ChatGPT, BARD,
and Other Large Language Models Meet Regulated Pharma

By Frederick Blumenthal, Martin Heitmann, Stefan Münch, and Brandi Stockton

ChatGPT and other large language models are positioned to change the world. They can also shift acceptance and prevalence of machine learning solutions in regulated industries in general. However, their arrival requires reconsiderations on risks, quality assurance, and validation from a GxP perspective.

ChatGPT [1] has taken the use of artificial intelligence (AI) to the next level, reaching a global user base beyond the established AI circles, and other large language models (LLMs) like LaMDA as in Bard [2] followed. But can we imagine applying a tool like ChatGPT in a safety-critical, product-related environment? To this end, we need to understand what ChatGPT is: As an LLM within the space of generative AI, it can create content, and therefore contrasts other objectives of machine learning (ML) to generate point estimate predictions or suggestions on categories or classifications. That said, is this technology also helpful in regulated areas of pharma? We will examine this question from quality, risk, and efficiency perspectives and discuss how such tools can shift the way we think about validation and quality assurance.

> ChatGPT is commonly used to refer to the chatbot web interface, and the underlying model. For the sake of readability, we will use it interchangeably for both interpretations throughout the article.

In sections where we applied ChatGPT for example cases, we indicate this as such:

- The question or remark from us is labeled as **Authors**.
- The response by ChatGPT is marked as **ChatGPT**.
- All respective content is enclosed in quotation marks.

For these ChatGPT examples, we use the web interface of ChatGPT: https://chat.openai.com/chat (recorded 9 February 2023).

Authors: "Is ChatGPT also useful for the use in regulated areas of pharma?"

ChatGPT: "Yes, ChatGPT can be useful in regulated areas of the pharmaceutical industry, such as providing information on drug products, their uses, dosages, side effects, and interactions with other medications. However, it's important to note that ChatGPT is a general language model, and its output should not be considered medical advice. In regulated areas, information should always be validated and verified by a qualified healthcare professional before being used in patient care or treatment decisions. Additionally, the use of AI in regulated areas is subject to specific laws, regulations, and guidelines, and it's important to ensure that any use complies with these requirements."

We note the following about this response. First, ChatGPT does not claim that it is useful: it just states that it *can* be useful. We share this view and will elaborate on that with respect to governance and quality assurance.

Second, ChatGPT provides a very broad, high-level overview of possible activities and tasks for which it may be used in the pharmaceutical industry. However, it does not provide further information on how it could be applied. Also, we think that ChatGPT can offer further use cases that are of equal importance in our view.

Third, we note that ChatGPT, as it mentions in its own response, is a general language model and accordingly uses terms such as "qualified healthcare professionals" and "medical advice." Specific aspects of pharmaceutical roles (e.g., quality assurance) are not mentioned here. We do acknowledge that ChatGPT points out that its answers should be verified by qualified healthcare personnel.

In summary, ChatGPT offers an interesting starting point to think about the use of such systems or other, tailored, generative AI systems in regulated areas. But, as we see in this first example, this can only be the start of critical thinking and careful evaluation of information.

We structure the rest of our article as follows: We provide an overview of how LLMs operate and how ChatGPT was constructed, describe use cases of ChatGPT as an example of generative AI that is optimized for chatting, and present use cases

ranging from the original idea of chatting and text creation to software development. Next, we discuss the risks related to the use of ChatGPT and LLMs in general. We transfer the capabilities of ChatGPT to specific use in context to the pharmaceutical value chain, and elaborate on the context of use in the regulated areas of pharma, considering ChatGPT and LLMs in general. Our article concludes with an outlook on how the use of LLMs may evolve, beyond the current hype.

OVERVIEW OF LLMs AND TRAINING OF LLMs

Simply put, a language model is a probability distribution over sequences of words, i.e., when given a text sequence, the model can predict what likely comes next. The term *large language model* is not clearly defined but usually refers to language models that are based on very large neural networks. By training to predict the next word, LLMs learn the underlying patterns and structure of the language as well as facts about the word. These models can subsequently be applied to a broad range of language processing and language understanding tasks.

Although the first neural language model was proposed over 20 years ago [3], there are some key innovations that led to the impressive capabilities of modern LLMs, in particular:

- Hardware advancements: The availability of specialized chips like graphics processing units (GPUs) and tensor processing unit (TPUs), which allow for faster processing of large data
- Software advancements: New neural network architectures, like transformer models that are more efficient for processing sequences of text, and more advanced numerical optimization techniques
- Larger and more diverse training datasets: Modern LLMs are trained on much larger and more diverse datasets than older models

For example, GPT-3 [4], considered one of the most powerful language models and what ChatGPT is based on, is a large transformer model with 175 billion parameters that was trained on about 45 terabytes (hundreds of billions of words) of multilingual text data (crawled websites, books, and Wikipedia).

The training of such language models is often referred to as pre-training because the pre-trained language model was usually tailored to the desired task on a supervised (labeled) dataset to perform a specific task: for example, sentiment classification or extraction of certain entities. The resource-expensive pre-training (which only has to be done once) gives the model the general linguistic capabilities, whereas the tailoring adapts the model parameters to the specific task. This approach is still reflected in the name GPT, which stands for Generative Pre-trained Transformer.

LLMs such as GPT-3 can be applied to natural language processing (NLP) tasks without further tailoring by only providing the input text to model together with an appropriately formulated prompt (i.e., a natural language instruction describing the task). This so-called zero-shot approach has the strong advantage that no data must be labeled for the tailoring step and that the same model can be used for many different NLP applications. However, this approach usually does not reach the model performance of tailored models. Putting a few input-output examples in front of the actual prompt often leads to significantly better results. This approach is called few-shot learning.

Thus, LLMs such as GPT-3 can be used to perform NLP tasks. However, these models, which are merely trained to predict the statistically most probable next word, are often not very good at following instructions and sometimes generate untruthful and toxic outputs. ChatGPT is a tailored derivative from GPT-3 to align the LLM better with the users' intentions, i.e., to generate responses that are more helpful and safer and to interact in a conversational way [5].

To this end, the model is provided with examples of text inputs and respective outputs written by human labelers. In addition, a relatively new approach around LLM called reinforcement learning with human feedback (RLHF) has been used to train the model, which optimizes the model based on outputs that have been ranked by human labelers. Ranking is a much easier task and much more efficient than writing outputs. As an extra layer of protection against undesired outputs, ChatGPT uses an algorithm to classify and filter out harmful content.

During the last few years, various LLMs have been developed by big tech companies, AI startups, and research initiatives. The most popular model is arguably the GPT-3 model family from OpenAI, a closed model that is commercially available via application programming interface (API) for inference and fine-tuning. The popular PaLM model by Google reportedly has very strong capabilities, but this closed model is not publicly available. Other popular models are OPT by Meta (an open model with use-based restrictions, e.g., noncommercial use only) and BLOOM by the research initiative BigScience (an open model with small use-based restrictions). It can be expected that many new foundational LLMs as well as specialized models (such as dialogue systems) will be developed soon.

Given this overview on how ChatGPT is constructed and trained, we will elaborate on specific quality and control implications for application in regulated areas in our article.

USING ChatGPT

Through our use of ChatGPT as one representative of LLM applications focused on chat tasks, we identified the following use cases:

- Brainstorming assistant: ChatGPT can provide further points for consideration and refine a suggestion based on an idea.
- Writing assistant: Provided an idea of content to be produced, ChatGPT can create a complete paragraph, adding details and further information. It can also tune the text to a specific audience or style.
- Fine-tuning assistant: Many language contributions in the pharmaceutical industry such as articles, evaluations, and reports are inherently complex due to the business content. However, complex grammar presented in these publications is often unnecessary. ChatGPT can be used to simplify the language, either as author or in the position of the audience.

- Knowledge base search: A topic can be explored and described in various tones from simple understanding to deeper and expert information. The chat nature of ChatGPT allows us to refine the insights gained and balance critical questions and exploration, as shown in the introductory example.
- Software engineering: ChatGPT can provide full implementations, including syntax, given a description of the intended use. It also can provide automated tests to verify the implementation.
- Proofing and robustification of software: ChatGPT can detect errors in implementation and provide hints on how to resolve them in the code.

Case 1: Creating Paragraphs from a Set of Keywords

Authors: "Please provide a section about data integrity in computerized systems from a pharma quality assurance perspective in 100 words."

ChatGPT: "Data integrity in computerized systems is critical for ensuring the quality and safety of pharmaceutical products. From a pharma quality assurance perspective, data integrity refers to the accuracy, consistency, and reliability of electronic data used in pharmaceutical operations. This includes data generated from laboratory equipment, manufacturing processes, and business systems. Quality assurance must ensure that data is protected from unauthorized changes and accurately reflects the intended state. This can be achieved through implementing effective controls, such as secure systems, data backup procedures, and access controls. Additionally, regular audits and inspections should be conducted to verify the accuracy of data and identify any discrepancies. Ensuring data integrity is essential to meet regulatory requirements and maintain the trust of patients and stakeholders."

Authors: "Please provide a section about data integrity in computerized systems from a software development perspective in 100 words."

ChatGPT: "In a pharmaceutical software engineering context, data integrity refers to the accuracy, consistency, and reliability of data stored in computerized systems. This is a crucial aspect of pharmaceutical software as it directly impacts patient safety. To ensure data integrity, pharmaceutical companies must implement strict validation processes, perform regular backups, and have in place robust security measures to prevent unauthorized access and data manipulation. Additionally, software systems must comply with regulatory standards such as 21 CFR Part 11, which lays down guidelines for the use of electronic records and signatures in the pharmaceutical industry. Maintaining data integrity requires a combination of technical, organizational, and procedural measures to ensure that data is protected at all stages of its lifecycle."

We evaluate these responses by ChatGPT as follows. Both responses provide an adequate description of data integrity and

its importance, but the responses are incomplete, as they do not cover all aspects of the ALCOA (attributable, legible, contemporaneous, original, and accurate) principles and lack some well-known extensions, such as traceability. None of the responses point to the relation of data integrity, product quality, and patient safety. However, they emphasize the importance of data protection and access control.

As expected, both responses are similar. The small, but subtle, difference in the request is well-reflected in the responses and the differences are obvious at first glance. The second response refers to some aspects of system design. However, crucial elements like a software development life cycle approach to ensure data integrity are not mentioned. Still, the response concentrates on business matters, which may be caused by "remembering" the context of the first request. Thus, it is important to consider that responses vary depending on a chat's history. Surprisingly, the number of words is greater than 100 in both cases (~120). The clearly formulated nonfunctional requirement is therefore not fulfilled.

Case 2: Refining Natural Language

Authors: "Please simplify the language in the following section of text" (example text is taken from [6]): Once the intention of use, the addresses and the acceptable risks have been identified, the solution should be designed in a way that it is able to fulfill its specific quality expectations. The impact of the steps before on the modeling decision cannot be underestimated: When users are confronted with the solution's results in every single case to build their decision on the AI output, a different modeling strategy may be pursued in comparison to a situation where the solution operates in a mode when only exceptions are handled by operators. However, it is important to critically reflect the decisions made so far, as more understanding of the use case will be gained once the first models are evaluated. Prototyping strategies for early feedback and alignment reduce the acceptance risks further down the development process stream. In this constructive process, documentation regarding the decisions made are important to justify the model and explainability mechanism selection in an audit context."

ChatGPT: "The solution should be made to meet its desired quality standards. The way it's made depends on how it will be used and who will use it. The impact of earlier choices can't be ignored. The solution's results will be different if users always use it to make decisions compared to if it only helps operators with exceptions. It's important to keep evaluating the choices made. Testing the solution early helps reduce risks. Keeping track of the decisions made helps explain why certain choices were made in the future."

We evaluate these responses by ChatGPT as follows. The text is considerably simplified in language. However, the inner logic of the section is not reflected anymore; for instance, "the decisions" are not adequately integrated into the text. ChatGPT does not identify AI as part of the key message. Thus, part of the context is

missing and leads to misunderstanding of the new text. In addition, it seems as if the role of "explain" is turned around; we interpret ChatGPT's response more in the sense that the process outcomes should explain the decisions made, whereas the original piece elaborated on the explainability of algorithms. However, we acknowledge that we see this paragraph from the authors' view and it might be difficult for a human to grasp the meaning of this isolated section, taken from a technical article handling a complex topic.

Case 3: Creating Software or Base Software Kits

Authors: "Please provide Java code that is able to run a linear regression on a tabular data set. Can you also provide a test case for your class?"

In response, we received the requested Java code with a short explanation text at the end, not shown here for the purpose of readability, while copies were kept.

We evaluate it as follows. We verified that the code compiles and runs in Java 17. The code is generally readable: common variable names for linear regression are used. The code style conforms to best practices, e.g., using the constructor for the actual estimation algorithm and providing a separate prediction method. The test case is a reasonable example of application of the code, but not tested in the software engineering sense, i.e., verifying the functionality is done by comparing the expected prediction with the actual prediction. In that, the methodological quality of the linear regression is assessed, but not the correctness of the implementation.

RISKS INHERENT TO THE USE OF LLMs

Although ChatGPT has its capabilities in various areas, as shown previously, it does not come without risks, which we need to consider in all use contexts, but even more so from a risk control perspective in the highly regulated pharmaceutical industry. We identified 10 key risks, which also apply to more general use of LLMs.

Quality and Correctness

ChatGPT can provide content quickly. However, as we already saw in the examples, the content needs to be reviewed and verified by a subject matter expert. The LLM may invent content ("hallucinate"), so it is the user's responsibility to decide which parts are of value for the intended use.

References and Verification

ChatGPT does not offer references per se to verify the information. The chat functionality can be used to elicit references on this topic; however, these "references" may be poor quality or may be fabricated. The LLM may invent titles and authors because it learned typical patterns in references, but it does not know how to verify the integrity of references. This raises concerns about the general traceability and trustworthiness of the generated responses and strengthens our assertion that each response must be carefully verified by a subject matter expert. However, our experience shows

that using the chat feature to request an explanation of the results sometimes reveals interesting background information.

Reproducibility

Responses by ChatGPT may change; a "regenerate response" option is deliberately provided. Hence, traceability of results and content is limited, and sensitivity to input is quite high, as was shown in previous examples. Retraining and updates of the hyperparameter configuration used for training may include other sources of variation in the responses, as in other applications of ML models, i.e., the model evolves over time.

Up-to-Date Information

When operating in a frozen mode regarding training data, language models only capture information to a specific time horizon (at the time of writing, training data for ChatGPT covered until about the fourth quarter of 2021 [7]). Information beyond this horizon cannot be known (or only to a limited amount) in a particular version of the model, which might lead to responses that do not reflect the current state of information.

Intellectual Property and Copyright

In a sense, the whole training data universe can be considered the source of each answer. Therefore, the exact information—or code—might already be created by a third party. Hence, double-checking the intellectual property of the results is important to mitigate legal risks. This also affects the input provided to ChatGPT, which may also fall under intellectual property and copyright considerations.

Bias

Bias may result from various steps in the training steps of ChatGPT: the pre-training on the GPT-3 model, fine-tuning, and moderation. The pre-training was trained on unsupervised data to predict the next word; therefore, the model learns biases from this data set. During the fine-tuning process, operators provided "golden responses" to chat input. Even if done to the standards of OpenAI principles, bias in this human opinion cannot be excluded. Eventually, ChatGPT is moderated to prevent harmful content, which is again trained on labeled data and may carry biases. In summary, the combined effects in this three-step process are not verifiable from the user's perspective. Therefore, users risk distorting their own content if they leverage ChatGPT's results without critical thinking.

Formation of Opinions

Regarding the previous point, particularly in evaluation exercises such as assessments of incidents, ChatGPT may form "an opinion" and hence influence a decision-making process. However, we like to note human decision-making is not bias-free either; thus, ChatGPT may contribute a second perspective on the question at hand, or even multiple perspectives if asked for different angles on a topic.

Figure 1: Application of artificial intelligence across the value chain.
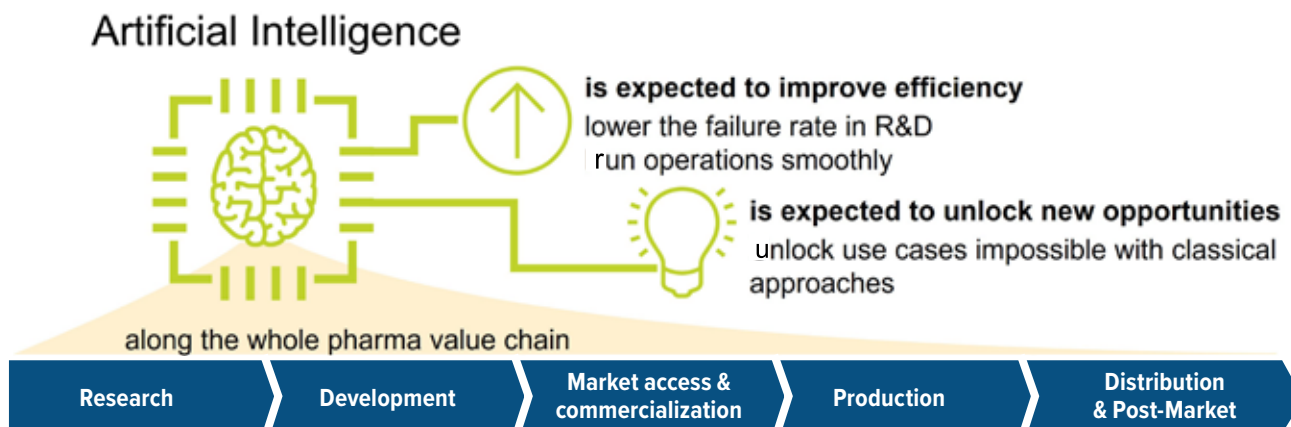
Service Availability
At the time of this writing, the free version of ChatGPT is not currently stable due to the heavy user load. However, ChatGPT also offers a pro version and API access for use in a professional environment.

Data Confidentiality
Every input entered into the ChatGPT interface is beyond the control of the users; hence, no internal or confidential information should be transmitted to mitigate legal risks. This limits the use of ChatGPT, especially in the areas of pharma and healthcare, where many use cases involve personalized identifiable information or business-critical information.

Regulation and Governance
It should be verified if the use of generative AI tools is permitted for the use case by regulatory or internal governance reasons. Due to these risks, some companies already block ChatGPT and similar systems in their network. However, is this the solution? In the next section, we will elaborate on specific use cases, with ChatGPT as a representative of what LLMs can support in the pharmaceutical industry.

USE CASES OF ChatGPT IN REGULATED AREAS OF PHARMA
Given their general areas of application and risk, where could LLMs like ChatGPT provide value to the pharmaceutical industry, and eventually to patients?

Along the pharmaceutical value chain, we think there are five use cases, discussed next and shown in Figure 1 and Table 1:
1. Research: ChatGPT may be used as a brainstorming assistant to generate ideas on a particular target mechanism.
2. Development: ChatGPT may be used to assist in the writing of study reports, provided case descriptions are given.
3. Market access and commercialization: ChatGPT may be used to

simplify the language of pharmaceutical descriptions to make it more accessible to patients.
4. Production: ChatGPT may be applied to create pieces of software required to control the production process.
5. Distribution and post-market: ChatGPT may be applied to summarize the insights from nonclassical pharmacovigilance sources such as social media.

In Table 1, we present our evaluation of the usage of LLM like ChatGPT in these situations. Based on the use cases, we do not think the use of LLM replaces process understanding and subject matter expertise. Even more, the application of these technologies requires a strong understanding of the use case and the inner mechanics to verify the results provided by the LLM, as well as a sensitivity for further risks arising from the black-box nature of these models.

Therefore, we see the most promising applications in the form of an auxiliary tool. Critical thinking is required to assess the output, regardless of whether this is technical content (code, automated tests) or subject matter content (e.g., created text, reports, or sections). To facilitate critical thinking and maintain quality, standard operating procedures need to define how to use and refine output, as well as to define areas where aid from LLM is ethical and where it is not. One example in which we consider the risk of bias to be critical is around evaluation or assessment, which may affect and shift the views of subject matter experts in the first place.

WIDER APPLICATION OF LLMs IN REGULATED PHARMA
Do we see the potential of ChatGPT and other language tools for critical tasks as more than an auxiliary tool? In our opinion, it will be difficult to apply an open, public model in a GxP-regulated context for the following reasons:
▪ The corporation has no direct control of and traceability over the model development process.

Table 1: Assessment of selected LLM applications along the pharma value chain.

Case	Potential Benefits	Risks	Assessment
1: Brainstorming assistant for research	Efficiency boost through leveraging already known information Gain new ideas on mechanisms	Information may be difficult to verify Lack of references Depending on the service model chosen, confidential information may be disclosed	Limited use, more in exploratory phases
2: Writing assistant during development	Efficiency boost in creation of study reports Possible harmonization of style and language in study reports	Misleading summarization of outcomes Depending on the service model chosen, confidential information may be disclosed	Limited use, only in exploratory phases
3: Fine-tuning assistant for market access and commercialization	Texts with high readability produced Texts specific to the audience	Possible loss of crucial information and relations	Requires careful review, but may improve on quality of communication
4: Software engineering assistant for pharma production purposes	Efficiency boost with syntactic and semantically correct blueprints or full code Developers can focus more on functionally valuable tasks Developers may let the LLM check the code	Incomplete business logic Errors in implementation Code that does not comply with usual software design patterns Copyright violation	High use, both from the developers' perspective and from a quality perspective; confidentiality should be maintained
5: Summarizing assistant for distribution and post-market	Additional input and ideation of structures and patterns	May involve the use of highly sensitive, patient-specific information Mass of specific input data for a pharma case may be difficult to process by LLM	In our view, this use case can be better handled by other language models that apply clustering techniques over a larger input set

- The corporation lacks control of the horizon of knowledge, i.e., the last point in time from which relevant training data entered the training set.
- The corporation lacks control on labels and feedback that were used for the training process.

However, ChatGPT demonstrated the power of AI in the field of NLP. This means that in a controlled development environment [8], the development of solutions specifically trained for the pharmaceutical or life sciences domain is promising. If such an LLM was integrated into a process controlled by subject matter experts and provided with evidence of improved human–AI–team performance [9] in a validation exercise, these approaches can add value to the process and ultimately to product quality. In terms of the AI maturity model [10], this would refer to a validation level 3 solution.

From a governance and control perspective, risks need to be monitored along the five quality dimensions [11] in the application life cycle, with a primary focus on retraining and refreshing the body of knowledge:

1. Predictive power: Is the output of sufficient quality and is it answering the question asked by the users?
2. Calibration: Does the language model perform sufficiently well within all relevant dimensions of use cases, or does it exhibit critical biases in some use cases?
3. Robustness: Are outputs sufficiently stable over time and are changes in the output comprehensible?
4. Data quality: Are the data used for training and the data provided as input to the system quality assured to best practices of data governance (e.g., labeling of data, analysis of representativeness of training and production cases)?
5. Use test: Do users understand the link between the problem at hand and the language model's results? Do they adequately react to the language model's output by questioning, verifying, and revisioning the responses?

Even in such controlled situations, we expect that review by a human is required to verify the output in a GxP-critical context. As far as we see, only a subject matter expert can truly evaluate whether a statement in pharma and life sciences in general is true or false.

CONCLUSION

LLMs and ChatGPT are here to stay. It is on us as users and subject matter experts to learn how to use this technology. This becomes even more important as this technology is more readily available in everyday business applications. In combination with critical thinking, results of these LLMs can also be helpful in the regulated environments of the pharmaceutical industry—not as a stand-alone solution, but as a work tool, boosting the efficiency of various operational units from software engineering to regulatory documentation. Using these services via web interface, business application, API, or commercial cloud services must be seen with the precondition of rigor, quality assurance, and mitigation of legal risks while we may need to be creative in finding suitable use cases.

Additionally, more specialized LLMs are likely to gain traction. As in other safety-critical, GxP-governed areas, these models must be developed under controlled quality and best practice conditions.

This rigor involves the quality assurance of input data, control of the development process and validation, and productive monitoring and quality risk management processes. This again can unlock two new powerful dimensions to true human–AI–team collaboration: boosting efficiency by focusing on the respective strengths of AI and cognitive intelligence with a suitable target operating model.

Therefore, it is paramount that corporations react to these developments by setting adequate standards and controls. In this article, we provided a general overview of potential use cases and risks inherent to the use of such LLMs. Building on this guidance, the application of such models always must be evaluated under the specific intention of use, either in the role of a working aid for subject matter experts, developers, and further staff, or as part of a computerized system itself. 🌀

Acknowledgments

References

1. ChatGPT: Chat Generative Pre-trained Transformer, a chatbot powered by GPT-3, developed by OpenAI. https://chat.openai.com/

2. Pichai, S. "An Important Next Step on Our AI Journey." The Keyword. Published 6 February 2023. https://blog.google/technology/ai/bard-google-ai-search-updates/

3. Bengio, Y., R. Ducharme, and P. Vincent. "A Neural Probabilistic Language Model." *Journal of Machine Learning Research* 3 (2003): 1137–55. https://papers.nips.cc/paper/2000/file/728f206c2a01bf572b5940d7d9a8fa4c-Paper.pdf

4. Brown, T., B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, et al. "Language Models Are Few-Shot Learners." *Advances in Neural Information Processing Systems* 33 (2020): 1877–901. https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

5. Ouyang, L., J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, et al. "Training Language Models to Follow Instructions with Human Feedback." *arXiv preprint* arXiv:2203.02155. 2022. https://arxiv.org/pdf/2203.02155.pdf

6. Altrabsheh, E., M. Heitmann, P. Steinmüller, and B. Vinco. "The Road to Explainable AI in GXP-Regulated Areas." *Pharmaceutical Engineering* 43, no. 1 (January/February 2023). https://ispe.org/pharmaceutical-engineering/january-february-2023/road-explainable-ai-gxp-regulated-areas

7. OpenAI. "ChatGPT General FAQ." Accessed 11 February 2023. https://help.openai.com/en/articles/6783457-chatgpt-general-faq

8. International Society for Pharmaceutical Engineering. *GAMP 5: A Risk-Based Approach to Compliant GxP Computerized Systems,* Second Edition. North Bethesda, MD: International Society for Pharmaceutical Engineering, 2022.

9. US Food and Drug Administration. "Artificial Intelligence and Machine Learning in Software as a Medical Device." April 2019. www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

10. Erdmann, N., R. Blumenthal, I. Baumann, and M. Kaufmann. "AI Maturity Model for GxP Application: A Foundation for AI Validation." *Pharmaceutical Engineering* 42, no. 5 (March/April 2022). https://ispe.org/pharmaceutical-engineering/march-april-2022/ai-maturity-model-gxp-application-foundation-ai

11. Altrabsheh, E., M. Heitmann, and A. Lochbronner. "AI Governance and Quality Assurance Framework: Process Design." *Pharmaceutical Engineering* 42, no. 4 (July/August 2022). https://ispe.org/pharmaceutical-engineering/july-august-2022/ai-governance-and-qa-framework-ai-governance-process

Corporate Partners

About the authors

Frederick Blumenthal is a Manager in the Applied AI group at d-fine, where he leads the development of innovative AI-based solutions for complex problems. He has MSc degrees in physics from Heidelberg University and financial mathematics from the University of Oxford, where he specialized in machine learning and scientific computing. Adopting a hands-on approach, Frederick oversees projects from inception to completion. His primary focus is on automating business processes using information retrieval and natural language processing techniques, particularly large language models. Additionally, he specializes in utilizing machine learning methods to develop powerful, yet inherently interpretable, forecasting, credit scoring, and anomaly detection models.

Martin Heitmann is a Manager at d-fine, where he started his career. His first projects were in the financial industry, where he designed large data pools to apply predictive analytics for rating assessments. As d-fine expanded into other industries, Martin helped found the pharma and healthcare unit. In his current role, he is responsible for transferring process design, data science, and large-scale implementation expertise to ambitious digital transformation projects. In the collaborative life sciences community, Martin is an active member in various knowledge-sharing platforms. He leads the GAMP® D/A/CH Special Interest Group on GxP-Compliant Use of AI. He has an MSc in business mathematics from the University of Mannheim. He joined ISPE in 2022.

Stefan Münch is Vice President of Validation and Qualification, responsible for the validation and qualification services of Körber Pharma Consulting. He studied computer science and has more than 25 years of experience in software development (manufacturing execution system [MES]) and as a consultant for the pharmaceutical industry, supporting organizations in their agile and digital transformation. In addition, Stefan is a speaker, moderator, trainer, and author for many topics, including computer system validation (CSV), computer software assurance (CSA), Pharma 4.0™, digitalization, risk management, AI/ML, software development, cloud computing, and data integrity. He is a member of Parenteral Drug Association (PDA) and the GAMP® D/A/CH Steering Committees. He has been a member of ISPE since 2007.

Brandi Stockton is co-founder of the Triality Group, LLC, providing quality, regulatory, and compliance consulting services for the life sciences industry. She has over 20 years of of GxP experience, with specialization in computer systems quality and data integrity. Brandi is a member of the GAMP® Global Steering Committee, serves as Secretary of the GAMP® Americas Steering Committee, and is Co-Chair of the GAMP® Global Software Automation and Artificial Intelligence Special Interest Group. She joined ISPE in 2014.