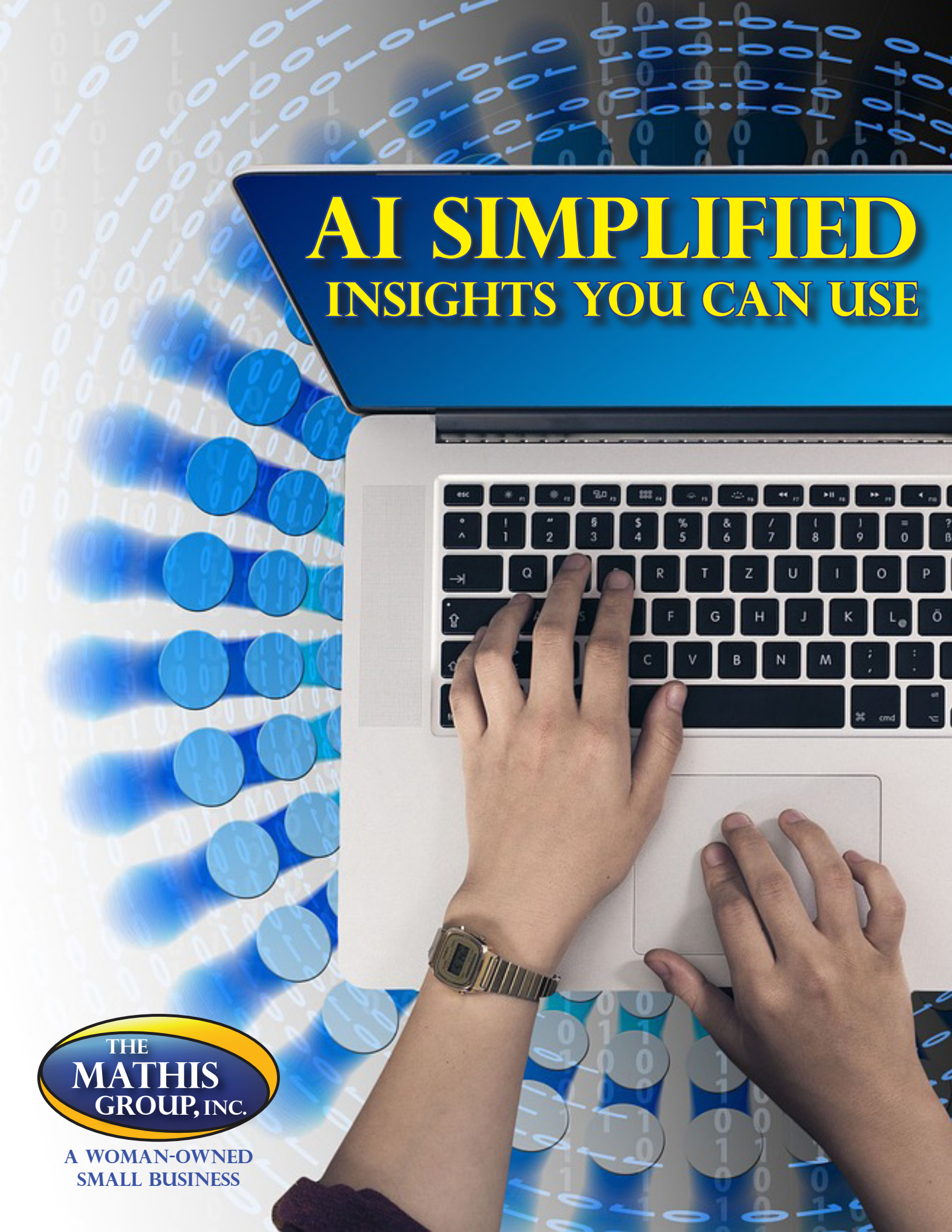

AI SIMPLIFIED
INSIGHTS YOU CAN USE
THE MATHIS GROUP, INC.
A WOMAN-OWNED SMALL BUSINESS

AUGUST 2026

THE DARK SIDE OF ARTIFICIAL INTELLIGENCE IN PROJECT MANAGEMENT

BIAS IN ARTIFICIAL INTELLIGENCE ALGORITHMS

Bias in artificial intelligence is often viewed as primarily a prompt or technical problem. If users wrote prompts more effectively, the technology would respond more closely to their desires. However, this is not always true. Depending on which AI tool a person or organization uses, it can affect the potential for bias. If the tool pulls information from the internet, the potential for bias is greater than if it focuses only on internal material from an organization.

For project management practitioners, the potential for bias is high because it depends on where the tool draws its information from and which algorithm it uses. This can include areas such as promotions, funding, and which risks are escalated for evaluation of one's level of expertise, or even something as simple as connecting workers to project assignments. If the analysis in these areas is wrong, it can become part of the organizational processes and appear to have legitimate foundations for all future recommendations and assignments. The problem is not that AI provides inaccurate information to project leadership; it's that project leadership believes that information to be true and makes decisions based on it.

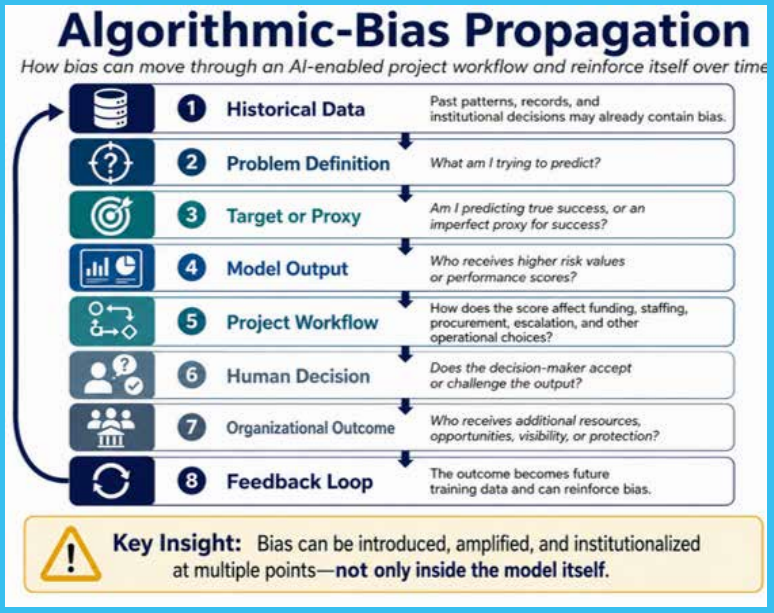
Bias can enter the AI environment in any project or process area through the selection of data, the definitions AI is solving, and the variables AI is tracking. For example, a project that uses metrics such as years of project experience, education, training events attended, hours, and the number of project assignments can appear to support operational goals and project key performance indicators, yet bring in many different types of data influenced by biases from other leadership and management. Remember, bias can be influenced by how others evaluate the project's progress and success factors that are not standardized across the organization. Different levels of risk, varying lifecycles, and individual interpretations can make one project successful while others might be easier to implement with lower risk.

Because of the potential for bias from so many external project initiatives, all AI projects must not only identify potential risks but also track AI bias as a risk component. Tracking AI bias as a risk should include monitoring its impact, both from internal and external sources, that could harm data and influence project and leadership decisions. One discussion is the level of harm bias can have on this decision, and whether the project team can reverse the decision once it is detected.

Other areas where bias can occur include vendor tracking. One of the most cost-effective ways to use AI is to track vendor work. AI can verify vendor information using reports, quality control plans, benefits realization reviews, or lessons learned documents. Using AI to verify and provide oversight can be one of the most important ethical considerations for organizations.

The graph below shows how bias can enter an AI system's workflow and remain undetected for some time. When this level of bias begins, data managers or data scientists may need to retrain the tool to ensure it works correctly. Leaving the tool in its current state can potentially hurt the project or decision and will tend to get worse over time, and not better.

Nazer et al. (2023) show that bias can enter multiple stages of the AI lifecycle, including problem formulation, data collection, preprocessing, model development, validation, and implementation (Nazer et al., 2023).



IMPACT OF DATA BIASES ON PROJECT DECISIONS

Each tool undergoes AI training on data sources, which may include the company's internal data or data from warehouses providing general information for AI training. The training data often reflects biases from past decisions and records. Each data point represents choices made at different times by various individuals discussing project successes and failures. Additionally, some project challenges are beyond the control of the project manager and team, caused by vendor disputes, incorrect data, or dissatisfaction. Due to issues caused by inaccurate or poor data, this must be considered in any project that uses the AI tool.

IBM (2024) explains that biased training data and evaluation practices can produce discriminatory outcomes. The 2024 report explains that biased training data and evaluation processes lead to discriminatory outputs that influence business and project decisions and stresses the need for AI governance to prevent this. (IBM, 2024). Part of the decrease in bias focuses on not initially accepting data as unbiased. Several questions should be part of the evaluation of data quality for training, such as completeness, consistency, and which data are included and which are excluded. This can show what is influencing the data and how the initial events are recorded.

Bias can also occur when teams make decisions that negatively affect protected classes of individuals. Over the years, data has reduced some of these biases, and attempts have been made to treat all individuals equally, but what about data before these standards? Old data can be a challenge for any AI tool and system. Because of the potential for bias or influence from its organizations, it must have a human in the loop to verify that all decisions are legal and ethical to support each person and reduce bias. The National Institute of Standards and Technology (NIST) noted this in a 2022 report. NIST has emphasized that AI bias originates partly in human and societal choices that shape data and technological systems (National Institute of Standards and Technology [NIST], 2022). This can affect

how one trains or retrains the AI tool after discovering areas of bias. The most important part is to anticipate that the AI tool could be wrong rather than accepting it as perfect in the early stages.

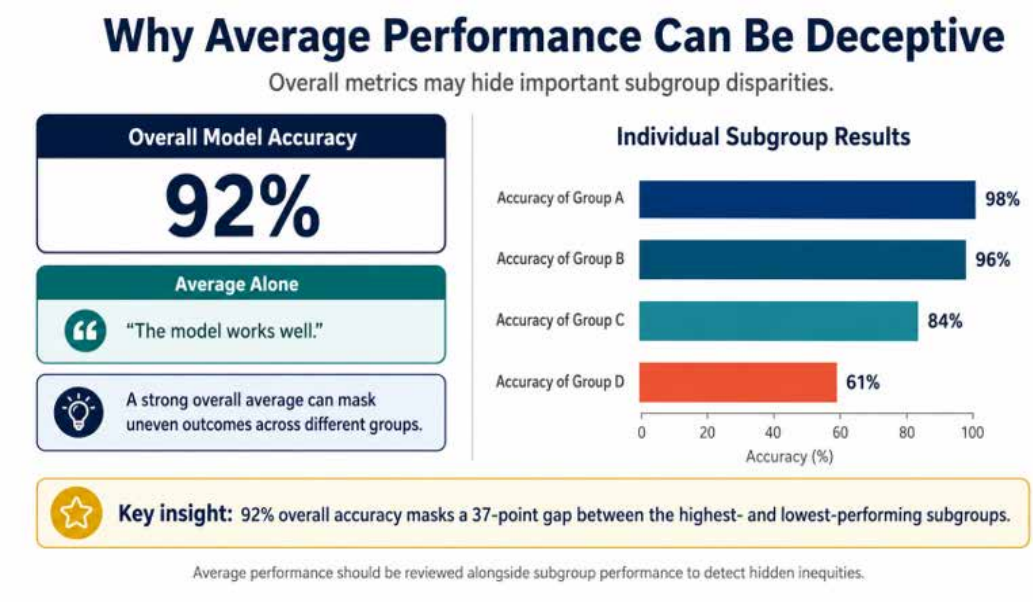
When the new AI tool is installed, part of the process is to train it initially and, over time, keep it up to date by retraining. Developing an AI solution also involves assessing the suitability of the training data. Evaluating the data, source, and how it was collected from various populations should be part of the process. What is the data distribution for the training, and how many months or years does it cover?

Unintentional Bias in Data Sets

Bias can also be unintentional and influence decisions without anyone deliberately trying to do harm. For example, in the hiring process, evaluating male candidates might lead to the conclusion that only a male should hold that position. This was not an intentional bias created by the company but rather what the AI tool learned from current research. Even if the hiring organization attempts to remove identifying information, the AI tool may still detect patterns that are unseen by humans involved in the process. Simply trying to eliminate opportunities for discrimination does not guarantee that the tool won't find ways to do it.

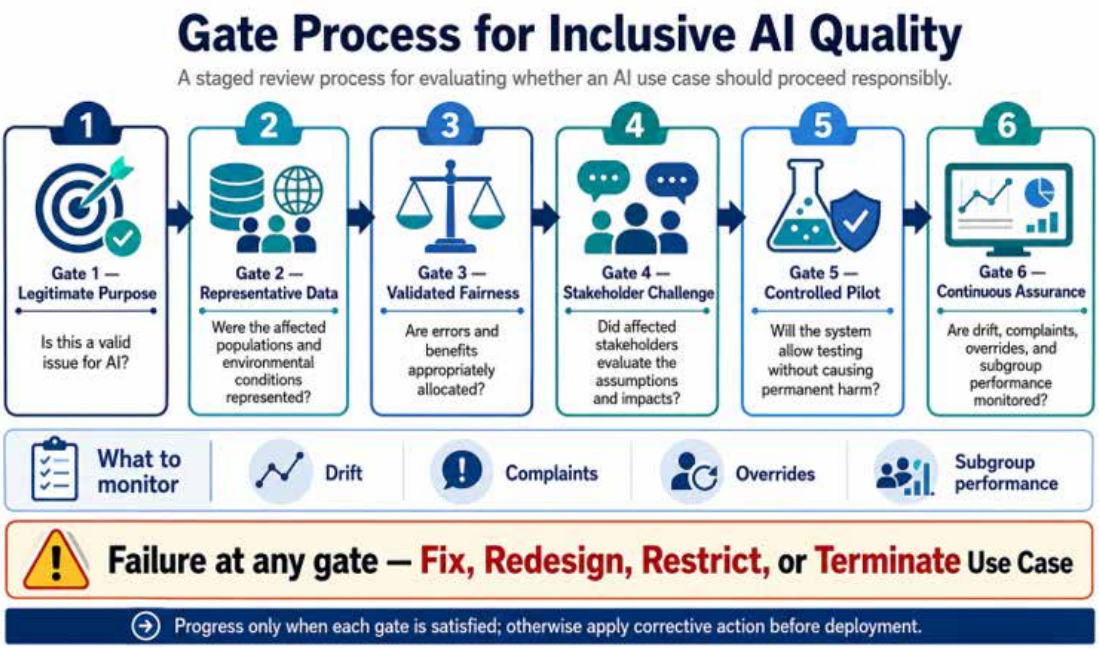
Project assignment bias can occur when an AI tool assigns resources and gives experienced people the newest projects, overlooking newer workers. This bias can carry over into other automated processes, such as vendor or performance evaluations. The tool will respond to and support the training data, so it must be specifically verified to ensure it does not sway the decision either towards or away from any group.

The following graphic can lead to a wrong conclusion, despite the model's 92% accuracy rating. Breaking the groups down, you see there are huge gaps in Groups C and D, with no clear data or evidence on what is happening in those situations. If the only metric everyone hears is 92% overall model accuracy, leadership could leave the meeting thinking AI is operating at high efficiency and with low bias.



Technological solutions must go beyond simply improving the balance in the training data or manually removing bias. Organizations must consider factors such as decision-making processes, performance under adverse conditions, and overall performance with edge tools.

Empirical research has identified unintended organizational consequences of AI-based decision-making, including workplace conflicts and disputes over responsibility (Papagiannidis et al., 2023). Placing fairness metrics alongside other metrics, such as demographic parity, equal opportunity, or predictive analysis, is often incompatible. Selecting the most important metrics must depend on the types of decisions being made and the severity of errors. Part of governance includes giving details on the concept of the type of error and why it fits this tool and action.



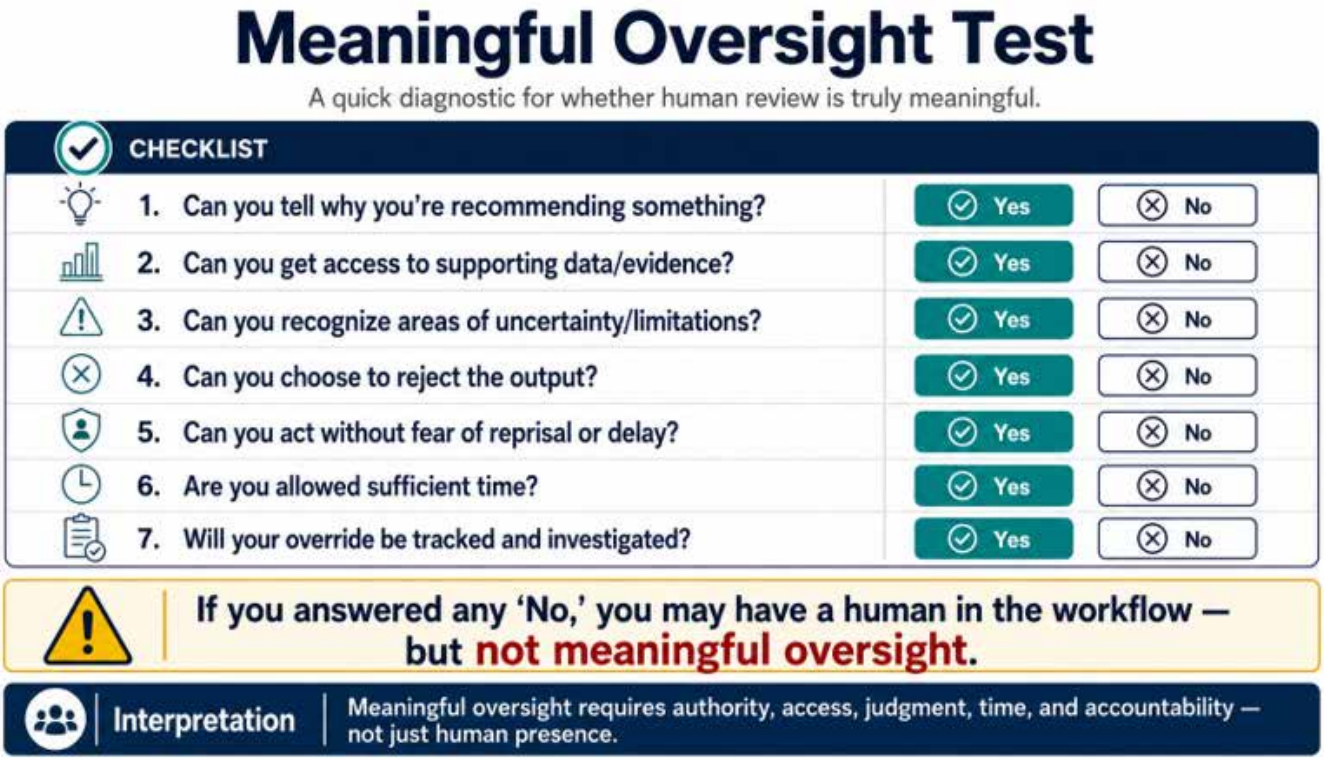
Using a human in the loop covers a variety of activities. In some cases, the human in the loop is actually approving and making recommendations based on what they see. Oversight will fail if the human reviewer is unable to spend time reviewing the output or that part of the process. Over time, mistakes will surface. Sometimes the lack of experience can undermine oversight, as the reviewer lacks context or details and approves something that should be rejected. Lastly, the challenge is that the organization is using humans and going through the review process. Still, if a deviation is found, the person lacks the authority to take corrective action. They must go through multiple decision points to fix something.

Keeping humans in the loop is critical, but in addition to this measure, the correct items, such as tracking overrides, acceptance rates, and the time it takes to review, should be considered. Each of these metrics helps organizations track the precision currently achieved. If the overrides are increasing, it could be a sign that the AI tool needs retraining, as it's producing products that don't align with the original settings. An extremely high acceptance rate could indicate that humans in the loop need additional training and should be taught other evaluation techniques. High acceptance rates could indicate that the reviewer is accepting the system's decision too quickly. Is it possible they approve items that should be rejected?

Keeping humans in the loop is not new. It is already in use, but the European Union's AI Act includes provisions related to this kind of problem. High-risk applications require a human in the loop to provide

oversight to access and monitor how well the system is running. However, intensive training does not guarantee the reviewer will catch all the flaws. Oversight needs to be built into the operating model, not just including it as a risk reduction.

The following graph shows foundational considerations an evaluator should consider when evaluating whether AI is providing solid recommendations or is biased. If the evaluator cannot answer these questions, they must proceed with caution, as the system may recommend incorrect solutions.



Failures Due to AI Misunderstanding

As good as AI is at predictive analysis and acting as an assistant, it can't interpret the project's context as a human evaluator would. AI identifies patterns and produces likely language or optimizes formal goals. The tool schedules differently from how a human would. A missing date might become the most important item on a schedule where the human understands this activity can be completed later while doing other work. These challenges happen between what is really happening and what AI thinks is happening. Misinterpretation can occur in any AI system and produce a detailed recommendation supported by a good explanation that is wrong to a human.

Huang et al. discuss that misinterpretations and harmful outcomes can stem from weak governance around data quality and documentation, not just from the model's construction or the lack of embedding ethics throughout the data lifecycle to reduce failure. They show that AI read the data, analyzed the content, but then just got it wrong (Huang, X., 2026).

Teams should classify all generated content into three categories based on its evidential status: Unvalidated Insight, Verified Analysis, and Approved Recommendation.

A Decision-Authority Matrix helps determine the best operational approach. Tasks that are low risk, reversible, and repeated multiple times can be substantially automated. Decisions that are high-risk, irreversible, uncertain, or that affect individual rights must have greater authority vested in humans in the institutional review process. The matrix should also address model uncertainty, sensitive data use, time constraints, legal liability, and provide avenues for legitimate appeals. Guidance for human-in-the-loop collaborative processes emphasizes working together, with AI assisting with execution. At the same time, humans maintain ownership of judging what is acceptable or unacceptable, handling ethical dilemmas, generating creative solutions, and engaging in contextual reasoning.

This matrix clearly defines sections for making specific decisions, including substantial automation, automatic with oversight, human authority, and institutional review. This graphic cuts the decision authority matrix into two parts, showing what does not need human oversight and what does. Low-risk, reversible areas are moved away from humans and handed over to automation. Anything time-sensitive is still automated, but with light oversight. When decisions are in moderate- or high-risk areas, a human and, later, the organization, are in the loop to verify the right decision.

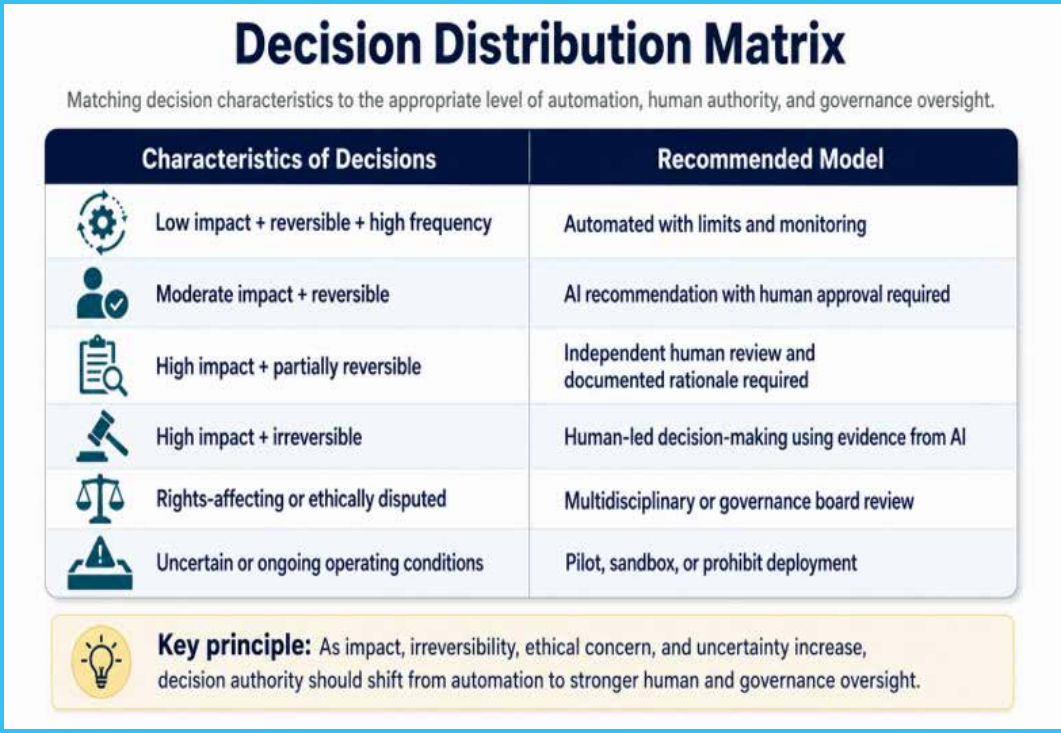
VALIDATING THE MODEL

Validating the model requires transparency so everyone knows how to build trust between humans and machines. Calibrating the tool involves ongoing education and evaluation, and adjusting any underperforming inputs or outputs to ensure reliable answers. All inaccurate acceptance and rejection areas of the model must be analyzed, and the model's gaps or missing outputs must be identified.

Research indicates that bias can enter model validation when datasets and evaluation metrics do not adequately represent the target population (Nazer et al., 2023). IBM (2024) similarly emphasizes the importance of examining the data and assumptions used to evaluate algorithmic performance. Validation needs more than a one-time accuracy check; it requires continuous evaluation to detect issues and any harmful effects. Validation can reduce bias through a checklist that focuses on bias awareness and anticipates potential harm to others.

Refinements of the model aim to align it with the goals and objectives for strategic use. The goal in these areas is to develop a new learning process to restore the model to the baseline. After refinement, the model will begin receiving good solutions and recommendations.

The following graphic guides decision-making. Not all decisions require the same level of commitment to time, research, and effort. Following the Decision Distribution Matrix helps show the characteristics of each decision and the recommended model or level of effort to use in making it.



Data Ethics Across the Project Lifecycle

All organizations collect and store data. Most have policies about what data to keep and how to manage it. The problem is that many organizations, even with standards and policies for handling internal data, lack clear guidelines. This results in archiving data without the ability to search or access it efficiently.

Huang et al. (2026) argue that effective AI governance must address data quality, documentation, accountability, and ethical considerations throughout the data lifecycle. Ethics integrates into all areas of AI use, from selecting the tool to ensuring it does not discriminate against anyone for any reason.

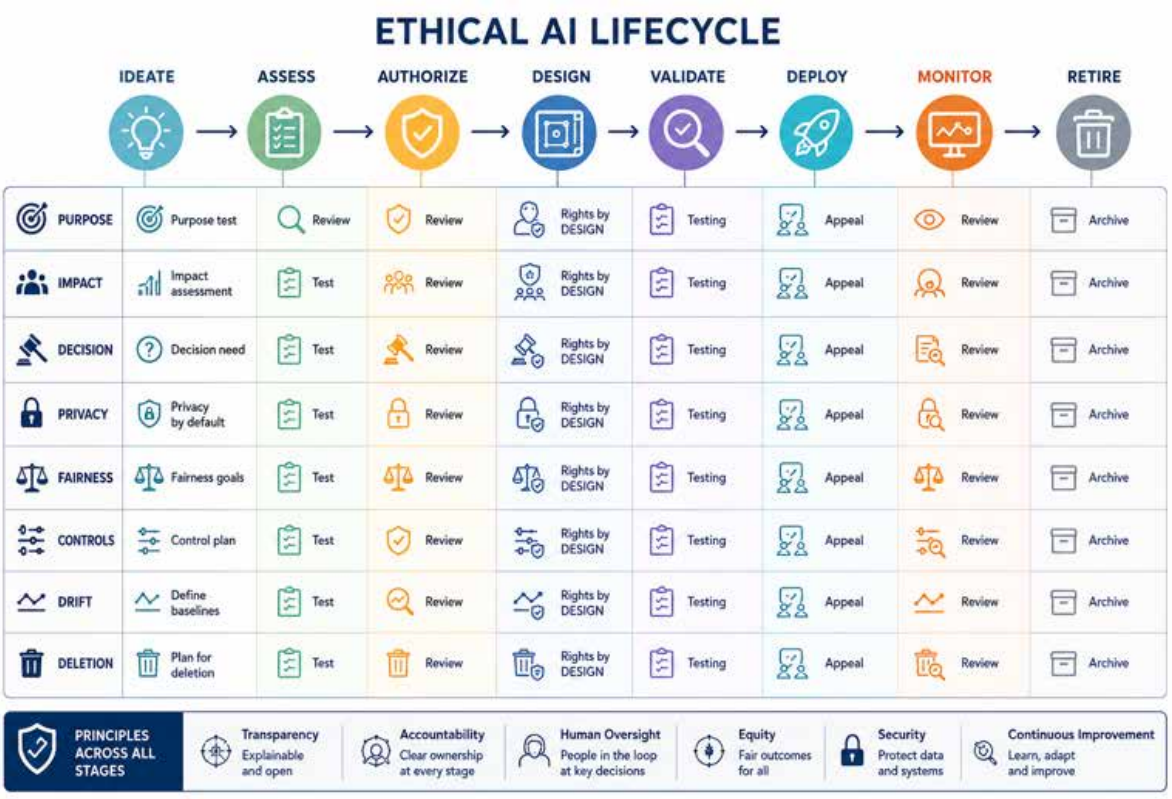
Using AI tools requires current and understandable data. The growing desire to let AI make decisions or give recommendations will expose gaps in the data, either because key parts are missing or because the data simply doesn't exist. When there are no standards for the data, anything goes, and organizations can generate information. However, it might not be usable by the AI tool and may need standardization before it can be utilized. For example, meeting minutes may be archived at the organization and can serve as evidence that decisions are made for changes and project directions. Yet, consider how many different meeting templates and styles can be used. Communication logs and email threads do not cover all topics. Some topics are routine updates and reminders, while others involve contractual matters like performance issues and missed delivery deadlines. The priority of each email may be considered the same, but its content might have little or no value for documentation.

The goal of the analysis is to gain usable insights that help leadership make decisions and reduce risk or other issues. Will the current data in the archives contain that information, and can we adjust it so it is readable by the AI tool? Unless those tools can read internal data, the project is currently running on a data warehouse that does not fully align with the organization’s culture and goals.
ISO/IEC 42001 describes an organizational management system that enables organizations to define, implement, maintain, and continuously improve responsible AI practices. This provides comprehensive governance models to incorporate ethical reviews alongside other oversight when implementing and supporting AI efforts.

There is a significant shift, moving from individual principles to repeatable processes, accountability, clear procedures, and ongoing improvement.

A clear example of applying ethical guidelines to tool use and all aspects of AI can be found in the nursing field. Ball Dunlap and Michalowski (2024) identify accountability, sustainability, fairness, transparency, privacy, and responsibility as key principles of AI data ethics.

This graphic shows the Ethical AI Lifecycle and outlines steps for moving from the beginning through implementation and use, and returning the AI tool when its usefulness is over. This graphic also provides guardrails for actions one should follow in each phase to reduce or prevent problems.

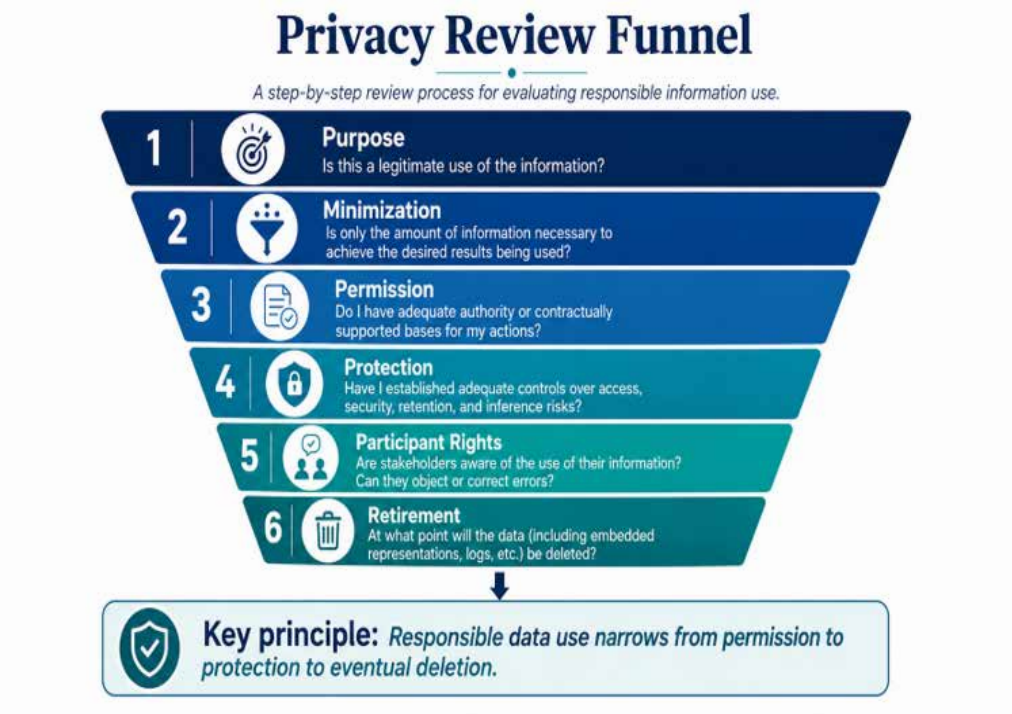


PRIVACY CONCERNS AND ETHICAL POLICIES FOR PROJECT DATA COLLECTION

Using AI systems to manage projects, generate recommendations, and handle client information carries risks. Data for everyone can take many forms. Creating guidelines and standards for privacy regulations focuses on where, how, and what types of data are collected, as well as the retention controls that govern them.

Data collection should be conducted only when required for an approved use case. This data should never be reused for unauthorized purposes. Verifying that everyone follows ethical guidelines and does not use any data without the owner's consent builds trust in others.

The privacy funnel below gives an overview of how to use information within the organization. Understanding the funnel's levels of engagement makes it clear what a person should and should not do with data. Only ethical use of data supports the organization's goals and objectives.



Ethical AI Policies

Ethical AI policies need to provide project teams with enforceable decision rights (the ability to decide) and not just a list of aspirational values (nice-sounding things). They need to define which applications are prohibited or restricted, assign approval levels to specific people, identify the responsible executives, clearly define the duties of the project manager, technical assurance roles, auditors, incident reporters, the appeals process, and the ramifications for non-compliance. The scope needs to cover both internal AI development and third-party AI use. There also needs to be a comprehensive inventory of your AI assets.

Measuring Effectiveness of Ethical AI Policy Guidelines

Measuring the effectiveness of ethical AI policies requires measurement and operational evidence. Setting quantitative metrics that measure the right behaviors and actions drives high-level performance and privacy. Contemporary AI-governance standards increasingly recognize the importance of impact assessment, testing, and auditing human oversight. Project managers should use these metrics to demonstrate if their ethical governance is driving behavior changes beyond additional documentation.

The graphic below shows the various roles and responsibilities within the organization for creating and maintaining ethical boards and leadership for oversight. This chart provides a clear view of others and their duties, while continuously supporting each role.

Red Flags Indicative of Poorly Developed AI Governance Models

Some organizations lack clear guidelines and governance, often noticing red flags only after issues occur. Some problems originate from the vendor's roles, which may struggle to fulfill their commitments.

Human oversight can and should be built into any AI system; however, some problems arise when humans do not provide the appropriate level of oversight. They assume the AI tool is working correctly and provide little or no checks to ensure the correct output is generated. Accuracy must be verified, and its origin.

Proving a human in the loop ensures that all AI-generated content is verified as it is entered into the project records. This level of oversight helps determine whether the AI tools are functioning properly and do not require additional training.

REFERENCES

Ball Dunlap, P. A., & Michalowski, M. (2024). Advancing AI data ethics in nursing: Future directions for nursing practice, research, and education. JMIR Nursing, 7, e62678. https://doi.org/10.2196/62678

European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX-%3A32024R1689&qid=1786550186208

Huang, X., Kou, T., & Zhou, Q. (2026). Embedding AI ethics in the data lifecycle: A framework for enterprise AI governance. Technology in Society, 86, 103261. https://doi.org/10.1016/j.techsoc.2026.103261

IBM. (2024). What is algorithmic bias? https://www.ibm.com/think/topics/algorithmic-bias

International Organization for Standardization. (2023). ISO/IEC 42001:2023 information technology—Artificial intelligence—Management system. https://www.a-lign.com/lp/iso-42001

Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., Hicklen, R. S., Moukheiber, L., Moukheiber, D., Ma, H., & Mathur, P. (2023). Bias in artificial intelligence algorithms and recommendations for mitigation. PLOS Digital Health, 2(6), e0000278. https://doi.org/10.1371/journal.pdig.0000278

Papagiannidis, E., Marikyan, D., Kundisch, D., & Bhattacharya, S. (2023). Uncovering the dark side of AI-based decision-making: A case study in a B2B context. Industrial Marketing Management, 115. https://doi.org/10.1016/j.indmarman.2023.10.003

THE MATHIS GROUP, INC.

A WOMAN-OWNED SMALL BUSINESS (WOSB)

Providing quality, customized training and consulting services that inspire, educate, and equip organizations to be better tomorrow than they are today.

DR. KEITH MATHIS, PMP, PMI-ACP, CSP-SM, CSP-PO
WANDA MATHIS, M.ED. PMI-ACP

PROJECT MANAGEMENT TRAINING

OVER 60 PROJECT MANAGEMENT COURSES REGISTERED WITH PMI

PRESENTATIONS THAT EDUCATE, MOTIVATE, AND INSPIRE

Since 1993, The Mathis Group has been helping organizations change worker productivity and behavior.

PROJECT MANAGEMENT
MARKETING
MOTIVATION
ORGANIZATIONAL BEHAVIOR
LEADERSHIP
CUSTOMER SERVICE

COMPANY MANDATE

The Mathis Group provides training and consulting that will impact the organization and individual while maintaining an outstanding reputation for success and integrity.

VALUES STATEMENT

Every person has worth and should be treated with respect.

AREAS OF EXPERTISE

- Curriculum Design
- Project Management
- Organizational Behavior and Development
- Management
- Agile Project Management
- Strategic Planning
- Executive Coaching
- Performance
- Team Building
- Emotional Intelligence
- Leadership
- Customer Service
- Supervisory Leadership
- Hybrid Project Management

9515 N Spring Valley Dr
Pleasant Hope, MO 65725
800-224-3731
417-759-9110
(voice/fax)

www.themathisgroup.com

keith@themathisgroup.com
wanda@themathisgroup.com

DUNS Number:
007722098
CAGE: 3C1N9
GSA Contractor Number:
GS-02F-0010V

GSA Contract Holder

ACCREDITED BUSINESS BBB