The Princeton 20

Assessing and Reducing Risks for AI Projects

PRINCETON CONSULTANTS

Information Technology and Management Consulting

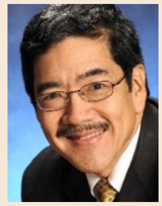
Executive Summary

Princeton Consultants has had the privilege of working with innovators across a wide range of industries to improve their decision-making using advanced analytics. For over 40 years, this work has worn many labels. What has historically been called expert systems, decision support, and operations research is today often simply called Artificial Intelligence (AI). During this time the core techniques have vastly expanded – from bases in statistics, simulation, and mathematical optimization, to include machine learning and new classes of hybrid algorithms. These advances have been multiplied by the exponential improvements in supporting hardware.

Yet, while the business press has focused on each fancy new label and academia has focused on improvements in techniques, we have found that there are consistent and repeatable factors that govern whether a given analytics project will make it into successful production. These factors also apply to the latest AI projects.

We have codified these learnings into what we call the "Princeton 20" – a list of 10 environmental and 10 technical risk factors that our experience has shown can be used to anticipate issues and by explicitly addressing risk areas, increase the likelihood of successful transition from promising research to production deployment and wide usage.

Contact

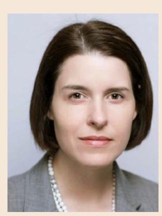


Steve Sashihara
CEO
ssashihara@princeton.com
609-987-8787 x214

Steve founded Princeton Consultants in 1981 to help organizations integrate optimization and new technologies into core business processes. Steve is the author of *The Optimization Edge* (McGraw-Hill), the first non-technical book to explain optimization to the busy business executive.

Dr. Irv Lustig
Optimization Principal
irv@princeton.com
609-987-8787 x148

An IBM/ILOG/CPLEX executive from 1993-2014, Irv led or managed technical services, sales and product development, and investigated forecasting and optimization integration at IBM Research. Irv is an award-winning mathematical programmer and a former professor at Princeton University.



Dr. Patricia Randall
Director
prandall@princeton.com
609-987-8787 x318

Patricia joined Princeton Consultants in 2007. She specializes in the design, development, and implementation of large-scale, high impact systems that help businesses optimize their decision-making at the strategic, tactical, and operational levels.

1. Overview

According to the Gartner IT Glossary (2020), “Artificial intelligence (AI) applies advanced analysis and logic-based techniques, including machine learning, to interpret events, support and automate decisions, and take actions.” This definition includes statistical forecasting, simulation, and optimization techniques that have been used in the operations research community for decades, as well as the more recent terms “advanced analytics” and “data science.” In its report on Data Science and Machine Learning (DSML) platforms, Gartner (2020) writes: “AI is overhyped and inescapable: All DSML can be classified as artificial intelligence (AI), but not all AI concepts should be called DSML.” Today, we are seeing an increase in promising early stage AI research in virtually every large business.

Despite this popularity and promise, most AI projects have failed to get into production environments. O’Neill (2019) summarizes findings from various sources indicating failure rates ranging from 74% to 87%, where failure consists of projects never making it into production or that business adoption of the AI initiatives does not occur. O’Neill’s citations include:

- July 2019: *VentureBeat AI* reports 87% of data science projects never make it into production.
- Jan 2019: *NewVantage* survey reports 77% of businesses report that “business adoption” of big data and AI initiatives continues to represent a big challenge for business. That means 3/4 of the software being built is apparently collecting dust. Ouch.
- Jan 2019: *Gartner* says 80% of analytics insights will not deliver business outcomes through 2022 and 80% of AI projects will “remain alchemy, run by wizards” through 2020.

Sadly, this is not a recent phenomenon: progressing from research study to production has plagued advanced analytics and operations research projects since the 1960’s.

Our firm, Princeton Consultants, has led and participated in successful AI projects spanning back to the early 1980s. During that time, our firm has built a framework that, in our experience, reveals 20 risk factors that have challenged or sunk these projects when not addressed, and been responsible for propelling enthusiastic adoption and positive business outcomes when properly covered. The Princeton 20 is a framework consisting of 10 environmental and 10 technical risk factors to evaluate and manage any potential AI project. For each factor, a score of +1, 0, or -1 is given, with a positive score indicating that there is high risk associated with that factor, 0 indicating neutral risk, and -1 indicating low risk. For each of the factors that have a high risk, we explicitly discuss and implement proven risk mitigation techniques to shore up each identified risk area. As opposed to applying weights to the factors, we treat all the factors equally in our project risk assessment and use the individual scores to identify the areas that require the most focus. Throughout the life of the project, all 20 factors are monitored, regardless of their initial score at project initiation.

THE PRINCETON 20

20 Scoreable Risk Factors that describe the core challenges of bringing AI from the Exploration phase into fully deployed Production

E10

20 ENVIRONMENTAL FACTORS

1. Sponsoring Business
2. Sponsoring Exec
3. Business Process
4. Decision Support Culture
5. Decision Scope
6. KPIs
7. Value Proposition
8. Resource Availability
9. Other Initiatives
10. User Engagement

T10

20 TECHNICAL FACTORS

1. Existing AI Systems
2. Application Novelty
3. New Technologies
4. AI Decision Scoring
5. Business Rules
6. Decision Data
7. Software Tools
8. Speed
9. UX
10. AI Process Maturity

In this paper, we detail each of these 20 risk factors and offer some mitigation techniques to manage them. The 10 environmental risk factors relate to the business conditions that increase or decrease the chance of success. The 10 technical risk factors relate to technology and related issues. For each factor, we will list the three characteristics that determine how the factor can be scored to measure the project risk and the mitigation strategies we use to reduce the risk. Understanding both sides – the environmental factors and the technical factors – is critical for project success.

2. Environmental Factors

Environmental risk factors arise in the business context. We believe that many of the reported AI project failures are due to insufficient attention to these environmental risk factors at the beginning of a project. Too many AI and Data Science projects devote almost all their attention to the technical side – the “if you build it, they will come” fallacy. The new breed of “data scientist” will often focus on new techniques for analyzing data—without proper consideration of the potential business value, the decisions that will be made, or how the project will be deployed and used to assist in making business decisions. The environmental factors address these issues.

E1. Sponsoring Business

This first environmental factor looks at the characteristics of the business that is sponsoring the project.

Score	Characteristic
-1	**Successful and growing**
0	**Consolidating or under-performing**
+1	**Startup or new business model**

Table E1: Sponsoring Business

A successful and growing business is generally a boon to a project, whereas applying AI to create a new business model or to a startup business has greater risk due to other external factors that may impact project success. We mitigate this risk by 1) carefully selecting the scope of the project and managing the scope of the project as a whole; 2) considering the organizational impact of centralized expertise versus resources embedded within teams; and 3) when speaking to management, drawing comparisons with similar projects that have a track record of tangible returns on investment (ROI).

E2. Sponsoring Executive

This environmental factor looks at the line of business executive responsible for the initiative.

Score	Characteristic
-1	Core General Manager of the line of business
0	Senior Manager for line business users
+1	Non line business exec (e.g. head of IT, R&D, Analytics)

Table E2: Sponsoring Exec

We have found that sponsorship driven from within the business unit usually outperforms sponsorship from external areas. Users will look up their chain of command to whether and how to change their behaviors. Also, in a world of ever shifting and competing priorities, the higher the sponsorship the less likely project resources will get pulled for other initiatives. When we are in situations where, instead of a business executive, a technical leader is driving the initiative, we mitigate the risk by 1) forming a steering committee, including people from the line of business who will be impacted by the deployment of AI, with regular reporting on the progress of the project; 2) creating explicit project tasks in our project plans to communicate and socialize progress on the project; and, 3) connecting with other transformation advocates in the business.

E3. Business Process

This environmental factor looks at the characteristics of the business process that will be affected by deployment of an AI system.

Score	Characteristic
-1	Business process is well-established / consistent, and new AI system will not represent a significant change
0	Business process is consistent, but AI system will represent a significant change
+1	In scope is to establish a brand-new standardized business process based on AI Table

E3: Business Process

In cases where there is an established business process and the new AI system will help improve that process, there is less risk than in cases where the AI system is being added, and the business process is changing at the same time. When AI is creating a new business process, success is linked to both the successful AI implementation and the required business transformation. The latter requires not just quants with high technical skills, but savvy people-oriented communicators with high emotional intelligence. We mitigate these risks by 1) including experienced change management resources as part of the core team; 2) creating RACI charts that indicate who is Responsible, Accountable, Consulted and Informed for various steps in the project plan; 3) using OKRs (Objectives and Key Results) as described by Doerr (2018) to align project participants towards success; and, 4) performing an up-front analysis of whether the AI system can support user performance metrics that will measure how well users are leveraging the AI system.

E4. Decision Support Culture

The next environmental factor relates to how well the organization has adapted to using AI systems to support decision making.

Score	Characteristic
-1	**AI-based decision making is used successfully for other adjacent decisions**
0	**User organization successfully uses AI-based decision making, but not for adjacent decisions**
+1	**This project will introduce AI-based decision making to this part of the organization**

Table E4: Decision Support Culture

When organizations are experienced and comfortable with using AI (broadly defined) to support making decisions, each new AI project is nurtured by existing organizational strengths – especially over the sometimes-rocky start. Conversely, when AI based decision making represents a new approach for an organization, the business has little experience navigating the superpowers and Achilles' heels of modern AI. We mitigate these risks by 1) paying careful attention to human-in-the-loop best practices; 2) including real-time data validation and cleaning into the AI system's user interface; 3) rewarding user champions; and, 4) telling stories about innovators in similar as well as other industries. The human-in-the-loop best practices make sure that we create systems with end users in mind by having those end users involved from the beginning in how they will interact with the AI system that we propose to build. As Sashihara (2011) describes, we always pay attention to the cultural patterns and the psychology of human decision making as these cultural aspects influence how well the organization can adapt to AI-based decision making.

E5. Decision Scope

The scope of decisions to be supported by an AI system is important to define for the success of the project. Taylor (2019) describes how any AI project must start with the definition of the decisions that will be impacted by AI.

Score	Characteristic
-1	**Decision scope is well defined, very crisp**
0	**Decision scope is generally defined, but has fuzzy areas**
+1	**Establishing the scope of decisions is required as part of the project**

Table E5: Decision Scope

When the decision scope is well defined, the impact of an AI system is easier to assess. When the project requires the determination of which decisions can be impacted by AI, there are risks associated with how well the organization will be able to transform its current business practices. AI which is too narrowly applied can fail to challenge the status quo of decision making, whereas AI that is too broadly defined may be impossible to effect changes in decision making processes. We mitigate these risks by 1) starting first with self-acknowledged pain areas and scenarios; 2) using agile development and deployment methodologies, including best DevOps (Kim, et. al. 2016) and DataOps (Bergh, et. al. 2019) practices; and, 3) creating a mid-term and long-term road map for how decision making can be transformed via AI. Focusing on the decisions to be impacted is critical for AI project success.

E6. KPIs

Key Performance Indicators (KPIs) are quantitative measures that evaluate the effects of the decisions that an AI system will impact. Importantly, these KPIs may measure effects that are not explicit in the AI system. For example, an AI system may explicitly reduce transportation costs and bring an additional benefit of reducing carbon dioxide (CO2) emissions.

Score	Characteristic
-1	**KPIs around the decisions have been recorded in sufficient detail to fairly evaluate the new system**
0	**The KPIs around the decisions have not been captured, but the actual decision has been recorded and can be easily re-simulated**
+1	**The organization does not have complete KPIs evaluating the decision and lacks the ability to replay and simulate decisions**

Table E6: KPIs

The benefits of an AI system can be measured in multiple ways. KPIs could be related to improved revenue or profits or lowering of costs. Other KPIs might quantify service levels. If the business has recorded the KPIs for the decisions that an AI system will replace, they can be used to benchmark the benefits of a new AI system. However, if these decisions have not been recorded, then it is necessary to establish a baseline set of KPI measures in order to be able to measure the benefits of AI. To mitigate the risks of a lack of established KPIs, we recommend 1) agreeing upon a quantitative experimental design; 2) using simulation and deriving insights from machine learning on past data representing past decisions; 3) engaging subject matter experts (SMEs) for their theories and insights about the decision process; and, 4) conducting thoughtful workshops to determine the right KPIs for the decision process.

E7. Value Proposition

The seventh environmental factor is the value proposition of AI. We believe many AI projects fail because the value propositions are not well articulated before the projects begin. While the value proposition and the KPIs almost always overlap, often the value proposition represents larger, and more strategic goals, such as resiliency, consistency, or nimbleness.

Score	Characteristic
-1	**Value proposition is well understood and quantifiable**
0	**Value is generally understood, but is hard to quantify**
+1	**In scope is to determine the value proposition in order to proceed**

Table E7: Value Proposition

A well-articulated value proposition makes it easier to find sponsorship and support for a project. Conversely, if the value proposition is not broadly accepted, it is necessary to establish that value before beginning. We mitigate the lack of a value proposition by 1) conducting a formal cost-benefit analysis (CBA); 2) searching for connections to other strategic goals that are currently well defined; and, 3) establishing a “contract” with the users where it is understood a priori that the benefits of the AI system will be measured, and that the support of the users is necessary to do that measurement.

E8. Resource Availability

A successful AI project requires participation of key stakeholders and other business resources that often span organizational boundaries.

Score	Characteristic
-1	**This is a key initiative, and the relevant personnel (execs, business SMEs, information technology (IT), Finance, human resources (HR)) are committed to providing the appropriate resources for the whole lifecycle of development, deployment, and production support**
0	**Some resourcing will be tight, but this is considered a key initiative and the organization is not overly committed to other initiatives**
+1	**This project will likely battle for key scarce resources or attention**

Table E8: Resource Availability

A successful AI project requires a commitment from the business to dedicate resources from the line of business that will be impacted by the new AI system and from many other divisions in the organization. We mitigate the risks of insufficient resources by 1) socializing and broadcasting the CBA; 2) giving high attention to the sponsoring executive; 3) creating a steering committee with project managers to identify critical resources; and, 4) building contingent resources into the project plan from the start.

E9. Other Business Initiatives

Other corporate initiatives may complement or conflict with the implementation of an AI system.

Score	Characteristic
-1	**The AI system will stand alone or be harvesting other already successful initiatives**
0	**The AI system is dependent upon or gated by other initiatives, but crisply scoped and resourced**
+1	**There are multiple overlapping or gapped initiatives with weak coordination**

Table E9: Other Initiatives

If the AI system is independent of other initiatives and/or can leverage past successful initiatives, it has a greater chance of success. On the other hand, other overlapping initiatives may starve necessary resources or create Catch-22s for deploying the new AI system. We mitigate these risks by 1) placing the most skilled project managers as core members of the team; 2) reducing and unlinking initiative dependencies where feasible; and, 3) providing fully transparent weekly status reports. Leveraging best project management practices and providing visibility to the progress of the AI initiative helps ensure success.

E10. User Engagement

The last environmental factor relates to identifying the end users of the AI system and understanding their current and future levels of engagement.

Score	Characteristic
-1	**Users are all employees or vendors under the sponsor's control**
0	**Users are not directly under the sponsor's control, but well understood**
+1	**Users of the system are indirectly involved, and experimentation will need to be performed to measure the impact on their engagement**

Table E10: User Engagement

When users are directly related to the project sponsor, it is easier to obtain their engagement and feedback. If the users are unknown or not part of the project team, there is risk that they will reject the new system once it is deployed. We mitigate this risk by 1) placing high focus on the AI system's user experience that mirrors best practices; 2) using iterative agile deployment techniques across multiple representative user groups; and, 3) analyzing demographic or situational factors that correlate with engagement.

3. Technical Factors

Technical factors relate to the software, methods, and IT infrastructure needed to implement the new AI system. Far too many businesses have been "sold" AI as almost a miracle cure-all. This section walks through 10 risk factors on the technical side and appropriate mitigation strategies.

T1. Existing AI System

The first technical factor evaluates whether the existing system uses AI for making the decisions defined in the decision scope (E5).

Score	Characteristic
-1	**The new AI system will improve/replace a system that is fully used today, and that system is fully documented**
0	**The new AI system will improve/replace a system that is used often, but not fully**
+1	**The scope is to create an AI system where none exists today**

Table T1: Existing AI System

If the incumbent is a heavily used AI system and the goal is to improve or replace it, this greatly simplifies the evaluation criteria for any new AI system. However, if AI usage has been inconsistent or non-existent, this calls for a careful examination of why AI has been less than successful. Technical tactics to help mitigate this risk include: 1) insisting that the AI user interface completely captures user behavior by recording acceptances and rejections of recommended decisions; 2) focusing on when users override decisions; 3) recording times when users engage the system and determining usage gaps; and, 4) focusing on improved robustness due to data failures.

T2. Application Novelty

The second technical factor relates to the novelty of the application. An application of AI in a new problem domain carries with it a significant risk.

Score	Characteristic
-1	AI systems like this application are widely available elsewhere
0	AI systems somewhat like this application are available, but this is a highly customized version
+1	This AI system will be a major innovation for decision making

Table T2: Application Novelty

It has been noted that in most corporations "success has many fathers, but failure is an orphan." If the AI techniques have been applied elsewhere in similar applications, that minimizes project risk, because success in the domain has been previously demonstrated and it encourages the organization to get behind what is likely to be a win. Conversely, if the AI techniques are relatively uncommon or unproven in the application domain, there is obviously more risk. We mitigate these risks by 1) finding legitimate and appropriate examples of successful AI that while not in exactly the same circumstances, fit many of the same and hardest characteristics of this problem, 2) decomposing the decision process into micro-decisions and providing a series of small special purpose/situation AI tools; and, 3) considering a broadening of the sponsorship to include other business units within the enterprise. By including other business units, it may be possible to identify other opportunities for AI projects with a higher likelihood of success.

T3. New Technologies

The third technical factor considers emerging technologies. New AI techniques and interdependent technologies (e.g. IoT, edge computing, new computation devices) are being rapidly introduced into the market, both from open source and commercial vendors.

Score	Characteristic
-1	**Does not use emerging or novel technology**
0	**Project may examine use of some emerging technology, but success is in no way dependent upon it**
+1	**In scope is to assess feasibility of a new technology**

Table T3: New Technologies

If the base AI project can be implemented using known and proven technologies, there is less risk compared to when there is a desire to evaluate whether a new technology is appropriate to address the business problem. We mitigate the risk by 1) using a Minimal Viable Product (MVP) approach to demonstrate business value; 2) consulting with non-competitive external groups to understand the benefits and limitations of the new technology; and, 3) investigating as backup the usage of components with tiers of different maturities as opposed to just using the latest state-of-the-art. In our experience, it takes a fair bit of time until a new technology has enough proof points and documentation of success until it is mature enough to be used in new applications by non-experts. Typically, the latest machine learning technologies require the expertise of the developers who created that technology in order to be successful, so we try to avoid using those technologies until the community of users is large enough to support our efforts.

T4. AI Decision Scoring

It is critical to be able to easily compute the value or score of a decision. In mathematical optimization, this equates to easily stating the objective function. In machine learning applications, this could correspond to measures such as , lift or classification measures such as precision, accuracy and the F-score. It is important to differentiate scores that are outputs of an AI system from scores that come from the deployment of that system. For example, once an AI system is deployed, one might measure how well users are adhering to the recommendations of the system. This is not a direct output of the AI system, but a result of the AI system being deployed.

Score	Characteristic
-1	AI Decision Scoring is well defined and easy to code
0	Scoring AI decisions is generally understood, but not completely crisp and may involve trade-offs between competing KPIs
+1	In scope is to create numerical scoring procedures for the first time

Table T4: Decision Scoring

Sometimes, the scoring of decisions is straightforward. In business there are often competing KPIs that need to be individually scored and then weighted to determine a final score for a decision process. When the scores are not well-defined, we mitigate risk by 1) conducting thoughtful workshops around scoring AI decisions; 2) using simulation techniques based on the idea of “digital twins” to guide the choice of a proper scoring mechanism; and, 3) using human-in-the-loop best practices, where we avoid black boxes and visualize the scores for the end users. One technique is to build a user interface that shows the scores of the current decision process that is not based on AI. Users then see how the scores of their process compare to the process that will be driven by AI.

T5. Business Rules

Business rules are the description of the boundaries for the decisions the AI system is being asked to operate within. Business rules are often a combination of physical reality, corporate policies, common industry practices, and regulations.

Score	Characteristic
-1	Business rules are well understood, static, and crisply numeric
0	Rules are generally static, but are relaxed in different circumstances
+1	In scope is to create consensus and codify rules

Table T5: Business Rules

When the business rules are well understood and can be numerically represented, there is little risk to building an effective AI system. However, if the rules are not well understood, then the project plan will need to include time to interview SMEs to better understand the rules used to operate the business that might constrain decisions. We mitigate this risk by 1) collecting a body of solved "hard cases" in a test suite; 2) preferring machine learning techniques to uncover the rules rather than relying 100% on SMEs; and, 3) building the user interface to allow users to understand the rules and to change the rules on the fly, although this can be a difficult implementation challenge.

T6. Decision Data

AI systems are data hungry, and a significant amount of time can be spent in making sure that the proper data is available and clean for the AI system. This technical factor looks at this critical requirement.

Score	Characteristic
-1	**All the data that drive the status quo decision process are known, captured, and reasonably clean**
0	**Major data are captured sufficiently, and our going-in view is that it can be sufficiently cleaned for a first cut decision**
+1	**In scope is to determine whether the data exists and can be sufficiently cleaned to outperform the status quo**

Table T6: Decision Data

As the term "data science" celebrates, AI is heavily dependent upon feeds of relevant data. If that data is understood, easy to access, and clean(able) from errors, it reduces the risk of building an AI system that lacks robustness. However, it is often the case that the data comes from multiple disparate sources and is subject to swings in timeliness and/or quality. In this case, we mitigate risks by 1) triaging meta algorithms (identifying areas where the AI system works, usually works, does not work); 2) generating datasets that extend beyond existing historical data (for instance, by using simulation runs or reinforcement learning); 3) exploring the value proposition in anomaly and fraud detection on source data; and, 4) finding exogenous data sources that are more complete in order to improve the signal. There are also other mitigation strategies that arise once data is made available and understood. This might involve changing the business process that collects certain data or building data lakes to coalesce different data sources into one "source of truth." Data quality problems at the client may not manifest until after the project kickoff, so this factor requires careful prior analysis, as much as possible.

T7. Software Tools

The availability of AI software tools encompasses both software (languages, packages, libraries, platforms, Software-as-a-Service) and hardware (CPUs, GPUs, CMOS accelerators), which are constantly evolving and expanding.

Score	Characteristic
-1	Selection of all software tools is completely in hands of AI/Data Science team, and budget is assured
0	Tools are largely prescribed, but AI/Data Science team has high experience and confidence they will be sufficient to complete the project
+1	Tools are prescribed, and in scope is to determine whether they will be sufficient or the best fit to complete the project

Table T7: Software Tools

In the current age of “AI salesmanship,” a risk factor is when a client has selected an AI suite on the promise that it will solve any problem anywhere – but has not yet used it for any deployment. In cases where the selection of tools is open and not mandated, the team is provided enough flexibility to choose the right tools for the project. We mitigate these risks by 1) adding experienced AI/Data Science tool architects early in the project; 2) abandoning tools early in the project if they are impeding progress; 3) designing a loosely coupled, micro-services architecture to isolate different parts of the AI system; and, 4) paying attention not just to the AI suite, but the specific analytic technique for each step. Junior AI practitioners are often tempted to start with the most recent and sexy techniques, where a more conventional approach would be more stable. Senior AI practitioners can be overly trapped in favoring the same tools they have always used and not keeping up with the field.

T8. Speed

The next technical factor relates to the speed of decisions. Some projects are using AI to improve the time it takes to make decisions, whereas others may be limited by the underlying business process in terms of how quickly decisions can be implemented.

Score	Characteristic
-1	**Improving decision speed will be a driving benefit of AI**
0	**The business is content with the current decision-making speed**
+1	**A fast decision is not possible in many cases (because of exogenous data or decision makers)**

Table T8: Speed

We look to understand the importance of improving decision speed when considering the potential value proposition of the project. If the goal is to improve the decision-making speed, it reduces the risk of the project, as it allows AI to play to one of its strengths. On the other hand, decision speed may be limited by external factors – for instance an external review by multiple parties across multiple timeframes. We mitigate these risks by 1) including a time-versus-value sensitivity curve in the Cost Benefit Analysis; and, 2) linking with in-house strategy or innovation teams to build a case for improving decision speed and a long-term data driven vision for making decisions. These mitigators allow the business to understand the tradeoffs of faster decisions versus the costs involved to remove the barriers to faster decisions.

T9. User Experience (UX)

How users interact with an AI system is vital and often the single greatest key to success.

Score	Characteristic
-1	**User experience (UX) technology and design are simple (e.g. simple "yes" / "no" recommendation)**
0	**The user experience involves substantial human-in-the-loop cooperation**
+1	**In scope is designing a user interface that will drive and attract usage from a diverse and/or highly optional user population**

Table T9: User Experience/Interface

For situations where the AI system recommendation is simple and easy to understand, the UX presents little risk. However, in the common case where the decision process is complex and success hinges upon user adoption, careful and thoughtful UX design is critical. Failing to get the UX right presents a risk of producing a system that is ignored or underused. We mitigate this risk by 1) designing an ideal AI user experience and make this experience drive the algorithm selection (rather than vice versa); 2) staffing the project with designers experienced in AI interaction (vs. generic application UX); and, 3) devoting sufficient time and budget to UX. We sometimes find that the costs of creating an appropriate user experience for AI is higher and more resource-intensive than the development and testing of AI models. We justify this cost by explaining how a good user experience will help drive the benefits of applying the AI system to make better decisions.

T10. AI Process Maturity

The final technical factor relates to the experience the AI team has in deploying and supporting AI systems in production. Transformative AI is not a one-time project. It is building and growing a constant stream of AI-driven improvements.

Score	Characteristic
-1	**Highly experienced and successful with production AI deployment and long-term support for this business area**
0	**Has successful production deployment experience of AI in other business areas, but not in this particular one**
+1	**Successful AI proofs-of-concept exist, but has limited experience with broad production deployment at the intended scale**

Table T10: User Experience/Interface

In businesses where AI systems have been consistently and successfully deployed, the business can leverage past experience in deploying the new system. Conversely, when there is limited experience in full AI deployments, there is increased risk. We mitigate that risk by 1) documenting and teaching best practices for each stage of the AI project lifecycle; 2) focusing primarily on usage and robustness; and, 3) conducting formal periodic independent reviews and/or audits of the system. Continuously monitoring the technical aspects of the AI system is one aspect of delivering success. Including reviews of how the business is leveraging the AI system as well as how the business is accepting the decisions made by the system becomes just as important.

4. Conclusion

We have found that a consistent and shared vocabulary of explicit risk factors, as well as best practices for remediation is a valuable asset as companies in all industries move themselves to improve their digital strategy and their adoption of AI for making better decisions. At Princeton Consultants, we use the Princeton 20 to score and prepare for each project. The factors that receive +1 scores will be discussed collaboratively with the business and technical staffs of the sponsoring client during this preparation stage – in other words, always before the project is staffed or formally starts. Perhaps counter intuitively, a project with a small number of risk factors is not necessarily better off than a project with more – failure along a single factor with high risk can threaten the entire project's success. It is not about the number of identified risks, but rather about how explicitly and candidly they are identified up front, and then thoughtfully monitored and managed during the project. Indeed, as the project unfolds, the stakeholders and steering committee should monitor these risk factors and make corrections as needed. It helps project management to focus on probable root causes and draw from proven successful approaches to solving them.

At the conclusion of each project, we produce a formal project After Action Review (AAR) that discusses the risk factors and examines, after the fog of battle subsides, whether we identified the right ones and effectively mitigated them. We also take this opportunity to add to our bank of "lessons learned," which we also organize by risk factor. These learnings include both stories of problems that arose in each AI project and the tactics that we found to be successful in solving them. These foundational stories form the basis of both our risk management and our mentoring programs.

While the buzz and excitement surrounding AI continues to expand, a sobering view is that even the most compelling AI is only valuable for a business to the extent to which it is successfully deployed into production, i.e., used to make actual business decisions. AI leaders who do not want to see their efforts (and careers) fall from Gartner's infamous "Peak of Inflated Expectations" into the "Trough of Disillusionment" need to find, adapt, and build methodologies that work for them. We believe the Princeton 20 provides a framework that may kickstart or complement that effort. We are pleased to share this work with the AI community at large for their use and feedback.

5. Bibliography

Bergh, C, Benghiat, G, Strod, E (2019) *The DataOps Cookbook: Methodologies and Tools That Reduce Analytics Cycle Time While Improving Quality* (DataKitchen, Cambridge, MA), Retrieved March 31, 2020, https://datakitchen.io/dataops-cookbook-second-edition.html .

Doerr, J (2018*) Measure What Matters: How Google, Bono, and the Gates Foundation Rock the World with OKRs* (Penguin Publishing Group, New York).

Gartner (2020a) IT Glossary, "Artificial Intelligence," Retrieved March 30, 2020, https://www.gartner.com/en/information-technology/glossary/artificial-intelligence .

Gartner (2020b) Magic Quadrant for Data Science and Machine Learning Platforms. P Krensky, P den Hamer, E Brethenoux, J Hare, C Idoine, A Linden, S Sicular, F Choudhary, February 11, 2020.

Kim, G, Humble, J, Debois, P, Willis, J (2016) *The DevOps Handbook: How to Create World-Class Agility, Reliability, and Security in Technology Organizations* (IT Revolution Press, Portland, OR).

O'Neill, Brian (2019) "Failure rates for analytics, AI, and big data projects = 85% - yikes!", Retrieved March 30, 2020, https://designingforanalytics.com/resources/failure-rates-for-analytics-bi-iot-and-big-data-projects-85-yikes/ .

Sashihara, Steve (2011) *The Optimization Edge: Reinventing Decision Making to Maximize All Your Company's Assets* (McGraw Hill, New York, NY).

Taylor, James (2019) *Digital Decisioning: Using Decision Management to Deliver Business Impact from AI* (Meghan-Kiffer Press, Tampa, FL).