International Journal of
Information Systems and Project Management

ISSN (print):2182-7796, ISSN (online):2182-7788, ISSN (cd-rom):2182-780X

Available online at ijispm.sciencesphere.org

The need for a risk management framework for data science projects: a systematic literature review

Sucheta Lahiri
School of Information Studies, Syracuse University
Syracuse city, New York State 13244
United States of America
sulahiri@syr.edu

Jeff Saltz
School of Information Studies, Syracuse University
Syracuse city, New York State 13244
United States of America
jsaltz@syr.edu

Abstract:
Many data science endeavors encounter failure, surfacing at any project phase. Even after successful deployments, data science projects grapple with ethical dilemmas, such as bias and discrimination. Current project management methodologies prioritize efficiency and cost savings over risk management. The methodologies largely overlook the diverse risks of sociotechnical systems and risk articulation inherent in data science lifecycles. Conversely, while the established risk management framework (RMF) by NIST and McKinsey aims to manage AI risks, there is a heavy reliance on normative definitions of risk, neglecting the multifaceted subjectivities of data science project failures. This paper reports on a systematic literature review that identifies three main themes: Big Data Execution Issues, Demand for a Risk Management Framework tailored for Large-Scale Data Science Projects, and the need for a General Risk Management Framework for all Data Science Endeavors. Another overarching focus is on how risk is articulated by the institution and the practitioners. The paper discusses a novel and adaptive data science risk management framework – "DS EthiCo RMF" – which merges project management, ethics, and risk management for diverse data science projects into one holistic framework. This agile risk management framework DS EthiCo RMF can bridge the current divide between normative risk standards and the multitude of data science requirements, offering a human-centric method to navigate the intertwined sociotechnical risks of failure in data science projects.

Keywords:
project management; risk management framework; data science; bias; methodology.

DOI: 10.12821/ijispm120403

Manuscript received: 26 September 2023
Manuscript accepted: 13 May 2024

1. Introduction

Data science is an interdisciplinary field that employs statistical, machine learning, and artificial intelligence (AI) tools to determine insights from complex and situated data [1]. Another definition stems from the knowledge insights that data science methods draw from data [2]. The often-heterogeneous data is never 'raw' [5,8], and rather cooked. The data represents a combination of ideas, negotiated decisions, contexts, individuals, protocols, macroeconomic factors, and institutional practices [5]. Due to the nature of data science projects, the processes, and practices of organizing and analyzing data are however rarely linear and often more exploratory. Although data science work practices bringing together data, practitioners, and tools, are heterogeneous and involve discretionary human-led decisions; paradoxically there is still a dominant theme of technology deterministic information systems (IS) literature to explain data science projects [3,9]. In other words, there is scant sociotechnical research that explains the messiness, tension, power conflicts, and 'how' the data science work is performed to create systems [4,5]. The case study of Google Maps and the politics of cartography provided in the second section of this paper is one such example of human biases and political influence in the making of the system. The cartography also testifies that the making is not objective and rather involves the discretion of various stakeholders who hold vested power.

Also, the research on data science work is studied mostly in the realm of restricted academic settings and not in real-time or practical corporate settings where the systems are deployed and distributed by data science practitioners [5]. The practitioners who are engaged in the making are driven by organizational goals, perceived incentives, institutional policies, and established processes. It is important here to recognize the collaborative human work of data science practitioners that is often invisible but essential in developing sociotechnical systems. In other words, data science as a sociotechnical system involves a great deal of human labor that collaboratively aids in data processing and analysis [6]. Methodological choices and contextual factors further affect the process of data analysis and consequently decision-making. In critical data studies, scholars characterize data science projects as messy and sociomaterial endeavors in which practitioners and technical skills together culminate sense-making and discretionary decision-making with data [5,7].

The collaborative work of data science in project execution raises several questions about data and system creation, such as how the data is produced, who is involved in data processing, what processes are used for managing data science projects, how is data managed and analyzed, how are business objectives translated into system development, how are knowledge insights interpreted, and who ultimately approves or rejects the final system. Additionally, questions about coercion, persuasion, and the legitimized and de-legitimized voices are important considerations in this process. Finally, understanding how a system is deemed a failure is crucial for improving the data science processes.

In a data-driven culture that values data practices, there are numerous areas of concern surrounding the process of creating data science knowledge insights. Specifically, questions linger around how a data science project's entire workflow is carried out and the risks of project failure emerge during the system creation process [10].

In this paper, we begin by understanding how failure is defined for data science projects, and why it is important to rather understand the risks before the failure of data science projects. We then define various types of risks of failure for data science projects with examples. Following the need to explore risk mitigation with risk management framework (RMF) for data science projects, we dedicate the next section to explain the methodology to conduct an in-depth SLR on RMF for data science projects. The last sections highlight the findings of the SLR and the proposal for an agile risk management framework DS EthiCo RMF to manage the risks of failure for data science projects.

The concept of failure has been defined in IS literature using the "iron triangle" of time, cost, and quality [11,12]. More broadly, project failure can occur when business objectives, processes, interactions, or user expectations are not met in the proposed timelines. However, failure is not a universal concept - its meaning and implications depend on the specific context. In other words, failure cannot follow a 'one size fits all' approach. For data science projects, one common definition is a failure to deploy the project into production, with studies showing up to 87% of projects end up before deployment [13,16,43]. However, this narrow view misses other important failure modes, such as successfully launched systems that break down in distribution or cause harm in practice [14]. Data science initiatives can also fail by producing unreliable predictions that undermine stakeholder trust or propagating biased results despite passing the testing phase. In essence, deployed models can fail by perpetuating human biases. Defining failure requires looking beyond the project deployment and before failure, the risk of failure to consider potential societal damages from biased, inaccurate, or unusable data science systems after the final launch.

While extensive research exists on utilizing data science for risk management, such as stochastic models driven by Bayesian Statistics to assess the systemic risk of the financial market [15], less attention is paid to managing the risks introduced by data science itself [10]. Despite the literature on the risks of software development projects, the research on data science risk is sparse [10,14,44]. This lack of research on managing data science project risk was determined via the results of an SLR that explored the use of RMF for managing data science project risk [10]. The SLR explored if there was any dedicated RMF for addressing the risk introduced when organizations use data science projects. The SLR explored published relevant conference papers, proceedings, and peer-reviewed journal papers from 2015 to 2019. The content analysis with title, abstract, and subsequent full-text analysis of filtered articles derived from generic or very specific standards that are not dedicated to data science was proposed to manage data science risk. In other words, no articles were found that explicitly highlighted RMF for data science projects. This study also highlighted that ethics was seen as a weak theme for RMF compared to other dominant themes such as big data execution challenges and sector-specific RMFs.

Due to the lack of research literature on the risks of data science while project execution deduced through the SLR conducted in 2019, this paper extends the SLR in 2022 to understand the trend of research done in data science risk of failure. The research probes the following overarching question:

RQ: How does the trend of literature on risks of data science failure show evidence of risk management frameworks?

The paper is divided into five sections. The first section provides a brief introduction to data science failure and associated research. The second section outlines various types of data science project failures. The third section describes both general and AI-specific RMFs that are recommended to be used for data science projects. This is followed by the fourth section describing the methodology for SLR to explore the risks of data science projects. The fifth section explains the results of a new SLR. Finally, the discussion section highlights the synopsis of existing RMFs for AI projects, followed by potential next steps and a summary section on agile-based risk management framework DS EthiCo RMF for managing the diverse risks of failure for data science projects.

2. Data science projects failures

Data science projects are unique from other IT projects such as software development due to their higher complexity [4], with a greater emphasis on data compared to typical software projects. In other words, data science projects possess certain characteristics that elevate the risk of failure when compared to other IT projects [4].

Data science projects involve a greater magnitude of structured and unstructured data, requiring due diligence to ensure data homogeneity and an emphasis on data lineage. Data science projects often have uncertain inputs and outputs due to unspecified objectives and project complexity, with the scope of these projects having expanded from descriptive to predictive and prescriptive with the introduction of AI tools [4]. The practitioners involved in data science projects aim not only to meet business objectives but also to investigate the extent to which knowledge insights add value to business decisions, emphasizing the multiple dimensions of failures relating to the complexity and scope of these projects.

Empirical studies on data science project failure highlight project challenges as data issues with access, quality and volume, budget, security, unstructured project execution, change management, unrealistic expectations, use case-related issues, talent, constraining regulations, lack of documentation, lack of transparency, scope creep, cyber-attacks, poor stakeholder management [17,18,19]. In other words, the failure of data science projects is linked to technological failure or regulatory breaches, as well as to potential societal harm through the misuse of AI.

2.1 Types of data science project failures

The Project Management Institute (PMI) defines a project as a "temporary effort to create a unique product, service or result" [20]. However, this definition does not fully capture the complexity of data science initiatives. While data science projects begin with a defined business objective and scope, these components often shift over time. Compared to IT projects, data science requires greater attention to detail, especially in data preparation. Unfortunately, failure rates are higher for data science than IT, stemming from three key challenge areas [21]:

- Data challenges arise from properties like variety, quality, availability, and volume across structured, unstructured, and social media sources.
- Process challenges relate to eliciting requirements, capturing/processing data, analysis, modeling, interpreting outputs, and more.
- Management challenges cover governance, privacy, ethics, regulations, policies, and infrastructure.

While process and management challenges affect any project, data issues are unique to data science. In particular, introducing bias through predictive models is a central data science risk of failure. Other drivers of failure include a lack of expertise and an inability to adapt to unexpected events. In summary, the mutable objectives, intensive data needs, and advanced data science techniques led to higher failure rates than traditional IT initiatives.

2.2 Risk of failure: types of data science project failure risk

Risk encompasses various hazards like social, climate, financial, legal, and reputational threats. Risk is commonly defined as a potential, anticipated danger before failure. Quantitatively, it represents the probability of an event and its positive or negative consequences. However, focusing only on the magnitude of risk fails to capture the full harm an event may inflict on society. Additionally, the subjectivity in determining disruptive algorithmic risks makes complete quantification elusive.

Existing project risk definitions, like that from the UK Association for Project Management, describe risk as uncertain circumstances that can impact objectives [22]. Per the Project Management Body of Knowledge (PMBoK), risks are uncertain events that, if occurring, sway objectives positively or negatively [23]. The PMBoK delineates known and unknown risks – those that are legitimized are managed while others stay opaque or left unidentified. Critically examining which data science risks are articulated versus overlooked, and developing techniques to manage failure risks, are vital as not all risks can be fully predicted or mathematically expressed. For instance, a notable example is Microsoft's Tay chatbot, which encountered issues with hate speech due to the presence of unfiltered public data. Despite data science professionals clearing the Tay bot project for user distribution, the overlooked risks associated with incorporating contaminated data had detrimental consequences [24]. Similarly, voice recognition models may encounter challenges in identifying higher-pitched voices, potentially perpetuating stereotypical gender biases towards more feminine voices. These risks can emerge even after rigorous evaluation checks, highlighting the complexity of ensuring a completely risk-free model.

Another example of risks of data science failure through cartographic politics is visible through Google Maps, which provides two different geospatial visualizations of contested areas between India and Pakistan in two locations [25]. The political incompatibility of over 70 years has given rise to a social construction of maps represented through Google Maps. For instance, Figure 1 shows the map available to citizens of Pakistan and the rest of the world and indicates areas of dispute through grey dotted lines.

On the other hand, the Google map in Figure 2 below is made available to the citizens of the Indian subcontinent. Unlike the map created for Pakistani and global citizens, India's map is different with the dotted lines of disputes missing. On the contrary, solid borders are shown to indicate that the disputed area is under the control of Indian territory.

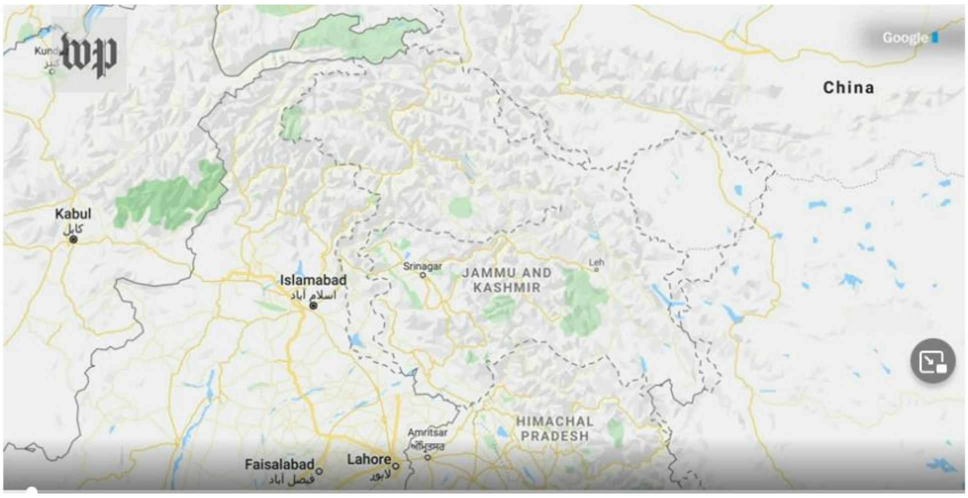

Fig. 1. From Pakistan's view, Kashmir looks disputed. The dotted lines show the areas of dispute.

*Source: G. Bensinger, “Google redraws the borders on maps depending on who’s looking,” Washington Post, 2020.

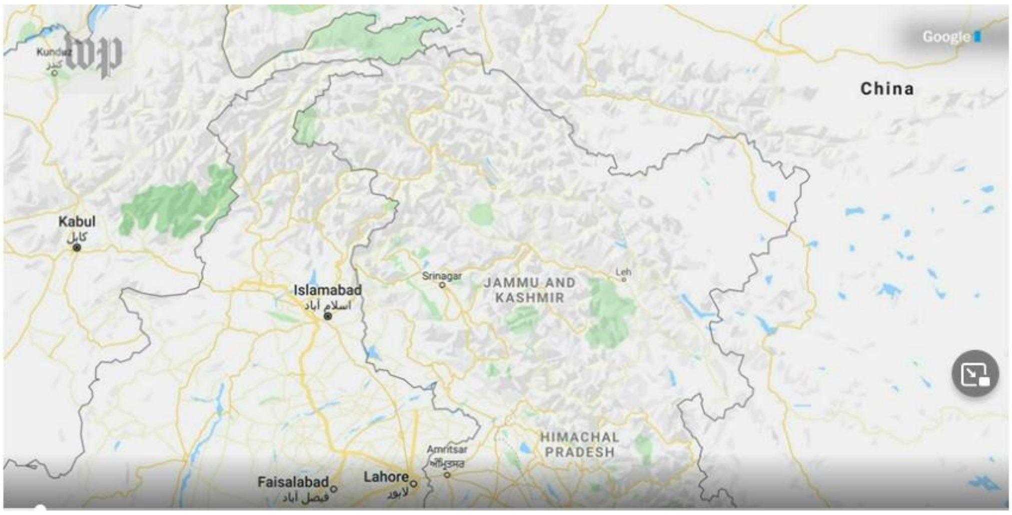

Fig. 2. From India, the entire region of Jammu and Kashmir looks like a part of India

*Source: G. Bensinger, “Google redraws the borders on maps depending on who’s looking,” Washington Post, 2020.

The slogan of Google is “to organize the world’s information”. Essentially, Google asserts that it relies on data sources that best depict borders as outlined in treaties and from highly revered entities such as the UN and ISO [25]. However, in this example, one can see that the mission bends to their discretion, or that they present “alternative facts” based on the content of the request.

As can be deciphered from the above-cited example, risk can be construed as a social and political concept. The factors used to classify events as risk-prone or not are negotiated by individuals and risk policy systems.

The mathematical definition of risk in the social world is not always accepted and rather criticized for several reasons. The mathematical representation of risk does not consider multiple definitions of risk either. The complex, multi-layered expected consequences defined with the mathematical definition cannot fully capture the average probabilities used in technical risk analyses. Limiting the scope of risk with a magnitude can increase the scope of actual risk. Additionally, defining risk as a single numerical quantity does not provide a broader picture of how harm is affected across individuals [26].

In the context of data science, one of the limitations is the difficulty in quantifying the uncertainty of risk while a project is ongoing. One example of the likelihood of a risk event is the usage of incomplete or irrelevant data for machine learning projects. The risk of uncertainty can be classified into two categories: uncertainties inherent to data that cannot be eliminated, and epistemic uncertainties arising when there is a lack of knowledge [27]. Furthermore, it is essential to recognize that these risks may not always occur in isolation but can be interconnected, creating a web of potential vulnerabilities that further complicate the project's health. Therefore, careful consideration and ongoing vigilance are crucial to mitigate risks throughout the entire life cycle of data science initiatives.

Many factors can lead to the failure of data science projects. A primary reason is the absence of a structured process model or methodology [14]. This indicates that the root of data science project failure is frequently not technical but rather tied to inadequate or misaligned project management approaches. Despite the existence of knowledge discovery in databases (KDD) as one of the methods for generating knowledge from vast amounts of data, the project management methodologies developed and motivated by KDD have not undergone any substantial material changes [4]. For example, one of the most extensively used project management methodologies is Cross Industry Standard Process for Data Mining (CRISP-DM) developed in the 1990s with funding from the EU. CRISP-DM was defined for managing data mining projects, and it became the most popular methodology for managing data science projects. CRISP-DM consists of five high-level steps: business understanding, data understanding, data preparation, modeling, evaluation, and deployment.

A paradoxical approach to data science project failure is observed with an iterative agile methodology where the data science practitioners are encouraged to drive projects to failure early so that they can learn from the failure and rectify it to avoid major material damage. The team is encouraged to “fail fast to learn fast”, which indicates that it is acceptable for an iteration to fail when a project uses an agile approach and breaks the work into smaller iterations [29]. This type of failure is encouraged to be realized and in turn, helps to better inform the practitioners how to execute the rest of the project without any further failures with more severe consequences. Failing fast enables the identification of solutions that show promise versus those that have the propensity to break down. Using an iterative agile process, the failures realized during an iteration are accepted as a means to prepare for a better output in the next phase. These failures are anticipated, and data scientists are prepared to accept the disappointment of that result. On the other hand, project failure is not anticipated and can occur at any phase of the project workflow, even after the approved artifact is distributed. Although CRISP-DM is not an agile process framework, this concept of acceptable iteration failure demonstrates the many types of project failures.

3. Potential risk management frameworks for data science projects

3.1 General purpose risk management framework

There are different international standards for risk management adopted and applied by organizations to introduce their own internal rules and guidelines. These standards also operationalize risk and solicit guidelines or general steps of mitigation. Unlike methodologies, standards do not provide a workflow to manage projects. While there are many risk management frameworks, such as ISO and COBIT [30], following are two of the more popular international risk management frameworks used by private and public organizations.

3.1.1 COSO (Committee of Sponsoring Organizations)

In 1985, the Committee of Sponsoring Organizations (COSO) was established through the collaborative efforts of leading institutions, including the American Accounting Association, American Institute of CPAs, Financial Executives International, Institute of Management Accountants, and Institute of Internal Auditors [31]. COSO backed the National Commission on Fraudulent Financial Reporting, an autonomous private-sector initiative aimed at identifying factors leading to deceptive financial reporting.

COSO’s primary mission encompasses thought leadership in three core areas: (1) enterprise risk management, (2) internal control, and (3) fraud deterrence. In 2017, COSO introduced an exhaustive ERM framework titled “Enterprise Risk Management - Integrating with Strategy and Performance.” This framework emphasizes consolidating isolated risks into a unified enterprise architecture. ERM is also characterized as “a process undertaken by an organization’s leadership, management, and other stakeholders, designed during strategy formulation and execution, to identify potential events affecting the organization, and manage risks within its tolerance, ensuring reasonable assurance in achieving organizational objectives” [32]. COSO’s ERM approach is top-down, spearheaded by management, with risk-focused deliberations occurring at strategic echelons, such as capital allocation discussions. A hallmark of the 2017 framework is its emphasis on corporate governance, organizational culture, and strategy formulation.

3.1.2 Three lines of defense (LoD)

The Institute of Internal Auditors (IIA) introduced the Three Lines of Defense (LoD) model in response to the financial crisis, aiming for a more structured and unified approach to risk management [32]. This model delineates the organization’s roles in managing risk across strategic, tactical, and operational tiers. The “3LoD” risk management model clarifies reporting obligations, metrics, and risk mitigation as follows:

- 1st Line of Defense: Encompasses both financial and non-financial activities that generate risk.
- 2nd Line of Defense: Covers roles responsible for risk oversight that aid in deploying the RMF. This tier also stipulates risk appetite, control measures, and independently supervises the risk management actions of the first line.
- 3rd Line of Defense: Comprises the group audit, ensuring the establishment and maintenance of internal control systems and risk management procedures.

While global standards like COSO serve as foundational frameworks that some companies leverage to construct data science risk management strategies, these standards often overlook specific risks inherent to data science projects [30]. In essence, while these standards provide a foundational scaffold upon which organizations - both private and public - can shape their risk appetite and thresholds, they fall short in addressing distinct risks pivotal to data science projects, which are central culprits of their failures. Nevertheless, a few RMFs have emerged that focus on addressing these data science-centric project risks.

3.2 Data science-specific risk management framework

While the previous SLR conducted in 2019 [10] did not find data science-specific risk management frameworks discussed or evaluated, there are several recent frameworks defined by organizations in the industry dedicated to model risk and AI. Although risks generated from forecasting models and AI might be different than the phenomenon of data science, we treat these concepts under the umbrella of data science.

3.2.1 Model risk management

Model risk management (MRM) has been employed to oversee and track AI model risks, especially within sectors such as financial services. MRM focuses on the management of risks stemming from potential negative outcomes due to decisions informed by flawed or inappropriately used models. The primary goal of MRM is to detect, quantify, and address risks associated with model inaccuracies or inappropriate application.

Nonetheless, recent observations suggest that traditional MRM approaches are becoming less effective due to the following reasons [33]:

- Evaluations under MRM are typically conducted once every one to five years, yet AI models frequently adapt based on ever-changing data. In the intervals between assessments, while MRM presumes model constancy, these models are, in reality, dynamic, evolving with data shifts.
- The review process of MRM can span 6 to 12 weeks post-model development, leading to deployment delays. The linear process of MRM does not seamlessly integrate with agile methodologies.
- MRM traditionally focuses on conventional risks, such as credit risk or capital sufficiency. However, modern challenges like reputational risk, consumer risk, employee-related risk, conduct risk, and biases often remain unaddressed.
- AI solutions are not homogenous. The landscape encompasses varied systems, from chatbots to HR analytics tools. These cutting-edge models differ fundamentally from traditional AI frameworks, like stress-testing or credit-risk models.

3.2.2 NIST artificial intelligence risk management framework 1.0

On January 26th, 2023, in collaboration with various public and private entities, NIST introduced the sector-neutral Artificial Intelligence Risk Management Framework (AI RMF 1.0) along with supplementary materials [34]. This Framework proposed as an elective tool, aims to assist organizations in effectively addressing societal risks posed by artificial intelligence. Moreover, it serves as a preventive measure against risks faced by both individuals and institutions. In other words, given the vast applicability of AI, the Framework extends its reach not just to organizations but also encompasses individual, communal, and societal dimensions. The AI RMF is crafted to bolster credibility and encourage the creation of ethically sound AI systems for both the public and private sectors. Formulated in compliance with the National AI Initiative Act of 2020, recommendations from the National Security Commission on Artificial Intelligence, and the Plan for Federal Engagement in AI Standards and Related Tools, the AI RMF is a consensual structure rooted in collaboration. NIST’s AI Framework characterizes risk as the likelihood and ramifications of adverse outcomes. Echoing ISO 31000:2018’s risk definition, it seeks not only to curtail AI-associated hazards but also to harness potential beneficial outcomes in harmony with societal well-being. It pinpoints three primary risk categories targeting (1) individuals, (2) institutions, and (3) broader ecosystems. These risks range from societal and reputational dangers to environmental threats. The NIST RMF advocates for a comprehensive test, evaluation, verification, and validation (TEVV) approach throughout an artifact’s lifecycle. While the Framework outlines broad risk management principles centered on the collaborative efforts of societal stakeholders, it lacks detailed guidance on navigating various model lifecycle stages - from conceptualization and data handling to deployment and impact assessment. Similarly, while it lists essential AI model attributes like trustworthiness, validity, and transparency, referencing ISO definitions, these descriptions are prescriptive without providing concrete steps to realize an optimal AI system. Future iterations of the NIST RMF plan to incorporate evaluations to gauge its efficacy in enhancing AI risk management [35].

Overall, there are four core functions of AI RMF: **governing** a culture of risk management which sets the control, depending upon the content the risk elements are **mapped**, **measuring,** or assessing the risk after identifying, and finally risks are prioritized and **managed**. The functions are described as high-level activities. Although the functions are claimed as actions with specific outputs, the guidelines are merely considered as checklists that require further research on deployment. For example, the playbook along with the guide to NIST AI RMF fosters the culture of risk management. However, it is not necessary that organizations operate with formal risk management business lines to manage risk. The risk management pillar can be embedded in an umbrella business line without explicitly working as a core function. It is necessary to then describe how the culture of risk management is defined on an organizational level where risk management practices are considered taken for granted. Additionally, social actors might have their intrinsic values or perceptions of defining risk. It is necessary to understand that in many organizations, documentation may not be created or in use for being labor intensive. A proxy measure should be advised where the suggested actions do not apply to the playbook. Another issue of transitioning to NIST AI RMF from the existing RMF or risk management process is not explained. In other words, change management remains an issue.

There are organizations where ad-hoc risk management processes are deployed. This poses the question if NIST AI RMF will apply to immature risk management processes. Also, as the implementation of the Framework lies at the discretion of mature processes, there are guidelines to follow and practices to adhere to, while working on AI projects. The intervention of guidelines and protocols with RMF might question the institutional aversion to acknowledge or bypass the protocol. For example, too many and too stringent guidelines if assigned to the practitioners, might be perceived as a roadblock to the AI innovation initiatives of the institutions. For example, according to Deloitte, the NIST AI risk management framework for artificial intelligence is recommended for identifying and managing the risk 'without hindering innovation with overly restrictive controls' [36]. In other words, the organization does not want to follow the framework at the cost of losing the opportunities that lead to AI innovation.

The categories and subcategories of the Framework call for human-led capabilities, processes, and procedures for deployment. It is still unclear if this Framework is only inclusive for the corporations that cannot afford resources and social actors for successful deployment. Clear guidelines around what constitutes 'third party' software are not defined which makes it difficult to ascertain if there should be a contingency procedure for third-party risk.

Amidst COVID-19, there are many changes around staff lay-offs, changes in the behavior of consumers, budget cuts, and data no longer in use for the model as the user sentiments have changed [37]. There is an ambiguity if the Framework can take the hit of climate risk.

Overall, the Framework is a steppingstone for the thinking process around risk management for artificial intelligence. However, it does not look at data science as a phenomenon that encompasses many more risks and contextual complexities around risk mitigation.

3.2.3 AI risk management framework by McKinsey

In the evolving landscape of AI, McKinsey emphasizes the integration of risk management within the AI innovation process to ensure continuous monitoring and adaptation in line with the fluid nature of data-driven cultures. In essence, the RMF should possess the ability to evolve. A salient challenge with a static framework is the protracted model risk management workflows, often spanning six to 12 weeks post-deployment for review. Such delays could lead to emergent use cases, necessitating another round of MRM workflow. Considering the multifaceted risks—encompassing model, operational, privacy, and reputational dimensions—it becomes intricate for stakeholders to efficiently oversee and coordinate risk. This scenario spurs a pivotal question: should risk management be centralized or decentralized?

To navigate these multifarious risk scenarios, McKinsey introduces an agile-centric risk management approach termed "derisking AI by design" [38]. This approach can be distilled into three primary stages: (1) endorsement for ideation, encompassing PoC/MVP/development; (2) green light for implementation, including data acquisition, evaluation, and model construction; and (3) authorization for deployment, covering inventory evaluation, monitoring, and review.

"AI by design" advocates for early integration of risk management measures during model formulation. By concurrently aligning risk management with model development, practitioners can circumvent prolonged waits post-model creation. Additionally, concurrent checkpoints for risk identification, evaluation, and control checks can yield efficiencies in terms of time, resources, and cost. This integrative approach, embedding controls within the development trajectory, proves especially advantageous for entities boasting advanced risk analytics divisions. Such a setup empowers risk management professionals to confidently embrace agile methodologies, facilitating swifter independent reviews.

To embark on this journey, irrespective of an organization's risk management maturity, senior leaders should undertake the following steps to centrally coordinate the "derisking AI by design" approach:

- Ethical Guideline Assimilation: Leadership should adopt a top-down perspective, critically appraising AI's ethical contours, recognizing guidelines, and understanding potential risks.
- Conceptual Framework Design: Based on established principles, an AI risk management framework should be crafted, encompassing stages from ideation to model review. Establish Governance & Define Roles: Subsequent steps involve identifying suitable professionals for analytics and risk management, ensuring their roles vis-à-vis

AI risk controls are lucidly articulated. Comprehensive training sessions should be organized to facilitate foundational analytics knowledge.
- Embrace Agile Risk Management: Analytics and risk management teams should collaborate, leveraging agile and sprint methodologies, fostering interdependence and efficient conflict resolution.
- Champion Transparency & Explainability: Cultivating a culture prioritizing tools with transparency and explainability features is paramount.
- Promote Awareness: Host awareness campaigns and workshops to familiarize risk and compliance professionals with intricate, risk-laden use cases.

It is pivotal to note that this framework predominantly adopts a technology-centric stance, overshadowing a more comprehensive sociotechnical perspective. Moreover, its design predominantly targets individual models, rather than the broader ambit of data science.

4. SLR for risk management framework for data science projects

4.1 SLR Methodology

Content analysis can be defined as “a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use” [28,30]. The content analysis for this SLR begins with the first step by identifying the research question. The overarching research question as mentioned in section two above is driven by the research questions of SLR in 2019, and the derived findings. The second step is to locate the appropriate data that can answer the research question. The scope of the data used for this SLR is restricted to six repositories – ACM Digital Library, IEEE Explore, Google Scholar, Science Direct, Scopus, and Web of Science. The third step is the random sampling of picking up the text articles based on keywords, titles, abstracts, and summaries that can represent the focused corpus of text on data science risk. The articles are picked up basis the relevance and the search are conducted till the articles are repetitive or irrelevant. The fourth step is to validate and reconcile the identified articles between the primary and secondary authors. The fifth step is to deep dive into the full text of the articles to check the relevance and derive themes. The sixth step is to validate, reconcile, and create a tree structure of the themes derived by both authors. In other words, broader themes are identified, and as a basis for the relationship between the two, a meaningful hierarchy of themes is clustered together. Finally, the frequency distribution of the themes is calculated to create a bar diagram of weaker and stronger themes.

To understand the current research and usage of risk management frameworks when organizations execute data science projects, a systematic literature review was conducted. This SLR was an extension of the SLR conducted in 2020 that focused on exploring the presence of a risk management framework for data science projects [10].

There were seven keyword search themes identified in the previous systematic literature search conducted in 2020 [30]. These same themes and the keyword combinations were taken as a starting point for this updated SLR (Figure 3).

The SLR started with a search of pertinent articles with six repositories – ACM Digital Library, IEEE Explore, Google Scholar, Science Direct, Scopus, and Web of Science. In terms of the search criteria, we restricted the period between 2021 and 2022 till October. Only English-language peer-reviewed conference proceedings, journal articles, and conference papers were added to the search results. The book chapters, theses, and inaccessible articles were discarded. In other words, the search criteria leveraged what was used for the 2019 SLR, except for the time bracket.

Specifically, to source relevant articles for the SLR, all keywords used in the initial SLR in 2019 were used, with one exception: the theme of “What is Big Data/Data Analytics/AI?” This exclusion stemmed from the theme’s emphasis on defining and elucidating the characteristics of big data, data analytics, and artificial intelligence, rather than addressing risk management frameworks for data science projects. A new keyword focusing on machine learning was added owing to the increasing use of that term on titles observed during article searches.

Once the articles were identified, the first step was to identify duplicates (the same article was found multiple times). Then, each paper was reviewed with title, abstract, and conclusion in detail by two reviewers, to determine if the paper should be included in the analysis. Next, via detailed content analysis, done independently by two researchers, each article was designated with a topic. After brainstorming the topics, the mismatch of topics was reconciled. New themes were derived. Throughout the process, regular discussions between the two researchers focused on differences in analysis. In other words, there were active brainstorming meetings to discuss mismatches of topics, which were then reconciled, to map the mutually agreed topics and themes.

Hence, the new literature review was both inductive and deductive. The SLR followed a deductive approach by analyzing if the patterns of the themes that had been uncovered in the previous SLR remained the same. In addition, the analysis adopted an inductive approach, as the authors remained open to identifying new themes without being confined to pre-existing ones found in the previous SLR.

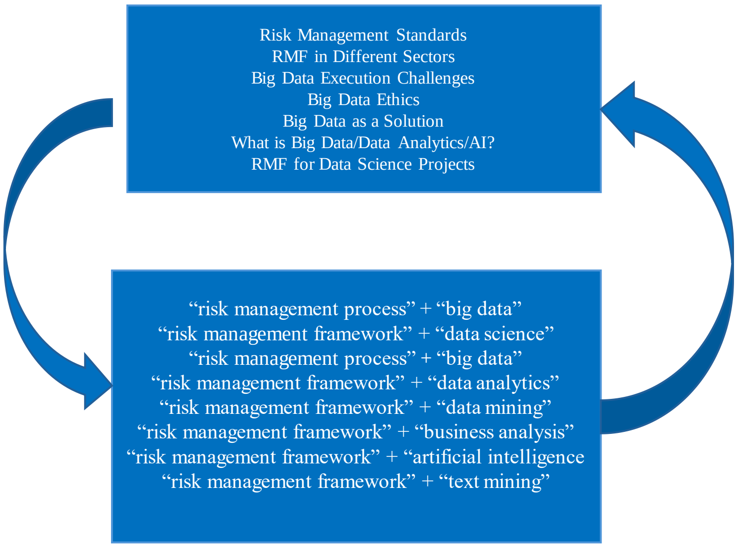
Risk Management Standards
RMF in Different Sectors
Big Data Execution Challenges
Big Data Ethics
Big Data as a Solution
What is Big Data/Data Analytics/AI?
RMF for Data Science Projects
“risk management process” + “big data”
“risk management framework” + “data science”
“risk management process” + “big data”
“risk management framework” + “data analytics”
“risk management framework” + “data mining”
“risk management framework” + “business analysis”
“risk management framework” + “artificial intelligence
“risk management framework” + “text mining”

Fig.3. Themes and Keyword Combinations used for SLR

5. Results and Discussion

5.1 SLR results

There was initially a total of 74 relevant articles identified for this SLR after scanning the title, abstract, keywords, and summary and weeding out duplicate items. The following frequency distribution of the articles was gleaned as per the figure 4. The keyword “Big Data” had the maximum frequency of appearance with the keyword strings risk management framework and risk management process. Out of 74 relevant articles, there were 34 unique articles identified after reconciliation with the secondary author. The 34 articles were then reviewed in depth with full-text analysis, which reduced the total number of articles to 14. All articles were then explored via content analysis, with the following dominant themes identified: *exploring AI Risks, governance for AI, RMF for AI, RMF for using AI, and RMF for Implementing AI in banking*.

As shown in Table 1, the *Need for an RMF specifically for the Big Data Science Projects* theme had the maximum frequency. This theme consisted of all the articles that stated the need for a risk management framework to minimize data science-specific risks (ex. bias, transparency, algorithm fairness). This observation highlighted growing research in risk management for data science projects. Specifically, two out of six published articles advocated for more research in risk management for data science projects. One of the other four articles highlighted the need to have risk-cognizant machine learning systems.

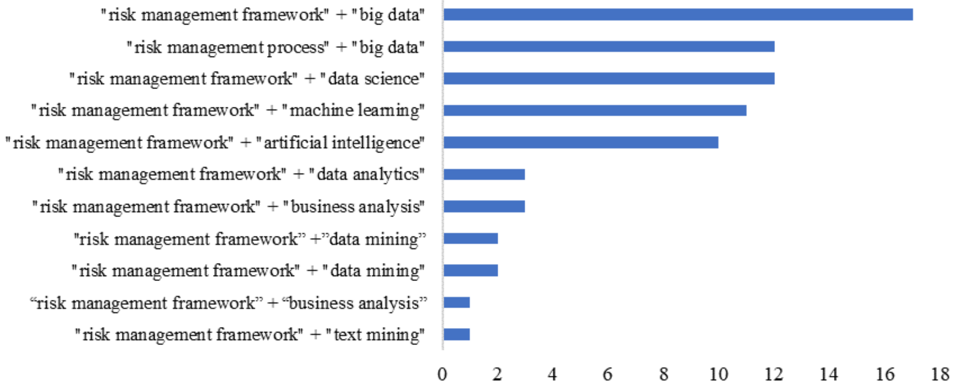

Fig. 4. Frequency Distribution of Keyword Combinations used in SLR

Table 1. Final themes gathered after full-text analysis of 34 unique articles

Keyword strings	**Number of articles obtained with SLR**	**Themes gathered**
“risk management framework" + “big data” “risk management process" + “big data” “risk management framework" + “big data”	4	Big Data Execution Challenges
“risk management framework" + "machine learning” “risk management framework" + "data science” “risk management framework" + "big data” “risk management framework" + "data analytics” “risk management framework" + "data science”	6	Need for an RMF specifically for Big Data Science Projects
“risk management framework” + “machine learning” “risk management framework" + “artificial intelligence” “risk management process"" + “big data”	4	RMF for specifically for Big Data Science Projects

Overall, the theme *Need for an RMF specifically for Big Data Science Projects* in Figure 5 stood out to encourage more research on finding solutions to managing risks for data science projects.

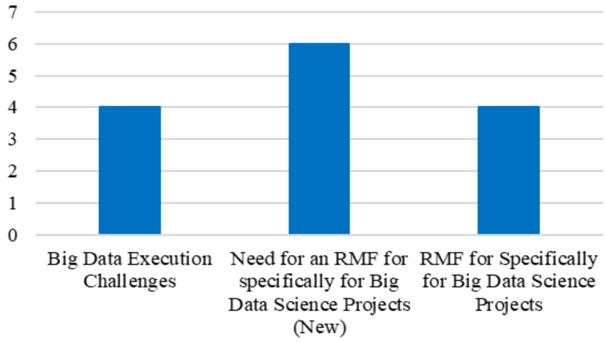

Fig. 5. 14 Articles Gleaned Through SLR in 2022

5.2 Key Insights

While established frameworks like Scrum and CRISP-DM offer methodologies to oversee data science projects, neither explicitly emphasizes risk management. Contemporary literature remains conspicuously silent on addressing specific data science risks through these project management methodologies. Aho et al. [39] noted this gap, and a systematic literature review (SLR) by Krippendorff [30] found a stark absence of standardized risk management frameworks catering to the distinct risks of data science projects.

NIST Risk management framework for AI and McKinsey's AI by design is aligned with AI and models respectively without looking at data science as a phenomenon. NIST AI RMF is a generic list of risk management control measures whereas Mckinsey's AI by design is more use-case specific.

On the other hand, existing risk management frameworks such as COSO do not explicitly integrate with current data science process frameworks. However, as noted by NIST [40, p23], the best RMF implementation is indistinguishable from the routine system development lifecycle processes carried out by organizations. That is, RMF tasks are closely aligned with the ongoing activities in the system development lifecycle processes. Hence, for risk management to be fully integrated into the team's process, the risk management framework must be explicitly integrated with the data science team's process.

Based on the clear potential risk of model bias in building predictive models, one question that arises is, what is the appropriate data science process used by the data science team while developing the artifact, and does that process reduce the risk of bias and fairness and more generally, the potential of harming certain groups or segments of the population?

This process needs to consider that, in the social world, the contemplation of risk emphasizes the mental models of different groups of people who interpret harm and hazard through their worldviews. If the risk reflects the standards set by a specific group of experts (e.g., white male data scientists), then this group might miss project risks that others might more easily identify and mitigate. In other words, due to the multiple views of risk management, different people, different organizations, and different cultures might try to identify and manage risk via different approaches.

In the context of project management for data science projects, the field needs to investigate how risk can occur, who might be responsible for mitigating that risk, how hazardous that risk might be, who might be impacted, and when the risk might be ignored (for example, due to minimal customer impact or compromising ethical risk in place of profits). Hence, research is needed to explore the viability of a risk management framework for data science to address project risks and ethical hazards when teams execute data science projects.

5.3 Potential next steps

Millstone et al. [41] defined three models of governance structure that can be considered for developing project risk management framework for data science projects:

- 'Technocratic' model in which scientists directly inform objectives through policy making. Scientists are the decision-makers who inform the policymakers on what to do.
- The 'Decisionistic' model follows the Red Book model in which the scientific aspects through risk assessment marry the political and value aspects through risk management through the overall process of risk analysis.
- 'Transparent (inclusive) governance' model in which actors from science, politics, economics, and science are invited to make contributions to risk assessment and risk management [42].

A data science project risk management framework should leverage this approach, as it will help identify and mitigate the full range of biases previously discussed. Having said that, due diligence must be done to make sure that not just scientists, but social scientists are involved in deciphering and borrowing the policymaking for risk management framework.

To start, one could explore everyday data science work practices, processes, existing measures of risk mitigation, and gold standards deployed in data science projects to manage risk. Furthermore, one could explore both the formal and informal work of data scientists that would help me locate vulnerable areas of risk and plausible solutions. One could also look at how data scientists participate in visible and taken-for-granted work to manage risk embedded in the day-to-day lives at the workplace. The goal would be to explore specific activities that are aligned with minimizing risk.

Future research could leverage specific milestones or structured approaches with incremental stages to explore a data science project risk management framework that carries not only the objectives of science but also the subjectivities of sensory experiences of civil society. Some other potential next steps are to create surveys to explore the thoughts of data scientists on project risk management frameworks for data science projects and to conduct ethnography studies to get more detailed insights.

5.4 Limitations

One of the limitations of the SLR was the usage of restricted keyword strings for deriving the search results. The extended SLR in 2022 was based on the keyword strings used in the earlier SLR study conducted in 2020. A broader scope of keyword search strings might have disclosed a wider scope for articles and diverse themes. Another limitation was the selection of time for considering the literature for the SLR. The SLR is from 2019 till 2022 which does not include the focused literature for 2023. The SLR also had a restricted set of sources to search for articles, this circumvented the study to only six repositories – ACM Digital Library, IEEE Explore, Google Scholar, Science Direct, Scopus, and Web of Science. As data science is a phenomenon that encapsulates machine learning, artificial intelligence, data and business analytics, and text mining, there is also a likely debate on how inclusive the risk management framework should be, which this SLR does not cover.

6. Conclusion

In essence, while data scientists should proactively consider risk management strategies for successful project outcomes, there remains an evident gap in frameworks that facilitate teams in identifying and addressing these risks. Consequently, focused exploration into project management methodologies tailored to pinpoint risk elements specific to data science projects is warranted. This imperative distinguishes itself from the emphasis in Saltz's 2015 study [4], which highlighted the necessity for well-defined data science processes and roles. Saltz's study addressed the prevailing trial-and-error or ad-hoc approaches adopted by practitioners and delved into recurrent challenges like data quality and source uncertainties. The current emphasis pivots towards addressing latent vulnerabilities in data science practices, which may inadvertently introduce unmanaged risks. The envisioned risk management structure constitutes a governance model specifically tailored for data science projects, encompassing comprehensive risk assessment, management, and communication.

The broader objective for the domain should be to analyze the multifaceted and global character of data science endeavors steered by process methodologies. The end goal is to empower teams to judiciously evaluate data science-associated risks and pitfalls, culminating in the formulation of an adaptive Data Science Ethical Collaborative Project Risk Management Framework (DS EthiCo RMF). DS EthiCo RMF is an agile-driven enterprise-level risk management framework that manages the risks of data science on a project level. As the project can be executed by multiple stakeholders, there also lies the possibility of the practitioners being co-located or sparsely distributed across multiple locations. Also, the requirements of the external stakeholders or clients are subject to change for data science projects, DS EthiCo RMF is agile driven to bring effective communication and flexibility. This framework, catering to a diverse and global audience, marries traditional project management methodologies with an ethical risk paradigm. The ethical dimension of this framework mandates active collaboration among data scientists, project managers, social scientists, and a varied user base. Summarily, the pivotal research inquiry to be addressed is:

How can we devise an adaptive risk management framework that seamlessly integrates ethical principles and project methodologies, tailored explicitly for geographically dispersed data science initiatives?

References

[1] T. Aho, O. Sievi-Korte, T. Kilamo, S. Yaman, and T. Mikkonen, “Demystifying data science projects: A look on the people and process of data science today,” in *Product-Focused Software Process Improvement: 21st International Conference*, PROFES 2020, Turin, Italy, 2020, pp. 153-167.

[2] S. Passi and P. Sengers, “Making data science systems work”. *Big Data & Society*, vol. 7, no. 2, p. 2053951720939605, 2020.

[3] E. Parmiggiani, T. Østerlie, and P. G. Almklov, “In the backrooms of data science,” *Journal of the Association for Information Systems*, vol. 23, no. 1, pp. 139-164, 2022.

[4] J. S. Saltz, “The need for new processes, methodologies and tools to support big data teams and improve big data project effectiveness,” *in 2015 IEEE International Conference on Big Data (Big Data)*, 2015, pp. 2066-2071.

[5] S. Passi and S. J. Jackson, “Trust in data science: Collaboration, translation, and accountability in corporate data science projects,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, CSCW, pp. 1-28, 2018.

[6] S. Passi and S. Barocas, “Problem formulation and fairness,” in *Proceedings of the conference on fairness, accountability, and transparency*, 2019, pp. 39-48.

[7] G. Neff, A. Tanweer, B. Fiore-Gartland, and L. Osburn, “Critique and contribute: A practice-based framework for improving critical data studies and data science,” *Big Data*, vol. 5, no. 2, pp. 85-97, 2017.

[8] R.Kitchin and, T.Lauriault, “Towards critical data studies: Charting and unpacking data assemblages and their work,”. *In The programmable city working paper*. Retrieved November 11, 2020, from https://paper s.ssrn.com/sol3/paper s.cfm?abstract_id = 24741 12 2014.

[9] S. T. Lai and F. Y. Leu, “A critical quality measurement model for managing and controlling big data project risks,” in Advances on Broad-Band Wireless Computing, Communication and Applications: *Proceedings of the 12th International Conference on Broad-Band Wireless Computing, Communication and Applications (BWCCA-2017)*, 2018, pp. 777-787.

[10] J. S. Saltz and S. Lahiri, “The Need for an Enterprise Risk Management Framework for Big Data Science Projects,” in *DATA*, 2020, pp. 268-274.

[11] S. Cicmil, D. Hodgson, M. Lindgren, and J. Packendorff, “Project management behind the façade,” *Ephemera: theory and politics in organization*, vol. 9, no. 2, pp. 78-92, 2009.

[12] J. McManus and T. Wood-Harper, “Understanding the sources of information systems project failure – a study in IS Project Failure,” *Management Services*, vol. 51, no. 3, pp. 38–43, 2007.

[13] S. Lahiri and J. Saltz, “Evaluating Data Science Project Agility by Exploring Process Frameworks Used by Data Science Teams,” *In 56th Annual Hawaii International Conference on System Sciences*, HICSS 2023 (pp. 6538-6547). IEEE Computer Society. 2023.

[14] S. Lahiri and J. Saltz, “The Need for an Enhanced Process Methodology for Ethical Data Science Projects,” in *2023 IEEE International Symposium on Ethics in Engineering, Science, and Technology (ETHICS)*, 2023, pp. 01-05.

[15] P. Cerchiello and P. Giudici, “Big data analysis for financial risk management,” *Journal of Big Data*, vol. 3, no. 1, pp. 1-12, 2016.

[16] VentureBeats, “Why do 87% of data science projects never make it into production?” [Online]. Available: https://venturebeat.com/ai/why-do-87-of-data-science-projects-never-make-it-into-production/, Accessed on: Jul. 17, 2023.

[17] T. Ermakova, J. Blume, B. Fabian, E. Fomenko, M. Berlin, and M. Hauswirth, “Beyond the hype: why do data-driven projects fail?” in *Proceedings of the 54th Hawaii International Conference on System Sciences*, p. 5081. 2021.

[18] J. Westenberger, K. Schuler, and D. Schlegel, “Failure of AI projects: understanding the critical factors,” *Procedia computer science*, vol. 196, pp. 69-76, 2022.

[19] C. Varela and L. Domingues, “Risks of Data Science Projects-A Delphi Study,” *Procedia Computer Science*, vol. 196, pp. 982-989, 2022.

[20] PMI, “What is Project Management?” [Online]. Available: https://www.pmi.org/about/learn-about-pmi/what-is-project-management, Accessed on: Mar. 5, 2023.

[21] J. Chen, Y. Chen, X. Du, C. Li, J. Lu, S. Zhao, and X. Zhou, “Big data challenge: a data management perspective,” *Frontiers of computer Science*, vol. 7, pp. 157-164, 2013.

[22] Association for Project Management, “What is Risk Management?” [Online]. Available: https://www.apm.org.uk/resources/what-is-project-management/what-is-risk-management/, Accessed on: Sep. 23, 2023.

[23] P. Weaver, “The meaning of risk in an uncertain world,” [Online]. Available: https://www.pmi.org/learning/library/project-risks-uncertain-world-8392, Accessed on: Sep. 23, 2023.

[24] E. Sengupta, D. Garg, T. Choudhury, and A. Aggarwal, “Techniques to elimenate human bias in machine learning,” in *2018 International Conference on System Modeling & Advancement in Research Trends (SMART),* 2018, pp. 226-230.

[25] G. Bensinger, “Google Redraws the Borders on Maps depending on Who’s Looking,” Washington Post, [Online]. Available: https://www.washingtonpost.com/technology/2020/02/14/google-maps-political-borders/, Accessed on: Sep. 23, 2023.

[26] A. Damodaran, Models of Risk and Return, *New York University Stern School of Business*, 2000.

[27] M. Abdar et al., “A review of uncertainty quantification in deep learning: Techniques, applications and challenges,” *Information Fusion*, vol. 76, pp. 243-297, 2021.

[28] White, Marilyn Domas, and Emily E. Marsh. “Content analysis: A flexible methodology.” *Library trends* 55, no. 1 (2006): 22-45.

[29] N. W. Grady, J. A. Payne, and H. Parker, “Agile big data analytics: AnalyticsOps for data science,” in *2017 IEEE international conference on big data (big data),* 2017, pp. 2331-2339.

[30] K. Krippendorff. *Content analysis: An introduction to its methodology*. Sage publications, 2018.

[31] K. Prewett and A. Terry, “COSO's Updated Enterprise Risk Management Framework—A Quest For Depth And Clarity,” *Journal of Corporate Accounting & Finance*, vol. 29, no. 3, pp. 16-23, 2018.

[32] T. Thabit, “Determining the effectiveness of internal controls in enterprise risk management based on COSO recommendations,” in *International Conference on Accounting, Business Economics and Politics*, 2019.

[33] J. A. Baquero, R. Burkhardt, A. Govindarajan, and T. Wallace, “Derisking AI by design: How to build risk management into AI development,” *McKinsey & Company*, 2020.

[34] “Artificial Intelligence Risk Management Framework (AI RMF 1.0),” [Online]. Available: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf, Accessed on: Mar. 5, 2023.

[35] “AI Risk Management Framework,” [Online]. Available: https://www.nist.gov/itl/ai-risk-management-framework, Accessed on: Mar. 5, 2023.

[36] “What you need to know about NIST’s AI Risk Management Framework published in January 2023,” [Online]. Available: https://ukfinancialservicesinsights.deloitte.com/post/102i6q5/what-you-need-to-know-about-nists-ai-risk-management-framework-published-in-janu, Accessed on: Mar. 5, 2023.

[37] QuantumBlack, “Machine Learning Models: Mitigating the Impact of COVID-19,” [Online]. Available: https://medium.com/quantumblack/how-to-continue-trusting-in-machine-learning-models-post-covid-19-5c5fd53fb83b, Accessed on: Mar. 5, 2023.

[38] J. A. Baquero, “Derisking AI by design: How to build risk management into AI development,” [Online]. Available: https://www.mckinsey.com/capabilities/quantumblack/our-insights/derisking-ai-by-design-how-to-build-risk-management-into-ai-development, Accessed on: Mar. 5, 2023.

[39] T. Aho, O. Sievi-Korte, T. Kilamo, S. Yaman, and T. Mikkonen, “Demystifying data science projects: A look on the people and process of data science today,” in *International Conference on product-focused software process improvement*, 2020, pp. 153-167.

[40] J. T. Force, “Risk management framework for information systems and organizations,” NIST Special Publication, vol. 800, p. 37, 2018.

[41] E. Millstone, P. van Zwanenberg, C. Marris, L. Levidow, and H. Torgersen, "Science in Trade Disputes Related to Potential Risks: Comparative Case Studies," European Science and Technology Observatory, Institute for Prospective Technological Studies, Seville, 2004. [Online]. Available: http://ftp.jrc.es/eur21301en.pdf.

[42] O. Renn, Risk governance: coping with uncertainty in a complex world, *Routledge*, 2017.

[43] D. Schlegel, K. Schuler, and J. Westenberger, “Failure factors of AI projects: results from expert interviews,” *International Journal of Information Systems and Project Management: IJISPM*, vol. 11, no. 3, pp. 25-40, 2023.

[44] C. P. Yu and W. Y. Lin, “Risks associated with the development process of in-house information system projects,” *International Journal of Information Systems and Project Management*, vol. 10, no. 2, pp. 66-78, 2022.

[43] Schlegel, D., Schuler, K., & Westenberger, J. (2023). Failure factors of AI projects: results from expert interviews. *International journal of information systems and project management: IJISPM*, *11*(3), 25-40.

[44] Yu, C. P., & Lin, W. Y. (2022). Risks associated with the development process of in-house information system projects. *International Journal of Information Systems and Project Management*, *10*(2), 66-78.

Biographical notes

Sucheta Lahiri

Sucheta Lahiri is a fifth-year PhD candidate at the School of Information Studies at Syracuse University. Her research focuses on understanding the risks faced by Global South data science practitioners during the execution of projects. Lahiri's work aims to shed light on how specific social agents legitimize and de-legitimize the risk of data science failure, employing methods such as persuasion, coercion, and leveraging power imbalances. Through her research, she seeks to develop a ground-up risk management framework that can effectively manage the risks of data science project failures and foster a more democratic work culture. Lahiri has 12 years of experience in the industry in India. She headed a risk management team in India before joining Syracuse.

Jeffrey Saltz

Jeffrey Saltz is an Associate Professor at Syracuse University, where he leads their graduate applied data science program. His research focuses on agile data science project management. Prior to joining Syracuse, Jeff reported to the global CIO at JPMorgan Chase, where he drove technology innovation across the bank. His previous roles at the bank included CTO consumer risk, Head of Risk Core Processing (within Chase Card Services) and the Chief Information Architect (across the consumer bank). Jeff holds a B.S. in computer science from Cornell University, an M.B.A. from the Wharton School at the University of Pennsylvania and a Ph.D. in information systems from the New Jersey Institute of Technology.