

# Directing AI: Charting a Roadmap of AI Opportunities and Risks

Mark Campbell , EVOTEK

Mladen Jovanović , Singidunum University

*While generative artificial intelligence's swift adoption has presented significant security, technical, and cultural concerns, the unprecedented opportunities it introduces cannot be ignored. New techniques are emerging to help companies benefit from these opportunities while mitigating risks.*

**W**hile artificial intelligence (AI) has continuously evolved for over five decades, the recent introduction of affordable, scalable, and readily deployable AI solutions has significantly disrupted all sectors of the global economy.<sup>1,2</sup> In particular, the swift rise of generative AI

capabilities presents both unprecedented opportunities and risks for today's corporate enterprises.<sup>3</sup>

## INTRODUCTION

Conventional AI solutions address tasks like recognition (such as image, facial, fingerprint, and voice recognition), decision-making (including risk profiling, route calculation, and workflow management), and anomaly detection (covering predictive maintenance, cybersecurity threats, and user behavior analysis). These solutions process extensive datasets, learn patterns and nuances, and then produce classifications when presented with new data.

Generative AI brings sophistication beyond just classification by

empowering machines to autonomously generate artifacts including images, language, data, code, and processes.<sup>4</sup> The speed and widespread impact of generative AI have taken many companies completely off guard, leaving them ill equipped to harness AI's potential benefits or safeguard against its potential pitfalls.<sup>5</sup> Moreover, generative AI introduces significant ethical, legal, and societal concerns that surpass the purview of any individual company or governing body.<sup>6</sup>



## LEVELS OF GENERATIVE AI DISRUPTION

Every emerging technology influences the market: some cause incremental shifts in the status quo, while others completely disrupt entire industries.<sup>7</sup> Generative AI is proving to be the latter,<sup>8</sup> disrupting enterprises across four dimensions: corporate, strategic, tactical, and operational.

### Corporate disruption

Corporate planning involves defining a company's "why": its mission, vision, market, customers, regulatory constraints, and shareholders.<sup>9</sup> Although crucial for examination, a comprehensive exploration of the disruption generative AI introduces at the corporate level exceeds the confines of this article.

### Strategic disruption

The strategic level of corporate planning addresses the "what" when accomplishing the corporate "why" including "What do we do?," "What do we need?," and "What must change?" Setting strategic objectives requires careful analysis of market dynamics, competitive landscapes, regulatory considerations, and an honest appraisal of the organization's own capabilities, limitations, and resources.<sup>10</sup> Moreover, strategic planning requires a keen awareness of the risks associated with maintaining the status quo versus embracing emerging technologies, such as generative AI, for a competitive edge.<sup>11</sup>

Generative AI introduces several opportunities companies must evaluate in their strategic planning including revolutionizing the customer experience, cost optimization realized through automated processes and staff realignment, and creating a competitive advantage over lagging competitors.<sup>11</sup>

Generative AI also presents strategic risks or existential perils to companies, such as employees not understanding generative AI application pitfalls and

corporate policies not addressing the copyright, data leakage, and security risks of using generative AI.<sup>12</sup> Currently, "shadow AI" is rampant and leaves many companies unaware of what AI tools and services are used or what risks they introduce.<sup>13</sup> However, there is also an overarching risk of not using generative AI and being eclipsed by a more adventurous competitor.<sup>14</sup>

To take advantage of generative AI opportunities while mitigating risks requires strategic objectives that create corporate cohesion, reasoned implementation, resource realignment, and competitive advantage.

### Tactical disruption

Tactical corporate planning defines the "who" component of the objectives described by strategic planning by answering "Who will do this?," "Who do we need?," and "Who are we missing?" Tactical objectives direct cross-functional initiatives requiring technical alignment, interorganizational processes, and coordinated programs spanning lines of business.<sup>15</sup>

Generative AI offers many tactical opportunities including the creation of generative applications and products, intelligent and adaptive processes, in-house large language models (LLMs) fine-tuned with the company's proprietary data,<sup>16</sup> or an AI center of excellence (COE) to concentrate AI expertise, governance, and processes.<sup>17</sup>

Several tactical risks are associated with generative AI, including the potential for creating applications with embedded bias or lacking transparency into their decision-making process.<sup>18</sup> These generative applications may also pose risks such as violating copyrights, leaking sensitive information, and generating libelous content.<sup>12</sup> Not only do in-house LLMs introduce new attack surfaces for malicious actors, but if not properly engineered, they can also contain bias,

produce inaccurate artifacts (referred to as hallucinations), and give a false sense of infallible authority.<sup>19</sup> Additionally, cybercriminals are leveraging generative AI to create adaptive and sophisticated cyberattacks, necessitating a large-scale revamping of cyberdefense strategies, tools, and skills.<sup>20</sup>

### Operational disruption

Operational corporate planning tackles the "how" of a company's tactical objectives and addresses questions such as "How will we do this?," "How will this change things?," and "How do we run, operate, and measure this?"

Generative AI is disrupting technical and business operations by introducing opportunities to automate the creation of text, code, voice, images, video, and data. For example, the emergence of LLMs in 2022 like ChatGPT, Bard, Anthropic, and GitHub CoPilot enabled companies to write content, produce applications, create legal documents, craft communications, develop test harnesses, and document procedures without human labor, freeing staff to pursue more profitable corporate tasks.<sup>21</sup>

However, this automation introduces substantial operational risks including making it difficult to understand, update, fix, or maintain code written by AI.<sup>22</sup> Many companies rely on external applications and services that expose them to all risks within the third party's ecosystem. In addition, introducing generative AI creates a much more nuanced attack surface vulnerable to new threats such as data poisoning, model hijacking, prompt injection, and adversarial inputs.<sup>23</sup>

## AI ROADMAPPING

As generative AI ushers in a flood of innovations and threats, companies must reappraise their corporate, strategic, tactical, and operational objectives, which can be a daunting task. Some jump right into generative application

development opportunistically for high visibility. Others take a more conservative “wait and see” approach until platforms, products, regulations, and features are more predictable. However, there are a few firms taking a measured, programmatic approach by charting a stratified roadmap to simultaneously create new opportunities and avoid risks as their resources, risk tolerance, constraints, market, and culture allow. This five-phase roadmapping technique (Figure 1) effectively sets and resets technical objectives in a structured yet nimble process that adapts to the rapidly changing generative AI landscape.

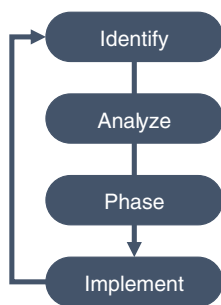


FIGURE 1. The AI roadmapping process.

**Identify**

Start by identifying a lattice of generative AI opportunities for each of the strategic, tactical, and operational layers. Although there are hundreds of possible use cases for each layer, identify the top four opportunities and risks for each layer: 12 opportunities and 12 risks. These use cases will change as technology matures, skills expand, regulations emerge, and resources adjust.

**Analyze**

Analyze each of the 24 use cases for its potential impact to the company’s objectives by classifying it as high impact, some impact, or no impact (Figure 2). Similarly, each use case is then evaluated for its feasibility considering existing technology, skills, budget, resources, time, regulations, and culture and rated as easy, hard, or intractable.

**Phase**

After analysis, group use cases into five categories regardless of layer or whether they are an opportunity or risk:

1. easy: high impact
2. easy: some impact

3. hard: high impact
4. hard: some impact
5. other: includes all use cases with no perceived impact or currently intractable.

Then select a set of use cases to implement immediately and table the rest for later development. Begin with category 1 use cases (that is, easy: high impact) and work down the ranking categories sequentially. The number of use cases selected for immediate implementation will vary depending on the company’s capabilities, resources, and risk appetite, but typically, first iterations have smaller scope, and later phases grow larger as proficiencies increase. Once the initial set of use cases is selected, schedule its start based on the company calendar, resource availability, and budget constraints (Figure 3).

**Implement**

Implementing a selected use case begins by assessing the current state of technology, market, skills, budget, timelines, regulations, and competition. The assessment may determine that the use case is more nuanced than expected and needs appropriate adjustments

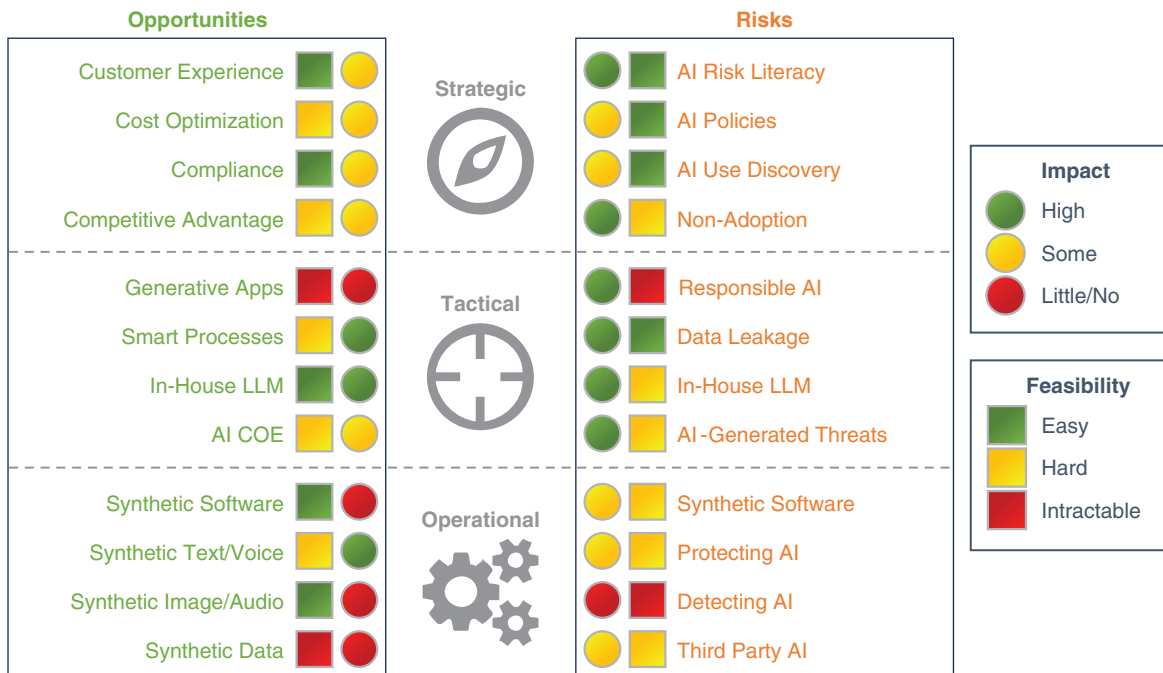


FIGURE 2. An AI roadmapping identify and analyze example.

to its impact and feasibility rankings. The assessment process also identifies the implementation approach for each use case (for example, buy versus build versus partner, proof of concept versus full deployment, etc.), as well as the expected results and success criteria.

Once assessed, the use case is chartered, kicked off, developed, and deployed using the company's defined development life cycle(s) and processes. It is important to note that generative AI development differs substantially from typical waterfall or agile software development life cycles, so newer processes may need to be adopted or adapted.

### Iterate

After initial use cases are deployed, repeat the process by reevaluating previous use cases and adding ones that have surfaced in the interim. However, after several iterations the AI roadmapping foundations are in place, and iteration cycle times will decrease while implementation use case sets will grow.

## FRAMEWORKS

Several frameworks have been released or are under development to identify risks, mitigations, and use cases and provide organizations with implementation and governance guidance, including the following:

- › *The IEEE AI Impact Use Cases Initiative*<sup>24</sup> provides a searchable


While the proposed AI roadmapping process leverages these and other

repository and taxonomy for risk-based AI use cases, their impacts, and a set of risk principles.

- › *The OWASP Top 10 for LLM Applications*<sup>25</sup> provides a list of risks to educate organizations on the potential security risks introduced when deploying LLMs.
- › *The NIST AI Risk Management Framework*<sup>26</sup> provides guidance to organizations on the risks associated with AI to “incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems.”
- › *The MITRE Adversarial Threat Landscape for Artificial Intelligence Systems (ATLAS) Framework*<sup>27</sup> provides a “living knowledge base of adversary tactics and techniques based on real-world attack observations and realistic demonstrations from AI red teams and security groups.”
- › *The Data Provenance Initiative*<sup>28</sup> provides a repository of over 1,800 popular textual datasets used to train and fine-tune generative AI models, along with comprehensive meta-information for each describing the origin, use, data sources, licenses, and creators.

emerging frameworks, it provides several advantages over using them in isolation:

- › *comprehensiveness*: equally considers opportunities, risks, and their connections to strategic, tactical, and organizational objectives
- › *structure*: provides a formal and systematized method to identify, analyze, and implement generative AI opportunities and risks in customizable phases
- › *effectiveness*: creates a solid grounding for practical implementation of prioritized use cases by focusing on immediate value delivery or risk mitigation and deferring all ancillary use cases to a later date
- › *nimble*: allows rapid maturation of underlying generative AI technologies and use cases by iterating back through the identification and analysis phases after each implementation cycle.

While generative AI's swift adoption poses many societal, competitive, security, technical, and cultural concerns, the unprecedented opportunities it introduces cannot be ignored. A new wave of processes, techniques, and frameworks is emerging to help companies take advantage of generative AI's promises while mitigating its perils. Aligning a company's strategic, tactical, and operational objectives into an evolving AI roadmap is the first step. 

## ACKNOWLEDGMENT

We acknowledge the substantial work of Dr. Rick Hubbard, whose “accelerated processes” and “plan on a page” methodologies underpin many of the concepts and recommendations offered here. The corresponding author is Mark Campbell.

## REFERENCES

1. “Get your LLM application from prototype to production.” LangChain.

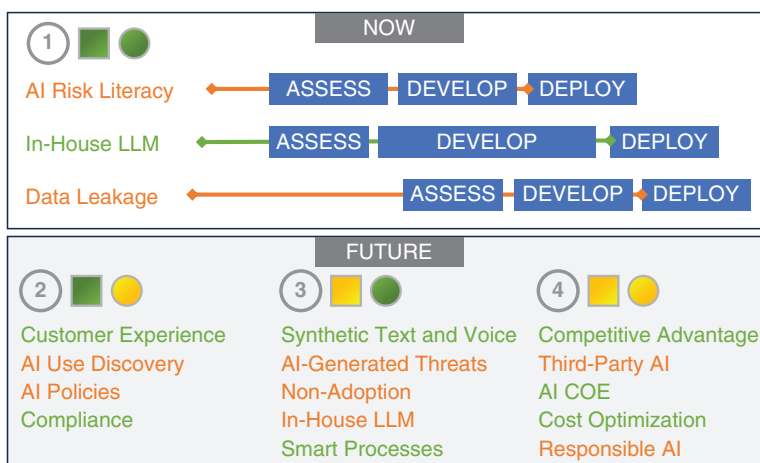


FIGURE 3. An AI roadmapping phasing example.



- Accessed: Nov. 27, 2023. [Online]. Available: <https://www.langchain.com/>
2. "Unleash the power of LLMs over your data." Meta. Accessed: Nov. 27, 2023. [Online]. Available: <https://www.llamaindex.ai/>
  3. A. Iyer and K. Kenthapadi. "Introducing Fiddler Auditor: Evaluate the robustness of LLMs and NLP models." Fiddler.ai. Accessed: Nov. 27, 2023. [Online]. Available: <https://www.fiddler.ai/blog/introducing-fiddler-auditor-evaluate-the-robustness-of-llms-and-nlp-models>
  4. M. Campbell and M. Jovanovic, "Generative artificial intelligence: Trends and prospects," *Computer*, vol. 55, no. 10, pp. 107–112, 2022, doi: 10.1109/MC.2022.3192720.
  5. S. Mohiuddin. "Reducing hallucinations with provenance guardrails." Guardrails.ai. Accessed: Nov. 27, 2023. [Online]. Available: <https://www.guardrailsai.com/blog/reduce-ai-hallucinations-provenance-guardrails/>
  6. B. C. Stahl and D. Eke, "The ethics of ChatGPT – Exploring the ethical issues of an emerging technology," *Int. J. Inf. Manage.*, vol. 74, Feb. 2024, Art. no. 102700, doi: 10.1016/j.ijinfomgt.2023.102700.
  7. A. Sigov, L. Ratkin, L. A. Ivanov, and L. D. Xu, "Emerging enabling technologies for industry 4.0 and beyond," *Inf. Syst. Frontiers*, pp. 1–11, Jan. 10, 2022, doi: 10.1007/s10796-021-10213-w.
  8. C. Bash, P. Paraboschi, E. Frachtenberg, P. Laplante, D. Milojevic, and R. Saracco, "Megatrends," *Computer*, vol. 56, no. 7, pp. 93–100, Jul. 2023, doi: 10.1109/MC.2023.3271428.
  9. M. Menz et al., "Corporate strategy and the theory of the firm in the digital age," *J. Manage. Stud.*, vol. 58, no. 7, pp. 1695–1720, 2021, doi: 10.1111/joms.12760.
  10. G. C. Kane, D. Palmer, A. N. Phillips, D. Kiron, and N. Buckley, "Strategy, not technology, drives digital transformation," *MIT Sloan Manage. Rev.*, Jul. 14, 2015.
  11. D. Kiron, M. Schrage, F. Candelon, S. Khodabandeh, and M. Chu, "Strategic alignment with AI and smart KPIs," *MIT Sloan Manage. Rev.*, pp. 1–6, Sep. 5, 2023.
  12. J. Mökander, J. Schuett, H. R. Kirk, and L. Floridi, "Auditing large language models: A three-layered approach," *AI Ethics*, vol. 3, no. 4, pp. 1–31, May 30, 2023, doi: 10.1007/s43681-023-00289-2.
  13. V. Uren and J. S. Edwards, "Technology readiness and the organizational journey towards AI adoption: An empirical study," *Int. J. Inf. Manage.*, vol. 68, Feb. 2023, Art. no. 102588, doi: 10.1016/j.ijinfomgt.2022.102588.
  14. C. Ebert and P. Louridas, "Generative AI for software practitioners," *IEEE Softw.*, vol. 40, no. 4, pp. 30–38, Jul./Aug. 2023, doi: 10.1109/MS.2023.3265877.
  15. S. Rahmanzadeh, M. S. Pishvaei, and M. R. Rasouli, "Integrated innovative product design and supply chain tactical planning within a blockchain platform," *Int. J. Prod. Res.*, vol. 58, no. 7, pp. 2242–2262, 2019, doi: 10.1080/00207543.2019.1651947.
  16. K. Wiggers. "Ikigai lands \$25M investment to bring generative AI to tabular data." TechCrunch. Accessed: Nov. 27, 2023. [Online]. Available: <https://techcrunch.com/2023/08/24/ikigai-lands-25m-investment-to-bring-generative-ai-to-tabular-data/>
  17. A. Mahurkar, "Four steps for building an AI center of excellence," *Forbes*, Dec. 9, 2022. [Online]. Available: <https://www.forbes.com/sites/forbes-techcouncil/2022/12/09/four-steps-for-building-an-ai-center-of-excellence/?sh=4d6d93c72ec5>
  18. D. Ganguli et al., "Predictability and surprise in large generative models," in *Proc. ACM Conf. Fairness, Accountability, Transparency*, New York, NY, USA, 2022, pp. 1747–1764, doi: 10.1145/3531146.3533229.
  19. V. Rawte, A. Sheth, and A. Das, "A survey of hallucination in large foundation models," 2023, *arXiv:2309.05922*.
  20. M. Gupta, C. Akiri, K. Aryal, E. Parker, and L. Praharaaj, "From ChatGPT to ThreatGPT: Impact of generative AI in cybersecurity and privacy," *IEEE Access*, vol. 11, pp. 80,218–80,245, Aug. 2023, doi: 10.1109/ACCESS.2023.3300381.
  21. E. Brynjolfsson, D. Li, and L. R. Raymond, "Generative AI at work," National Bureau of Economic Research, Cambridge, MA, USA, 2023. [Online]. Available: <https://www.nber.org/papers/w31161>
  22. S. Bankins, A. C. Ocampo, M. Marrone, S. L. D. Restubog, and S. E. Woo, "A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice," *J. Org. Behav.*, early access, Aug. 2, 2023, doi: 10.1002/job.2735.
  23. Y. Liu et al., "Prompt injection attack against LLM-integrated applications," 2023, *arXiv:2306.05499*.
  24. "The IEEE AI impact use cases initiatives," IEEE, Piscataway, NJ, USA, 2023. Accessed: Nov. 25, 2023. [Online]. Available: <https://standards.ieee.org/industry-connections/ai-use-cases-initiative/>
  25. "OWASP Top 10 for large language model applications," OWASP, Wakefield, MA, USA, 2023. Accessed: Nov. 25, 2023. [Online]. Available: <https://owasp.org/www-project-top-10-for-large-language-model-applications/>
  26. "AI risk management framework," NIST, Gaithersburg, MD, USA, 2023. Accessed: Nov. 25, 2023. [Online]. Available: <https://www.nist.gov/itl/ai-risk-management-framework>
  27. "Atlas." MITRE. Accessed: Nov. 25, 2023. [Online]. Available: <https://atlas.mitre.org/>
  28. "Data provenance explorer." Data Provenance Initiative. Accessed: Nov. 25, 2023. [Online]. Available: <https://www.dataprovenance.org>

**MARK CAMPBELL** is the chief innovation officer at EVOTEK, San Diego, CA 92121 USA. Contact him at [mark@evotek.com](mailto:mark@evotek.com).

**MLADAN JOVANOVIĆ** is an associate professor of computer science at Singidunum University, 11000 Belgrade, Serbia. Contact him at [mjovanovic@singidunum.ac.rs](mailto:mjovanovic@singidunum.ac.rs).