

The Quest for Responsible AI

Given the increasing popularity of Generative AI (GenAI), questions of ethical and responsible AI development have become imperative. Risk mitigation is key during such an uncertain chapter in human history. Sure, the phrase “*may you live in interesting times*” sounds like a wish of good fortune, but the global pandemic has recently taught us that living in interesting times may be anything but fortunate. Today, just a few years after the onset of COVID-19 — in the face of AI-generated articles, art, and deepfakes — it’s hard to discern whether these will be times of good fortune or deep concern for the future of humanity.

According to Gartner® [1], responsible AI can be described as follows:

“...Responsible AI is an umbrella term for aspects of making appropriate business and ethical choices while adopting AI in the organization’s context. These include business and societal value, risk, trust, transparency, fairness, bias, mitigation, explainability, accountability, safety, privacy, regulatory compliance and others.”

Whether as individuals, corporate players, investors, or employees, it is no longer a question of whether we should or should not use AI to support us in our lives. AI has the potential to empower communities, combat climate change, and increase efficiency within businesses across industries, and it is being built into the devices and software that we use from all our major providers. The quest must now be to find ways of working with AI that preserve the rights of enterprises, employees, public users, and other related stakeholders. Beyond that, we have a responsibility to consider the adverse effects of our AI creations and their deployment to ensure that it is safe in as many ways as possible including privacy of information, use of resources, and the mitigation of harm to people, the planet, and all other life forms on it.

“We’re change optimists, and so we always think technology can change the world, and we believe that generative AI is like that as well. But now more than ever, and with this particular technology, we have to make sure we get it right [...] ensuring we have the guardrails [...] in place.”

- Sam Sebastian, VP Google Cloud Canada, at [The Walrus Talks Artificial Intelligence](#) [2]

As a Canadian Indigenous company, kama.ai has considered these questions and design goals since its inception and the initial development of our solution. In the creation of a transparent, contextually aware intelligence guided by human values, our goal of developing a responsible AI has been at the heart of our human design-centred process. In the following sections, we will review what we believe to be the relevant issues guiding responsible AI, all of which we hold ourselves accountable to.

Responsible AI Tool Use

AI tools for personal and enterprise use are expanding at an unprecedented rate; the pace is difficult for enterprise policy and administration staff to stay ahead of.

Enterprises are encouraged to choose vendor and business advisory partners that have already assessed the various tools available on the market and/or integrated tools into their platforms in ways that support responsible use.

Explainability – How does the AI you are considering work? Can your vendor clearly and meaningfully explain their solution? Does that explanation give you confidence that the solution will deliver predictable results for your organization?

Transparency – Where and what is the data that AI uses? Can you review and edit that data so that you have confidence in the outcomes?

Diversity – Does the AI solution allow your organization to include and/or service diverse user groups? Can it represent certain interests that may be important to minority communities? Are there risks of outcomes which may be driven by biased data? Is it open to be ‘programmed’ by diverse groups to suit their culture or community interests?

Accessibility – Does the solution you are considering adopt accessibility methods which allow persons with all abilities to use it for their needs? Is the solution accessible from a technical perspective or does it require a high degree of technical skills and other resources to operate or program?

Corporate Alignment and Governance – Are you sure that your AI solution is aligned, and will remain aligned, with your organizational values and stay true to your brand? Are your decision makers and strategists able to contribute to the voice and shape the personality of your digital intelligence? Are there effective ‘human-in-the-loop’ checks and balances that create a healthy and safe balance between the advantages of generative tools and the oversight of your organization’s policy, brand, equity, and related subject matter experts or other relevant stakeholders?

Moral/Ethical Compass – Is there any opportunity for your AI solution to have some societal grounding in morals or ethics that is generally appreciated by humans? Are there open communication channels between society and your AI through which it can uncover new concerns within the market or society at large?

Social Good – Are there applications of the technology that provide benefits besides those based on shareholder wealth and return on investment? Is the technology commercially accessible to not-for-profits or other social organizations with moderate or limited budgets?

Environment & Planet – Is the platform efficient to operate on an inquiry-by-inquiry basis, or does it consume large amounts of precious resources placing additional strain on our planet?

kama's Model for Responsible and Ethical Artificial Intelligence

Guiding the development of its own form of artificial intelligence, kama.ai has outlined the requirements necessary to achieve an ethical, responsible, and sustainable conversational AI platform. This model is inspired by the Gartner '[Best Practices for the Responsible Use of Natural Language Technologies](#)' and the International Association for Privacy Professionals' '[AI Governance Professional Body of Knowledge](#)', in addition to the newly released '[Canadian Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems](#)'.

kama.ai's Model for Responsible and Ethical AI

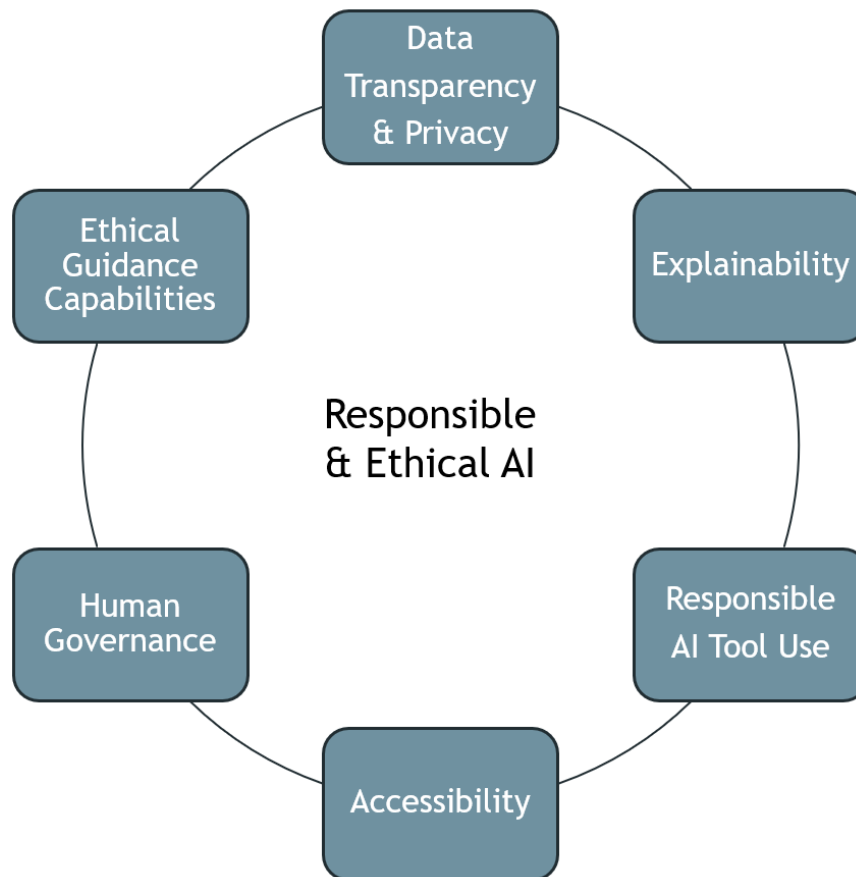


Figure 1: Responsible & Ethical AI Model

kama's Designed Experiential Intelligence® Responsible Conversational AI Platform

While many forms of AI can be valuable to enterprises or the broader society, kama's mission is focused on a conversational intelligence platform that powers virtual assistants in the service of humanity. Conversation and communication are at the heart of the human experience and the basis of an informed and progressive society.

At kama, we believe that an AI must be contextually aware of the human condition. To support this, we adopted a [human-centred design](#) approach and based our foundation on a proprietary [knowledge graph](#) that operates much like our human minds. It supports an infinitely interconnected web of related knowledge and information that can be informed by both external sources and your own organization's information about products, services, and organizational values. Together, this information creates a human behavioural model that can be sensitized to a user's context, but it is equally informed by your organization's 'personality' to create the best possible automated engagement between an organization and its respective stakeholders.

With the recognition that human experience, not algorithms, will design and govern the user journeys and information or knowledge delivered to users through the platform, we called our platform a **Designed Experiential Intelligence®**, or **kama DEI**.

To provide ethical guidance to the kama DEI behavioural model, we introduced a patented control and evaluation layer of over 150 human values that cross all cultures and societies (e.g. awareness, communication, compassion, safety, security, style etc.). We then developed overarching natural language understanding (NLU) and problem-solving layers that can recognize human issues and related values across contextual issues detected from user inputs or inquiries.

In operation, the system decodes user inputs into one or more intentions or 'intents'. The NLU can detect and decode multiple intents within a user input. An example of this is *"I'm hungry but I'm too tired to go out"* or *"We really need a vacation but we don't want to go far"*. The NLU prepares the input into separate or combinations of intents that can be compared against the knowledge graph to determine the contextual situation(s) the user is facing. The value-rating evaluation method provides human-like prioritization of either problems, solutions or even potential inferences to be validated with the user through dialogue or intelligent triage.

Application of the Responsible AI Model within the kama DEI Platform

Even with kama DEI's particular human-like form of intelligence, we acknowledge the need to provide an AI that can be trusted by our enterprise customers, and by their members, consumers or other stakeholders. The following outlines the application of the kama's Responsible and Ethical AI Model to the kama DEI Platform.

Responsible AI Tool Use

We recognize that AI solution vendors, including kama, may incorporate 3rd party or Open Source AI tools within their platforms. Whether it is the core AI solution, or 3rd party sub-system AI enablers embedded in the overall system, we incorporated a design requirement ensuring that AI components are used responsibly according to the model described in Figure 1 above.

In particular, using generative tools inside kama DEI must ensure that prompts and information used to generate response drafts for kama DEI Knowledge Managers must not cross-pollinate sensitive information from one of our customers to another. Special care is taken to use generative tools inside kama DEI to protect enterprise data from distribution. This extends to human-in-the-loop governance that supports the enterprise in ensuring the proper review of draft-generated information before securing it in our graph knowledge base for safe and governed distribution.

Additionally, the responsible use of AI tools means that tools should also be used efficiently from an energy standpoint to not further endanger our environment by wasting precious energy resources. To us at kama, that means that if there are energy-intensive enablers used, such as generative models, they should be used from time to time to assist organizational knowledge curators or kama DEI Knowledge Managers. However, these highly energy-consuming tools should not be used in every interaction of the end-user inquiry and response flow. Once draft responses are reviewed and saved in our knowledge base, they are intelligently and contextually retrieved for end-user consideration with energy consumption that is orders of magnitude less than that used to support real-time generative responses.

Accessible AI

Ensuring that an AI platform or technology is accessible to persons with all physical abilities, and by persons without deep technical knowledge, is also a form of responsible and ethical AI.

Providing a zero-code and accessible approach to 'programming' AI allows organizations to deliver authentic and truthful responses that reflect the '[voice of \(the\) society](#)' (or micro-society) using the platform.

The accessible zero-code kama DEI platform allows communities, cultural groups, or other communities of practice, to 'voice' their own authentic responses in the culture, language and values that they align with.

Data Transparency & Privacy

Having a transparent AI means that an organization has a full view of how the AI operates and where its response data or information comes from.

The data should be able to be fully discoverable within the platform in clear language so that the organization's AI management, subject matter experts, or other stakeholders, can view and edit the data and the user messaging deliberately and efficiently to deliver and reflect their particular desired outcomes and values respectively.

With kama DEI, data is curated by subject matter experts, working as, or in conjunction with, kama DEI Knowledge Managers using simple natural language input resulting in transparent and clear data that can be reviewed and edited at any time without the need for highly technical skilled resources.

Tightly coupled with the way that we use AI tools within our platform, and the way we design our own AI platform, is the issue of data privacy. Personal information privacy is not only important, it is becoming more and more regulated in global jurisdictions. Regulations such as the General Data Protection Regulation ([GDPR](#)) in the EU, The Personal Information Protection and Electronic Documents Act

([PIPEDA](#)) in Canada and various federal and state privacy regulations in the United States all seek to protect the public from identity fraud or other undesired outcomes. Our product management is continually reviewing new privacy regulations and best practices to enhance our product and roadmap to keep user data safe and secure.

Human-in-the-Loop Governance

Due to limitations with current AI tools, it is kama.ai's opinion that enterprises must incorporate human review before information delivery to employees, clients, or other stakeholder end-users. Human reviewers can ensure that information is truthful, unbiased, on brand, and in line with the organization's values.

kama.ai's platform began from the perspective that all enterprise information is curated by the enterprise. Since then, advances in generative AI have encouraged us to leverage these generative capabilities. Still, our philosophy remains that only enterprise 'governors' can provide enterprise governance. Whether that governance comes from a deep understanding of the response material via subject matter experts, keeping messaging on-brand with marketing expertise, or aligning with organizational policies of equity and inclusion, it is humans who should make the final call on the end-user payload.

With the introduction of generative AI into the kama DEI platform, we leverage the industry's advanced approach to Retrieval Augmented Generation or 'RAG'. This approach produces truly amazing results, but it must be done with safeguards, and it currently must only be used for drafting purposes; the results must be received by real humans, every time. There may come a time when generative responses are trustworthy enough to interact directly with stakeholders; in the meantime, the humans within enterprises will approve content for distribution within the kama DEI platform.

Explainable AI

Explainability adds confidence to organizations looking to produce predictable and truthful automated responses with AI. If an AI's process is more understandable, the organization can have greater confidence that the AI will behave in ways that are predictable, acceptable and ethical, and that it will not introduce risk to the organization's brand, reputation, users or user communities accessing the system/service.

With kama DEI, our knowledge base, our values-based inference algorithm, and our response data are clear and understandable to non-technical personnel. If there is any question or need for change, response or payload information is immediately discoverable and editable.

Our human value-rating algorithm, which can detect one or more problems in user inquiries and deliver one or more prioritized solutions, is also very simple, straightforward, and explainable to non-technical laypeople.

Emotion AI as an Ethical Enabler

Emotional Intelligence or Emotion AI is a class of AI technologies that incorporate or detect some forms of human emotion.

Common forms of Emotion AI are Sentiment Analysis, which looks for words that convey positive or negative emotional states, (video) Facial Analysis, which detects emotional states by looking at the gestures in facial expressions, and Tonal Analysis, which detects emotions by listening to spoken voice.

kama DEI's Emotional Intelligence can detect multiple states of emotion by analyzing the combinations of words in phrases and evaluating these in the priority of the communicating user's personal values. The human value-rating approach facilitates several key conversational functions including: i) multiple prioritized recommendations to a particular 'problem', ii) prioritizing multiple problems or issues found in a single utterance from a user, and, iii) clarifying the most important potential issue in a conversational 'triage'.

Not all forms of Emotion AI are inherently responsible AI, but the ability to gain a human-like understanding of issues arising from/underlying a user's conversation can inform the natural direction of a conversation. Furthermore, recognizing or confirming human contextual issues or situations allows for more empathetic conversations that are 'responsible' to the human condition.

Creating a Designed Experiential Intelligence®

Experiential Intelligence' (or 'XQ'), is a term coined by the past president of the American Psychological Association, Robert Sternberg. Sternberg recognized practical experience as an asset that can predict job success better than intelligence. People with XQ have gained tacit knowledge or 'posteriori (self-learned) knowledge', that allows them to become trusted subject matter experts or knowledge-keepers in specific domains. This combination of knowledge, experience, and their ability to share this experience can lead to personal success, the success of those around them, and the success of the organizations they work for.

At kama, we believe that digital intelligence, informed by knowledge-keepers, product managers, or other subject matter experts, is anything but 'artificial'; no more artificial than manuals, books, educational videos, or other resources designed to impart knowledge on the persons who study these resources. For this reason, we call our platform Designed Experiential Intelligence®. This term is meant to imply that it is people who are delivering knowledge through an intelligent understanding of inquiries and prioritized delivery of information that aligns with the users' personal values. We believe that this is ultimately a much more responsible and ethical form of digital intelligence, especially when delivered alongside all of the other more technical elements of responsibility discussed in this paper.

Summary

- Ensure that your organization uses AI tools responsibly. This not only includes choosing responsible tools, but also the institution of internal policies on how tools should be used within the enterprise.
- Strive for Accessible AI. Ensure accessibility to persons of all abilities. Reduce the technical requirements of building and managing conversational AI solutions.
- Seek and/or implement systems that provide data transparency (to control responsible inclusive messaging) and properly secure any private client information or conversations.
- Understand how your AI solutions leverage human-in-the-loop governance to control information delivery and mitigate or eliminate bias and misinformation.
- Speak with your teams and/or your vendors about how your conversational AI functions. The process of delivering information, recommendations, or solutions should be understandable and therefore predictable to reduce risk to your enterprise and your brand.
- Lastly, look for opportunities to leverage Emotion AI as part of an ethical framework or guidance system to support authentic and prioritized conversations that align better with the personalities of the people that you seek to serve.

References

[1]Choudhary, F., Sicular, S. Gartner. *A Comprehensive Guide to Responsible AI*, 2022, July 28.
<https://www.gartner.com/document/4017080?ref=solrAll&refval=383266494&>

[2] The Walrus. (2023, May 16). *[WATCH] The Walrus Talks Artificial Intelligence*.
<https://thewalrus.ca/the-walrus-talks-artificial-intelligence/>

Bryant, J. (2023, June 20). IAPP releases AI Governance Professional Body of Knowledge. *International Association of Privacy Professionals*.
<https://iapp.org/news/a/iapp-releases-ai-governance-professional-body-of-knowledge/>

Gartner. *Best Practices for the Responsible Use of Natural Language Technologies, 2023, April 26*,
<https://www.gartner.com/document/4303799?ref=solrAll&refval=382311194&>

IBM. *What is a Knowledge Graph?* | IBM.
https://www.ibm.com/topics/knowledge-graph?mhsrc=ibmsearch_a&mhq=knowledge%20graph

Stanford Institute for Human-Centered Artificial Intelligence. <https://hai.stanford.edu/>

Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems. (2023, October 20).
<https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems>

GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.

About kama.ai and the kama DEI Solution

kama.ai is the creator of the Designed Experiential Intelligence™ platform, kama DEI, that allows curated information to be rated and distributed through a conversational agent based on personality value profiles. This establishes a level of personalization between the consumer/client and an enterprise's information that has not been achieved previously. The result is an automated Virtual Assistant that works around the clock to address inquiries with the right information for the right reasons.

kama DEI has an underlying knowledge base that unifies common information and an Enterprise Portal through which non-technical users can curate product and service information using simple natural language methods. The platform also allows for the setup of various target market demographic profiles (kama DEI Personas) to form the basis for various consumer or employee personality types.

kama.ai also offers a rapid-launch, front-end chatbot that can be configured and integrated within hours to your enterprise's web and mobile web pages. As an alternative, the kama DEI Chatbot API allows integration to a current chat facility that you may already be using for live chat, and it can also power other channels such as Messenger, SMS text, or WhatsApp. For more information on kama DEI, please chat with "Kady" on our website at [kama.ai](https://www.kama.ai) or fill out an [inquiry form](#) on our site.