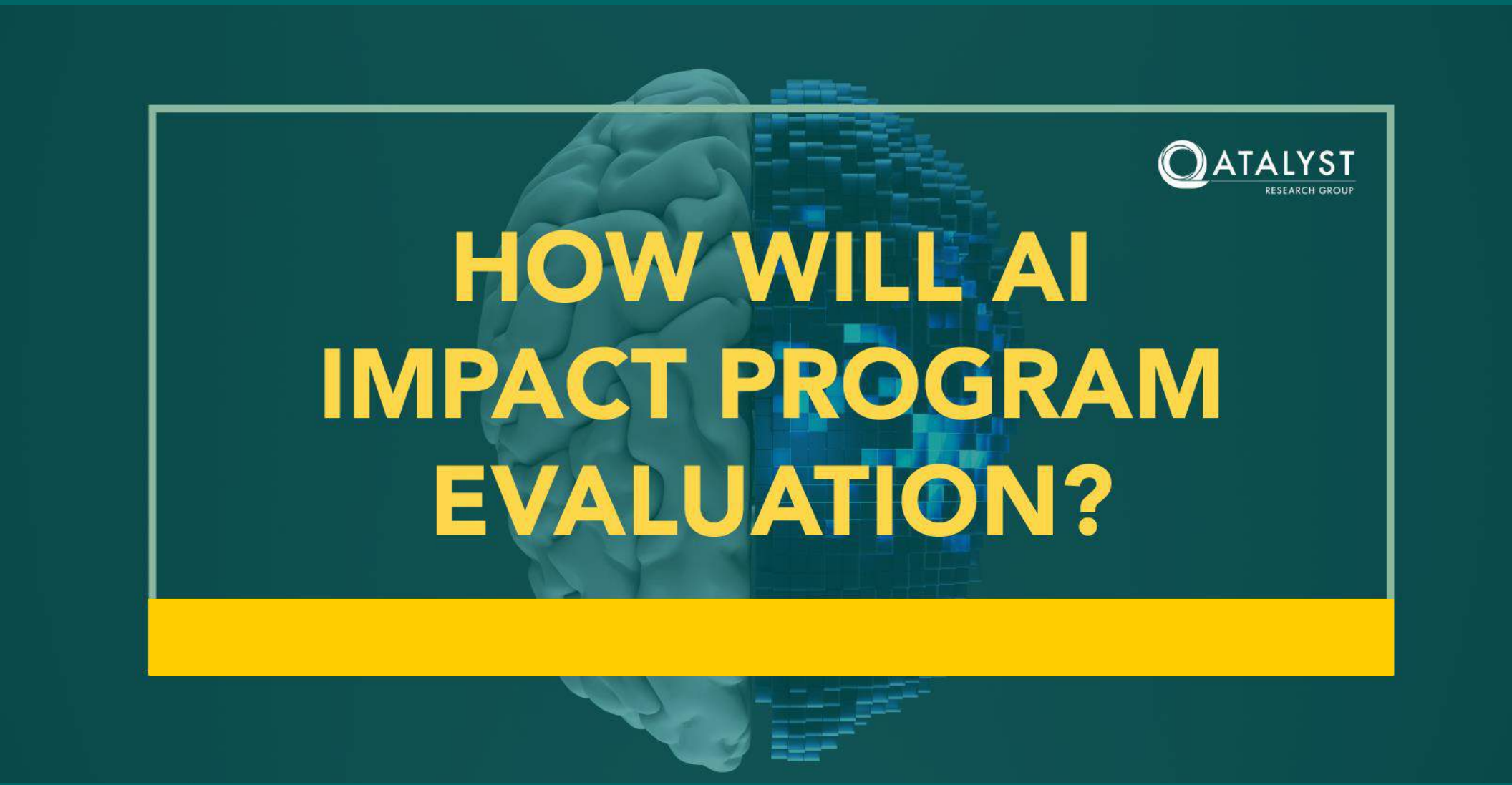

QATALYST
RESEARCH GROUP

HOW WILL AI IMPACT PROGRAM EVALUATION?

Table of CONTENTS

Introduction

Program evaluators are facing an uncomfortable reality: AI will rapidly automate core evaluation tasks, reshape client expectations, and introduce new competitors into a field that has remained relatively stable over the years. Some estimates suggest that 40% to 50% of program evaluation work could be automated in the next few years through AI.

At Qatalyst Research Group, we're deeply immersed in this transformation - integrating AI across our evaluation practice, building custom tools, and studying AI's broader impact on our profession. This article shares what we've learned about the disruption AI poses to program evaluation and, more importantly, what evaluators and evaluation firms can do to maintain and strengthen their competitive position in this rapidly evolving sector.

This blog addresses four critical questions:

- How amenable is program evaluation to AI automation?
- How will that impact the profession?
- How should you respond?
- What are your technology options?

Who Is This For?

We've written this blog primarily with program evaluators in mind. If you're a client of evaluation services or work in a related field, you'll also gain insight into how AI is reshaping the evaluation

1. How Amenable Is Evaluation to AI Automation?

A Perfect Storm of Automation-Friendly Characteristics

Management consulting is commonly cited as one of the professions most vulnerable to disruption from AI. Our analysis indicates that, within consulting, evaluation ranks near the upper end in terms of where disruption will be most profound. Why is that?

- **Evaluation draws on standardized methodologies and processes**. While approaches are tailored to each evaluation, evaluators draw from a common playbook in areas such as program profiles, logic models, evaluation matrices, analytical frameworks, and reporting. Large repositories of evaluation frameworks, workplans and reports provide rich training data for AI systems.

- **Evaluation is intensely data-driven**. AI excels at processing structured and unstructured data, including literature, program documents, administrative data, survey responses, and interview transcripts.

- **Evaluation involves repeatable workflows**. The process often follows predictable patterns: developing evaluation questions and matrices, designing data collection instruments, gathering

data, analyzing findings, and reporting results. These repeatable processes are exactly what AI automation targets.

- **Many workflows require high effort but are relatively simple to verify**. This creates the strongest business case for AI automation. AI can automate coding and analysis of interview transcripts while creating audit trails that make results easy to review. One AI app can automate extensive literature reviews that otherwise require many staff hours, while another validates citations and confirms linkages between findings and source documents. AI reduces time requirements while shifting focus from processing data to verifying and validating results.

AI can be applied across virtually the entire evaluation process:

- **Planning Phase**: AI develops program profiles and sketches approaches and methodologies. Tools like Qualtrics, Survey Monkey, and Typeform automate questionnaire development. Based on what we've learned. we could create now an app that prepares comprehensive evaluation frameworks and workplans. Even if we don't, others will.

- **Implementation Phase**: AI enables data collection and analysis at scale. Firms like Outset automate interviews while transcription apps handle recordings. NVivo, ATLAS.ti, and MAXQDA have added AI capabilities, while companies like Insight7 and CoLoop offer AI-first analysis tools. Research Rabbit, Elicit, Consensus, and Connected Papers facilitate literature reviews. At Qatalyst, we've developed our own custom tools for detailed qualitative and quantitative analysis tailored to our methodologies.

- **Reporting Phase**: AI drafts technical reports, synthesizes findings, validates citations, produces executive summaries, and creates presentations in PowerPoint or Canva.

What This Means for You

The automation potential of AI in evaluation is real and substantial. The field's characteristics - standardized processes, text-heavy workflows, data intensity, and verification-effort asymmetries - make it particularly vulnerable to disruption. However, this vulnerability also represents opportunity. Evaluators who understand where and how AI can be applied can dramatically increase productivity, take on more complex projects, and deliver greater value to clients. The key is recognizing that AI won't eliminate evaluation work - it will transform what evaluators spend their time doing.

2. How Will AI Impact the Profession?

There has been broad but not deep adoption of AI in program evaluation. While most evaluators use tools like ChatGPT for occasional tasks, few have integrated AI extensively. However, that will change sharply over the next 24 months, as client demands combined with increasing competition force deeper, more strategic adoption.

Key Changes We Expect

- **Rising Client Expectations**: Clients will expect faster turnarounds, greater breadth and depth of analysis, new deliverables (interactive dashboards, videos, real-time monitoring, predictive analytics), and custom tool development. The bar for high-quality evaluation will rise.

- **Demand and Supply Imbalance**: By reducing time required for evaluations, AI effectively increases evaluator capacity. While demand may grow following the Jevons Paradox (when resources become more efficient, overall demand increases), we anticipate demand growth will lag behind supply growth, creating market pressure. As efficiency improves and competition intensifies, evaluation firms will experience margin compression on work that previously commanded premium fees.

- **Arrival of New Competitors**:
  - AI-First Evaluation Startups: New firms built from the ground up around AI capabilities, designing business models and methodologies around AI-augmented work
  - Larger Consulting Firms: Major consultancies with heavy AI investments are expanding into evaluation markets previously considered too specialized
  - Internal Evaluation Capacity: Organizations developing internal AI-enabled evaluation capabilities, reducing outsourcing
  - SaaS Platforms: Automated evaluation platforms offering AI-powered survey design, qualitative coding, and comprehensive management systems

- **Commoditization**: Standard evaluation services (logic model development, survey design, literature reviews, qualitative coding) will become commoditized as evaluators use similar tools.

- **Push for New Forms of Differentiation**: Evaluators must find new ways to deliver value-added services: deeper subject matter expertise, specialized methodological capabilities, exceptional client service, capacity building, custom tool development, or strategic advisory services beyond traditional evaluation scope.

- **Redefinition of Roles and Skills**: Technical skills in prompt engineering, verification, AI tool selection, data pipeline design and development will grow in importance. Subject matter expertise becomes even more critical as technical execution becomes less differentiating. Client relationship skills, strategic thinking, and change management capabilities will command premium value.

- **Reduced Entry-Level Hiring**: Firms will reduce entry-level hiring as AI handles tasks formerly assigned to junior staff. This means fewer traditional entry points and a need to rethink how the field trains the next generation of evaluators.

- **Re-Engineering of Methodologies and Processes**: Over time, evaluation methodologies, outputs, and processes will be re-engineered specifically to leverage AI capabilities. We may see new evaluation approaches that weren't practical before AI - such as real-time, continuous evaluation; simultaneous large-scale qualitative analysis across a myriad of programs; or predictive evaluation that forecasts program outcomes before they occur.

What This Means for You

The evaluation profession is entering significant disruption and transformation. The comfortable equilibrium of past decades is giving way to a more turbulent environment. AI is reshaping who delivers evaluation services, how they're delivered, what clients expect, and how evaluators is valued. Professionals who will thrive recognize that adaptation isn't optional - it's essential.

3. How Should You Respond?

Adapting to AI in program evaluation requires taking deliberate steps along a continuum, choosing actions that fit your situation while accounting for broader changes in practice, people, and culture.

The Four Levels of AI Adoption

- **Laggards: No AI adoption**. These firms may be unaware, skeptical, or overloaded. The essential action is beginning experimentation with free AI tools to build comfort and familiarity.

- **Experimenters**: Individual staff use AI casually without formal approach. The next step is shifting from ad hoc use to basic integration with informal standards and practices.

- **Integrators:** Organizations embed AI in selected workflows, set quality processes, and formalize use cases and guidelines. The challenge is scaling up - customizing, refining, and connecting tools to use AI as a differentiator.

- **Innovators**: Firms with proprietary AI solutions and redesigned evaluation processes built around digital and human strengths. They lead the field but must stay alert as AI advances rapidly.

Steps You Can Take to Advance Along the Continuum

- **Focus on High-Impact Use Cases:** Evaluate cases using criteria like frequency, intensity, verification ease, data readiness, and risk level. Prioritize tasks that are frequent, time-intensive, and easily verifiable. Examples include literature reviews, coding open-ended responses, and transcription.

- **Build Your AI Strategy:** Ask key questions: What unique edge do you want from AI? Which use cases come first? What technical and governance skills are needed? How will you measure progress and impact?

- **Establish Governance:** Develop policies for use, safeguards for bias and fairness, privacy and security protocols, transparency standards for clients, and robust quality checks. These must evolve as you scale and as regulations develop.

- **Implement**: Technology readiness is essential, but it’s just one part of the journey. Effective AI deployment requires strong prompt engineering capabilities, reliable well-labeled data, active human-in-the-loop quality assurance, staff literacy in AI capabilities and limitations, auditable processes, effective governance frameworks, and rigorous results tracking.

What This Means for You

The evaluators and organizations who thrive in this AI-driven landscape will be those who take deliberate, informed action now. Waiting for clarity or perfect timing is a luxury you don't have - the competitive window is closing. Start with experimentation and honest assessment of where you stand on the adoption continuum. The path forward isn't about adopting every AI tool or becoming a technology firm; it's about strategically integrating AI to amplify your unique value, protect your competitive position, and deliver measurable impact for your clients. Your response to AI is as important as AI itself.

4. What are the Technology Options?

We see four distinct tiers of AI technology, each with its own advantages, limitations, and appropriate use cases. Most successful organizations won't choose just one tier - they will strategically deploy multiple tiers depending on the specific need, the sensitivity of the data, and their competitive objectives.

- **Tier 1: Direct Large Language Model (LLM) Interfaces**. These are consumer platforms like ChatGPT, Claude, Gemini, or Perplexity through web or API interfaces. It's ideal for quick starts, experimentation, and individual tasks - drafting questions, summarizing documents, brainstorming. These tools are cheap, accessible, and require no setup, but have limits on security, workflow integration, and cannot access your organizational knowledge. *Best fit: Individual evaluators, freelancers, and small firms testing AI without infrastructure investment.*

- **Tier 2: Commercial AI-Powered Applications**. This covers established platforms (e.g., NVivo, MAXQDA, Qualtrics, Outset, Insight7, CoLoop) with AI integrated into workflows. They offer collaboration features, audit trails, and improved data security. However, most offer narrow functionality (coding or literature reviews) and limited customization. *Best fit: Teams that value compliance, user management, and predictable costs and who already use these platforms.*

- **Tier 3: AI Orchestration Frameworks and Scaffolding**. These systems connect LLMs to your proprietary data, methodologies, and workflows (using scaffolding and agents). They can embed your organization's knowledge, automate stages, and facilitate custom analytics and dashboards. While requiring some technical skill and setup costs, those barriers are declining as software coding itself is being automated. *Best fit: Firms with unique approaches, existing technical capacity, and a need for secure, custom integration.*

- **Tier 4: Fully Custom Platforms**. Here, organizations build AI tools and platforms from scratch - complete control over models, interfaces, and algorithms. This is resource-intensive, but enables long-term scaling, high security, and IP creation. Best fit: L*arge firms with advanced technical teams looking to productize their capabilities or deliver new market offerings.*

Factor	Tier 1: Direct LLM	Tier 2: Commercial Apps	Tier 3: Orchestration	Tier 4: Custom
Cost	Low	Medium	Medium-High	Very High
Time to Value	Days	Days-Weeks	Weeks-Months	Months-Years
Technical Skill Required	None	Minimal	Moderate	High
Data Security	Low	Medium-High	High	Very High
Customization	None	Limited	Extensive	Complete
Scalability	Limited	Good	Excellent	Excellent
Organizational Learning	Minimal	Low	High	Very High
Best Fit	Individuals/ freelancers	Small/ medium teams	Firms with proprietary methods	Larger orgs with technical expertise

At Qatalyst, we have focused largely on Tier 3 although we regularly use Tier 1 and have developed an internal platform to integrate our tools (Tier 4).

What This Means for You

Understanding these four tiers provides a framework for strategic technology choices. You don't need to jump immediately to custom development - Tier 3 orchestration offers a middle path between capability and feasibility. Start with Tier 1 to build AI literacy, adopt Tier 2 tools for high-volume standardized tasks, consider Tier 3 for differentiation opportunities, and reserve Tier 4 for specialized needs or long-term strategic investments. The key is moving deliberately and strategically rather than ignoring AI entirely or investing beyond current needs and capabilities.

Our Upcoming Webinar

This fall, we're hosting a webinar for evaluators on how you can leverage AI to strengthen your services and operations. If you'd like an invitation, let us know and we'll share the details.

If you prefer a one-on-one conversation to discuss your specific situation, we're happy to do that as well. We can help you assess your AI readiness, identify high-priority use cases, and develop a practical strategy for AI adoption in your evaluation practice.

Later this year, we'll be launching an online service where you can assess your organization's AI readiness and develop a strategy to improve your performance. We can also provide consulting services to help you review your operations, prioritize use cases, and develop or integrate the AI tools you need.

If you're interested, reach out to Sushant Tare at sushant.tare@qatalyst-research.ca

About Qatalyst Research Group

Qatalyst Research Group is a management consulting firm headquartered in Vancouver, Canada, specializing in program evaluation, policy research, and strategic advisory services for government and non-profit organizations. We've been at the forefront of integrating AI into evaluation practice, developing custom tools and approaches that combine AI capabilities with deep evaluation expertise.

This blog post is based on materials prepared for our November 2025 webinar, "How Will AI Impact Program Evaluation?"

QATALYST
RESEARCH GROUP

CONSULTING EXPERTISE, AMPLIFIED BY AI

Blending human insight and advanced technology for faster, more effective results.

At Qatalyst Research Group, we are integrating AI into every aspect of our practice —from sales to research, analysis, and output generation. Combining senior consulting expertise, an in-house development team, and a suite of AI tools, we can assist non-profits, government and other firms in integrating AI into their operations. We offer end-to-end AI transformation services that meet organizations where they are in their digital journey.

Go to https://qatalyst.ca/ or reach out to Sushant Tare at sushant.tare@qatalyst-research.ca to find out more about how we might work together.