DOMINO

WHITEPAPER

Responsible Generative AI

Reap the Rewards, Manage the Risk

CONTENTS

Responsible Generative AI: Reap the Rewards, Manage the Risk

Trillions of dollars. According to McKinsey,[1] that's the potential economic impact of generative AI (GenAI). That's just for the 63 use cases they studied.

Analysts disagree about the numbers. But everyone agrees: GenAI will affect every industry. Companies that embrace it stand to gain market share; those who don't – well, you know what happens to laggards.

You've heard all about the rewards of GenAI. You've also heard about the risks. There are documented cases of bias:

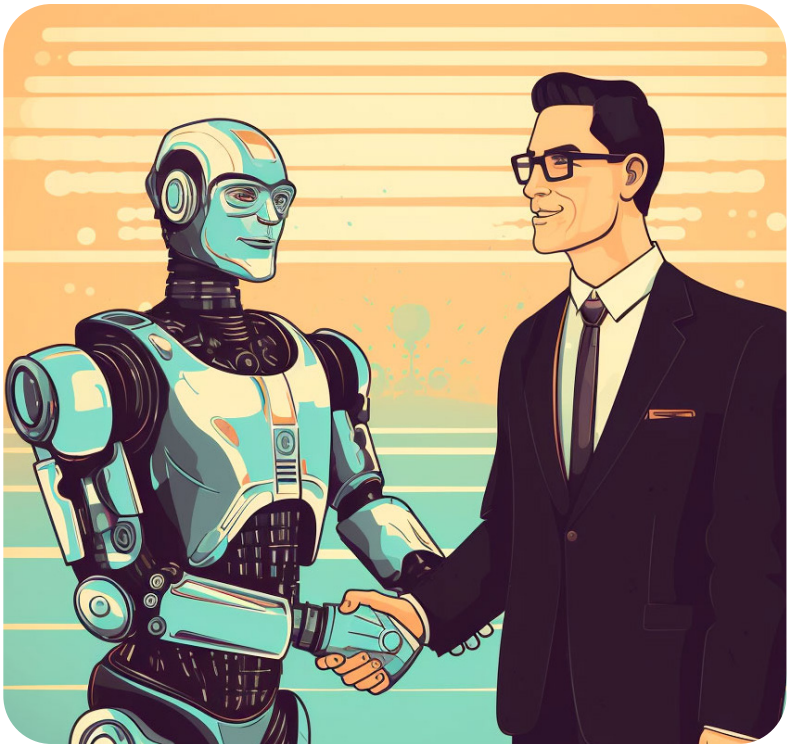
- A study of an AI-driven image generation service found several kinds of systematic bias, including ageism and sexism.[2]
- An Asian-American graduate student wanted GenAI to make her LinkedIn photo more "professional." The application made her look Caucasian.[3]
- The algorithm in an "AI beauty contest" systematically ranked light-skinned contestants higher than those with dark skin.[4]

And you may have heard about some notable GenAI FAILs:

- OpenAI's ChatGPT made up facts about The New York Times.[5]
- In its first demonstration, Google's AI chatbot Bard embarrassed[6,7] the company with factual errors.[8] Microsoft's Bing demo also included mistakes.[9]
- ChatGPT invented six case citations. The lawyer who included them in his case filings is now in trouble with the court.[10] A librarian at Duke University cautioned[11] students and faculty: do not trust ChatGPT for citations; it will simply make them up.

Not surprisingly, there are lawsuits.[12] Axios describes[13] GenAI as "a legal minefield."

Growing pains? Remember: the Wright Brothers crashed and wrecked the Wright Flyer on their third flight. Every technology has teething problems. Bugs are good. Bugs mean people are engaged.



Keep reading. We'll discuss the benefits and risks of GenAI in more detail and then help you define a risk management strategy.

The Power of Generative AI

You've heard about the immense potential of generative AI (GenAI). Let's look at what leading companies do with GenAI *now*.

- **JP Morgan & Co.** accelerates[14] research and model development with GenAI. Security and privacy barriers prevent the bank's data scientists from using some internal data sources. JP Morgan AI Research uses GenAI techniques to create synthetic data sets representing inaccessible data. The synthetic data supports numerous applications, including anti-money laundering, customer journey analysis, and fraud detection.
- **Airbnb** actively uses[15] GenAI. A 'Real-Time Agent Assistant' provides customer service agents with real-time guidance that helps them improve user issue resolution. Another model rephrases user problems for AirBnB's chatbot, ensuring the bot understands customer issues correctly.
- **Adidas** creates[16] novel designs for its footwear using generative AI. The company's research team created AI Archive, an image-generating model based on diffusion

models. The team trained the model on the entire archive of Adidas sneakers dating back to the 1950s. The tool generates new and innovative sneaker designs that draw inspiration from the library and pay homage to the rich history of the Adidas brand.

- **Eli Lilly** leverages[17] generative AI in drug discovery and robotics drug compound delivery. Researchers feed project-specific R&D data into generative AI models through design, making, testing, and analysis cycles. This process continues until researchers identify a drug candidate with optimal properties. With autonomous robotic workstations, the company performs precise and energy-efficient parallel chemical synthesis and assays around the clock.

- **Profluent**, a life sciences startup, creates[18] tools that generate protein sequences. Profluent trained the company's model (ProGen) on 280 million protein sequences. Researchers interact with ProGen with English-like prompts. The model produces novel sequences that are not naturally recurring. These provide a path to new therapies and cures.

- **Insilico Medicine**, a drug discovery company, uses[19] GenAI to accelerate research. Using GenAI, Insilco advances drugs to Phase Two clinical trials in a third of the time and a tenth of the cost, saving more than **$300M.** The company's Chemistry42 engine, built on GenAI, designs new potential drug compounds in days. The tool develops drug-like molecular structures from scratch.

- **Skin+Me**, a cosmetics company, uses[20] image analysis of customer-uploaded photos to prescribe skincare treatments. A traditional AI model identifies conditions and matches relevant products. A GenAI model delivers a personalized image demonstrating the Skin+Me product.

GenAI solves problems across many different use cases:

Software Development. A McKinsey study revealed[21] that software developers complete coding tasks twice as fast with GenAI. GenAI tools expedite manual and repetitive work, jump-start first drafts, accelerate updates, and increase the developer's ability to take on new challenges. **Redfin**, a national real estate broker, uses GenAI for software development. The models assist with internal tasks like translation, analysis, and conversion.[22]

GenAI doesn't have the "potential" to deliver value. It delivers now.

Content Generation. GenAI creates engaging marketing materials, personalized product descriptions, and compelling social media posts, streamlining and enhancing brand communication. A study[23] performed by industry analyst **Gartner** found that Marketing leads other sectors in GenAI adoption. **Omnicom** partnered[24] with Microsoft to create Omni, a generative AI assistant to help ad agency employees with tasks across the advertising process, such as compiling audience insights and building media plans.

Data Augmentation. GenAI efficiently creates synthetic data, helping data scientists build better machine learning models when "real" data is unavailable. **Amazon** uses synthetic data to train Alexa; Google's **Waymo** trains self-driving cars; **Roche** uses it for clinical research.[25]

Customer Service. Businesses employ GenAI for interactive chatbots that support, resolve queries, and deliver personalized recommendations. US-based airline JetBlue partnered with ASAPP to automate and augment its chat channel using GenAI. The contact center has been able to save 73,000 hours of agent time in a single quarter.[26]

In the banking industry, **Accenture** estimates[27] that 90% of all working hours will be affected by GenAI. Accenture says that early adopters use GenAI to transform the front office, develop hyper-personalized marketing content, streamline operations, and improve data management.

Bankers say they use GenAI today to summarize customer conversations and analyze customer data. Other use cases in banking include refactoring code, helping developers with code testing, and investigating financial crime.[28]

GenAI accelerates drug discovery in the pharmaceutical sector, according to a BCG survey.[29] It improves clinical trial planning and execution, leading to more precision medicine therapies. BCG also found early-stage use cases for GenAI in developing treatments for rare cancers, developing protective antibodies, and generating protein sequences with specific functions.

Pharmaceutical researchers use[30] a GenAI technique called Generative Adversarial Networks (GANs) for drug development. GANs generate novel chemical structures and optimize their properties. Other pharmaceutical applications include "de novo" drug design and targeted product development.

GenAI doesn't have the "potential" to deliver value. It delivers now.

The Challenges of GenAI

But there's a catch.

GenAI poses many critical challenges for your organization.

Your most pressing problem is talent: people with the skills to work with large language models (LLMs) are scarce and expensive. LLMs are a new technology introduced about five years ago and mainly under the radar until OpenAI launched ChatGPT. Many data scientists need skills in deep neural networks and transformer models. Big Tech companies compete aggressively for those with the required know-how, increasing salaries.

GenAI makes extreme demands on your computing infrastructure. Data scientists use enormous computing clusters of expensive hardware to train foundation models. You can customize or fine-tune a model with a smaller cluster, but it's still more costly than most conventional AI projects. Inference also requires expensive computing resources. Unlike model training or customization, you will continue to incur these costs as long as your application operates.

Most organizations with a mature data science program have a process to support the workflow. A business process helps the organization evaluate rewards and risks for each project, supports consistent working methods, and assures quality deliverables. GenAI is much more complex than "traditional" AI, and it amplifies the challenge in several different ways:

Proliferating models. A single large language model can spawn many special-purpose models for tasks like image recognition or translation. That means the sheer volume of models your team produces will proliferate. If your team fine-tunes or incrementally trains an open-source model, your organization owns the retrained model, so you must maintain many more models.

Tracking model information. There's much information to track for simple models, including data sources, lineage, experiments, and accuracy statistics. Now imagine keeping track of a GenAI model: the models are vastly more complex, with more experiments to track and more information about each trial. Imagine managing all of that information at scale.

Workflow control. It's embarrassing when an inaccurate or biased "traditional" AI model goes live. You do everything possible to prevent that. However, a flawed GenAI model can be a disaster for your brand. Immature AI programs place most of the burden for quality assurance on the individual data scientist. Your organization will need a different approach as your volume and throughput scale.

Testing and validation. Everyone understands the importance of testing. But do you know how GenAI will affect your testing regime? Testing and validation for GenAI is an order of magnitude more complex than for traditional AI. The testing approach and metrics are entirely different and may require more contributors. Due to the high cost of inference for GenAI, operational testing is critical.

Monitoring and observability. We evaluate "traditional" AI models with familiar statistics, such as accuracy, specificity, sensitivity, AUC, and ROC. The metrics for GenAI models are entirely different and more complicated. Data scientists require richer metrics to assess arithmetic, learning, and reasoning model skills. They also need to evaluate bias, hallucinations, and offensive language.

Auditability. If you work in a regulated industry, like banking, insurance, or pharmaceuticals, you already know how regulatory compliance affects data science practice. It's like dealing with the tax authorities – you need well-organized records of everything you do. There are growing concerns about the impact of GenAI and calls for regulation. That means your team must be able to reproduce any project exactly.

The unique risks of GenAI require process controls built directly into your machine learning platform. To profit from GenAI, you can't simply hope for the best. Your platform must have integrated governance, so that no model can wreck your business.

The Domino Solution

You understand the challenges of generative AI. Here are seven things you can do to mitigate those risks.

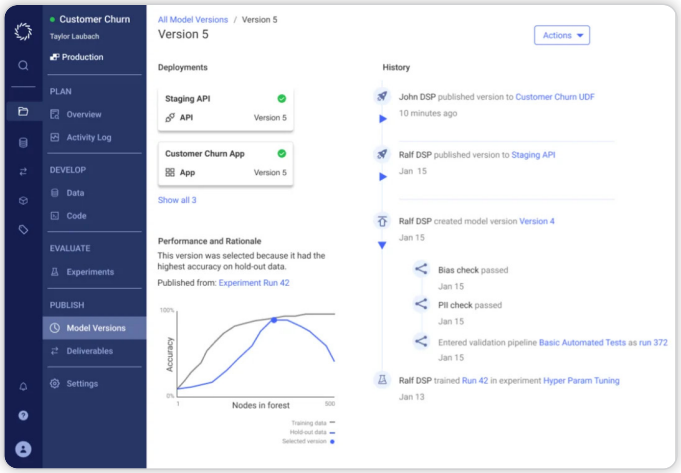
Domino Model Registry

Every model your team builds is a valuable company asset. Your team invests valuable time in research, data prep, training, tuning, and exhaustive evaluation. You work closely with stakeholders to secure their approval for production use.

The stakes are higher for GenAI. GenAI projects are more expensive and complex than "traditional" projects.

With Domino Model Registry, you can view and manage all models in your company, including those your team builds outside of Domino. Powered by MLFlow, Model Registry integrates seamlessly with Domino's Experiment Manager.

Authorized users can search the Model Registry by name and tags. This feature helps you reduce costs by adapting existing models to new applications.

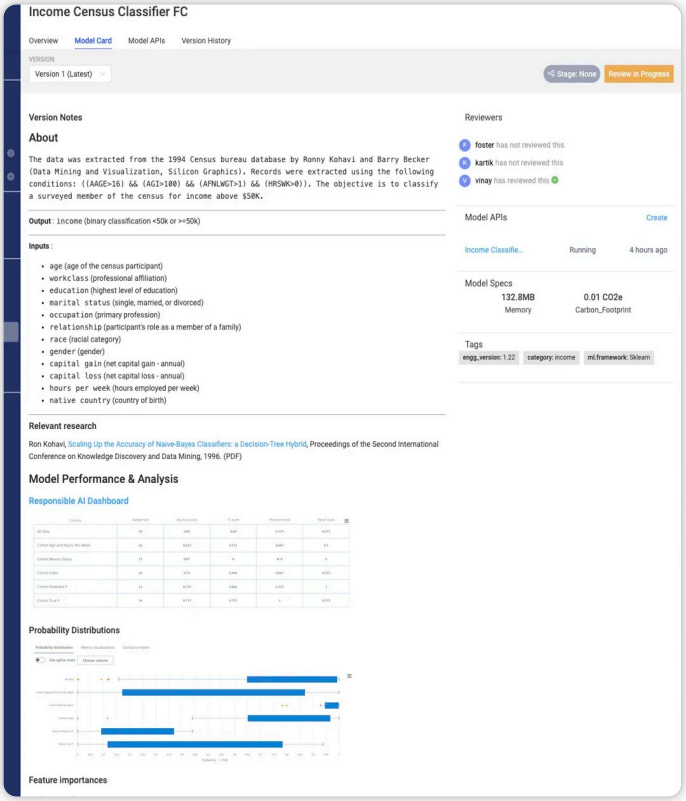


Domino Model Cards

You subject every model to an exhaustive review. That's great. New methods and higher risks for generative AI make this even more critical than ever. However, formal reviews are only helpful if the reviewers have the information they need to make good decisions.

Domino Model Cards capture all the information pertinent to a model. Domino automatically collects critical information, such as data sources, datasets, features, experiments, training environment, code commits, model metadata, accuracy metrics, and deployment endpoints.

Users can append additional information as required. For example, your team can run a model fairness check with one of the many available open-source fairness frameworks[31] and then append results to a model card.



Domino Model Review

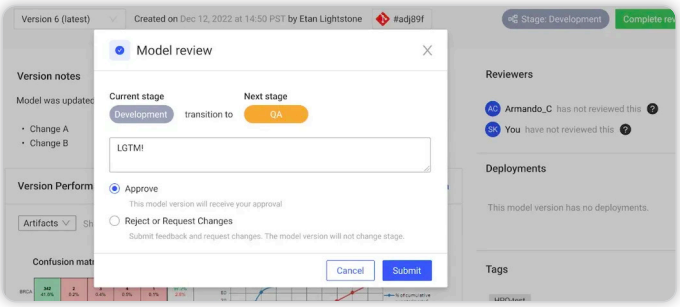
Some vendors say their platform supports "one-click" model deployment. Does that expression send chills up your spine? A flawed model can wreck your business. Do you want anyone to deploy a model without prior review and approval?

You can ask your team to follow review procedures; that approach only works for small teams. Your best bet is a platform with built-in workflow controls so users can only deploy a model with the required approvals.

Generative AI makes this requirement more urgent than ever.

Domino Model Review is a built-in model approval workflow. It routes models to designated stakeholders for domain-specific reviews. For example, data scientists evaluate model quality metrics; IT evaluates model security and cost; Risk evaluates a model for bias; Compliance teams test for sensitive fields.

Domino tracks the model review status and sends notifications. Reviewers evaluate the model and either approve or request changes. When all reviewers approve, Domino updates the model status to permit deployment.



Domino Model Staging

Your team used best practices to develop a generative AI model. Expert data scientists tested the model comprehensively in the development environment.

What happens now? Straight to production?

Not so fast. Before that model goes into production, you must test and validate it independently. Sometimes, you may want to use "red teaming" – an independent team validating the work for high-risk generative AI models.

Your IT organization needs to perform operational testing to define inference infrastructure. Inference workloads differ significantly from training workloads, requiring a different computing setup. This optimization step is imperative for generative AI models due to the high unit cost of GPU-accelerated computing.

Domino Model Staging supports a controlled deployment process. Users deploy models seamlessly to a separate environment for quality, compliance, and performance testing. Upon approval, an authorized user transfers the model to a production environment, isolated for better security.

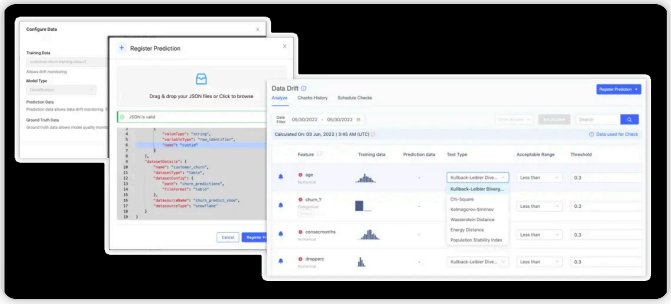
Domino Model Monitoring

Your generative AI model is in production. Time to relax?

Unfortunately, models do not work well forever. Even the best model degrades over time.

Domino Model Monitoring provides a comprehensive view of model traffic, drift, and quality trends for your models, including those your team deploys outside Domino. Domino delivers out-of-the-box metrics for model drift and quality. Automatically alerts when drift, divergence, and data quality checks exceed configurable thresholds.

Data scientists generally agree on model monitoring standards for "traditional" AI models. Not so for generative AI. Due to rapid innovation in the field, metrics for generative models are in flux. Domino Model Monitoring is flexible and extensible, so your team can monitor GenAI models with one of the growing libraries of open-source monitoring frameworks for LLMs.



Domino Reproducibility

The auditors are here. Please show us exactly how you built this model.

Can your team survive an audit? There is growing concern about how AI affects people. Legislators actively discuss imposing tighter regulations. Regulation is familiar in some industries, like banking, insurance, and pharmaceuticals. Generative AI will bring more outside scrutiny to the practice of data science in every industry.

Domino's best-in-class Reproducibility Engine automatically captures every step in the process and versions all assets:

- Data
- Code
- Environments
- Experiments
- Models
- Deployments
- Batch Jobs
- Applications

You can reproduce any project with Domino, even if the tools and users change. Since Domino captures every change and every state, you can roll a project back when necessary. You can trace backward from a document to its development steps or trace forwards from a data source.

Best of all, Domino Reproducibility runs in the background and works automatically. Users do not need to invoke the Reproducibility Engine or upload assets. Consequently, you can rest assured that your team's work is reproducible.

Customer Stories

Artificial intelligence helps researchers discover novel treatments. **Janssen Pharmaceutical**, a division of **Johnson & Johnson**, uses deep learning to analyze whole-slide images of biopsy and surgical specimens (called Histopathology images). Domino helps Janssen accelerate model development and makes research reproducible and reusable so that Janssen can apply results from one cohort of patients to others. With Domino, Janssen reduces the use of ad hoc code, so models are explainable and interpretable.[32]

Evidation, a life sciences company, continuously measures the health of individuals (at their behest and with their consent) using patient-generated health data (PGHD) they share from apps and wearable technologies, such as smartphones, activity trackers, and smartwatches. Evidation uses Domino to develop, deploy, manage, and monitor AI. Domino provides Evidation with reproducible workflows that reduce the time to spin up new projects. Evidation uses Domino Model Monitoring to monitor data drift and model performance.

Topdanmark, a large European insurance company, uses data science to provide consumers with a better, faster insurance experience. The company uses the Domino platform for model-driven workflows. With Domino, Topdanmark can detect data and model drift in real-time, a task that took up to three months previously.

Do This Now

When it comes to GenAI, there's a lot on the line. The potential for new revenue, cost savings, and improved products is a great motivation to adopt the technology. On the other hand, GenAI increases the number of risks and unknowns you need to mitigate to avoid disaster. But how do you get started? Is there a suitable approach to incorporating responsible AI practices into your GenAI projects?

Since no two companies are alike, there is no foolproof method anyone can point you to that will guarantee success. At the same time, we have seen successful patterns emerge.

- **Establish cross-functional project teams.** Such cross-functional teams will shepherd models throughout the model lifecycle. Team members will gain a thorough understanding and ownership of models, their data dependencies, and operational impact and benefits. This reduces risk and improves the validation process before a model scales to production.

- **Identify risk by working backward.** Consider using the inversion technique, having team members create a list of potential model-related risks across their respective domains. Identifying risks simplifies isolating gaps and determining the steps to address and monitor them. These steps can act as the foundation for establishing a concrete model review process and serve as the basis for the organization's responsible AI charter. You can learn more about this approach here.

- **Use technology to make responsible AI real.** Putting plans into action takes time and effort. Responsible practices are virtually impossible to implement by combining meeting minutes, code, and the occasional data snapshots. You need a cohesive solution with a repeatable framework to build on. Even if you are not operating in a regulated industry, the risks are significant enough for you to consider this investment.

This is by no means an exhaustive Responsible AI plan. Nonetheless, it should offer you some food for thought. As we demonstrated, Domino is built from the ground up to provide you with the framework, tools, and technology to adopt GenAI and deliver ROI.

We look forward to helping you plan your GenAI strategy. Please contact us and speak with one of our experts to get started.

Additional Readings

Dignum, V. (2022). Responsible Artificial Intelligence -- from Principles to Practice. https://arxiv.org/abs/2205.10785v1

Östreich, T. (2022). Deep Learning Reproducibility and Explainable AI (XAI). https://arxiv.org/abs/2202.11452

Bickford, J., Gupta, A., Mills S., Roselund, T., (2023). Responsible AI Belongs on the CEO Agenda. https://www.bcg.com/publications/2023/ceo-agenda-must-include-responsible-use-of-ai

Bell, A., Nov, O., & Stoyanovich, J. (2022). Think About the Stakeholders First! Towards an Algorithmic Transparency Playbook for Regulatory Compliance. https://arxiv.org/abs/2207.01482

References

1 https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#key-insights

2 https://theconversation.com/ageism-sexism-classism-and-more-7-examples-of-bias-in-ai-generated-images-208748#:~:text=How%20does%20bias%20happen%3F,return%20certain%20kinds%20of%20outputs

3 https://www.masslive.com/news/2023/07/an-asian-mit-grad-asked-ai-to-make-her-photo-professional-it-made-her-white.html

4 https://www.theguardian.com/technology/2016/sep/08/artificial-intelligence-beauty-contest-doesnt-like-black-people

5 https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucination.html

6 https://www.theverge.com/2023/2/8/23590864/google-ai-chatbot-bard-mistake-error-exoplanet-demo

7 https://time.com/6254226/alphabet-google-bard-100-billion-ai-error/

8 https://www.theguardian.com/technology/2023/feb/09/ai-blunders-google-chatbot-chatgpt-cause-trouble-more-firms

9 https://www.cnn.com/2023/02/14/tech/microsoft-bing-ai-errors/index.html

10 https://arstechnica.com/tech-policy/2023/05/lawyer-cited-6-fake-cases-made-up-by-chatgpt-judge-calls-it-unprecedented/

11 https://blogs.library.duke.edu/blog/2023/03/09/chatgpt-and-fake-citations/

12 https://techcrunch.com/2023/01/27/the-current-legal-cases-against-generative-ai-are-just-the-beginning/

13 https://www.axios.com/2023/02/24/chatgpt-generative-ai-legal-minefield

14 https://www.jpmorgan.com/technology/technology-blog/synthetic-data-for-real-insights

15 https://medium.com/airbnb-engineering/how-ai-text-generation-models-are-reshaping-customer-support-at-airbnb-a851db0b4fa3

16 https://dl.acm.org/doi/10.1145/3587421.3595416

17 https://www.pharmaceutical-technology.com/news/xtalpi-ai-drug-discovery-eli-lilly/

18 https://www.nature.com/articles/s41587-022-01618-2

19 https://blogs.nvidia.com/blog/2023/06/27/insilico-medicine-uses-generative-ai-to-accelerate-drug-discovery/

20 https://econsultancy.com/beauty-industry-generative-ai/

21 https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/unleashing-developer-productivity-with-generative-ai

22 https://www.geekwire.com/2023/ai-and-coding-how-seattle-tech-companies-are-using-generative-ai-for-programming/

23 https://martech.org/marketing-leads-other-sectors-in-generative-ai-acceptance/

24 https://www.wsj.com/articles/omnicom-group-to-debut-tool-using-generative-ai-for-advertising-employees-4976424?mod=Searchresults_pos1&page=1

25 https://www.statice.ai/post/types-synthetic-data-examples-real-life-examples

26 https://www.bcg.com/publications/2023/how-generative-ai-transforms-customer-service

27 https://bankingblog.accenture.com/how-banks-scale-generative-ai-for-growth

28 https://www.cnbc.com/2023/06/13/banks-are-talking-up-ai-amid-chatgpt-buzz-but-keeping-its-use-limited.html

29 https://www.bcg.com/publications/2023/how-generative-ai-is-transforming-health-care-sooner-than-expected

30 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10385763/

31 https://murat-durmus.medium.com/an-overview-of-some-available-fairness-frameworks-packages-ff22fde9d2f4

32 https://domino.ai/customers/janssen


DOMINO