], it does not yet bridge the gap between current technological breakthroughs in DGL and their constructive functional capabilities that foster innovation. In particular, the constructive characteristic of DGL raises new and yet unanswered questions. Leaving this gap open may result in inefficiently allocated investments or missing valuable opportunities. However, to allow organizations to utilize AI effectively, an in-depth understanding of the specific underlying technology is vital [53]. Hence, to break ground for creating business value using DGL, our paper aims at fostering the general understanding of DGL within the IS domain. To this end, we strive to answer the following research question:

What are DGL's constructive functional capabilities and how do they support innovation?

To address this question, we conducted an extensive and structured literature review. With our paper, we bridge the gap between DGL's technical characteristics and its potential for innovation. Thereby, we understand innovation as the "multi-stage process whereby organizations transform ideas into new/improved products, service or processes, in order to advance, compete and differentiate themselves successfully in their marketplace" [3, p. 1334]. We contribute to the academic discourse by identifying and describing six constructive functional capabilities as well as discussing DGL's potentials for innovation.

2. Deep Generative Learning

Rai et al. [52] refer to AI as machines performing cognitive functions that are typically associated with humans. Currently, the dominant procedure to implement AI relies on machine learning, a technique that enables machines to improve by learning from experience and data [24]. As a type of machine learning, deep learning models [24] represent artificial neural networks consisting of several deep, hidden layers [34]. As such, deep learning models can explore complex, non-linear patterns within data [34].

Based on deep learning, early endeavors of implementing DGL models date back to 1992 [33] and have recently gained traction by the rise of cost function-based models such as VAEs [39] or GANs [25] and energy-based models such as Deep Belief Networks (DBN) [30, 50]. These models significantly contribute to achieving AI's final objective of adding "intelligence into machine[s]" [70, p. 1]. Thus, researchers view the rise of DGL models as an important milestone in the field of AI [70].

Despite this trend, discriminative models have been preferred to generative ones in recent times. Generative models learn the joint distribution $p(x, y)$ of the input x and the label y . To predict y , they can calculate $p(y|x)$ using the Bayes rule. In contrast, discriminative models directly calculate $p(y|x)$ [1] and do not have to solve the more general problem of modeling the distribution of input features x . Thus, discriminative models tend to be more accurate and less computationally-intensive [46]. But by combining probabilistic generative models with deep learning architectures [39, 54], DGL models can achieve high efficiency [50]. Specifically, DGL models are generative neural networks consisting of a recurrent neural network (RNN) and a generative subsystem [15]. Consequently, DGL can implicitly as well as explicitly [50] capture the inner probabilistic distribution of data over several variables [24] even for complex, high-dimensional, imperfect, or noisy data [69]. Therefore, DGL models are able to sample

unseen data efficiently [54] even in the case of minimal amounts of labels [18] or entirely unsupervised. We understand such generation of data as the key ability of DGL models and consider it as DGL's "constructive" characteristic. As mentioned above, we understand DGL algorithms as a class of algorithms combining RNNs with a generative subsystem. Given this, we focus on DGL's constructive functional capabilities within our paper.

3. Research Method

We conducted a structured literature review as an established method in IS research to identify, synthesize, and analyze the, in most cases, technical papers from an application perspective [e.g., 65, 68].

After an initial literature search to become familiar with the established concepts and terms, we developed the following search string: "deep generative model" OR "deep generative learning" OR "generative deep learning" OR "constructive machine learning" OR "generative neural networks" OR „generative machine learning“. When developing the search string, we deliberately addressed the different terms used in the literature. We applied the search string on title, abstract, and keywords at five established databases to identify journal or conference articles. In this search, we received a total of 217 papers without duplicates (AIS eLibrary: 4; EBSCOhost: 13; Web of Science: 218; IEEE Xplore Digital Library: 122; ScienceDirect: 11). After screening titles, abstracts, and, if necessary, full texts, we deemed 181 papers as relevant for our study. We included English full papers that develop a DGL model and apply it to solve at least some kind of real-world problem. We excluded papers that only discuss technical improvements without focusing on a real-world problem since they do not allow us to bridge the gap between technological opportunity and the organizational impact. Based on the insights from the literature sample, we conducted another manual search to account for missing search results in the domain of text-to-speech as well as music generation. We opted for a manual search instead of a forward or backward search, as this allowed us to complete our already large literature body in a targeted manner. Based on this manual search, we could enhance our body of literature by another 14 papers. In total, we identified 195 relevant papers for our analysis.

We followed a concept-centric approach and iteratively developed and discussed a concept matrix in our paper team to analyze the relevant papers. In doing so, we mapped all 195 relevant papers to concepts (i.e., constructive functional capabilities). We further recorded processed data types. We accepted a constructive functional capability as a

concept as long as we could assign more than two papers from our sample. Due to the submission’s page restriction, we cannot insert a table covering all articles and their assignments. Instead, we provide a selection of papers and information on relative proportions of our literature sample when introducing a constructive functional capability.

4. Deep Generative Learning’s Constructive Functional Capabilities

As a result of our structured literature review, we propose six constructive functional capabilities of DGL models, namely *creating*, *transforming*, *modifying*, *reconstructing*, *reducing dimensionality*, and *predicting*. Table 1 provides an overview of the observed constructive functional capabilities and their shares within the sample. We consider *predicting* as a partly constructive functional capability since its main output (i.e., labels or predictions) does not necessarily require a constructive approach. We include it nevertheless, as several papers constructively perform *predicting* tasks by creating data using DGL models. Besides, we observe papers using DGL models for a clustering task as well as a pretraining task improving a non-DGL model. Since both observations are only covered by one paper, we do not include them in our results. For every constructive functional capability, we generalize its core capability and illustrate it by means of specific applications provided within our sample. Eventually, we elaborate on the constructive functional capability’s rationale for deploying specific DGL models. We excluded papers that do not fall into any of the constructive functional capabilities resulting in 155 papers as our total sample. We specify the percentage of constructive functional capabilities based on the total sample of 155 papers. Multiple assignments of papers are possible, such that

percentages do not add up to 100%. To foster readability throughout the upcoming sections, we only name the model’s bases (e.g., VAE or GAN) and do not specify any of their special types or specific deformations.

4.1 Creating

Our extensive literature review reveals that DGL models can autonomously create data that is similar to known instances. Within our total sample, 43% of the papers belong to the respective constructive functional capability. *Creating* processes audio, text, visual (i.e., images and videos), and domain-specific data. Domain-specific data includes several data types that depend on the particular application area such as molecular graph, 3D, or gene expression data. Across these formats, *creating* generates data that is new, previously unseen, and generally interpretable for humans (e.g., images or music).

In the vital domain of visual data, GANs can, for instance, generate pictures of a specific type of plant to increase the volume of annotated training data and thereby improve other models’ classification accuracy [23]. Further, Despraz and Gomez strongly activate individual neurons or groups of neurons of a convolutional neural network (CNN). They let a DGL model illustrate the effects on the CNN’s internal representations by producing respective output images to foster model transparency and understanding [16]. Proceeding to audio data generation, DGL models frequently generate music. Especially GANs fulfill this task [e.g., 10, 27] due to their capability of creating clear and realistic musical fragments. But music generation is not limited to GANs as, for example, van den Oord et al. [63] invented a versatile method called WaveNet, to handle multiple audio generation tasks such as music generation. Besides the prevalent

Table 1. DGL’s constructive functional capabilities

Functional capability	Definition	Percentage of sample
Creating	Create data that is similar to known data	43%
Transforming	Generate data by transforming data into another data type	12%
Modifying	Modify input data without changing its data type	17%
Reconstructing	Restore data quality by embedding to and decoding from a latent space	12%
Reducing dimensionality	Provide basis for further processes by reducing dimensionality	14%
Predicting	Learn a mapping from input to output [46]	8%

domains of visual data and music creation, an encoder-attention-decoder model can, for instance, suggest new and less expensive advertising keywords of high quality [78]. Additionally, Shu et al. [60] make GANs create realistic headlines that augment training data sets to improve clickbait detection models. Further, we find promising work on various domain-specific applications of DGL models. For example, some papers [e.g., 19, 35] investigate how DGL can improve and even generate new, highly efficient metasurfaces. Metasurfaces are micro-structured artificial surfaces that affect the behaviors of electromagnetic waves [35]. Further, Samanta et al. [56] implement a VAE that creates novel, valid and diverse molecules aiming to benefit patients therapeutically. As the last example, GANs can support physicists by generating new configurations of quantum fields that are in line with physical observables [79].

Despite the variety of deployed models within *creating*, all applications use DGL models due to their ability to capture complex probability distributions over multiple input variables. DGL models can draw new samples of the learned data distribution [24].

4.2 Transforming

Like *creating*, *transforming* (12%) generates new and previously unseen data. But in contrast, *transforming* translates data from one data type into another. The constructive functional capability processes audio, text, visual, and domain-specific data.

Within audio data, *transforming* converts written text sequences to voices reading the text. We find several papers conducting such text-to-speech syntheses, which often deploy GANs [e.g., 37, 55] due to their ability to generate high-fidelity speech [72]. Besides text-to-speech and attributing to domain-specific data, GANs can infer voxel-based 3D models from several 2D images of specific objects like chairs or airplanes [20]. Further, visual data generation is, amongst others, relevant to the medical area. As an example, VAEs can translate an electrocardiography signal into an activation map indicating the health of specific heart regions [22]. Also, the above-mentioned example of translating brain activity signals into pictures of faces [64] belongs to *transforming*.

Transforming deploys predominantly VAEs and GANs: VAEs can create meaningful latent representations [e.g., 64] and GANs are able to generate high quality, realistic data [e.g., 55].

4.3 Modifying

The constructive functional capability *modifying* (17%) processes audio, text, visual, and domain-specific data. To clearly differentiate from *creating*, we understand *modifying* as modifications of input data without changing its data type.

For instance, the sound of a speaking or singing voice can be changed using a VAE [61]. Within the vital domain of visual data, DGL models can anonymize faces by exchanging them with unseen ones without losing the original facial expression [e.g., 28, 45]. VAE and GAN models can manipulate human poses and their outfits [4, 12], facial expressions [21, 43], their gaze [59], or their face age [75]. Further, Xu et al. implemented a GAN-based model for domain-transfer that can translate face avatars into images of human faces or even simple line drawings into colored pictures [71]. An encoder-decoder network can edit the color of voxel-based 3D objects like cars [73] or a GAN can modify 3D models of buildings such that these models satisfy multiple physical constraints [62].

Our sample indicates that *modifying* DGL models predominantly deploy VAEs and GANs. Again, VAEs benefit from their ability to construct a structured as well as interpretable latent space and GANs are able to generate high-quality data [4].

4.4 Reconstructing

The constructive functional capability *reconstructing* (12%) restores data quality by embedding inputs to a latent representation and subsequently reconstructing them. This allows for, e.g., imputing missing parts of data, reducing noise, or sharpening data. *Reconstructing* processes mainly visual, but also audio and domain-specific data.

For instance, Deep Boltzmann Machines are able to complete shapes of several objects from noisy synthetic aperture radar pictures [17]. Further, a mixture of CNN- and flow-based models can fuse two pictures with different foci to obtain one picture with a higher level of detail [66]. Further, GANs and models combining properties of both GANs and VAEs can infer randomly missing parts of images reliably. This mainly includes human faces [e.g., 42, 77], but also landscapes, buildings, or other scenes [47, 74]. Next to visual data, for instance, VAEs can infer missing values within electronic health records of patients to predict their in-hospital mortality [36] and transform noisy speech (e.g., due to background sounds) to clean speech [e.g., 2, 58].

Our sample indicates that *reconstructing* DGL models also focus on VAEs and GANs. For one thing, VAEs provide a principled way of inferring the latent

representations of both training and observed data [2]. For another thing, GAN's ability to generate pictures that are both globally and locally semantically consistent [77] as well as visually realistic [42] leads to its deployment for many *reconstructing* tasks. Some applications even deploy a mixture of both to benefit from the upsides of the two models.

We also find some papers that only try to reconstruct the initial input data as accurately as possible. The rationale for such reconstructions lies in validating the model's performance [49]. However, as such reconstructions generate at best data that is very similar to the input [49], we do not consider them as a part of *reconstructing*.

4.5 Reducing Dimensionality

The constructive functional capability *reducing dimensionality* (14%) maps the distribution of input data to a lower-dimensional feature space by generating a meaningful latent representation. Since this representation is the output of the DGL model and serves as input for further processes without any decoding, *reducing dimensionality* is clearly separated from *reconstructing*. Further, these latent representations are not interpretable for humans, such that *reducing dimensionality* differentiates from *creating*. The constructive functional capability processes audio, text, visual, and domain-specific data and reveals two different patterns.

First, *reducing dimensionality* is relevant to approaches that combine DGL models with further, more conventional machine learning algorithms. Specifically, our sample shows that latent representations computed by DGL models can serve as a basis for subsequent non-DGL classification, clustering, and quantification algorithms. As an example of classification, Ko and Kim [40] map the complex relationships within high-dimensional industrial process data to latent features deploying a VAE. Then, they analyze these latent features with a k-nearest neighbors and a support vector machine algorithm to detect process faults. Additionally, Chen et al. [11] use an autoencoder to extract dense and low-dimensional features from text documents and, on that basis, a spectral clustering [48] technique to measure the similarity of these text documents.

Second, we find research that focuses on deploying a DGL model to compress data volume such that the information loss is minimized. Specifically, Gupta and Kumar [29] reduce ultrasound images to their latent representations using a VAE to transfer less data volume to a cloud-system that reconstructs the images without significant loss of resolution.

All models within *reducing dimensionality* outperform more conventional baseline models as they use DGL models' ability to learn complex mappings from input data to their compact but precise latent representation [8]. Papers belonging to the constructive functional capability *reducing dimensionality* predominantly implement VAEs, other encoder-decoder networks, or DBNs.

4.6 Predicting

Following Murphy [46], we understand *predicting* (8%) as learning "a mapping from inputs x to outputs y " [46, p. 2]. In other words, *predicting* determines y given x . Consequently, and in contrast to the previous constructive functional capabilities, *predicting* focuses on y , not on x . In general, the predicted label y can be either real-valued (i.e., continuous) or categorical. Accordingly, we find two separate ways of *predicting*: Regression for continuous and classifying for categorical labels [46]. In general, classification and regression tasks can be performed without generating data. However, we observe models that predict by creating data for both regression and classification. Thus, we only include those constructive parts of *predicting*, making it a partly constructive functional capability that processes audio, visual, and domain-specific data.

We describe the pattern regression as "estimating future events and conditions on a continuous scale" [31, p. 9]. To achieve such constructive predictions through regression, DGL models estimate future events or conditions by creating data describing these predicted states. Since *predicting* is their primary objective, we do not allocate these papers to the constructive functional capability *creating*. To illustrate the pattern, a VAE is, for instance, capable of generating images of activation maps indicating future human brain topographies to predict aging and neurodegeneration of the brain [14]. As another example, Watter et al. [67] create images describing dynamical systems like future movements of a robot arm.

In contrast to regression, we interpret the pattern classification as annotating data with known and discrete labels. DGL models can constructively perform such annotations by reconstructing data from a latent space. As an example within visual data or video classification, Kim et al. [38] implement a model that can decide whether Virtual Reality (VR) videos are likely to cause sickness while watching or not. To do so, the authors train a GAN to reconstruct VR videos that do not cause sickness. After training, the GAN's generator can reconstruct such videos realistically. But, when reconstructing videos that are

likely to cause sickness, the GAN fails to reconstruct the videos accurately. Thus, for videos causing sickness, the reconstructed version significantly differs from the original and the GAN inherently labels the VR video as causing nausea. We find similar reconstructive classification approaches, e.g., for detecting spoofing [13], cyberattacks [9], or production anomalies in images of screw holes [44]. Further, we observe another way of constructively classifying images in a paper by Xu et al. [69]. They train a VAE to map images and several words that describe their content in a joint distribution. After training, the model can infer discrete labels for new images that serve as a classification.

Both patterns reveal differences regarding their mainly deployed models. Regression uses VAE's ability to efficiently learn low-dimensional representations of complex data [14] and the fact that these latent representations can be easily and robustly controlled [67]. Besides, classification focuses on GANs and VAEs, since both can effectively learn features over several areas of the input, even for large-scale datasets with a high amount of free parameters [5]. Additionally, Bhat et al. [5] state that deploying DGL models for classification tasks minimizes the cost of feature modeling, since DGL models deeply understand input data distributions and do, therefore, not require separate statistical tests to extract the most important features.

5. DGL's Potential for Innovation

When revisiting DGL's constructive functional capabilities, we identify three major potentials for innovation.

First, DGL may enable new or enhance existing product or service features. Enhanced product or service features can become a differentiating factor of competitive importance. For example, enhanced text-to-speech syntheses could take conversational agents to the next level. Another example constitutes software-based noise reduction to enhance picture quality. Further, DGL can help humans to improve their interaction with AI systems by adjusting human inputs such that these systems can better understand them [57]. While in some industries, improving a product or service feature depended on years of overcoming an engineering challenge, DGL can be a differentiator based on software instead of hardware. Thus, software-based capabilities could jeopardize companies' competitive edge through years of experience in mastering engineering challenges. DGL could further substantiate complementary value-adding offerings such as changing outfits in an online fashion store. As an example, SUPERPERSONAL

(www.superpersonal.com) allows customers to watch a video of themselves wearing clothes they can choose. Further, DGL has the potential to complement prediction results. For instance, DGL allows for predicting integrated activation maps that reflect predicted states. Besides, DGL could further break innovation barriers by providing solutions to problems that enable other ideas to work. For example, DGL can automatically anonymize faces in video-streams and, hence, contribute to solving the data protection issues of innovative business models. For instance, BrighterAI (www.brighter.ai) already offers face anonymization services. Moreover, companies could sell the creation of DGL models (e.g., music) or provide services that enhance a user's creation (e.g., conversion of scribbles into pictures). Thereby, our results demonstrate a wide variety of possible output types. However, this potential also leads to a bunch of new questions, ranging from economic considerations such as the value of DGL-made creations to rather philosophical considerations like the extent of DGL's creativity.

Second, DGL has the potential to augment or automate processes within organizations. For instance, autoRetouch (www.autoretouch.com), a startup founded in early 2020, automates the workflow of retouching fashion images using GANs. Besides, our results include applications generating context-specific headlines or creating 3D models from 2D images. We also find several DGL applications that disrupt the predominantly costly trial and error process of molecular design by automatically suggesting biologically effective molecular structures. These examples indicate the potential to rethink the way companies approach processes well beyond simple process automation. Taking applications as an example that convert simple scribbles or drawings into complex colored pictures raises the question of who the designer of a new product could be or to what extent DGL mediates the ability of people expressing their creativity. Accordingly, future research could, for example, address and analyze the collaboration between humans and DGL applications from a capability perspective. Especially for constructive tasks, it seems promising to analyze how DGL could improve or complement human capabilities.

Third, we also observe an indirect way of DGLs creating potential. Here the potential does not arise directly through the application of DGL but through the creation of the technological prerequisites. For example, DGL demonstrated the potential to reduce data size or dimensionality as a basis for other algorithms. Also, DGL could enhance the transparency of machine learning models. Besides, DGL incorporates the potential to improve the

performance of other machine learning models by creating synthetic training data or removing data quality issues. Also, the reconstruction of missing data (e.g., missing parts of images) has great potential to boost image recognition models' performance. These examples indicate that DGL's potential does not necessarily have to be direct; there are also indirect potentials that one should not neglect when considering DGL's potential for innovation. Nevertheless, it remains unclear so far whether this potential can become a competitive factor.

However, along with these three superordinate ways of generating innovation potential, many application-, organization- and market-related questions remain unanswered motivating future research in this field. Exemplary questions include: How can DGL's constructive functional capabilities disrupt business models? What are the pitfalls of data generated by DGLs? How can you prevent the misuse of DGL? Will DGL change innovation processes in organizations? How do DGL-based innovations transform the market of creative goods? Will tech companies expand into new markets?

We further critically point out that DGL does not only offer new functional capabilities. While we can observe new possibilities for innovation when dealing with functions such as *creating*, functions such as *predicting* have already been frequently covered by many other algorithms. However, even a functional capability such as *creating* could be accomplished with other AI algorithms. Consequently, DGL is not the only approach for product, service, or process innovation. Furthermore, a product's technical feasibility does not necessarily translate into its beneficial exploitation.

6. Conclusion

In conclusion, our paper sheds light on the emerging field of DGL models to better understand their practical potential and, therefore, their consequences for organizations. By conducting an extensive structured literature review, we synthesized DGL's constructive functional capability from 195 papers to form a solid basis for a business-related discourse. We find that the relatively narrow notion of DGL as being able to create unseen but plausible data [41] translates into a remarkable variety of constructive functional capabilities. By discussing DGL's potential for innovation, we establish the link between DGL's technical characteristics and the arising potential for innovation. Without closing this gap, discussions about DGL's organizational impact would remain superficial and poorly comparable. We contribute to the literature in two ways: First, we

identified and described six constructive functional capabilities of DGL. By conceptualizing DGL's functional profile, our results guide both practice and future research by providing a solid base to evaluate and discuss DGL's use in products, services, and processes. Second, we provide an initial discussion on DGL's potential for innovation that underpins the relevance of further research on DGL from an innovation perspective. In this context, we demonstrated that DGL is multifunctional and not limited to creating some type of output, such as the media-effective deepfakes.

Our study is not without limitations. First, the identification of DGL's constructive functional capabilities solely relied on academic literature. In this regard, it is possible that academic literature does not yet cover all constructive functional capabilities of DGL. However, in a research-driven field such as AI, we are confident that the academic literature includes DGL's hitherto known constructive functional capabilities. Moreover, it is in the inherent nature of the subject that functional capabilities evolve over time. Second, although conducting an extensive literature search, there is a minor chance that our search does not cover all relevant papers in an emerging research field where concepts can develop rapidly. However, our paper offers a good springboard for further research. Third, we do not incorporate empirical evidence from practice when analytically discussing how DGL models translate into potential for innovation. However, our current state of research synthesis regarding DGL's constructive functional capabilities allows future research to assess the actual potential for innovation in or across domains. Notwithstanding the limitations, our conceptualization is an essential step for research and practice to complement the business perspective to DGL's technological breakthroughs.

References

- [1] Andrew Y. Ng and Michael I. Jordan, "On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naive Bayes", *Advances in Neural Information Processing Systems*. 2002.
- [2] Bando, Y., M. Mimura, K. Itoyama, K. Yoshii, and T. Kawahara, "Statistical Speech Enhancement Based on Probabilistic Integration of Variational Autoencoder and Non-Negative Matrix Factorization", in *2018 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Calgary, Canada. 2018.
- [3] Baregheh, A., J. Rowley, and S. Sambrook, "Towards a multidisciplinary definition of innovation", *Management Decision*, 47(8), 2009, pp. 1323–1339.

- [4] Bem, R. de, A. Ghosh, A. Boukhayma, T. Ajanthan, N. Siddharth, and P. Torr, "A Conditional Deep Generative Model of People in Natural Images", in 2019 IEEE Winter Conf. on Applications of Computer Vision (WACV), Hawaii, USA. 2019.
- [5] Bhat, R.R., V. Viswanath, and X. Li, "DeepCancer: Detecting Cancer via Deep Generative Learning Through Gene Expressions", in 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC / PiCom / DataCom / CyberSciTech), Orlando, FL, USA. 2017.
- [6] Brynjolfsson, E. and A. McAfee, "The business of artificial intelligence", *Harvard Business Review*, 2017, pp. 1–20.
- [7] Brynjolfsson, E. and T. Mitchell, "What can machine learning do? Workforce implications", *Science*, 358(6370), 2017, pp. 1530–1534.
- [8] Chaidaroon, S., D.H. Park, Y. Chang, and Y. Fang, "node2hash: Graph aware deep semantic text hashing", *Information Processing & Management*, in press, 2019.
- [9] Chandy, S.E., A. Rasekh, Z.A. Barker, and M.E. Shafiee, "Cyberattack Detection Using Deep Generative Models with Variational Inference", *Journal of Water Resources Planning and Management*, 145(2), 2019.
- [10] Chen, H., Q. Xiao, and X. Yin, "Generating Music Algorithm with Deep Convolutional Generative Adversarial Networks", in 2019 IEEE 2nd International Conference on Electronics Technology (ICET), Chengdu, China.
- [11] Chen, H., J. Xu, Q. Wang, and B. He, "A Document Modeling Method Based on Deep Generative Model and Spectral Hashing", in 9th International Conference on Knowledge Science, Engineering and Management (KSEM), Passau, Germany. 2016.
- [12] Chen, Y., S. Xia, J. Zhao, Y. Zhou, Q. Niu, R. Yao, and D. Zhu, "Appearance and shape based image synthesis by conditional variational generative adversarial network", *Knowledge-Based Systems*, 193, 2020.
- [13] Chettri, B., T. Kinnunen, and E. Benetos, "Deep Generative Variational Autoencoding for Replay Spoof Detection in Automatic Speaker Verification", *Computer Speech & Language*, 63, 2020.
- [14] Choi, H., H. Kang, and D.S. Lee, "Predicting Aging of Brain Metabolic Topography Using Variational Autoencoder", *Frontiers in Aging Neuroscience*, 10, 2018.
- [15] Danihelka, I., J.D. Rezende, and S. Mohamed, "Generative Neural Networks"(US10657436B2), 19.05.2020.
- [16] Despraz, J., S. Gomez, H.F. Satizabal, and C.A. Pena-Reyes, "Exploring Internal Representations of Deep Neural Networks", in *Computational Intelligence*, C. Sabourin, J.J. Merelo, K. Madani, and K. Warwick, Editors. 2019. Springer Nature Switzerland.
- [17] Dou, F., W. Diao, X. Sun, Y. Zhang, and K. Fu, "Aircraft Reconstruction in High-Resolution SAR Images Using Deep Shape Prior", *ISPRS International Journal of Geo-Information*, 6(11), 2017.
- [18] Durk P. Kingma, Shakir Mohamed, Danilo Jimenez Rezende, and Max Welling, "Semi-supervised Learning with Deep Generative Models", in *Neural Information Processing Systems 2014*, Montreal, Canada. 2014.
- [19] Fan, J.A., "Generating high performance, topologically-complex metasurfaces with neural networks", in 2019 Conference on Lasers and Electro-Optics (CLEO), San Jose, CA, USA. 2019.
- [20] Gadelha, M., S. Maji, and R. Wang, "3D Shape Induction from 2D Views of Multiple Objects", in 2017 International Conference on 3D Vision (3DV), Verona, Italy. 2017.
- [21] Geng, J., T. Shao, Y. Zheng, Y. Weng, and K. Zhou, "Warp-Guided GANs for Single-Photo Facial Animation", *ACM Transactions on Graphics*, 37(6), 2018.
- [22] Ghimire, S., J. Dhamala, P.K. Gyawali, J.L. Sapp, M. Horacek, and L. Wang, "Generative Modeling and Inverse Imaging of Cardiac Transmembrane Potential", in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, A.F. Frangi, J.A. Schnabel, C. Davatzikos, C. AlberolaLopez, and G. Fichtinger, Editors. 2018.
- [23] Giuffrida, M.V., H. Scharr, and S.A. Tsafaris, "ARIGAN: Synthetic Arabidopsis Plants using Generative Adversarial Network", in 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, Italy. 2017.
- [24] Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
- [25] Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets", in *Neural Information Processing Systems 2014*, Montreal, Canada. 2014.
- [26] Grisoni, F., C.S. Neuhäus, G. Gabernet, A.T. Mueller, J.A. Hiss, and G. Schneider, "Designing Anticancer Peptides by Constructive Machine Learning", *ChemMedChem*, 13(13), 2018, pp. 1300–1302.
- [27] Guan, F., C. Yu, and S. Yang, "A GAN Model With Self-attention Mechanism To Generate Multi-instruments Symbolic Music", in 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary. 2019.
- [28] Guo, S., S. Feng, Y. Li, S. An, and H. Dong, "Integrating Diversity into Neural-Network-Based Face Deidentification", in 37th Chinese Control Conference (CCC 2018), Wuhan, China. 2018.
- [29] Gupta, S.K., K. Kumar, C.S. Seelamantula, and C. Singh Thakur, "A Portable Ultrasound Imaging System Utilizing Deep Generative Learning-Based Compressive Sensing On Pre-Beamformed RF Signals", in 2019 41st

- Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany. 2019.
- [30] Hinton, G.E., S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets", *Neural Computation*, 18(7), 2006, pp. 1527–1554.
- [31] Hofmann, P., J. Jöhnk, D. Protschky, and N. Urbach, "Developing Purposeful AI Use Cases – A Structured Method and Its Application in Project Management", in 15th International Conference on Wirtschaftsinformatik (WI), Potsdam, Germany. 2020.
- [32] Hofmann, P., S. Oesterle, P. Rust, and N. Urbach, "Machine Learning approaches along the Radiology Value Chain: Rethinking Value Propositions", 2019.
- [33] Hu, Z., Z. Yang, R. Salakhutdinov, and E.P. Xing, *On Unifying Deep Generative Models*, 2017.
- [34] Jiang, F., Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, and Y. Wang, "Artificial intelligence in healthcare: past, present and future", *Stroke and vascular neurology*, 2(4), 2017, pp. 230–243.
- [35] Jiang, J. and J.A. Fan, "Global Optimization of Dielectric Metasurfaces Using a Physics-Driven Neural Network", *Nano Letters*, 19(8), 2019, pp. 5366–5372.
- [36] Jun, E., A.W. Mulyadi, and H. Suk, "Stochastic Imputation and Uncertainty-Aware Attention to EHR for Mortality Prediction", in 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary. 2019.
- [37] Juvela, L., B. Bollepalli, J. Yamagishi, and P. Alku, "Waveform Generation for Text-to-speech Synthesis Using Pitch-synchronous Multi-scale Generative Adversarial Networks", in 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Brighton, UK. 2019.
- [38] Kim, H.G., H.-T. Lim, S. Lee, and Y.M. Ro, "VRSA Net: VR Sickness Assessment Considering Exceptional Motion for 360° VR Video", *IEEE Transactions on Image Processing*, 28(4), 2019, pp. 1646–1660.
- [39] Kingma, D.P. and M. Welling, "Auto-Encoding Variational Bayes", in The 2nd International Conference on Learning Representations (ICLR), Banff, Canada. 2014.
- [40] Ko, T. and H. Kim, "Fault Classification in High-Dimensional Complex Processes Using Semi-Supervised Deep Convolutional Generative Models", *IEEE Transactions on Industrial Informatics*, 16(4), 2020, pp. 2868–2877.
- [41] Kos, J., I. Fischer, and D. Song, "Adversarial Examples for Generative Models", in 2018 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA. 2018.
- [42] Li, Y., S. Liu, J. Yang, and M.-H. Yang, "Generative Face Completion", in The 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, Hawaii, USA. 2017.
- [43] Lindt, A., P. Barros, H. Siqueira, and S. Wermter, "Facial Expression Editing with Continuous Emotion Labels", in 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), Lille, France. 2019.
- [44] Matsubara, T., R. Tachibana, and K. Uehara, "Anomaly Machine Component Detection by Deep Generative Model with Unregularized Score", in 2018 International Joint Conference on Neural Networks (IJCNN), Rio, Brazil. 2018.
- [45] Meden, B., R.C. Malli, S. Fabijan, H.K. Ekenel, V. Struc, and P. Peer, "Face deidentification with generative deep neural networks", *IET Signal Processing*, 11(9), 2017, pp. 1046–1054.
- [46] Murphy, K.P., *Machine learning: A probabilistic perspective*, MIT Press, Cambridge, Mass., 2012.
- [47] Nandy, J., W. Hsu, and M.L. Lee, "Normal Similarity Network for Generative Modelling", in 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece. 2018.
- [48] Ng, A., M. Jordan, and Y. Weiss, "On Spectral Clustering: Analysis and an algorithm", in *Advances in Neural Information Processing Systems 14 (NIPS 2001)*, T.G. Dietterich, S. Becker, and Z. Ghahramani, Editors. 2001.
- [49] Nugraha, A.A., K. Sekiguchi, and K. Yoshii, "A Flow-Based Deep Latent Variable Model for Speech Spectrogram Modeling and Enhancement", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2020.
- [50] Oussidi, A. and A. Elhassouny, "Deep generative models: Survey", in 2018 International Conference on Intelligent Systems and Computer Vision (ISCV2018), Fez, Morocco. 2018. IEEE: Piscataway, NJ.
- [51] Pan, Z., W. Yu, X. Yi, A. Khan, F. Yuan, and Y. Zheng, "Recent Progress on Generative Adversarial Networks (GANs): A Survey", *IEEE Access*, 7, 2019, pp. 36322–36333.
- [52] Rai, A., P. Constantinides, and S. Sarker, "Next-Generation Digital Platforms: Toward Human–AI Hybrid", *Management Information Systems Quarterly*, 43(1), 2019, pp. iii–ix.
- [53] Ransbotham, S., D. Kiron, P. Gerbert, and M. Reeves, "Reshaping Business With Artificial Intelligence: Closing the Gap Between Ambition and Action", *MIT Sloan Management Review*, 59(1), 2017.
- [54] Rezende, D.J., S. Mohamed, and D. Wierstra, "Stochastic Backpropagation and Approximate Inference in Deep Generative Models", in *International Conference on Machine Learning*, Beijing, China. 2014.
- [55] Saito, Y., S. Takamichi, and H. Saruwatari, "Text-to-Speech Synthesis Using STFT Spectra Based on Low-/Multi-Resolution Generative Adversarial Networks", in

- 2018 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Calgary, Canada. 2018.
- [56] Samanta, B., A. De, G. Jana, P.K. Chattaraj, N. Ganguly, and M.G. Rodriguez, "NEVAE: A Deep Generative Model for Molecular Graphs", in Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19), Honolulu, Hawaii, USA. 2019.
- [57] Schneider, J., "Human-to-AI Coach: Improving Human Inputs to AI Systems", in Advances in Intelligent Data Analysis XVIII, M.R. Berthold, A. Feelders, and G. Krempel, Editors. 2020. Springer International Publishing: Cham.
- [58] Sekiguchi, K., Y. Bando, K. Yoshii, and T. Kawahara, "Bayesian Multichannel Speech Enhancement with a Deep Speech Prior", in Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Honolulu, Hawaii, USA. 2018.
- [59] Seo, M., T. Kitajima, and Y. Chen, "High-Resolution Gaze-Corrected Image Generation based on Combined Conditional GAN and Residual Dense Network", in 2020 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA. 2020.
- [60] Shu, K., S. Wang, T. Le, D. Lee, and H. Liu, "Deep Headline Generation for Clickbait Detection", in The 19th IEEE International Conference on Data Mining (ICDM), Beijing, China. 2019.
- [61] Tanaka, K., H. Kameoka, and K. Morikawa, "Vae-Space: Deep Generative Model of Voice Fundamental Frequency Contours", in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, Canada. 2018.
- [62] Ueda, T., M. Seo, and I. Nishikawa, "Automatic Correction by Deep Generative Model and its Application to Building Construction", in 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Huangshan, China. 2018.
- [63] van den Oord, A., S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, WaveNet: A Generative Model for Raw Audio, 2016.
- [64] VanRullen, R. and L. Reddy, "Reconstructing faces from fMRI patterns using deep generative neural networks", *Communications Biology*, 2, 2019.
- [65] Vom Brocke, J., A. Simons, B. Niehaves, K. Riemer, R. Plattfaut, and A. Cleven, "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process", in 17th European Conference on Information Systems (ECIS 2009), Verona, Italy. 2009.
- [66] Wang, M., X. Liu, and H. Jin, "A generative image fusion approach based on supervised deep convolution network driven by weighted gradient flow", *Image and Vision Computing*, 86, 2019, pp. 1–16.
- [67] Watter, M., J.T. Springenberg, J. Boedecker, and M. Riedmiller, "Embed to Control: A Locally Linear Latent Dynamics Model for Control from Raw Images", in Advances in Neural Information Processing Systems 28 (NIPS 2015), C. Cortes, N.D. Lawrence, D.D. Lee, M. Sugiyama, and R. Garnett, Editors. 2015.
- [68] Webster, J. and R.T. Watson, "Analyzing the Past to Prepare for the Future: Writing a Literature Review", *MIS Quarterly*, 26(2), 2002, pp. 13–23.
- [69] Xu, C., Y. Dai, R. Lin, and S. Wang, "Social image refinement and annotation via weakly-supervised variational auto-encoder", *Knowledge-Based Systems*, 192, 2020, p. 105259.
- [70] Xu, J., H. Li, and S. Zhou, "An Overview of Deep Generative Models", *IETE Technical Review*, 32(2), 2015, pp. 131–139.
- [71] Xu, W., K. Shawn, and G. Wang, "Toward learning a unified many-to-many mapping for diverse image translation", *Pattern Recognition*, 93, 2019, pp. 570–580.
- [72] Yamamoto, R., E. Song, and J.-M. Kim, "Parallel Wavegan: A Fast Waveform Generation Model Based on Generative Adversarial Networks with Multi-Resolution Spectrogram", in ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing, Barcelona, Spain. 2020.
- [73] Yang, Z., L. Liu, and Q. Huang, "Learning Generative Neural Networks for 3D Colorization", in Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA. 2018.
- [74] Yu, J., Z. Lin, J. Yang, X. Shen, X. Lu, and T.S. Huang, "Generative Image Inpainting with Contextual Attention", in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA. 2017.
- [75] Zeng, J., X. Ma, and K. Zhou, "Photo-realistic face age progression/regression using a single generative adversarial network", *Neurocomputing*, 366, 2019, pp. 295–304.
- [76] Zhang, C. and Y. Peng, "Stacking VAE and GAN for Context-aware Text-to-Image Generation", in 2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM), Xi'an, China. 2018.
- [77] Zhang, J., R. Zhan, Di Sun, and G. Pan, "Symmetry-Aware Face Completion with Generative Adversarial Networks", in *Computer Vision – ACCV 2018, C.V. Jawahar, H. Li, G. Mori, and K. Schindler, Editors. 2019.*
- [78] Zhou, H., M. Huang, Y. Mao, C. Zhu, P. Shu, and X. Zhu, "Domain-Constrained Advertising Keyword Generation", in The Web Conference 2019, San Francisco, CA, USA. 2019.
- [79] Zhou, K., G. Endrodi, L.-G. Pang, and H. Stoecker, "Regressive and generative neural networks for scalar field theory", *Physical Review D*, 100(1), 2019.