Evaluating and prioritising artificial intelligence projects: A guide for better decision making and investment outcomes

By Stefan Hajkowicz, Chief Research Consultant, Analytics and Decision Sciences, CSIRO

June 2025

Executive Summary - AI Investment Decision Checklist

Use this checklist to help achieve rigorous evaluation and better outcomes when making AI investment decisions. A "No" or "Unsure" response indicates areas maybe needing deeper analysis before proceeding.

1. Strategic alignment and problem definition

- ☐ Clear problem focus: Is the project centred on solving a well-defined business problem or societal challenge, rather than being technology-driven?
- ☐ Strategic fit: Does the AI initiative clearly align with your organisation's broader mission, vision and strategic goals?
- ☐ Value proposition: Is there a compelling business case articulating the expected benefits (financial, operational, societal) and can these benefits be measured/assessed with performance metrics?

2. Investment and resources

- ☐ Holistic investment view: Does the investment plan account not just for technology, but also for essential human capital (skills, training), data needs and necessary organisational changes?
- ☐ Realistic costing: Have both upfront development costs and ongoing operational/maintenance expenses been reliably estimated?

3. Data foundations

- ☐ Data availability and quality: Is the required data available, accessible, and of sufficient quality to support the AI solution?
- ☐ Data governance and compliance: Are data ownership, privacy, security and regulatory requirements clearly understood and manageable?

4. Analytically robust evaluation, prioritisation and project selection

- ☐ Financial and economic appraisal: Where financial returns or economic impacts are relevant, have appropriate metrics (e.g., ROI, NPV, Payback Period) and broader economic analysis (e.g., Benefit-Cost Analysis) been applied to assess the project's value?
- ☐ Cost effectiveness and cost utility analysis: When targeting specific measurable outcomes, have you assessed whether this is the most resource-efficient approach to achieve those goals?
- ☐ Multiple criteria analysis: Where there are multiple objectives, multiple options, or mixed monetary and non-monetary outcomes, have you used systematic comparison frameworks like weighted scorecards or multiple criteria analysis?

- Qualitative approaches: When quantitative analysis alone is insufficient, have qualitative factors, stakeholder perspectives, and non-quantifiable benefits been systematically considered and documented?
- Strategic flexibility: For high-uncertainty projects, has the value of managerial flexibility to adapt in response to new information been considered (e.g., using Real Options Analysis)?

5. Risk management and ethics

- Comprehensive risk assessment: Have diverse risks – technical, operational, financial, societal, and especially ethical (e.g., bias, fairness, transparency, accountability) – been identified, assessed, and documented? Consider established AI ethics frameworks.
- Mitigation strategies: Are there clear strategies in place to mitigate identified risks, particularly ethical and societal concerns?
- Uncertainty planning: Have methods like scenario planning or iterative evaluation been incorporated to manage the inherent uncertainties in AI projects?
- Technology uncertainty: Have you considered changes in technology capability during the investment timeframe? Could advances in AI (or other technology) render the project obsolete?

6. Portfolio perspective

- Balanced portfolio: Is this project considered within a broader portfolio of AI initiatives, balancing risk, reward and strategic objectives?
- Non-AI approaches: Is there an equally (or more) effective and efficient lower-tech non-AI solution?

7. Stakeholder engagement and qualitative factors

- Broad consultation: Have relevant stakeholders (e.g., end-users, domain experts, IT teams, legal, affected communities, ethics experts) been consulted throughout the evaluation process?
- User needs: Do you know who the users are, their needs from the AI system and the likely ways they'll interact with it?
- Qualitative justification: Beyond numbers, is there strong qualitative reasoning supporting the investment decision?

8. Systems Integration and implementation

- Systems thinking: Is the AI solution evaluated not in isolation, but as part of a larger system of people, processes, and existing technologies?
- Integration & change management: Is there a clear understanding of integration challenges, necessary workflow changes and the human effort required for successful adoption?

9. Iterative approach and governance

- Adaptive evaluation: Is there a plan for ongoing monitoring and an iterative evaluation process, allowing for adjustments as the project evolves and new information emerges?
- Clear governance: Are there robust governance structures in place for responsible AI development, deployment and ongoing management? Are there plans for discontinuing the project if needed?
- Culture and readiness: Does your organisation have the culture and change management capability? Is there leadership buy-in? Is there training and skill development plans for affected staff?

Introduction

Artificial intelligence (AI) holds transformative potential for both private companies seeking to achieve business objectives and public entities aiming to improve societal well-being. However, AI projects can involve significant investment, technical complexity, ethical considerations, and uncertain outcomes. Many AI initiatives, especially those exploring novel applications, resemble early-stage Research and Development (R&D) with less predictable cause-effect relationships and outcomes compared to traditional investments.

Therefore, rigorous *ex ante* (before the investment) evaluation is important for choosing and prioritising AI initiatives that are most likely to deliver intended value and align with strategic goals while mitigating potential risks. This evaluation must be practical and adaptable, avoiding analysis paralysis while still providing sufficient insight for robust decision-making in the face of ambiguity.

This guide outlines a range of evaluation frameworks and approaches that public and private sector entities can employ to make informed decisions about which AI models to build. Effective evaluation typically involves a combination of quantitative analysis, qualitative reasoning, and stakeholder consultation, tailored to the specific characteristics and uncertainties of AI.

What’s an AI project?

This guide is about investments in AI projects which are also called “AI use cases”. Identifying, prioritising and choosing AI projects is one of the main challenges at the beginning of an AI journey for most organisations. So, what is an AI project? An AI project is a structured set of activities that applies artificial intelligence technologies and leverages multiple organisational resources to solve a specific business problem or achieve a defined objective. Unlike casual experimentation with AI tools, these projects have defined scope, timelines, success metrics, and resource allocation designed to deliver measurable business value.

AI projects are fundamentally problem-focused rather than technology-focused. They begin with a clear business challenge such as reducing customer wait times, improving diagnostic accuracy, or optimising supply chains, and then determine whether AI represents the most effective solution approach. This problem-first methodology distinguishes successful AI implementations from technology-driven initiatives that often fail to deliver practical value.

Successfully executing AI projects requires the convergence of four critical resources: data, hardware, software, and wetware (people). High-quality, relevant data represents the foundation, as AI systems require access to comprehensive datasets for training and ongoing operation. Organisations frequently discover that their data quality, accessibility, or governance systems need substantial improvement before AI implementation can succeed, making data infrastructure development a prerequisite rather than an afterthought.

The hardware component encompasses the computational infrastructure necessary for AI processing, including specialised processors like graphics processing units, cloud computing resources, and storage systems capable of handling large datasets. Many organisations underestimate the hardware requirements for AI projects, particularly the processing power needed for model training and the ongoing computational demands of deployed AI systems.

Software requirements include both development platforms and operational systems, encompassing machine learning frameworks, AI development tools, specialised AI applications, and integration software that connects AI capabilities to existing business systems. The software landscape for AI is rapidly evolving, requiring careful selection of platforms that balance current functionality with future scalability and vendor stability.

The

AI factories are here – We need to decide what they'll build

- Specialised production hubs: AI factories are evolving beyond traditional data centres, becoming dedicated facilities engineered for the end-to-end production and deployment of AI models and intelligence.
- Streamlined AI lifecycle: They integrate crucial components – data pipelines, algorithm development tools, experimentation platforms, and powerful AI-specific infrastructure (often using accelerated computing like GPUs) – to optimise the entire process from raw data to operational AI.
- Efficiency and scalability: The focus is on automation and streamlined workflows to efficiently process vast data volumes, enabling faster innovation, optimised resource use, and reduced operational costs.
- Driving competitive advantage: The rise of AI factories signals a maturing AI industry, offering organisations the capability to scale AI initiatives and secure a competitive edge through accelerated development and deployment of AI-driven solutions.

wetware component—the people involved—proves equally critical and often most challenging to secure. AI projects require skilled teams that combine technical expertise in data science and machine learning with deep domain knowledge about the business problems being addressed.

This includes not only AI specialists and data scientists, but also project managers, analysts and subject matter experts who understand both the technical possibilities and business requirements. The scarcity of qualified AI talent makes wetware often the most expensive and difficult resource to acquire, requiring organisations to invest heavily in recruitment, training or partnerships to build necessary capabilities.

Types and applications

AI projects typically fall into several categories, each with distinct characteristics and applications. Machine learning projects represent the most common type, involving either supervised learning where models are trained on labelled data to make predictions about new cases, or unsupervised learning where algorithms discover patterns in unlabelled data. Examples include fraud detection systems that learn from historical transaction patterns or customer segmentation tools that identify distinct groups within large datasets.

Foundation model applications have become increasingly popular, involving the adaptation of pre-trained AI models like large language models for specific business uses. Rather than building AI systems from scratch, these projects fine-tune existing sophisticated models for particular domains or tasks. This approach often proves faster and more cost-effective than custom development while still delivering specialised capabilities. In some cases organisations may build their own foundation AI models; this is typically a costly exercise.

AI agent development focuses on creating autonomous software entities that can perceive their environment, make decisions and take actions. These agents range from customer service chatbots that handle routine inquiries to complex automated trading systems that execute financial transactions based on market conditions. Reinforcement learning systems represent another category, where AI learns optimal behaviours through trial and error in specific environments, commonly applied to optimisation problems and resource allocation challenges.

Beyond technology implementation

What distinguishes AI projects from simple technology purchases is their comprehensive approach to organisational integration. An AI project encompasses not just the deployment of AI technology, but also the development of supporting infrastructure, the training of personnel, and the modification of business processes to accommodate new capabilities. This systems-level thinking recognises that AI tools deliver value only when effectively embedded within existing organisational workflows.

The resource requirements for AI projects extends beyond computational power and software licenses. Successful projects require skilled teams that combine technical expertise in data science and machine learning with domain knowledge about the business problems being addressed. Project management capabilities become crucial for coordinating complex implementations that often involve multiple departments and external vendors.

Governance frameworks represent another component, establishing ethical guidelines, security protocols and compliance procedures that ensure responsible AI development and deployment. These frameworks become important as AI systems increasingly influence business decisions and customer interactions.

Integration and Reality

The reality of AI projects is that they succeed not because the AI technology works well in isolation, but because they're effectively integrated into larger systems of people, processes and technologies. An AI diagnostic tool in a hospital, for example, delivers value not merely through technical accuracy, but through seamless integration with patient record systems, effective training of clinical staff, and adaptation of hospital workflows to leverage the AI's capabilities.

This integration requirement means that AI projects typically involve significant change management components. Organisations need to prepare their workforce for new ways of working, establish new decision-making processes that incorporate AI insights, and develop performance monitoring systems that track both technical performance and business outcomes. The AI component might be technically sophisticated, but the project's ultimate success depends on how well it enhances and integrates with existing operations to solve real business problems.

How much are we investing?

According to the 2025 AI Index Report by Stanford University total corporate investment in AI, which includes private investment and mergers and acquisitions, reached AU$392.27 billion/year in 2024. The Stanford report also finds that global private investment in AI specifically climbed to AU$101.99 billion/year in 2024 in the United States alone, with global private investment in generative AI attracting AU$57.21 billion in 2024.

The amount of expenditure is set to rise over the coming years. Global spending on artificial intelligence is projected to reach AU$982.36 billion/year by 2028, according to the International Data Corporation (IDC) - a market intelligence firm with Headquarters in Needham, Massachusetts, USA. This forecast, detailed in IDC's Worldwide AI and Generative AI Spending Guide, highlights a compound annual growth rate of 29% for the 2024-2028 period.

The amount spent on individual AI projects ranges widely. Small and simple projects like building an AI agent to summarise document information stored in PDFs could cost a relatively modest amount. But a large, complex general purpose AI foundation model trained on large datasets could cost an extraordinary sum - hundreds of millions. Overall, the costs of making models delivering the same levels of intelligence is declining.

However, governments and companies are spending larger sums to achieve higher levels of intelligence, as the benefits of increased intelligence, and potentially achieving artificial general intelligence are practically boundless. A training run alone for the world's top performing AI foundation models could exceed AU$150 billion by 2027 according to Dario Amodei, CEO of Anthropic (Tom's Hardware, 8 July 2024).

The realities of AI project implementation: And a call for rigorous evaluation

Despite the immense promise of AI, the current landscape is marked by a high rate of failure. Studies indicate that 80% or more of AI projects fail to deliver on their intended goals or be successfully deployed. For instance, a

RAND research report titled “The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed” (2024) cites estimates placing the failure rate above 80%, significantly higher than traditional IT projects. This same 80% failure rate for AI projects is given in the book “Why Data Science Projects Fail: The Harsh Realities of Implementing AI and Analytics, without the Hype” (2024) by Douglas Gray and Evan Shellshear. This significantly exceeds the failure rate of traditional IT projects and highlights the unique challenges inherent in AI development and implementation.

While various factors contribute to this high failure rate, a primary driver is the initial decision-making process – specifically, choosing the wrong AI projects in the first place. This can stem from a misunderstanding of the problem AI can realistically solve, a lack of necessary data or infrastructure, overambitious expectations, or a failure to align the project with strategic needs and practical implementation realities. Investing time and resources into ill-conceived or poorly defined AI initiatives is a direct path to wasted investment and contributes significantly to the high rate of projects that never see successful deployment.

However, this high failure rate should not deter investment in AI. The risks of doing nothing could be greater. As with R&D and any form of innovation, risk-taking is the only way to succeed in the long term, and in today’s rapidly changing world, often the short term. The 20% of AI projects that succeed often deliver substantial returns, frequently paying off the investment made in many failed projects many times over. For both private companies and public entities, successfully leveraging AI can unlock significant value, drive innovation, improve efficiency, and provide a crucial competitive advantage or enhanced public service delivery. Organisations that figure out how to effectively identify, evaluate and implement AI projects will gain a significant edge over those that do not. The high failure rate underscores not a reason to avoid AI, but rather the critical importance of robust and tailored *ex ante* evaluation to increase the probability of selecting projects that fall into the successful minority.

Core evaluation frameworks and their application to AI

Working out how to spend money wisely isn’t a new problem. Over many decades frameworks have emerged to help us judge whether an investment is worthwhile and to prioritise a set of alternative projects based on their overall desirability. These established frameworks for investment appraisal can be adapted to assess the potential of AI projects, though with specific considerations for the unique characteristics of AI. Here we briefly summarise the main frameworks and explore how they can be used to evaluate AI investments:

1. Financial Appraisal Frameworks

- Primary Applicability: Predominantly private sector, but also crucial for public sector projects requiring financial sustainability.
- Core Objective: To assess the direct financial viability and profitability of an AI investment from the perspective of the investing entity.
- Key Techniques:
    - Return on Investment (ROI): Measures the profitability of an investment relative to its cost.
    - Net Present Value (NPV): Calculates the difference between the present value of future cash inflows and the present value of cash outflows, accounting for the time value of money.
    - Internal Rate of Return (IRR): The discount rate at which the NPV of all cash flows from a particular project equals zero; represents the project's expected rate of return.
    - Payback Period: Determines the time required to recoup the initial investment.

  - Discounted Cash Flow (DCF) Analysis: An umbrella term for valuation methods (including NPV and IRR) that estimate the value of an investment based on its expected future cash flows, discounted to their present value.

- Application to AI Investments: These techniques are employed to forecast and evaluate the financial consequences of developing and deploying an AI model. This involves:
  - Projecting Cash Inflows: Estimating potential revenue from new AI-driven products or services, cost savings achieved through task automation, and efficiency gains leading to increased output.
  - Projecting Cash Outflows: Quantifying development costs (data acquisition, preparation, labelling, model training, infrastructure), alongside ongoing operational expenses (maintenance, updates, monitoring).
  - DCF methods like NPV and IRR are particularly favoured as they incorporate the time value of money, which is critical for AI projects where costs and benefits are typically spread over several years. The Payback Period offers a simpler, quicker measure of investment recovery.
- AI-Specific Considerations:
  - Benefit Quantification Challenges: Accurately forecasting the financial benefits of novel AI applications can be difficult due to uncertain cause-and-effect relationships, unpredictable market adoption rates, and the inherent complexities in predicting performance improvements.
  - Cost Estimation Uncertainties: Upfront estimation of costs related to data governance, ensuring model explainability, and adhering to potential regulatory compliance specific to AI can be challenging.
  - Rapid Technological Obsolescence: The fast pace of AI development can significantly impact the expected operational lifespan and ongoing maintenance costs of an AI model, complicating long-term financial forecasting.

2. Economic efficiency and societal value frameworks

- Primary Applicability: Primarily public sector, but increasingly relevant for private sector entities focusing on Environmental, Social, and Governance (ESG) criteria and demonstrating broader societal impact.
- Core Objective: To assess the overall impact of an AI project on societal welfare or economic efficiency, extending beyond direct financial returns to the investing entity.
- Key Techniques:
  - Benefit-Cost Analysis (BCA): A systematic process for identifying, quantifying and comparing the total societal benefits and costs of a project, often expressed in monetary terms. A positive net benefit or a benefit-cost ratio greater than one suggests the project enhances overall welfare.
  - Social Return on Investment (SROI): A framework for measuring and accounting for a broader concept of value, incorporating social and environmental outcomes alongside economic ones. It involves stakeholder engagement to identify experienced changes and values these outcomes (often using proxies or non-market valuation techniques) to calculate a ratio of social value generated per unit of investment.
- Application to AI Investments: These frameworks adopt a comprehensive perspective to capture the full spectrum of value created (benefits) and diminished (costs) by an AI project.
  - BCA for AI: Involves identifying and monetising (where feasible) all significant societal or economic impacts. For AI, this could include enhanced economic productivity, improved public safety, reduced environmental impact (e.g., AI-optimised energy consumption), but also potential costs such as job

displacement, retraining needs, or the economic consequences of privacy violations or algorithmic bias.

  - SROI for AI: Focuses on the social and environmental changes resulting from AI implementation. This is particularly pertinent for AI applications in sectors like health, education, social services and environmental sustainability, where stakeholder-defined outcomes are paramount.
- AI-Specific Considerations:
  - Valuation of Non-Market Impacts: Assigning monetary values to the non-market impacts of AI (e.g., the value of enhanced fairness in decision-making, the societal cost of reduced human autonomy, or the benefit of improved public health outcomes) is methodologically complex and often relies on assumptions and specialised valuation techniques.
  - Uncertainty and Unintended Consequences: The novelty of many AI applications means that precise outcomes can be uncertain, and identifying all potential unintended social, ethical and economic consequences is a critical yet challenging aspect of a comprehensive analysis. Acknowledging inherent ambiguity in valuing complex social or ethical outcomes is crucial.

3. Cost-effectiveness frameworks for specific outcomes

- Primary Applicability: Relevant for both public and private sectors when the primary goal is to achieve a specific, often non-monetary, outcome in the most resource-efficient way.
- Core Objective: To compare the relative costs of different interventions (e.g., various AI models, or AI versus non-AI solutions) in achieving a common, clearly-defined effect or outcome.
- Key Techniques:
  - Cost-Effectiveness Analysis (CEA): Compares the costs of alternative ways to achieve a specific non-monetary outcome (e.g., cost per accurately identified defect, cost per successfully automated customer inquiry, cost per life saved in a defined context). The output is typically expressed as a ratio (e.g., cost per unit of effect).
  - Cost-Utility Analysis (CUA): A specific form of CEA predominantly used in healthcare, where the outcome is measured in terms of health utility, often using Quality-Adjusted Life Years (QALYs). It compares the cost per QALY gained from an intervention.
- Application to AI Investments: These frameworks are useful when decision-makers need to choose the most efficient method among several options for achieving a predefined objective, especially when benefits are difficult to monetise directly.
  - CEA for AI: Helps identify the most efficient AI model or AI-based approach for attaining a particular operational or strategic goal by comparing their cost per unit of a specific outcome.
  - CUA for AI: Applicable to AI models used in healthcare for diagnosis, treatment recommendations, or health management, allowing comparison of interventions based on their cost relative to the improvement in health-related quality of life.
- AI-Specific Considerations:
  - Outcome Definition and Measurement: Requires a clear, consistent and measurable definition of the outcome metric. This can be challenging for nuanced AI outputs where the causal link between the AI and the outcome may not be fully understood or easily quantifiable.
  - Data for Utility Assessment (CUA): CUA specifically necessitates robust data on how an AI intervention impacts health states and the associated utility values. For novel AI applications, such evidence may be limited, requiring careful estimation or reliance on preliminary data.

4. Strategic flexibility and uncertainty valuation Frameworks

- Primary Applicability: Valuable for both public and private sectors, especially for innovative and high-uncertainty projects typical of many AI initiatives.
- Core Objective: To value the managerial flexibility inherent in an investment, allowing for adaptation as new information emerges and uncertainties are resolved over time.
- Key Technique:
    - Real Options Analysis (ROA): Treats an investment not as a singular, irreversible “go/no-go” decision, but as a series of potential future choices or “options” (e.g., to expand, defer, abandon, or switch strategies). It uses principles from financial option pricing to quantify the value of this strategic flexibility.
- Application to AI Investments: ROA is particularly well-suited to AI projects due to significant uncertainties surrounding technical development pathways, market acceptance, regulatory landscapes and evolving use cases.
    - It acknowledges that initial AI investments can create opportunities for future actions as uncertainties unfold. For example, a pilot AI project might provide the option to scale up if successful, pivot to a different application, or halt further investment if early results are unpromising.
- AI-specific Considerations:
    - High Upfront Investment vs. Potential Upside: AI projects often involve substantial upfront research and development costs, but also offer significant potential for future benefits if successful. Conversely, the option to abandon can limit downside risk. ROA helps capture the value of this asymmetric payoff structure and the strategic flexibility it affords, which traditional discounted cash flow methods like NPV might undervalue by assuming a single, predetermined path of future outcomes.
    - Identifying Key Uncertainties and Decision Points: Effective application of ROA requires careful identification of the major uncertainties affecting the AI project and defining the potential decision points where managerial flexibility can be exercised based on how these uncertainties are resolved.

Handling risk, uncertainty, and scenario planning

Investing in AI is inherently subject to significant risk and uncertainty. This can stem from the evolving nature of the technology, data availability and quality, regulatory changes, market acceptance, ethical considerations and the potential for unforeseen consequences. The R&D nature of many AI projects means there is often substantial ambiguity about cause-effect relationships and the precise nature and magnitude of both monetary and non-monetary outcomes. Effective *ex ante* evaluation must explicitly address these factors to provide a realistic picture of potential outcomes without falling into analysis paralysis. Techniques include:

1. Qualitative risk Assessment: As mentioned previously, this involves identifying potential risks (technical, ethical, societal, operational, etc.), assessing their likelihood and potential impact, and developing mitigation strategies. For AI, this often requires expert judgment and stakeholder input to identify less obvious risks like algorithmic bias, unintended system interactions, or the impact of future technological shifts that could render a chosen approach obsolete.
2. Sensitivity analysis: This technique examines how the outcome of a quantitative analysis (like NPV or BCA) changes when key input variables are varied within a plausible range. For an AI project, this could involve testing the impact of changes in development costs, the rate of adoption, the level of efficiency gains, the cost of data acquisition, or the estimated value of a non-monetary outcome. It helps understand which variables have the biggest impact on the project's viability and where more data or analysis might be

needed, guiding where to focus evaluation efforts.

3. Scenario planning: This involves developing several distinct, plausible future scenarios based on different assumptions about key uncertainties (e.g., rapid technological advancement vs. slow progress, favourable regulation vs. strict controls, high vs. low market adoption, successful vs. unsuccessful resolution of a technical challenge). The potential outcomes (e.g., NPV, societal impact, ethical performance) of the AI project are then evaluated under each scenario. This helps decision-makers understand the range of possible results and the project's performance under different future conditions, informing more robust decisions and strategies that are resilient to uncertainty.
4. Real options analysis (ROA): This is specifically designed to value the flexibility to adapt decisions in response to how uncertainty unfolds. By identifying and valuing the “real options” embedded in an AI project, evaluators can account for the potential upside from favourable outcomes and the ability to limit losses from unfavourable ones, providing a more accurate measure of value in uncertain environments. This is particularly useful when the path and outcomes of the AI development are highly contingent on future events.
5. Iterative evaluation: Given the R&D nature and high uncertainty, a practical approach is to adopt an iterative evaluation process. Initial *ex ante* evaluation can be less detailed, focusing on high-level strategic fit, major risks, and order-of-magnitude estimates of potential value. As the project progresses through stages (e.g., proof of concept, pilot, scaled deployment), uncertainty is reduced, and more detailed evaluation can be conducted, refining cost and benefit estimates and reassessing risks based on new information. This prevents analysis paralysis upfront and ensures evaluation is proportionate to the level of investment and certainty at each stage.

Portfolio approach to AI investment

Given the high failure rate of individual AI projects, a strategic approach is to view AI investment not as a series of isolated decisions, but as building and managing a portfolio of AI initiatives. This approach acknowledges the inherent risk and uncertainty and seeks to balance it across multiple projects to optimise overall value creation.

This strategy aligns with the principles of Modern Portfolio Theory (MPT), a financial theory that suggests investors can construct portfolios to maximise expected return for a given level of market risk. While MPT was developed for financial assets, its core concepts are applicable to AI project portfolios:

- Diversification: Investing in a variety of AI projects with different risk profiles, application areas and technical approaches can help mitigate the impact of individual project failures. A failure in one high-risk project is less detrimental if balanced by successes in other parts of the portfolio.
- Risk-Return Trade-off: A portfolio approach explicitly considers the relationship between the potential return of a project and its associated risk. Organisations can strategically allocate resources across projects that range from lower-risk, incremental improvements (e.g., automating a well-understood process) to higher-risk, potentially transformative moonshots (e.g., developing a novel generative AI application).
- Correlation: In financial MPT, the correlation between assets in a portfolio is important. For AI projects, this translates to considering how the success or failure of one project might impact others. Projects that rely on the same data infrastructure or technical expertise might be more correlated, while projects in different business units or with distinct technical stacks might be less so.

Applying MPT principles to AI project portfolios doesn't necessarily require complex mathematical optimisation, especially given the difficulty in precisely quantifying all AI outcomes. However, the *mindset* of portfolio management is crucial. It encourages organisations to:

- Categorise projects by risk and potential reward: Group potential AI projects based on their technical uncertainty, data availability, ethical complexity and potential for financial or societal impact.

- Allocate resources strategically: Ensure a balance of projects across different risk-return profiles, aligning with the organisation's overall risk appetite and strategic objectives. Don't just chase the highest potential return; consider the probability of success and the impact of failure.
- Manage the portfolio actively: Regularly review the performance of the AI project portfolio, re-evaluate projects based on new information (iterative evaluation), and make decisions about continuing, pivoting, or stopping projects based on their contribution to the overall portfolio goals.

By adopting a portfolio approach guided by MPT principles, organisations can make more resilient investment decisions in the face of AI's inherent uncertainties, increasing the likelihood of achieving overall strategic success despite the high failure rate of individual projects.

The indispensable role of qualitative approaches, reasoning and consultation

While quantitative frameworks offer valuable numerical insights, the decision to invest in Artificial Intelligence (AI) extends far beyond mere numbers. A comprehensive evaluation, especially when dealing with ambiguous outcomes and significant non-monetary factors, necessitates a strong reliance on qualitative approaches, structured reasoning and extensive consultation. These elements are indispensable for navigating the complexities of AI adoption and ensuring that such initiatives are not only technically sound but also strategically aligned, ethically responsible and widely accepted. The following three crucial considerations underscore this qualitative dimension:

1. Strategic alignment and business justification: At the forefront of any AI investment is the fundamental need to develop a clear and compelling business case or investment thesis. This involves articulately defining the strategic problem or opportunity that the AI model is intended to address. It requires a logical explanation of why AI stands as the most appropriate solution and a clear demonstration of how the project aligns with the broader mission, vision and overarching AI strategy of the organisation. Furthermore, this justification must encompass well-defined objectives, which are often non-financial, such as enhancing customer experience, improving the quality of decision-making, or building internal AI capabilities. Crucially, AI projects must harmonise with the organisation's existing data strategy, technology infrastructure and available talent pool. The business case should therefore thoroughly address the organisational changes required to successfully adopt, integrate and derive value from the AI model.
2. Qualitative risk and feasibility assessment: A rigorous and thorough assessment of a wide array of risks is essential when considering AI projects. Beyond straightforward financial risks, AI initiatives inherently involve significant considerations across several domains. Technical feasibility is a primary concern, questioning whether the proposed model can realistically be built with available data, suitable algorithms and the necessary expertise. Ethical risks are also paramount, encompassing potential issues such as bias in algorithms, a lack of fairness in outcomes, transparency challenges and gaps in accountability. Societal risks further broaden the scope, considering potential impacts on employment, equity, individual privacy and the potential for misuse of the technology. Finally, operational and integration risks address the challenges associated with deploying, maintaining and seamlessly embedding the AI model into existing workflows. Ethical and societal risks are particularly prominent and complex in the context of AI, often demanding qualitative assessment, expert judgment and the proactive development of robust mitigation strategies before any development proceeds. Similarly, technical feasibility can often be uncertain, requiring a careful and honest evaluation of the current state-of-the-art technology and the organisation's internal capabilities.
3. Consultation and stakeholder engagement: Engaging with a diverse range of stakeholders is invaluable throughout the entire AI evaluation and development lifecycle. This includes interacting with end-users to deeply understand their needs, gather detailed requirements and assess the potential impact of the AI on

their workflows and overall experience. For public sector AI projects, in particular, consultation with affected communities and the general public is vital to understand concerns regarding fairness, privacy and trust, and to gather input on desired outcomes. Subject matter experts offer critical insights into technical feasibility, data quality and domain-specific nuances. Ethics experts play a crucial role in identifying and helping to address potential ethical pitfalls early in the process. Internal teams, including IT, operations and HR, are also key for assessing operational readiness, integration challenges and any required training. Given AI's potential for significant societal impacts, broad and inclusive consultation is not just beneficial but vital to ensure that AI is developed and deployed responsibly and gains the necessary social license to operate. Furthermore, such consultation is key for identifying and qualitatively assessing non-monetary outcomes, especially when precise quantitative valuation proves difficult or inadequate.

Integrating approaches: Multiple Criteria Analysis for prioritisation

Given that AI projects involve a diverse set of criteria – including financial, economic, social, ethical, technical and strategic factors, often with significant uncertainty – Multiple Criteria Analysis (MCA) provides a structured framework for integrating insights from various quantitative and qualitative assessments to aid in prioritisation. In some organisations this is referred to as a “weighted scorecard” approach. An MCA can be anything from a simple back-of-envelope calculation to an advanced and complex computer model. The stages of an MCA typically include:

1. Define Decision Criteria: Based on the business case, strategic alignment, risk assessment and consultation, identify all relevant criteria for evaluating potential AI projects (e.g., NPV or range of NPVs from scenario analysis, expected societal benefit or qualitative assessment of impact, technical feasibility score, ethical risk rating, alignment with strategic pillars, ease of implementation, value of strategic flexibility from ROA). Include criteria that capture the important non-monetary outcomes identified.
2. Assign Weights: Determine the relative importance of each criterion, often through discussion and consensus among decision-makers and key stakeholders. This process can help clarify priorities when facing trade-offs between different types of value (e.g., financial vs. social).
3. Score Options: Evaluate each potential AI project against each criterion. This will involve using the results from the quantitative analyses (e.g., the calculated NPV or a range, the ROA value) and the outcomes of qualitative assessments and consultations (e.g., a score for ethical risk based on the assessment, a rating for strategic alignment, a qualitative ranking of expected social impact). For uncertain outcomes, scores might represent expected values or performance under different scenarios.
4. Aggregate and Rank: Combine the weighted scores for each project to arrive at an overall score, allowing for ranking and comparison.
5. Sensitivity Analysis: Conduct analysis to see how the ranking changes if the weights or scores for certain criteria are adjusted, highlighting the robustness of the decision and the impact of different assumptions or priorities.

It is worth noting that many of the techniques outlined above such as BCA, ROI, CEA, CUA, sensitivity analysis, scenario planning and qualitative approaches can be used within an MCA framework. The techniques are not mutually exclusive. And the MCA model can be stress tested via sensitivity analysis and scenario analysis.

MCA is particularly useful for AI because it allows decision-makers to explicitly consider the trade-offs between different desirable outcomes (e.g., balancing high financial return with lower ethical risk, or prioritising a project with high social benefit over one with a quicker financial payback). It provides a transparent and defensible process for making complex decisions involving diverse and often uncertain criteria. It also allows for the inclusion of criteria derived from qualitative assessments and stakeholder input, which are crucial for AI.

AI in context: considering a systems approach to evaluation

An AI model, agent, or tool, however advanced, typically delivers business benefits or solves complex problems when working in concert with other elements. The true value of AI can often be more fully realised when it's viewed and implemented as an integral part of a larger, intricate system. This broader system usually includes human actors, institutional frameworks, organisational processes and other interacting technologies.

For example, an advanced AI diagnostic tool in a hospital might be technically proficient. However, its real-world impact could depend on how smoothly it integrates with existing patient record systems, how effectively clinicians are trained to use and interpret its outputs, whether hospital workflows can adapt to leverage its speed and how patients perceive its role in their care. Without considering these systemic elements, even a highly sophisticated AI might not reach its full potential. Viewing AI projects through a systems lens can be helpful for evaluation because:

- It may reveal a fuller scope of benefits and costs: A systems approach encourages looking beyond the direct outputs of the AI model to the wider effects on efficiency, productivity, decision-making quality and stakeholder experience. Similarly, it can help identify a broader range of costs, including those associated with integration, training, change management, and ongoing maintenance of the entire interconnected system.
- It can highlight important interdependencies: AI projects often rely on data inputs from various sources, produce outputs that feed into other processes, and involve human oversight or intervention. A systems view can help map these dependencies, potentially identifying bottlenecks or points of challenge that lie outside the AI model itself but could be relevant to its success.
- It might surface indirect and emergent effects: Complex systems can sometimes produce emergent behaviours and unintended consequences. Evaluating AI within its systemic context may allow for a more thorough anticipation of these effects, including ethical considerations, impacts on workforce roles and organisational culture shifts.

A city infrastructure analogy: the potential network effect

Infrastructure planning offers a potentially useful analogy. Imagine a city investing in a sophisticated new traffic management AI designed to optimise signal timing and reduce congestion. The AI itself might be an impressive piece of engineering. However, its success isn't solely determined by the cleverness of its algorithms. It could also depend on factors like the quality of road sensor data it receives, its integration with public transport information systems, the responsiveness of maintenance crews to system alerts and even public acceptance of dynamically changing routes.

Furthermore, much like in urban planning where a strategically placed, relatively small investment – such as a new transit hub or even a pedestrian walkway – can unlock disproportionately large benefits by connecting disparate parts of the city and improving overall flow, AI projects can sometimes have similar network effects. A well-integrated AI tool, even if not the largest or most complex, might act as a critical connector or enabler within an organisational system. This could streamline processes, help break down information silos and create cascading benefits that extend beyond its standalone capabilities. An evaluation process might aim to capture this potential for systemic leverage.

Integrating a systems view into AI project evaluation

To bring this systems perspective into the practical evaluation of AI projects, you might consider the following:

1. Mapping the ecosystem: Before assessing the AI model in detail, it could be useful to identify and map the broader system in which it will operate. This might include the people, processes, data flows and other technologies that will interact with or be affected by the AI.

2. Extending evaluation criteria: It may be beneficial to ensure that evaluation frameworks (such as those discussed earlier, including financial appraisals, BCAs, or MCAs) explicitly incorporate systemic factors. This could mean assessing integration complexity, the potential need for process re-engineering, the requirements for workforce training and adaptation, and the robustness of data governance across the system.
3. Broadening stakeholder consultation: Engaging not only with the direct users or beneficiaries of the AI but also with those involved in upstream and downstream processes, IT infrastructure, and organisational change management can provide valuable insights. Their perspectives can be helpful for understanding systemic impacts and requirements.
4. Exploring systemic risks and opportunities: Beyond the technical and ethical considerations specific to the AI model, a systems view encourages an evaluation of both risks and new opportunities. This includes assessing risks related to system integration, data dependencies and the organisation's capacity to adapt. Equally, it involves looking for opportunities for synergistic benefits, emergent capabilities, or enhanced value that arise from the AI's interaction with the wider operational context.

Ultimately, the transformative potential of AI is often most fully realised not just from the technology in isolation, but through its thoughtful and effective integration into a dynamic system of people, processes and other technologies. By considering a systems view in evaluation, organisations may make more informed investment decisions, better anticipate challenges, identify novel opportunities, and thereby increase the likelihood of AI projects delivering their intended business and societal benefits.

Conclusion

The investment opportunity in artificial intelligence is unprecedented, with global spending projected to exceed AU$980 billion annually by 2028. Yet this massive capital deployment occurs against a sobering backdrop: 80% of AI projects fail to deliver their intended outcomes. The primary cause is not technical inadequacy but poor initial decision-making—choosing the wrong AI projects to pursue in the first place.

This represents both an enormous challenge and a critical opportunity. Organisations that master rigorous *ex ante* evaluation can dramatically improve their odds of success, potentially capturing disproportionate value while competitors waste resources on ill-conceived initiatives. The frameworks outlined in this guide—from financial appraisals and cost-effectiveness analysis to stakeholder consultation and portfolio management—provide proven methodologies for making these crucial decisions.

Effective evaluation requires matching analytical depth to investment scale and complexity. Simple projects may need only basic ROI calculations, weighted scorecards and risk assessments, while transformative initiatives warrant comprehensive analysis incorporating multiple frameworks, scenario planning, and extensive consultation. The key is proportionate evaluation that enables informed decisions without creating paralysis.

Success demands viewing AI not in isolation, but as part of an integrated system involving people, processes and existing technologies. By applying these evaluation principles systematically, organisations can identify projects most likely to deliver genuine value while avoiding the costly failures that plague this field.