

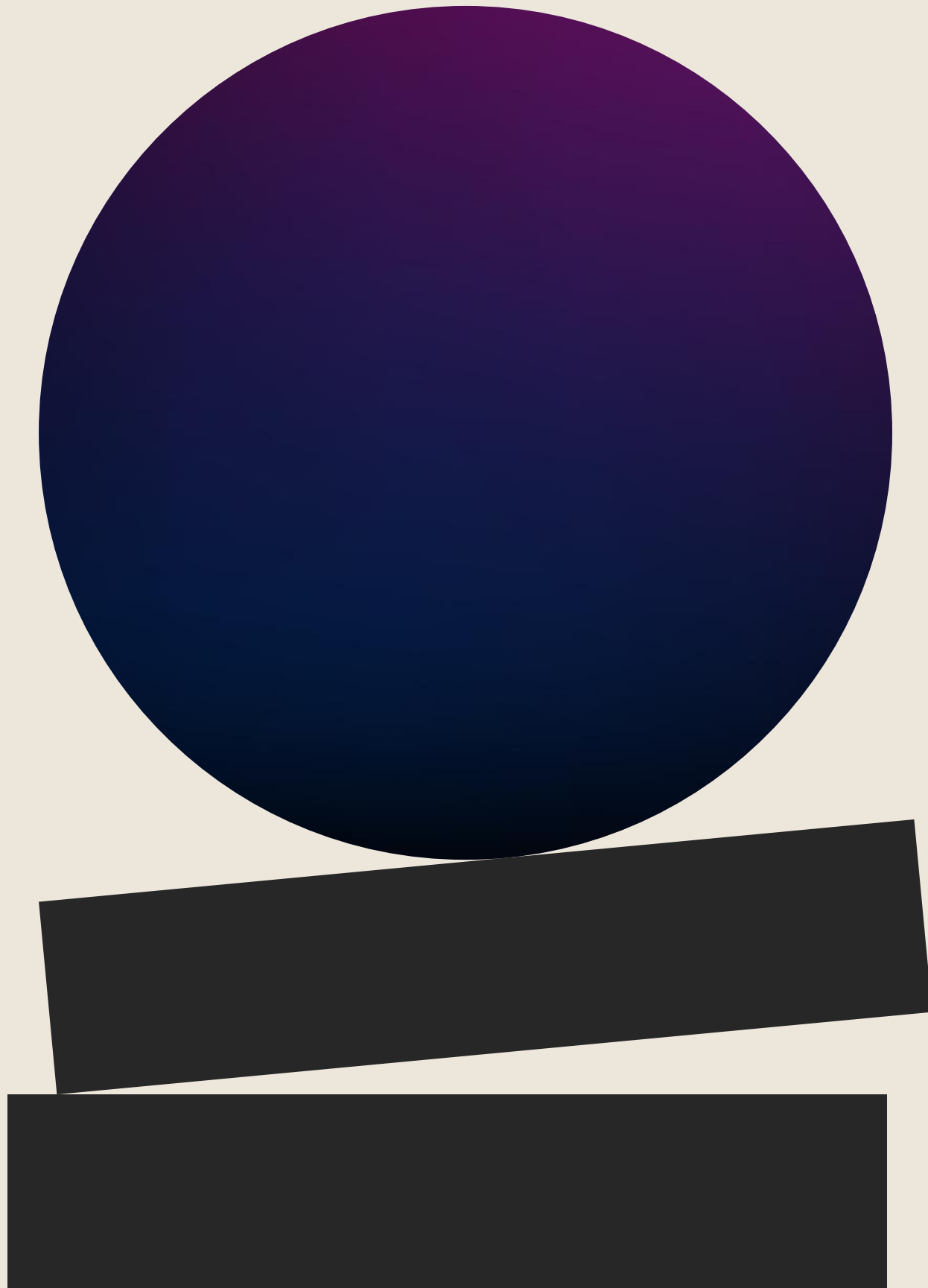
Debiasing Gen AI for Enterprise Applications

AI use cases are here to stay - how do we ensure they are useful, but not harmful?

Rahul Zende

Principal Data Scientist at Navy Federal Credit
Union (NFCU)





Status Quo:

Many say Gen AI is a
mirror - is it really?

AI is capable of distorting reality, no matter
how equitable or skewed it is

We need to understand better, before going mainstream with
use-cases



Identifying Bias

Acknowledging and educating ourselves, is the first step.

Gen AI models are being used and tested en masse - this tech has successfully made early adopters of the global population, on a scale never seen before.



Data Diversity for the rescue

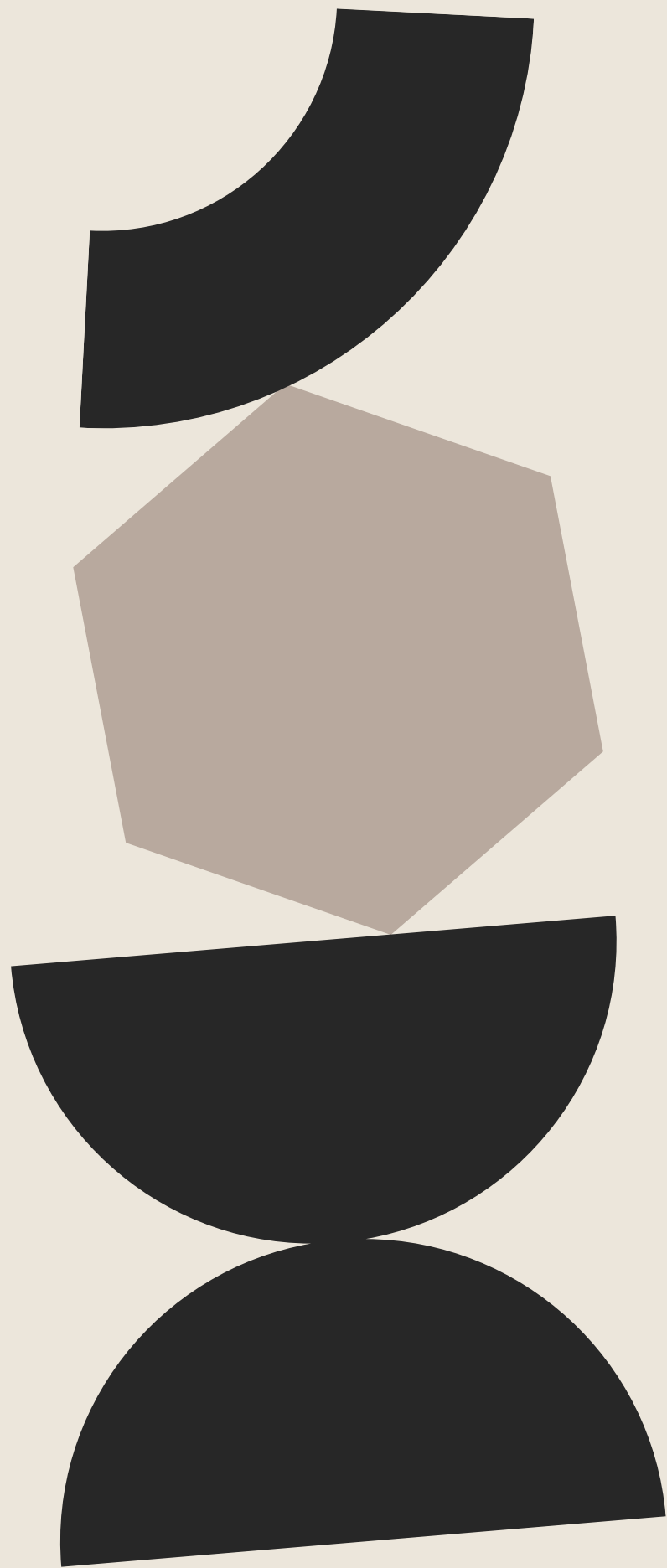
Covering all bases is, ironically enough, both easier and more difficult than ever before.

Frameworks are essential

Many respected entities have
proposed/published frameworks

As always, probably none are perfect - but most
are useful.





Monitoring is key

MLOps is a fairly established function within enterprises - can we say the same for LLMOps?

Lots of entrepreneurs getting into this space

Focus on compatibility, scalability, and fast updates to your overwatch systems

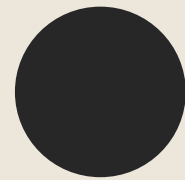
Key Concepts Revised



Implementing algorithms in unbiased ways for fair decision-making processes



Identifying scope and ways for bias to creep in



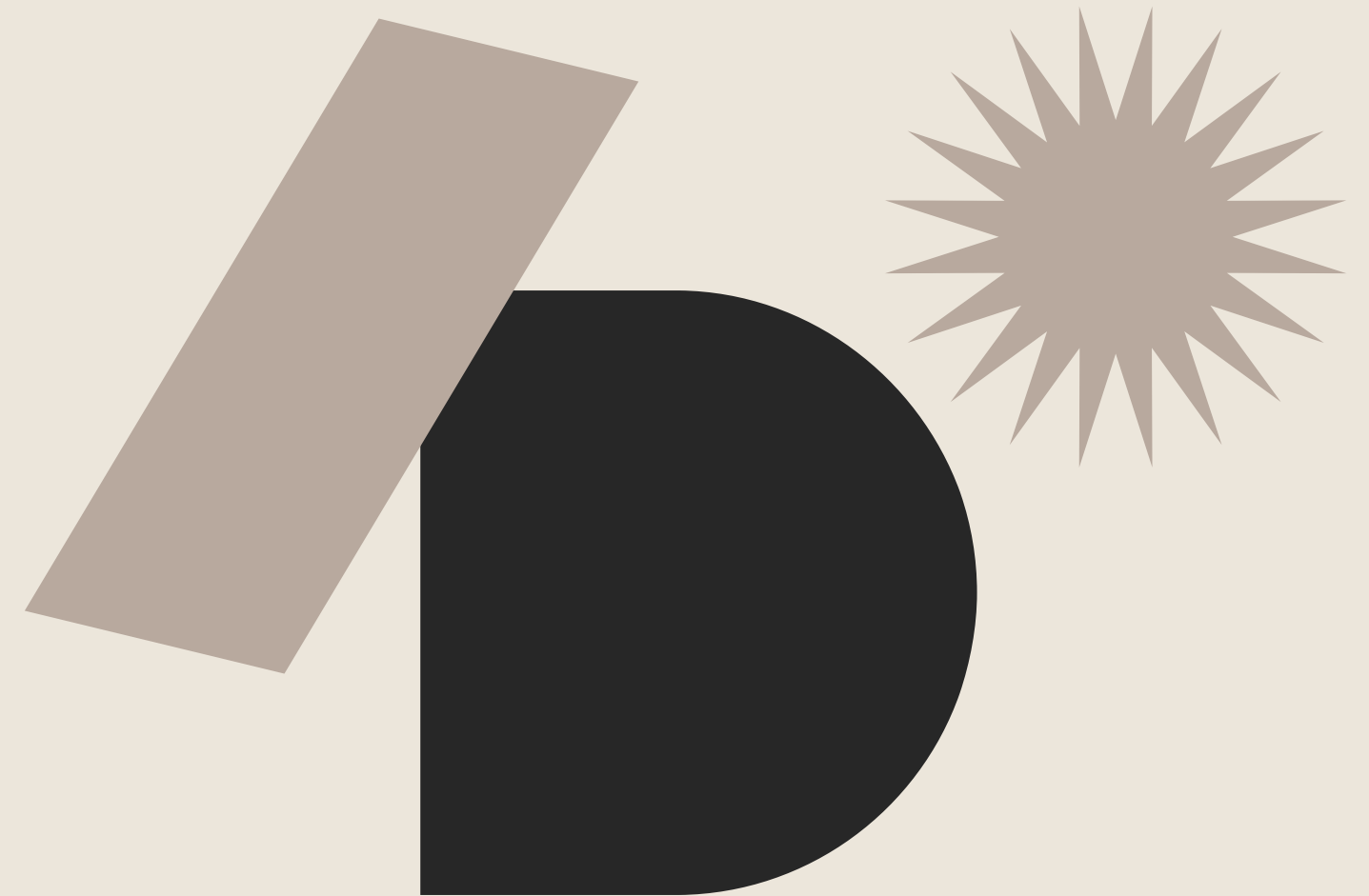
Data Selection and Diversification



Ethical Considerations Addressed



Monitoring Model Outputs



- 
- (1) Five strategies to mitigate bias when implementing generative AI. <https://www.telusinternational.com/insights/ai-data/article/mitigating-genai-bias>.
 - (2) Reducing bias in AI-based financial services | Brookings. <https://www.brookings.edu/articles/reducing-bias-in-ai-based-financial-services/>.
 - (3) Bias in Generative AI - arXiv.org. <https://arxiv.org/pdf/2403.02726.pdf>.
 - (4) Bias Mitigation in Generative AI - Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2023/09/bias-mitigation-in-generative-ai/>.
 - (5) Digital ethics and banking: A strong AI strategy starts with ... - Deloitte. <https://www.deloitte.com/global/en/Industries/financial-services/perspectives/digital-ethics-and-banking-a-strong-ai-strategy-starts-with-customer-trust.html>.
 - (6) Overseeing AI: Governing artificial intelligence in banking.
https://impact.economist.com/perspectives/sites/default/files/eiu_temenos_article_2_overseeing_ai_june_2020_1.pdf.
 - (7) From ethical AI frameworks to tools: a review of approaches. <https://link.springer.com/article/10.1007/s43681-023-00258-9>.
 - (8) Eliminating AI biases in Fintech & Banking - VALID Systems. <https://validadvantage.com/insight/eliminating-ai-biases-in-fintech-banking/>.
 - (9) The Challenges of AI Algorithm Bias in Financial Services. <https://www.techopedia.com/ai-algorithm-bias-in-financial-services>.
 - (10) Shedding light on AI bias with real world examples - IBM Blog. <https://www.ibm.com/blog/shedding-light-on-ai-bias-with-real-world-examples/>.
 - (11) What Do We Do About the Biases in AI - nwscrs.org.
http://www.nwscrs.org/uploads/4/1/5/1/41511209/biases_detected_by_ai.pdf.
 - (12) The Ethics of AI Addressing Bias, Privacy, and Accountability in <https://www.cloudthat.com/resources/blog/the-ethics-of-ai-addressing-bias-privacy-and-accountability-in-machine-learning>.
 - (13) What Is AI Bias? | IBM. <https://www.ibm.com/topics/ai-bias>.
 - (14) Explained: Generative AI | MIT News | Massachusetts Institute of Technology. <https://news.mit.edu/2023/explained-generative-ai-1109>.
 - (15) AI image generators often give racist and sexist results: can ... - Nature. <https://www.nature.com/articles/d41586-024-00674-9>.
 - (16) The future of AI in banking: Choosing the right model | McKinsey. <https://www.mckinsey.com/industries/financial-services/our-insights/scaling-gen-ai-in-banking-choosing-the-best-operating-model>.
 - (17) CBI | The ethics of AI - Chartered Banker Institute. https://www.charteredbanker.com/resource_listing/knowledge-hub-listing/the-ethics-of-ai.html.
 - (18) AI in Finance: Revolutionizing the Future of Financial Management. <https://www.datacamp.com/blog/ai-in-finance>.
 - (19) Ethical Considerations of AI in Finance - redresscompliance.com. <https://redresscompliance.com/ethical-considerations-of-ai-in-finance/>.