**Evaluating Generative AI for Regulated Enterprise Environments**

Submitted to Northcentral University

School of Computer Science

in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OFPHILOSOPHY

by

Nate Bachmeier

La Jolla, California

June 2023

Abstract

Begin writing here…

Checklist:

Briefly introduce the study topic, state the research problem, and describe who or what is impacted by this problem.

Clearly articulate the study purpose and guiding theoretical or conceptual framework of the study.

Provide details about the research methodology, participants, questions, design, procedures, and analysis.

Clearly present the results in relation to the research questions.

State the conclusions to include both the potential implications of the results on and the recommendations for future research and practice.

Do not include citations and abbreviations or acronyms, except those noted as exceptions by the American Psychological Association (APA).

Do not exceed 350 words. Strive for one page.

Table of Contents

[Chapter 1: Introduction 1](#_Toc51929210)

[Statement of the Problem 2](#_Toc51929211)

[Purpose of the Study 2](#_Toc51929212)

[Introduction to Theoretical or Conceptual Framework 3](#_Toc51929213)

[Introduction to Research Methodology and Design 4](#_Toc51929214)

[Research Questions 4](#_Toc51929215)

[Hypotheses 4](#_Toc51929216)

[Significance of the Study 5](#_Toc51929217)

[Definitions of Key Terms 6](#_Toc51929218)

[Summary 6](#_Toc51929219)

[Chapter 2: Literature Review 7](#_Toc51929220)

[Theoretical or Conceptual Framework 7](#_Toc51929221)

[Subtopic 8](#_Toc51929222)

[Summary 8](#_Toc51929223)

[Chapter 3: Research Method 10](#_Toc51929224)

[Research Methodology and Design 10](#_Toc51929225)

[Population and Sample 10](#_Toc51929226)

[Materials or Instrumentation 11](#_Toc51929227)

[Operational Definitions of Variables 12](#_Toc51929228)

[Study Procedures 13](#_Toc51929229)

[Data Analysis 13](#_Toc51929230)

[Assumptions 14](#_Toc51929231)

[Limitations 14](#_Toc51929232)

[Delimitations 14](#_Toc51929233)

[Ethical Assurances 15](#_Toc51929234)

[Summary 15](#_Toc51929235)

[Chapter 4: Findings 16](#_Toc51929236)

[XXX of the Data 16](#_Toc51929237)

[Results 17](#_Toc51929238)

[Evaluation of the Findings 18](#_Toc51929239)

[Summary 18](#_Toc51929240)

[Chapter 5: Implications, Recommendations, and Conclusions 19](#_Toc51929241)

[Implications 19](#_Toc51929242)

[Recommendations for Practice 20](#_Toc51929243)

[Recommendations for Future Research 20](#_Toc51929244)

[Conclusions 20](#_Toc51929245)

[References 22](#_Toc51929246)

[Appendix A XXX 23](#_Toc51929247)

[Appendix B XXX 24](#_Toc51929248)

List of Tables

Begin list of tables here…

List of Figures

Begin list of figures here…

# Chapter 1: Introduction

Generative AI (GAI) is a class of machine learning algorithms that can learn from and creates content such as text, images, video, audio, and code (Wu et al., 2023). GAI transforms documents into summarizations and user-defined classification labels (Pengfei et al., 2023). It also lowers the barrier to entry for software engineering tasks, increasing developer velocity (Ozkaya, 2023). Generative behaviors contrast with discriminative machine learning algorithms, which learn decision boundaries (Sun et al., 2022). For example, traditional logistical regression strategies are appropriate for predicting housing prices.

Implementations like Open AI’s ChatGPT capture the business community’s imagination and raise questions about its applications within regulated enterprises (Daugherty, Wilson, & Narain, 2023; Rozado, 2023). This combination necessitates research into the understanding of the technology and practical use cases. For instance, where are the delineations between what one can versus should do today? Equally important, how can developer operations (DevOps) teams leverage these capabilities securely, reliably, and safely? Since there are multiple technology choices, when should enterprises choose one over the other? How can enterprises methodically select the right option?

This chapter aims to frame these questions and establish the purpose of this constructive research project. It briefly describes the business context and essential drivers to assess the technical state of Generative AI and map this information to business architectures. Chapter 2, *Literature Review*, expands on core concept overviews and examines the academic perspective to leverage GAI successfully. Chapter 3, *Research Method*, documents the research approach and mechanisms for quantifying success. These instructions aim to be concise, specific, and reproducible by the broader engineering community. Chapter 4, *Findings*, contains an analysis of the study and observations regarding the research questions. Chapter 5, *Implications, Recommendations, and Conclusions*, interprets the data’s meaning and provides opinions for future implementation.

But first, what exactly is Generative AI, and where did it come from? Frank Rosenblatt (1958) proposed the Mark I Perception as the first neural network architecture. This construct attempts to explain animals’ biological networks to *perceive* the world around them. The network consists of a collection of weighted sensors that converge into *one* learning circuit. Mechanical devices can replicate that implementation and train the connected weights to emulate animal intelligence with noisy data. Since this seminal paper, researchers have expanded deep neural networks to incorporate hundreds to thousands of connectivity layers. However, the specific configuration remains more art than science, with researchers manually iterating through trial-and-error experimentation (Ünal & Başçiftçi, 2021). The research field has two eras, with the first being the evolution of the artificial neural network (ANN) from 1989 to 2015. Then, the deep neural network (DNN) evolution became the primary focus from 2015 to the present. The industry is moving so fast that information beyond two years is becoming outdated.

Solutions like EMLo are Long-Term, Short-Term (LTSM) architectures that provide *context-sensitive* through self-attention mechanisms (Peters et al., 2017). Intuitively, this means that recently used words (tokens) should have a higher probability of reuse. Devlin et al. (2018) introduced the Bidirectional Encoder Representation from Transformers (BERT) architecture that *pre-trains* representations of data that maps left-to-right context (*inputs to outputs*) and vice versa (*outputs to inputs*). Their work fundamentally differs from decades of previous research into natural language processing (NLP) that follows the pattern of fine-tuned, purpose-built unidirectional task solvers. Specifically, BERT doesn’t require fine-tuning and labeled data for many scenarios because it learns these self-attentions through large cohorts of unstructured data.

The ubiquity of public cloud providers enables researchers to create more extensive and sophisticated models. For example, GPT-3, an autoregressive language model, has over 175 billion parameters, a 10x increase over GPT-2 (Brown et al., 2020). More recently, Google Brain demonstrated an NLP translation model with over a trillion parameters (Fedus et al., 2022). This exponential parameter growth is likely to continue into the foreseeable future. As predictive capabilities mature, it democratizes machine learning by lowering the barrier to getting started and removing the need for extensive labeled training sets.

Businesses can consume and extend these *large language models* (LLMs) to solve unsupervised learning tasks. While the domain is rapidly maturing, numerous open research problems remain regarding prediction accuracy, explainability, suitability, safety, and security. These challenges become more pronounced in regulated enterprise settings such as financial, health care, and workplace safety. Suppose a generative response creates an insurance policy and offers it to end-users. In that case, the insurance provider becomes responsible for the terms of that policy. How can the insurer guarantee and represent its principles and business practices within that offering? Failures in this regard could hurt the business’s reputation, cause financial hardship, and alienate customers.

## Statement of the Problem

The problem to be addressed in this study is an exploratory analysis of Generative AI use cases is the viability and practicality of mechanisms for ensuring norms and expectations of enterprise businesses within regulated environments. Researchers continuously discover new characteristics and challenges within LLMs that could misalign corporate standards. For instance, OpenAI’s ChatGPT-3.5 has safeguards and content moderation checks that ensure its responses are politically neutral. However, prompt engineering techniques bypass that validation and statistically demonstrate it’s broadly a left-leaning Libertarian (Rozado, 2023). While that’s appropriate for many settings, these implicit biases could be at odds with other organizations. Another set of challenges comes from international requirements, such as the European Union proposing copyright rules and China exploring a regulatory framework (International Business Times, 2023; Mondaq Business Briefing, 2023).

Technology hype cycles transition through five key phases: the innovation trigger, the peak of inflated expectations, the trough of disillusionment, the slope of enlightenment, and the plateau of productivity (Gartner, 2023). It can be challenging in the present to understand which stage enterprise businesses are experiencing, and that impacts their ability to prioritize today’s investments versus tomorrow’s.

## Purpose of the Study

The purpose of this constructive design study is to examine the state of Generative AI capabilities within enterprise businesses and create a framework for prioritizing use cases based on risk and practicality. For example, which business challenges are more economically solvable through statistical modeling and discriminative machine learning? Many organizations use automatic speech recognition (ASR) to transcribe audio-to-text economically. To what extent do LLMs enhance this performance? Since business and commercial communities operate at different speeds, defining the suitability of those performance gains is essential. These questions are not unique to audio use cases and span images, videos, rich text documents, and code scenarios. This research project will establish a multi-modal use case taxonomy as described in Chapter 3, *Research Method*.

## Introduction to Theoretical Framework

The design of experiments research creates purposeful artifacts and applies them to study a phenomenon (Hevner et al., 2004). Academic and business communities employ this method as a standard approach to Information Technology and Communication (IT&C) problems (Bryar & Carr, 2021; Peffers et al., 2007). It has well-defined guidelines (see Table 1) to implement a three-phased procedure. First, the researcher(s) must identify a domain-specific challenge. Next, that researcher creates artifacts that study this phenomenon. Third, those artifacts assess the topic and communicate answers to the research questions.

**Table 1**  
Design-science Guidelines (Hevner et al. 2004)

|  |  |
| --- | --- |
| Guideline | Description |
| Design as an Artifact | Design-science research must produce a viable artifact as a construct, a model, a method, or an instantiation. |
| Problem Relevance | Design-science research aims to develop technology-based solutions to important and relevant business problems. |
| Design Evaluation | A design artifact’s utility, quality, and efficacy must rigorously demonstrate well-executed evaluation methods. |
| Research Contributions | Effective design-science research must provide transparent and verifiable contributions to design artifacts, foundations, and/or design methodologies. |
| Research Rigor | Design-science research relies on rigorous methods to construct and evaluate the design artifact. |
| Design as a Search Process | The search for a compelling artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment. |
| Communication of Research | Design-science research must be presented effectively both to technology-oriented and management-oriented audiences. |

This study uses these guidelines and conceptual steps to identify a research-worthy topic and an actionable aspect. Next, it defines an abstract approach and implements a concrete proof-of-concept, the simulation process, to assess patient monitoring (via CV) and remediation (via CPS) technologies. Third, the artifacts expand the body of knowledge through the research questions. See Chapter 3: *Research Method* for more information.

## Research Questions

Begin writing here...

### RQ1

Text…

### RQ2

Text…

## Hypotheses

### H10

Text…

### H1a

Text…

### H20

Text…

### H2a

Text…

Checklist:

Present research questions directly answerable, specific, and testable within the given timeframe and location identified in the problem and purpose statements.

Include the exact same variables/constructs, participants, and location mentioned in the problem and purpose statements. No new variables/constructs should be introduced.

## Significance of the Study

Begin writing here…

Checklist:

Describe why the study is important and how it can contribute to the field of study.

For applied studies, explain how the results might both be significant to leaders and practitioners in the field and contribute to the literature. For PhD studies, explain how the results advance the guiding framework and contribute to the literature.

Describe the benefits of addressing the study problem, achieving the study purpose, and answering the research questions. Whereas the problem statement should articulate the negative consequences of not conducting the study, this section should highlight the positive consequences of completing the study.

Do not exceed one page.

## Definitions of Key Terms

### Term 1

Text…

### Term 2

Text…

Checklist:

Alphabetize and bold terms directly related to the dissertation topic and not commonly used or understood.

Paraphrase the definitions of the terms using complete sentences and provide a citation for each one.

Do notdefine theories, conceptual frameworks, statistical analyses, methodological terms, or the variables/constructs under examination.

## Summary

Begin writing here…

Checklist:

Briefly restate the key points discussed in the chapter. Review the headings and/or table of contents to ensure all key points are covered.

# Chapter 2: Literature Review

Begin writing here…

Checklist:

Begin with the first sentence of the purpose statement and problem statement that leads to a brief explanation of the organization of the literature review. Do not simply cut and paste the Purpose Statement section from Chapter 1.

Provide an overview of the sub-headings in the literature that will be discussed.

At the end of this section, indicate the databases accessed and the search engines used. Discuss all the search parameters, including the search terms and their combinations (with more detailed search terms located in an appendix, if appropriate), range of years, and types of literature.

Devote approximately 30 to 60 pages to this chapter to include citations to at least 50 relevant sources.

## Theoretical or Conceptual Framework

Begin writing here…

Checklist:

Describe the guiding theoretical/conceptual framework of the study, including the definitions of all the concepts, an explanation of the relationships among the concepts, and a presentation of all the assumptions and propositions.

Explain the origin and development of the framework. Demonstrate detailed knowledge of and familiarity with both the historical and the current literature on the framework.

Identify existing research studies that used this framework in a similar way. Mention alternative frameworks, with a justification of why the selected framework was chosen.

Describe how and why the selected framework relates to the present study and how it guided the development of the problem statement, purpose statement, and research questions.

## Subtopic

Begin writing here…

### Level 3 Heading

Text...

Level 4 Heading. Text...

Checklist:

Critically analyze (i.e., note the strengths and weaknesses) and synthesize (i.e., integrate) the existing research. Rather than reporting on each study independently, describe everything known on the topic by reviewing the entire body of work.

Present a balanced integrative critical review of the literature, ensuring all points of view are included. Cover all the important issues with a discussion of areas of convergence (i.e., agreement) and divergence (i.e., disagreement). Provide potential explanations for areas of divergence.

Address issues of authority, audience, and/or bias/point of view in the sources used.

## Summary

Begin writing here…

Checklist:

Briefly restate the key points discussed in the chapter. Review the headings and/or table of contents to ensure all key points are covered.

Highlight areas of convergence and divergence as well as gaps in the literature that support the need for the study. This discussion should logically lead to Chapter 3, where the research methodology and design will be discussed.

# Chapter 3: Research Method

Begin writing here…

Checklist:

Begin with an introduction and restatement of the problem and purpose sentences verbatim.

Provide a brief overview of the contents of this chapter, including a statement that identifies the research methodology and design.

## Research Methodology and Design

Begin writing here…

Checklist:

Describe the research methodology and design. Elaborate upon their appropriateness in relation to the study problem, purpose, and research questions.

Identify alternative methodologies and designs and indicate why they were determined to be less appropriate than the ones selected. Do not simply list and describe research methodologies and designs in general.

## Population and Sample

Begin writing here…

Checklist:

Describe the population, including the estimated size and relevant characteristics.

Explain why the population is appropriate, given the study problem, purpose, and research questions.

Describe the sample that will be (proposal) or was (manuscript) obtained.

Explain why the sample is appropriate, given the study problem, purpose, and research questions.

Explain the type of sampling used and why it is appropriate for the dissertation proposal methodology and design. For qualitative studies, evidence must be presented that saturation will be (proposal) or was (manuscript) reached. For quantitative studies, a power analysis must be reported to include the parameters (e.g., effect size, alpha, beta, and number of groups) included, and evidence must be presented that the minimum required sample size will be (proposal) or was (manuscript) reached.

Describe how the participants will be (proposal) or were (manuscript) recruited (e.g., email lists from professional organizations, flyers) and/or the data will be (proposal) or were (manuscript) obtained (e.g., archived data, public records) with sufficient detail so the study could be replicated.

## Materials or Instrumentation

Begin writing here…

Checklist:

Describe the instruments (e.g., tests, questionnaires, observation protocols) that will be (proposal) or were (manuscript) used, including information on their origin and evidence of their reliability and validity. OR as applicable, describe the materials to be used (e.g., lesson plans for interventions, webinars, or archived data, etc.).

Describe in detail any field testing or pilot testing of instruments to include their results and any subsequent modifications.

If instruments or materials are used that were developed by another researcher, include evidence in the appendix that permission was granted to use the instrument(s) and/or material(s) and refer to that fact and the appendix in this section.

## Operational Definitions of Variables

Begin writing here...

### XXX

Text…

Checklist:

For quantitative and mixed methods studies, identify how each variable will be (proposal) or was (manuscript) used in the study. Use terminology appropriate for the selected statistical test (e.g., independent/dependent, predictor/criterion, mediator, moderator).

Base the operational definitions on published research and valid and reliable instruments.

Identify the specific instrument that will be (proposal) or was (manuscript) used to measure each variable.

Describe the level of measurement of each variable (e.g., nominal, ordinal, interval, ratio), potential scores for each variable (e.g., the range [0–100] or levels [low, medium, high]), and data sources. If appropriate, identify what specific scores (e.g., subscale scores, total scores) will be (proposal) or were (manuscript) included in the analysis and how they will be (proposal) or were (manuscript) derived (e.g., calculating the sum, difference, average).

## Study Procedures

Begin writing here…

Checklist:

Describe the exact steps that will be (proposal) or were (manuscript) followed to collect the data, addressing what data as well as how, when, from where, and from whom those data will be (proposal) or were (manuscript) collected in enough detail the study can be replicated.

## Data Analysis

Begin writing here…

Checklist:

Describe the strategies that will be (proposal) or were (manuscript) used to code and/or analyze the data, and any software that will be (proposal) or was (manuscript) used.

Ensure the data that will be (proposal) or were (manuscript) analyzed can be used to answer the research questions and/or test the hypotheses with the ultimate goal of addressing the identified problem.

Use proper terminology in association with each design/analysis (e.g., independent variable and dependent variable for an experimental design, predictor and criterion variables for regression).

For quantitative studies, describe the analysis that will be (proposal) or was (manuscript) used to test each hypothesis. Provide evidence the statistical test chosen is appropriate to test the hypotheses and the data meet the assumptions of the statistical tests.

For qualitative studies, describe how the data will be (proposal) or were (manuscript) processed and analyzed, including any triangulation efforts. Explain the role of the researcher.

For mixed methods studies, include all of the above.

## Assumptions

Begin writing here…

Checklist:

Discuss the assumptions along with the corresponding rationale underlying them.

## Limitations

Begin writing here…

Checklist:

Describe the study limitations.

Discuss the measures taken to mitigate these limitations.

## Delimitations

Begin writing here…

Checklist:

Describe the study delimitations along with the corresponding rationale underlying them. An example of delimitations are the conditions and parameters set intentionally by the researcher or by selection of the population and sample.

Explain how these research decisions relate to the existing literature and theoretical/conceptual framework, problem statement, purpose statement, and research questions.

## Ethical Assurances

Begin writing here…

Checklist:

Confirm in a statement the study will (proposal) or did (manuscript) receive approval from Northcentral University’s Institutional Review Board (IRB) prior to data collection.

If the risk to participants is greater than minimal, discuss the relevant ethical issues and how they will be (proposal) or were (manuscript) addressed.

Describe how confidentiality or anonymity will be (proposal) or was (manuscript) achieved.

Identify how the data will be (proposal) or were (manuscript) securely stored in accordance with IRB requirements.

Describe the role of the researcher in the study. Discuss relevant issues, including biases as well as personal and professional experiences with the topic, problem, or context. Present the strategies that will be (proposal) or were (manuscript) used to prevent these biases and experiences from influencing the analysis or findings.

In the dissertation manuscript only, include the IRB approval letter in an appendix.

## Summary

Begin writing here…

Checklist:

Summarize the key points presented in the chapter.

Logically lead the reader to the next chapter on the findings of the study.

# Chapter 4: Findings

Begin writing here…

Checklist:

Begin with an introduction and restatement of the problem and purpose sentences verbatim and the organization of the chapter.

Organize the entire chapter around the research questions/hypotheses.

## XXX of the Data

Begin writing here…

Checklist:

For qualitative studies, clearly identify the means by which the trustworthiness of the data was established. Discuss credibility (e.g., triangulation, member checks), transferability (e.g., the extent to which the findings are generalizable to other situations), dependability (e.g., an in-depth description of the methodology and design to allow the study to be repeated), and confirmability (e.g., the steps to ensure the data and findings are not due to participant and/or researcher bias).

For quantitative studies, explain the extent to which the data meet the assumptions of the statistical test and identify any potential factors that might impact the interpretation of the findings. Provide evidence of the psychometric soundness (i.e., adequate validity and reliability) of the instruments from the literature as well as in this study (as appropriate). Do not merely list and describe all the measures of validity and reliability.

Mixed methods studies should include discussions of the trustworthiness of the data as well as validity and reliability.

## Results

Begin writing here…

Checklist:

Briefly discuss the overall study. Organize the presentation of the results by the research questions/hypotheses.

Objectively report the results of the analysis without discussion, interpretation, or speculation.

Provide an overview of the demographic information collected. It can be presented in a table. Ensure no potentially identifying information is reported.

### Research Question 1/Hypothesis

Text…

Report all the results (without discussion) salient to the research question/hypothesis. Identify common themes or patterns.

Use tables and/or figures to report the results as appropriate.

For quantitative studies, report any additional descriptive information as appropriate. Identify the assumptions of the statistical test and explain how the extent to which the data met these assumptions was tested. Report any violations and describe how they were managed as appropriate. Make decisions based on the results of the statistical analysis. Include relevant test statistics, *p* values, and effect sizes in accordance with APA requirements.

For qualitative studies,describe the steps taken to analyze the data to explain how the themes and categories were generated. Include thick descriptions of the participants’ experiences. Provide a comprehensive and coherent reconstruction of the information obtained from all the participants.

For mixed methods studies,include all of the above.

## Evaluation of the Findings

Begin writing here…

Checklist:

Interpret the results in light of the existing research and theoretical or conceptual framework (as discussed in Chapters 1 and 2). Briefly indicate the extent to which the results were consistent with existing research and theory.

Organize this discussion by research question/hypothesis.

Do not draw conclusions beyond what can be interpreted directly from the results.

Devote approximately one to two pages to this section.

## Summary

Begin writing here…

Checklist:

Summarize the key points presented in the chapter.

# Chapter 5: Implications, Recommendations, and Conclusions

Begin writing here…

Checklist:

Begin with an introduction and restatement of the problem and purpose sentences verbatim, and a brief review of methodology, design, results, and limitations.

Conclude with a brief overview of the chapter.

## Implications

Begin writing here…

Checklist:

Organize the discussion around each research question and (when appropriate) hypothesis individually. Support all the conclusions with one or more findings from the study.

Discuss any factors that might have influenced the interpretation of the results.

Present the results in the context of the study by describing the extent to which they address the study problem and purpose and contribute to the existing literature and framework described in Chapter 2.

Describe the extent to which the results are consistent with existing research and theory and provide potential explanations for unexpected or divergent results.

Identify the most significant implications and consequences of the dissertation (whether positive and/or negative) to society/desired societal outcomes and distinguish probable from improbable implications.

### Research Question 1/Hypothesis

Text…

## Recommendations for Practice

Begin writing here…

Checklist:

Discuss recommendations for how the findings of the study can be applied to practice and/or theory. Support all the recommendations with at least one finding from the study and frame them in the literature from Chapter 2.

Do not overstate the applicability of the findings.

## Recommendations for Future Research

Begin writing here…

Checklist:

Based on the framework, findings, and implications, explain what future researchers might do to learn from and build upon this study. Justify these explanations.

Discuss how future researchers can improve upon this study, given its limitations.

Explain what the next logical step is in this line of research.

## Conclusions

Begin writing here…

Checklist:

Provide a strong, concise conclusion to include a summary of the study, the problem addressed, and the importance of the study.

Present the “take-home message” of the entire study.

Emphasize what the results of the study mean with respect to previous research and either theory (PhD studies) or practice (applied studies).

# References

Reference 1

Reference 2

Author, A., & Author, B. (year). Article title. *Journal title, X*(X), xxx-xxx.

https://doi.org/xxxxx

# 

# Appendix A XXX

Insert Appendix A content here…

# Appendix B XXX

Insert/type Appendix n content here…