COALITION FOR INNOVATION

AI
Blueprint for the Future

Coalition for Innovation, supported by LG NOVA

Jami Diaz, Director Ecosystem Community & Startup Experience
William Barkis, Head of Grand Challenges & Ecosystem Development
Sokwoo Rhee, Executive Vice President, LG Electronics, Head, LG NOVA

Coalition for Innovation Co-Chairs

Alex Fang, CleanTech Chair
Sarah Ennis, AI Chair
Alfred Poor, HealthTech Chair

Authors

Adrien Abecassis, Johnny Aguirre, John Barton, Ann M. Marcus, Olivier Bacs, Taylor Black, Micah Boster, Mathilde Cerioli, Carolyn Eagen, Sarah Ennis, Annie Hanlon, Christina Lee Storm, Andrew Yongwoo Lim, Jess Loren, Refael Shamir, Svetlana Stotskaya

Senior Editor, Alfred Poor
Editor, Jade Newton

October 2025

Preamble

The Coalition for Innovation is an initiative hosted by LG NOVA that creates the opportunity for innovators, entrepreneurs, and business leaders across sectors to come together to collaborate on important topics in technology to drive impact. The end goal: together we can leverage our collective knowledge to advance important work that drives positive impact in our communities and the world. The simple vision is that we can be stronger together and increase our individual and collective impact on the world through collaboration.

This "Blueprint for the Future" document (henceforth: "Blueprint") defines a vision for the future through which technology innovation can improve the lives of people, their communities, and the planet. The goal is to lay out a vision and potentially provide the framework to start taking action in the areas of interest for the members of the Coalition. The chapters in this Blueprint are intended to be a "Big Tent" in which many diverse perspectives and interests and different approaches to impact can come together. Hence, the structure of the Blueprint is intended to be as inclusive as possible in which different chapters of the Blueprint focus on different topic areas, written by different authors with individual perspectives that may be less widely supported by the group.

Participation in the Coalition at large and authorship of the overall Blueprint document does not imply endorsement of the ideas of any specific chapter but rather acknowledges a contribution to the discussion and general engagement in the Coalition process that led to the publication of this Blueprint.

All contributors will be listed as "Authors" of the Blueprint in alphabetical order. The Co-Chairs for each Coalition will be listed as "Editors" also in alphabetical order. Authorship will include each individual author's name along with optional title and optional organization at the author's discretion.

Each chapter will list only the subset of participants that meaningfully contributed to that chapter. Authorship for chapters will be in rank order based on contribution: the first author(s) will have contributed the most, second author(s) second most, and so on. Equal contributions at each level will be listed as "Co-Authors"; if two or more authors contributed the most and contributed equally, they will be noted with an asterisk as "Co-First Authors". If two authors contributed second-most and equally, they will be listed as "Co-Second Authors" and so on.

The Blueprint document itself, as the work of the group, is licensed under the Creative Commons Attribution 4.0 (aka "BY") International License: https://creativecommons.org/licenses/by/4.0/. Because of our commitment to openness, you are free to share and adapt the Blueprint with attribution (as more fully described in the CC BY 4.0 license).

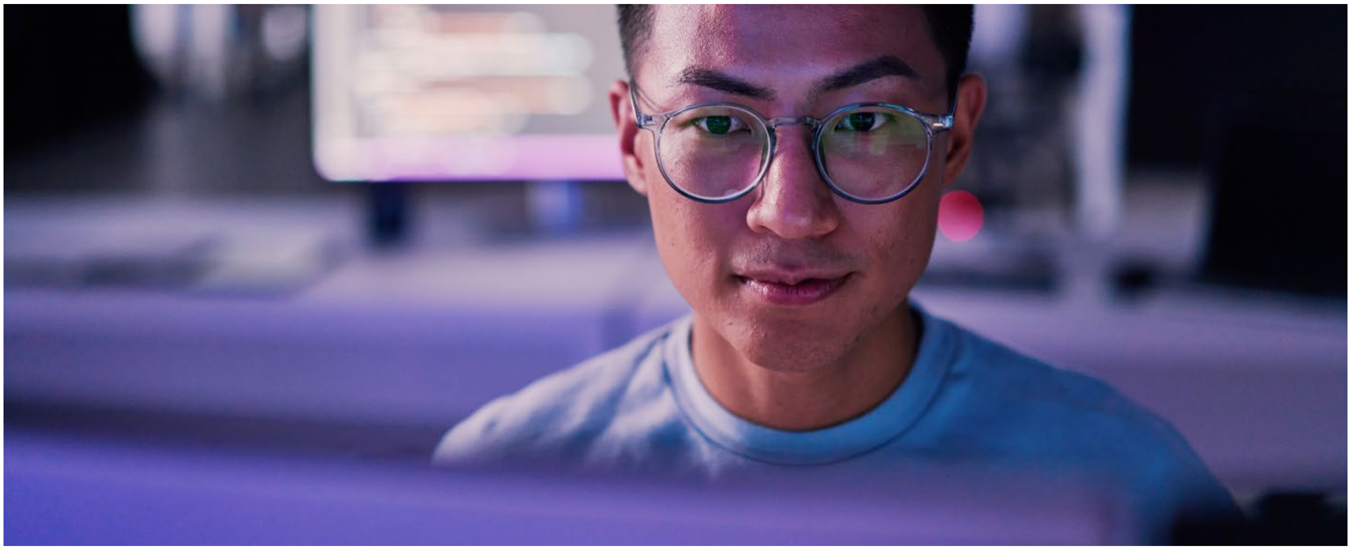
The Coalition is intended to be a community-driven activity and where possible governance will be by majority vote of each domain group. Specifically, each Coalition will decide which topics are included as chapters by majority vote of the group. The approach is intended to be inclusive so we will ask that topics be included unless they are considered by the majority to be significantly out of scope.

We intend for the document to reach a broad, international audience, including:

- People involved in the three technology domains: CleanTech, AI, and HealthTech
- Researchers from academic and private institutions
- Investors
- Students
- Policy creators at the corporate level and all levels of government

Chapter 6:
Human Factor Contributions in the Development of GenAI Applications

Authors: Refael Shamir, Ann M. Marcus



Why "Human Factors" Sits (or should) at the Center of GenAI Development

Human factors, as defined by Human Factors 101, encompass the interaction between individuals, their work, and the organizational environment. This involves understanding the demands of the task, the capabilities of the people performing it, and the characteristics of the organization.

These principles are crucial for designing artificial intelligence (AI) interfaces and interactions that are usable, safe, and effective. By considering human cognitive limitations, decision-making processes, and natural technology interactions, designers can create more intuitive systems. For instance, recognizing automation bias in humans allows for the development of AI systems that foster appropriate user trust and encourage critical evaluation of AI outputs, rather than unquestioning acceptance.

Ergonomics is another term for human factors, though the two terms are sometimes differentiated according to the physical and psychological aspects of the human. Psychological capabilities are more commonly associated with human factors, while physical aspects are more commonly associated with ergonomics. Generally, though, the two terms can be considered synonyms.

Generative AI (GenAI) introduces novel usability and safety considerations that traditional human factors work had not previously encountered. These include issues such as prompt engineering (how to communicate effectively with AI), managing over- or under-reliance on AI assistance, and understanding when AI outputs might be hallucinations or biased. The conversational nature of many GenAI systems creates expectations about intelligence and capability that may not align with actual system limitations.

GenAI systems may automate text, code, images, or decisions, but every step of their lifecycle is shaped by human judgment, cognition, and culture. Research on the "automation paradox" shows that the more capable the automation, the more crucial the human role becomes for safe, reliable performance. Human-factors engineering therefore asks: *How do people's abilities, biases, limits, and values influence — and become influenced by — GenAI?*

The answer may depend on a number of complex considerations: The nature of the "problem" for which AI is being employed, the stage at which AI is being consulted, the role that humans play in the process of identifying, describing, querying, interpreting, verifying, applying, integrating, and acting upon their interactions with AI. It must also consider the associated risk(s) should errors, hallucinations, misunderstandings, or other unexpected actions or outcomes by either the tool or the human take place in using AI tools, to name just a few.

Stage in the GenAI Lifecycle	How People Make the Difference	Human-factor Risks If Neglected
1. Problem framing & goal-setting	Stakeholders articulate real user needs, define the purpose of the model, set success metrics, and surface social/ethical constraints.	Misaligned objectives, "solutionism," products nobody needs.
2. Data stewardship	Humans choose sources, label data, set inclusion/exclusion rules, and document provenance. Diverse teams catch blind spots and steer data toward representativeness.	Embedded biases, privacy breaches, colonial data extraction. Studies show GenAI can amplify hidden cultural or religious bias when curation is weak (Nature).
3. Model & prompt design	Architects translate goals into model size, context windows, and guard-rails; prompt engineers encode domain knowledge and mental models of users.	Brittle behavior, hallucinations, cognitive overload if outputs don't match user mental models.
4. Evaluation & alignment	Human raters run red-team tests, provide Reinforcement learning from human feedback (RLHF) judgments, and iterate on "design principles" for good UX (e.g., IBM's	Safety gaps, opaque behavior, distrust.

	six principles: clarity, context, control, etc.) (IBM Research).	
5. Interface & experience design	UX and accessibility specialists craft affordances, explainability cues, and recovery paths. Human-centered guidelines help build trust (Medium).	Unsuitable mental workload, exclusion of low-vision, low-literacy, or non-English users.
6. Governance, oversight & continuous operations	Policy teams define accountability maps ("who is on the hook for what?" (Mozilla Foundation)), set escalation paths, and keep humans-in-the-loop for high-stakes use. (Mozilla Foundation).	"Shadow" AI, regulatory non-compliance, erosion of public confidence.

Members of the Northwestern University Robotics Club outlined the following eight actions required to use GenAI effectively and ethically during its early stages:

1. **Keep technology honest and accurate**. Many leaders have been surprised by the inaccuracy of generative AI tools as well as their "hallucinations." The tools seemingly respond to prompts with imagined "facts" that are not true and produce confident, incorrect statements. The statistical predictive nature of models mean hallucinations can occur when there is little training data relevant to a required piece of generated content. They can also occur if prompts are poorly phrased. Effective leaders coach human users to create quality prompts and verify the accuracy of content.
2. **Keep technology ethical and legal**. Effective leaders establish usage standards and use cases to facilitate employees' ability to consistently respect privacy and copyrights, cite sources, and only use information obtained with the creators' consent. Because GenAI tools are often trained on large amounts of data, it can be difficult for users to determine the source of the training data. Effective leaders employ techniques where content is generated from a known set of verified documents that are searched and incorporated into the context (input) of the model rather than a model containing "all" the knowledge.
3. **Keep confidential information safe**. Additionally, effective leaders establish usage standards that include guidelines and procedures to keep confidential organization, employee/volunteer, and customer/user data from being exposed publicly. They develop and implement policies and checks to prevent proprietary information from being inadvertently released via AI platforms, including through AI learning and training. They also ensure legal and security reviews of AI services. While some services provide data security and privacy guarantees, others make it clear that users are responsible for protecting sensitive data and cede rights for any entered data.
4. **Maintain transparency**. Effective leaders inform users how models work and educate them on their limitations. The very nature of AI generation and function means that tools can act as "black boxes," making it difficult to accurately evaluate what the models will produce and what sources, if any, they are referencing. Effective leaders maintain transparency wherever possible for all stages

of models and generated outputs. They select AI tools that list the sources (links) they have used when generating content, which helps address transparency challenges.

5. **Provide context**. Modern AI effectively synthesizes content on which it is trained, but it is less effective with situational awareness and analysis. While AI tools learn more each day, human users have the vital role of providing context. Effective leaders hire users into new roles such as AI prompt engineers – individuals highly skilled at leveraging generative AI tools and their output. These leaders deploy individuals in such roles to effectively use models, including information about context and desired output in a way that is both efficient for the tool and clarifying for those who use the output.
6. **Provide authentic empathy, compassion, and connection**. GenAI can be trained or prompted to provide output in a style that mimics empathy and compassion. Thus, AI is being used today to generate seemingly human interaction, develop "relationships," and provide emotional support. Effective leaders help employees understand that while AI models may create outputs that appear to provide emotions, they are not real. These leaders help their people interpret messages with a correct, healthy framework and set of expectations. They also provide real human connection in an increasingly digital and virtual age.
7. **Address bias**. Because GenAI models are trained on content that naturally includes the biases of the human users who created it, historical biases (including **analytical biases** such as recency bias and **social biases** such as discrimination) become built into models and replicated by tools as they relearn. Effective leaders operate knowing that AI is only as good as the data it is trained on. They take steps to ensure objectivity and fairness on data input and interpretation of output.
8. **Complete the work**. Examples of different uses of generative AI tools include writing job descriptions, creating computer code, writing sales plans, developing marketing messages, creating operations task lists, generating research, and answering routine employee and customer questions. Effective leaders understand that for some jobs, the tool may do the majority of the busywork, but people must still complete tasks by adding their insight and shaping outputs based on their skills and experience. In virtually all cases, it is up to the human user to finish the job.

Testing Human Factors When Designing GenAI

The following are some design deep-dives to undertake in assessing the testing for the human factors aspects when using GenAI:

- **Cognitive ergonomics**: Outputs should match the way humans scan, remember, and reason. Chunking long answers, surfacing sources, and allowing drill-down to reduce cognitive load.
- **Bias mitigation as a sociotechnical task**: Technical debiasing must be paired with diverse human review panels and clear bias taxonomies.
- **AI-literacy & upskilling**: Experiments show that training users in prompt strategies and judgment skills markedly improves Human-AI collaboration outcomes.
- **The "automation paradox" playbook**: As capabilities grow, raise *human* requirements: scenario training, simulation drills, and fallback procedures.

Effective leaders should evolve the role of human users in parallel with GenAI technology to maximize the benefits of these new and developing tools while mitigating their associated risks.

- Humans (should) set the goals, supply the data, critique the outputs, and govern the consequences; **every GenAI success or failure is fundamentally sociotechnical**.

- Investing in **diverse, AI-literate teams** and **robust human-factors processes** is cheaper than remediating biased, unsafe, or unusable products later.
- Treat GenAI not as a "black-box oracle" but as a **power tool whose safety relies on skilled operators, clear interfaces, and systemic oversight**.

Here is a practical checklist for teams developing and/or using GenAI to ensure that they are maximizing the benefits and reducing their risk.

Teams building GenAI can start by:

KEY QUESTION	QUICK TEST
Human-in-the-loop?	For every failure mode, can a qualified person detect, override, or audit it in time?
Diverse voices?	Does your data-curation and red-team roster include domain experts *and* historically under-represented groups?
Explainability fit-for-purpose?	Can an *average* end-user understand why the model gave a recommendation and what to do next?
Skills plan?	Have you budgeted for AI-literacy programs for developers, reviewers, and end-users?
Ongoing governance?	Who owns model updates, monitors drift, and reports incidents — and by which cadence?

- Embedding human factors experts early in the development process
- Red-teaming for usability and bias *before* launch, not after
- Piloting interfaces with diverse user groups to surface usability, trust, or comprehension gaps
- Budgeting for AI literacy training across roles, not just for developers

- Defining governance paths early such as who monitors drift, owns outputs, and updates protocols

These steps reduce downstream failure, improve user trust, and support safer, more responsible GenAI deployment.

Conclusion

The successful development and implementation of GenAI applications are intrinsically linked to a deep understanding and integration of human factors. From the initial stages of problem framing and data stewardship to the ongoing governance and oversight, human judgment, capabilities, and limitations profoundly influence every aspect of the GenAI lifecycle. Neglecting these human elements can lead to misaligned objectives, biased outputs, safety gaps, and a significant erosion of trust.

As GenAI technologies continue to advance, the role of humans is evolving from merely interacting with the tools to becoming crucial "skilled operators" who set goals, supply diverse data, critically evaluate outputs, and ultimately govern the consequences. This requires a proactive approach to identifying the potential risks presented by inaccuracies, ethical dilemmas, data confidentiality, and inherent biases and addressing them early. Effective leadership across this changing landscape demands a commitment to transparency, the provision of adequate context, and the fostering of genuine human connection, recognizing that AI-generated empathy is not a substitute for the authentic humankind.

By consciously incorporating cognitive ergonomics, implementing robust bias mitigation strategies, investing in AI literacy and upskilling, and developing an "automation paradox" playbook, organizations can maximize the benefits of GenAI while mitigating its associated risks. Ultimately, treating GenAI as a powerful tool that requires skilled human operators and systemic oversight rather than treating it as a "black-box oracle" is paramount for fostering safe, reliable, and truly impactful AI solutions. The core principle remains: every GenAI success or failure is fundamentally sociotechnical, underscoring the indispensable role of human factors at the center of its development.

Author (In order of contribution)

Refael Shamir, Founder, Letos

Refael Shamir, is a seasoned entrepreneur in the field of affective neuroscience and is working towards introducing a new medium for gaining insights into spontaneous human reactions based on seamless integrations of devices in everyday environments. Refael is also a renowned speaker having presented his learnings in highly acclaimed conferences such as NVIDIA GTC, MOVE Mobility Re-Imagined, NeurotechX, among others.

Ann M. Marcus, Director, Ethical Tech & Communications, WeAccel

Ann M. Marcus is a Sonoma-raised, Portland-based communications strategist and ethical technology analyst focused on smart cities, community resilience, and public-interest innovation. She leads the Marcus Consulting Group and serves as director of ethical technology and communications at WeAccel.io, a public-good venture advancing mobility, communications, and energy solutions for communities. Ann has advised public and private organizations—including Cisco, the City of San Leandro, Nikon, AT&T, and InfoWorld—on trust-based data exchange, digital public infrastructure, resilience strategy, AI and more. Her current projects include a California senior evacuation program, a Portland robotics hub, and digital energy resource initiatives with utilities in Portland and the Bay Area.

For more information about the Coalition for Innovation,
including how you can get involved, please visit coalitionforinnovation.com.

View the Next Chapter