



AI TASK FORCE REPORT

Inform USA AI Task Force: Report

*These materials are subject to ongoing review and revision. **Please comment!***

What can be clarified, elaborated upon, etc? What is missing?

Table of Contents

Preface: About This Document and The AI Task Force Our	3
Values and Principles	8
Use Case Analysis and Policy-making: an example	11
Guidelines for Ethical AI Procurement and Policymaking	16
Toward Shared AI Infrastructure for the Community	18
Navigation Industry	20
Appendix A: Glossary	23
Appendix B: Types of Use Cases	25
Appendix C: Types of Risks	30
Appendix D: Recommended Readings	

Thanks to:

Task Force Members –

Emily Aubele, Pennsylvania 211
Jeff Cary, United Way Austin TX
Samantha Gardner, ADVancing States
John Gonzales and Sara Tribe Clark, USAging
Carla Lundeen, Great Rivers 211
Richard Mugford, Distress Centre Calgary
Sarah Pottelberg, Inform USA
Sasha Reinoso, Do Good Data
Courtney Smith, United Way 211 Rhode Island

Advisors:

Jennifer Abels of Inform USA
Kevin De Liban of TechTonic Justice
Neil McKechnie of Yanzio
Kevin Claybon of Everest Effect
Val Elefante of MetaGov

Facilitated by: Greg Bloom

Preface: About This Document and The AI Task Force

As technologies associated with "Artificial Intelligence" proliferate, so do questions about how – and whether – they should be used in the context of human service Information and Referral (aka "Community Navigation"). In response to an increase in inquiries from funders, partners, and members about the adoption of "AI" in their work, Inform USA [chartered its AI Task Force](#) to support the Community Navigation sector in developing clear, appropriate, and effective terms of engagement for the responsible use of AI tools.

The Task Force's Objectives:

- Analyze the risks and opportunities presented by emerging AI technologies for the information and referral sector, our clients and communities.
- Educate Inform USA members about the types, uses, and risks of AI, including common terms and key concepts that can support informed decision-making.
- Recommend a set of principles, practical guidelines, evaluation frameworks, and proposed standards for the design and use of AI in Community Navigation.
- Advise the Inform USA Board on any strategies that our organization can undertake to promote best practices and responsible development of AI across the sector.

This document synthesizes an initial set of outputs from the Task Force's work – over six months of research, analysis, and deliberation. These outputs include:

- [An executive summary](#) of our findings and recommendations.
- [Values and principles](#) that guide the work of Inform USA and our membership.
- [Analysis of specific use cases](#), including one shared here in full.
- [Example of a specific "acceptable use policy"](#) for a given use case
- [Proposed Guidelines for Ethical AI Procurement and Policymaking](#)
- [Draft strategic framework for development of industry-wide shared infrastructure](#)
- **Appendix A: [Definition of terms](#)** referenced in the discussion of AI technologies.
- **Appendix B: [Taxonomy of risks](#)** associated with use of AI tools, especially in a Community Navigation context.
- **Appendix C: [Taxonomy of use cases](#)** relevant to Community Navigation and their associated benefits.
- **Appendix D: [References and suggested readings](#).**

This content reflects a sample of the output produced by the Task Force.

- [Our Knowledge Base](#) includes an index of the key materials that we referenced.
- [This Airtable](#) aggregates our analysis of use cases and risks.
- Additional materials (including collateral such as Task Force meeting notes) available in [this Shared Drive Folder](#).

Executive Summary of the AI Task Force's findings

Inform USA's members produce and use context-sensitive, localized knowledge to connect people – often in circumstances of vulnerability – to resources and support that can improve their quality of life. “Artificial Intelligence” is increasingly a topic of interest for Community Navigation providers, so Inform USA chartered a Task Force to assess the challenges and opportunities associated with these technologies.

The AI Task Force began by seeking to clarify the definition of key terms, as aggregated in [the Glossary](#). We observed that AI is often associated with a range of ambiguous and potentially misleading definitions; we've acknowledged several definitions, though the shortest might also be the most useful: **AI refers to “a mathematical method for prediction”** (Broussard, 2019).

Many (though not all) of the examples of AI analyzed by the Task Force involve **“Large Language Models” (LLMs) that mathematically predict which word is likely to follow a sequence of other words in response to a given prompt**. In other words, LLMs produce outputs that are *plausible*, but not necessarily *factual*. This presents dilemmas for an industry like Community Navigation, which is critically concerned with the reliability of information and the effectiveness of outcomes.¹

With such concerns in mind, the AI Task Force set out to assess the circumstances in which such tools might be useful and appropriate.

Our Findings: People and processes should come first and throughout

The AI Task Force has compiled an array of resources – including direct research, field surveys, and guidance from government and nonprofit leadership – [available in our Knowledge Base](#). Several resources consist of aggregation and analysis of hundreds or even thousands of studies to assess common patterns of failure and identified risks. The Task Force referenced these aggregate analyses and prioritized the risks that are most relevant to our field – see [Appendix C](#) and [here in our Airtable](#).

¹ Many aspects of the work of Community Navigation reflect known weaknesses of LLMs outputs: topics that aren't commonly documented on the internet; users who have marginalized identities and complex needs; and demand for highly localized, nuanced information that changes frequently.

Prioritized risks can be sorted into two broad meta-categories:

The outputs of AI tools are not reliable: In some contexts (such as marketing, creative activities, or even customer support for megacorporations) this risk may not be considered significant; in the context of Community Navigation, however, these risks must be taken seriously. Specific risks include biases in data inputs and design, incorrect or misleading outputs, and inequitably variable performance.

Adoption of AI tools can result in a diminishment of capacities: from de-skilling of professionals to institutional "forgetting," to a lack of control over data inputs and outputs, and broader operational dependence on technology vendors whose interests might not align with the communities we serve, AI tools can result in various forms of dependence and diminishment.

These broad kinds of risk involve different timeframes and contexts: the former, immediate and operational; the latter, long-term and diffuse. Both, however, can be addressed with the same approach: **invest in people (especially Community Navigation staff) and processes (Community Navigation operations), and apply them with care, before and after implementation of any AI tool.**

Our Recommendations: Principled process for careful decision-making

After defining terms and assessing risks, the Task Force articulated [a set of principles](#) that should guide Community Navigation providers in decision-making processes. We then applied these principles to develop a shared understanding of what careful, responsible decisionmaking entails. Practically, this entails analyzing the context of particular scenarios in which specific AI tools might be used for a given purpose (["use cases"](#)) in order to weigh the [prospective benefits and risks](#) and develop ["acceptable use policies."](#) One example of analysis and policy is [included in this document](#); more available in [our shared folder](#).

We only scratched the surface of this work; such analysis should be conducted in ongoing iterative processes at both industry-wide and organizational levels. We cannot anticipate every risk of every scenario; however, through principled practice, we can learn and act responsibly amid uncertainty. Toward this end, we produced a draft set of [Guidelines for Ethical AI Procurement and Policymaking](#).

In the meantime, we offer this broad summary of our recommendations:

We encourage careful experimentation with uses of AI that can automate routine and repetitive administrative tasks in ways that enhance the work of Community

Navigation staff – promoting human resources to focus on the production of local, specialized, context-sensitive knowledge, the safeguarding of information integrity and data quality, and the promotion of human agency.

We recommend against uses of AI that make subjective decisions that might directly impact people's lives. Any such uses of AI tools – such as user-facing chatbots, risk-scoring, etc. – should be considered highly risky and carefully mediated by responsible humans or even refused altogether.

We further qualify these recommendations with these resolutions for careful use of any such tools:

The quality of outputs depends on the quality of inputs. "Organizational readiness for AI" entails a responsible and effective culture of data management, which entails trained and empowered *humans* – i.e. Community Navigation staff – working together through accountable processes.

Success depends on humans in the loop. In any effective use of AI tools, trained and responsible humans must be actively engaged in assuring reliability. Furthermore, given the known risks of "automation bias" (in which people tend to suspend their own judgment when presented with computed output), any individual "human in the loop" ought to be situated within a broader evaluative loop through which humans assess the aggregate quality of operations over time.

Finally, shared data infrastructure can help mitigate the risks and ensure equitable benefits of AI for the entire industry. This Task Force primarily focused on questions that Community Navigation providers can ask to develop effective practices for their own organizations; however, we observe that many of the challenges and opportunities presented by these technologies may be best addressed at the sectoral level. Inform USA is uniquely positioned to set standards and build capacities that might benefit its entire membership – along with the human service sectors at large, and the communities we serve. [This might entail development of shared infrastructure that all Community Navigation providers and associated partners could use in development and operation of AI tools; we offer this memo as a starting point for discussion.](#)

Proposed next steps

With this document, the Task Force concludes its second phase of work. We encourage the Inform USA Board to review these materials and recommend a path forward.

Prospective future steps that the Task Force suggests for consideration:

- Present these findings in a webinar to the entire Inform USA community.
- Elicit feedback, real-world examples, and additional reference material from Inform USA membership, and iteratively develop these materials accordingly.
- Produce “onboarding materials” to orient members to these concepts and processes.
 - Develop a standardized “decision tree” specifically designed to address considerations of AI use in Community Navigation.²
- Establish standing capacity to conduct this analysis and policy-development over time, such as a technical committee.
- Consider [the draft Guidelines for Ethical AI Procurement and Policymaking](#) as an input to the Inform USA standards development process.
- Consider emerging strategies for licensing of data, [such as Creative Commons Signals](#).
- Develop proposals for funding in order to build shared capacity and industry infrastructure for more ambitious development.

How to get involved:

- Comment on this document with your feedback
- Email technology@informusa.org and/or Sarah Pottelberg (sarah@informusa.org) to inquire and/or make suggestions.

² For now, we recommend that Inform USA members reference this generic decision tree for non-profits considering adoption of AI tools: <https://techmatters.org/should-i-be-using-ai-for-this>

Our Values and Principles

In developing our strategic framework for decisionmaking, we start with reflection on our values and principles. The AI Task Force first leveraged Inform USA's official values – *i.e.*, the core qualities that we aspire to as a community. Then we reflected on our principles – *i.e.*, the ways we agree to do our work – in part by considering these values in light of the risks and opportunities evident in the field of AI tools. We also consulted a wide array of already-existing statements of principles on this subject that we found in an array of fields, from government to healthcare to civic technology. [[See our Knowledge Base for more.](#)]

The following statements, taken together, help us to draw boundaries around what kinds of actions and outcomes are desirable – and, in contrast, what kinds should be considered unacceptable. Note that these Principles should be considered a draft; we welcome any feedback from Inform USA members that might help us to evolve and further clarify these principles. Please leave a comment!

Values of Inform USA

Solidarity: We value and embody the diversity of the sector. Our interconnectedness is our power. Our shared commitment to the standards is our guide.

Empowerment: We help our members to meet people where they are and help them get where they need to go. We empower those who empower the public.

Accountability: We set standards and hold each other to them. We elicit input from those we serve and act on it. We serve with excellence, integrity, and transparency.

Critical Thinking: We are always learning and growing. We do not shy away from challenging conversations. We analyze problems, welcome all ideas, and test assumptions.

Accessibility: Inform USA members guide people through complex systems, breaking down barriers to resources that help them survive and thrive. Inform USA supports this work by creating access to and education around the industry standards for Community Navigation.

Principles for the use of AI in Community Navigation [DRAFT]

(working from [this tab of our shared knowledge base](#))

1. We assume responsibility for the well-being of our workforce and the people we serve – and this responsibility includes the technologies we choose to use.

- We carefully consider the potential risks of any algorithmic technologies, weighing them alongside their potential benefits. The greater the risks, the greater the care with which we make decisions, test assumptions, and monitor results.
- We ensure with a high degree of reliability that any given tool is appropriate for its task – which entails taking steps to anticipate and mitigate the risks of undesirable outcomes.
- We don't believe the hype; we do due diligence to assess the veracity of claims and test our assumptions.
- We monitor the performance of algorithmic tools over time, to ensure that they are safe, effective, and equitable.

2. We elicit the input of people whose interests are at stake in our decision-making process.

- Workers, helpseekers, and communities at large should be involved in processes of design, implementation, and evaluation – especially with new technologies that impact them.
- Ethical decision-making involves a commitment to listening, learning, and repeatedly applying our values in processes of quality improvement over time.

3. We ensure that users of AI-enabled systems have transparency into, and agency within, the working of those systems.

- People should know when an automated system is being used, and understand how and why it contributes to outcomes that impact them.
- The functionality of algorithmic technologies should be clearly and accessibly documented, including the processes by which people can provide feedback, articulate concerns, escalate grievances, etc.
- Feedback is an essential aspect of an appropriate system – it should be clear how users can provide feedback on the function of an algorithmic system, and there should be sufficient and appropriate means of receiving, processing, and acting on that feedback.

- We should have the ability to correct and adapt – or, if necessary, refuse to use – an AI technology to ensure it is effective for its purpose and appropriate for its context.

4. AI systems should augment and support humans rather than replace them.

- Users should be able to opt out of the use of algorithmic technologies.
- Users should have access to a person who can quickly consider and remedy problems they encounter.
- Algorithmic tools should help humans do more, higher-quality work; an organization's staff should be involved in the process of considering, designing, and evaluating whether they achieve these objectives.
- Most algorithmic technologies enable us to “do more with more”; we don't assume that it is always possible to “do more with less.” This means we consider the staffing costs that will be necessary to appropriately use an algorithmic technology as part of the design and implementation process.

5. The output of AI systems should be subject to audit and correction.

- Community Navigation providers are ultimately responsible for the quality and reliability of the information that we provide. We cannot outsource that responsibility to unaccountable technologies.
- “A computer can never be held accountable, therefore a computer must never make a management decision.”
- *Should we further specify this or any of the others? Leave a comment with your suggestion!*

Use Case Analysis and Policy-making: an example

The AI Task Force identified a broad range of use cases that may be relevant to Community Navigation providers – some of which are merely hypothetical, and some of which are already being implemented today. (The full taxonomy of these use cases is in [Appendix B](#).) Task Force participants then selected several prioritized use cases – based on immediacy of relevance and/or level of interest – and analyzed the prospective benefits and [risks](#) of each, developing a set of questions that a Community Navigation provider might ask to apply [our principles](#) in responsible decision making over the use of such tools.

In one instance ([see below](#)), we worked with Inform USA staff to draft [a prospective “Acceptable Use Policy”](#) that applies the analysis to organizational procedures pertaining to a given use case. We are including this example of a use case analysis and acceptable use policy in this document.

Some **additional use case analyses have been drafted and can be accessed through the hyperlinked entries in the Appendix taxonomy**. We recommend that Inform USA implement an ongoing process of analysis and policy guidance to address additional use cases in the future.

Transcribing Inform USA Accreditation Reviews: Risk/Benefit Analysis

Copied from:  AI Use Case analysis: Summarization/Transcription

User: Inform USA Accreditation Reviewer

Task: Recording and transcribing “accreditation reviews” with applicants for Inform USA accreditation *Note: This use case could be similar to transcribing training performance for frontline staff evaluation in Community Navigation organizations.*

Details: Accreditation reviews are conducted by Inform USA Accreditation contractors in dialogue with Inform USA members that are applying for accreditation. There are two types of accreditation reviews.

Consultative/Site Reviews

Typically, each site review takes about [x] time, resulting in transcripts around 10 pages in length. Site reviewers want to be able to take their own assessment notes, but taking their own notes makes it difficult for them to also keep transcription of the specific words that are said by the applicant, which is helpful for later review.

Database Reviews

Typically, a database review takes about 4 hours and is broken into two meetings. Reviewers ask the applicant to share their screen to demonstrate their use of the database and ask questions about policies and procedures. The Reviewers take notes on what they see to help them complete a report.

Benefits:

- Recording and transcription of the interview enables the Reviewer to focus on the context of the interview and take their own notes while having a verbatim record of what was said for later review.
- Reviewer notes could be enhanced by supplementing them with the transcription.
- Questions that may arise after the interview could be answered by a review of the recording or transcription without having to ask the applicant for more information.
- Recordings, with consent, could be used for training purposes for new Reviewers.
- Recordings, with consent, could be shared with other Reviewers if there are questions or concerns about how an organization meets a specific requirement.

Risks

- **False or misleading information** can be generated in transcription and summaries
- **Loss of information** in cases when the transcription tool glosses over context.
- **Compromise of privacy** by obtaining, leaking or inferring sensitive information
- **AI system security** vulnerabilities and attacks
- **Over-reliance and Unsafe use in** staff who may become dependent on the tool and no longer skilled at synthesizing meaning out of notes.

Questions:

- What is the intention/goal behind using AI? How is this expected to improve upon our existing processes?
- What steps can/should staff take to mitigate the risk of errors or lost information in the summary?

- What kinds of privacy protections do we need to ensure that this content is not accessed or otherwise used as input by the technology vendor?
- How should the transcript be used? Under what circumstances can the outputs be used to make decisions?
- How should the transcript be stored/retained/shared? How long should the transcript be stored, and at what point should it be deleted?
- How will the quality of this tool's output be evaluated over time?
- How is consent requested before using the tool?
-

Possible policy requirements:

- Notify the interviewed organization of the transcription service before beginning, and confirm consent.
- Preserve recording to cross-reference with transcription [for x amount of time].
- Human review of transcription immediately after interview, identifying any missed, fabricated, or unclear information – and follow-up with interviewee to seek clarification as necessary.
- Confirm that the privacy policies of the given tool sufficiently address our specific concern about information capture.

DRAFT Policy: Accreditation Review Record and Transcription

Copied from:  DRAFT Policy / Procedure: Site Review Transcription Use Case

The following is an example of a prospective organizational policy that can be developed based on the analysis conducted in the previous section. This policy is developed using [Inform USA's policy template](#). NOTE: This is not an official policy, but it is an example of a proposal that may be submitted for consideration as a proposed policy.

An “acceptable use policy” like this could be developed for every use case identified in our taxonomy.

Policy Statement: Inform USA is committed to ensuring accountable and fair processes for our Accreditation Review Process. This policy addresses the use of recording and transcription tools during Accreditation Review interactions.

Scope: Accreditation Reviewers using AI tools to record and/or transcribe their interviews.

Purpose: To ensure that Accreditation Reviewers can efficiently generate reliable transcriptions of their interviews with Accreditation applicants, building their capacity to conduct thorough and attentive evaluations while minimizing the risk of inaccurate transcriptions.

Procedures:

Before selecting a tool to use: Confirm that the privacy policy of the tool sufficiently affirms that the content generated by the tool cannot be directly accessed by anyone outside of the organization that is using the tool [and *other concerns about information capture that may be specified in the future*].

Before conducting an interview:

- Create and document the existence of a designated folder in which transcriptions will be stored.
- Confirm that settings for the tool adhere to the necessary conditions for this activity.

At the beginning of an interview:

- Notify the organization of the recording and transcription service and how it will be used before beginning, and confirm their consent.
 - If they do not consent, then the tool can not be used.

After the interview:

- Conduct a human review of the transcription immediately after the interview, identifying any missed information, fabricated information, or unclear information. If necessary, follow up with the interviewee to ask for clarification on any unclear points or causes for doubt in the transcription.
- Preserve recording to cross-reference with transcription [for *x amount of time*].

Definitions:

- *Accreditation Review:* An online meeting (conducted via video conference software) for specific phases of the Inform USA Accreditation process (ie. Database Review and Virtual Site Visit).
- *Accreditation Reviewer:* The Inform USA contracted staff who conducts the review.
- *Transcription:* A written text record created by capturing spoken words during a video call.
- *Accreditation Applicant:* The organization going through the accreditation process.

Related Policies: *List any other policies that are related or relevant.*

Responsibilities: *The Accreditation Reviewer is responsible for following this policy and informing the Accreditation Applicant of their options.*

Penalties/Disciplinary Action: *(Optional) Outlines consequences for non-compliance.*

Contact Information: Accreditation Director, Jennifer Abels

Policy/Procedure History: *Lists the dates approved, revisions and/or updated (and optional, by whom)*

Guidelines for Ethical AI Procurement and Policymaking

This outline below has been synthesized by Task Force members with reference to resources such as [the 2024 White House memorandum on agency procurement of AI technologies](#), and subsequently revised in accordance with our analysis of various use cases and risks. *Feedback welcome: What is missing? For these guidelines to be upheld, what capacities and processes should be established?*

1. Ensuring collaboration across your organization and among partners: Representatives of all relevant parts of an organization should be involved in consideration, design and deployment, and management of an AI system or service.

- a. This may involve a standing cross-functional workgroup including representatives of different parts of the organization, partner organizations – and (when appropriate) the community served by the organization.
- b. This process should involve documented articulation of the purpose of the intended technology, assessment of risks and open questions, gathering of information and community feedback, through iterative multi-stakeholder review.
- c. This process should continue through the life-cycle of a technology, including transparent processes by which performance issues can be identified, addressed, and resolved through policy change and/or decommissioning.

2. Establishing rights and responsibilities: fiduciary duties should be clearly established before deployment, and upheld through the life-cycle of a technology.

- a. Agencies should establish clear duties of care, *i.e.*, designated responsibilities as to who represents the interests of users of these technologies, to ensure that harm is prevented when possible and/or remediated when discovered.
- b. Agencies should ensure that they retain the ability to opt out of and/or prohibit the contractor from using agency data to train AI without an agency's consent.
- c. Agencies should prioritize interoperability in vendor selection and contract development, and take steps to prevent vendor lock-in.

3. Managing AI Risks and Performance: Equitability and risks should be proactively considered when planning, managing, and evaluating deployment of AI technologies.

- a. Standard processes should be established to determine whether risks from the use of AI are acceptable and will be sufficiently managed – and to evaluate and improve the reliability and quality of such tools over time.

- b.** AI vendors should demonstrate how they can address these risks, including through privacy-enhancing technologies.
- c.** Responsibility for risk mitigation should be clearly assigned through contractual agreements that stipulate what will happen if something goes wrong.
- d.** Agencies should establish clear processes to evaluate technologies – and potentially decommission them – in order to ensure that performance justifies use and that benefits outweigh costs compared to low-tech low-risk alternatives.
- e.** Agencies should establish clear processes for incident reporting, including accessible user-facing means, and transparent internal processes by which reports will be received, assessed, and acted upon.

Toward Shared AI Infrastructure for the Community Navigation Industry

The Challenge: using guess-machines in an answer business

Amid all of the excitement about the potential benefit of AI, these tools present Inform USA's members with significant dilemmas. AI outputs tend to be unreliable (in that they are statistical predictions, not factual claims), unpredictable (in that the same question asked in slightly different ways might yield very different results), and opaque; these weaknesses are in tension with Community Navigation sector's core values of reliability and accountability. There are known methods of coping with such weaknesses and mitigating key risks, but these methods may require expertise and capacity that few Community Navigation providers can establish themselves.

Opportunity: Infrastructure for knowledge stewardship

Inform USA is uniquely positioned to establish the infrastructural capacities that would be necessary for algorithmic technologies to be ethically designed, reliably deployed, and accountably governed. As the industry association for over a thousand organizations, Inform USA facilitates the development of standards, best practices, and shared knowledge about the ways and means of human service informatics. This role already entails knowledge stewardship; with some support and development, we can build data infrastructure that would equitably distribute the benefits of AI while mitigating its inherent risks.

Proposed path forward: develop the foundational elements of an information ecosystem

Leveraging our Community Navigation Standards, in partnership with stewards of the 211HSIS service taxonomy, Inform USA can establish shared infrastructure for the Community Navigation sector – such as libraries, models, evaluations, and governance processes – that can serve as the foundation for a thriving information ecosystem that is committed to the public interest. By establishing shared data infrastructure, Inform USA can make it easier to design, test, and deploy responsible uses of algorithmic technologies.

This infrastructure might include:

- A “specialized language model” and/or “open weight model” tailor-made in accordance with Inform USA standards, service taxonomies, best practices and associated knowledge.
- Training systems that can leverage the distributed knowledge of Inform USA members
- Evaluation processes that can test, assess, and improve algorithmic tools, such as <https://weval.org>
- Governance processes that can receive feedback from across the field and develop new standards and protocols over time

Key questions:

- What are the key components of infrastructure that can foster innovation in algorithmic technologies while preserving accountability, reliability, and community control?
- What can we learn from other precedents for shared industry-wide data infrastructure?
- What capacities and processes can facilitate cooperation across our membership to ensure community-based control and effective use of these tools?

Appendix A: Glossary

AI (artificial intelligence)	A technical and scientific field devoted to the engineered system that generates outputs such as content, forecasts, recommendations, or decisions for a given set of human-defined objectives. A mathematical method for prediction.	International Organization for Standardization Artificial Unintelligence, by Meredith Broussard
AI ethics	The set of guiding principles that stakeholders use to ensure artificial intelligence technology is developed and used responsibly. This means taking a safe, secure, humane, and environmentally friendly approach to AI.	Coursera
Algorithm	A process designed to solve a specific problem. It's a set of instructions that end up in a desired conclusion.	Forbes
Application programming interface (API)	Mechanisms that enable two software components to communicate with each other using a set of definitions and protocols.	Amazon Web Services
Augmented Reality	Techniques for overlaying the physical world with computer-generated content. This content may be displayed on a smartphone screen, or through more specialized apps or hardware.	Allied Media Projects
Bias	Output errors caused by skewed training data. Such bias can cause models to produce inaccurate, offensive, or misleading predictions.	MIT Sloan Teaching & Learning Technologies
Chatbot	A computer program that simulates and processes human conversation (either written or spoken), allowing humans to interact with digital devices as if they were communicating with a real person.	Oracle
Data mining	The use of machine learning and statistical analysis to uncover patterns and other valuable information from large data sets.	IBM
Generative AI	A technological approach that enables the creation of content including text, images, and videos. By analyzing and discerning patterns within extensive training datasets, generative AI can autonomously construct material that shares comparable characteristics to its training input. This capability stems from the AI's understanding of data patterns and its ability to replicate or innovate based on these patterns.	MIT Sloan Teaching & Learning Technologies

Guardrails	Restrictions and rules placed on AI systems to make sure that they handle data appropriately and don't generate unethical content.	Coursera
Hallucination	Incorrect or misleading results that AI models generate. These errors can be caused by a variety of factors, including insufficient training data, incorrect assumptions made by the model, or biases in the data used to train the model.	Google Cloud
Information Technology	Using computers to work with, store, access, and send information.	Allied Media Projects / A People's Guide to Tech
Large language model	A category of foundation models trained on immense amounts of data, making them capable of understanding and generating natural language and other types of content to perform a wide range of tasks.	IBM
Machine learning	The subset of artificial intelligence (AI) that focuses on building systems that learn—or improve performance—based on the data they consume.	Oracle
Model	A program that has been trained from specific datasets to recognize, identify, and/or reproduce certain patterns.	Allied Media Projects / A People's Guide to Tech
Natural language processing	A subfield of artificial intelligence and computational linguistics that focuses on enabling machines to understand, interpret, and generate human language to be understood by humans.	MIT Sloan Teaching & Learning Technologies
Pattern recognition	The ability of machines to identify patterns in data, and then use those patterns to make decisions or predictions using computer algorithms.	Viso
Predictive analytics	The process of using data to forecast future outcomes. The process uses data analysis, machine learning, artificial intelligence, and statistical models to find patterns that might predict future behavior.	Google Cloud
Prescriptive analytics	The practice of analyzing data to identify patterns, which can be used to make predictions and determine optimal courses of action.	IBM

Prompt	The input you provide to the model to elicit a specific response. This can take various forms, ranging from simple questions or keywords to complex instructions, code snippets, or even creative writing samples.	Google Cloud
Structured data	Data that fits neatly into data tables and includes discrete data types such as numbers, short text, and dates.	Amazon Web Services
Token	The smallest unit of text that an AI model processes and understands; this is typically 4 characters in English, or about 3/4 of a word. Tokens may include whole words, parts of words, individual characters, punctuation marks, and special characters.	MIT Sloan Teaching & Learning Technologies
Training	The process of teaching an AI system to improve and learn from existing data so that it can accurately make sense of any new data that it is given.	Allied Media Projects / A People's Guide to Tech
Unstructured data	Data doesn't fit neatly into a data table because its size or nature: for example, audio and video files and large text documents. Sometimes, numerical or textual data can be unstructured because modeling it as a table is inefficient.	Amazon Web Services

Appendix B: Types of Use Cases

This section includes an enumeration of a set of prospective “use cases” – i.e. beneficial uses of AI tools – which have been either already implemented or at least imagined by members of the Task Force and the Community Navigation community at large.

The Task Force prioritized a set of Use Cases (specifically focusing on instances that are already being discussed) for which we conducted risk/benefit analysis. These use cases have been highlighted in blue below.

One analysis – for [Inform USA’s Accreditation Site Review Transcription](#) – was built upon with [a draft proposal for an “acceptable use policy” as highlighted in green](#), included in this document above.

Types of Users

- **People we serve** – callers, clients, etc.
- **Staff** – community resource navigators and/or specialists
- **Organizational Leadership** – executives, analysts
- **Partner Organizations**

Types of Uses

- **Administrative Tasks**
 - Transcription [\[Notes\]](#)
 - [Accreditation reviews for Inform USA \(Draft use case analysis\)](#)
 - [Draft “acceptable use policy” for this use case](#)
 - [Call / Client Interaction Transcription](#)
 - [Meeting Notes Transcription](#)
 - [Agency Interaction Transcription](#)
 - Summarization [\[Notes\]](#)
 - Meeting Notes Summarization
 - Agency Interaction Summarization
 - Client Interaction Summarization
 - Email Composition
 - Project Planning
 - Employee recruitment

- **Algorithmic Assessment / Predictive Analytics**
 - Risk Assessment
 - Rating people's likelihood of negative health/social outcomes
 - Eligibility determination
 - Screening for benefits access
- **Automating Workflows**
- **Chatbots** [Notes]
 - Help Seeker-Facing Chatbots – Helping users via front-end chatbot interfaces
 - Interface for decisiontree (pre-determined set of possibilities)
 - Generative predictive text like ChatGPT (open-ended set of possible – and essentially unpredictable – responses)
 - Chatbots for Training Scenarios – supporting trainees via chatbot simulations
- **Data Quality**
 - Styleguide review and improvement –
 - Aligning Different Style Guides into a Shared Standard
 - Resource Data Quality Assessment
 - Resource Data Style/Standards Evaluation
 - Resource Data Quality Improvement [Notes]
 - Data Cleaning
 - Producing Structured Data from Unstructured Free Text
 - Web Scraping
 - Language Translation
- **Sentiment Analysis**
 - Individual Interactions
 - Post-Call QA
 - Aggregate Pattern Analysis
- **Text Suggestions**
 - Resource Recommendations
 - For Helpseekers
 - For Resource Navigators
 - Pathway navigation
 - Taxonomy Term Management and Mapping
- **Community needs assessment**
 - Analysis of resource allocation patterns to understand assets and gaps
 - Analysis of patterns in requests for help to understand community needs

Appendix C: Types of Risks

The AI Task Force assessed a broad array of literature about the risks involved with algorithmic technologies. See [this table in our shared knowledge base](#). We identified at least two ‘taxonomies’ of risks that themselves were produced from large-scale literature review – *i.e.*, analysis of hundreds or thousands of studies each. (See: [The MIT AI Risk Repository](#), and [A Collaborative, Human-Centred Taxonomy of AI, Algorithmic, and Automation Harms](#).) We assessed the overlap of these taxonomies, and prioritized among the risks that seem most relevant to Community Navigation. The below list of risks reflects the broad array of material aggregated in [our Knowledge Base](#), and has been used to guide our initial cycle of “use case analysis” – such that each use case was considered in light of this set of generically framed risk possibilities. (More detailed “use case analysis” should further refine these generic risk types into particular descriptions of context-specific scenarios.)

We also want to note **one prominent risk that is not independently enumerated in any of these taxonomies: the risk that the benefit of an AI project might be outweighed by its total cost** – in licensing fees, workflow design, staff time, evaluation, and etc. – especially when compared to other alternative approaches that might achieve a similar objective. (This risk, of course, should be considered as a possibility in the initial stages of any technology project of any kind.)

Discrimination & Toxicity

1.1 Unfair discrimination and misrepresentation

Unequal treatment of individuals or groups by AI, often based on race, gender, or other sensitive characteristics, resulting in unfair outcomes and representation of those groups.

1.2 Exposure to toxic content

AI exposing users to harmful, abusive, unsafe or inappropriate content. May involve AI creating, describing, providing advice, or encouraging action. Examples of toxic content include hate speech, violence, extremism, illegal acts, child sexual abuse material, as well as content that violates community norms such as profanity, inflammatory political speech, or pornography.

1.3 Unequal performance across groups

Accuracy and effectiveness of AI decisions and actions is dependent on group membership, where decisions in AI system design and biased training data lead to unequal outcomes, reduced benefits, increased effort, and alienation of users.

Privacy & Security

2.1 Compromise of privacy by obtaining, leaking or correctly inferring sensitive information

AI systems that memorize and leak sensitive personal data or infer private information about individuals without their consent. Unexpected or unauthorized sharing of data and information can compromise user expectation of privacy, assist identity theft, or loss of confidential intellectual property.

2.2 AI system security vulnerabilities and attacks

Vulnerabilities in AI systems, software development toolchains, and hardware that can be exploited, resulting in unauthorized access, data and privacy breaches, or system manipulation causing unsafe outputs or behavior.

Misinformation

3.1 False or misleading information

AI systems that inadvertently generate or spread incorrect or deceptive information, which can lead to inaccurate beliefs in users and undermine their autonomy. Humans that make decisions based on false beliefs can experience physical, emotional or material harms

Malicious actors

None of the items in this category were rated as a high risk by any members of the Task Force, though this may be re-assessed in the future.

Human-Computer Interaction

5.1 Overreliance and unsafe use

Users anthropomorphizing, trusting, or relying on AI systems, leading to emotional or material dependence and inappropriate relationships with or expectations of AI systems. Trust can be exploited by malicious actors (e.g., to harvest personal information or enable manipulation), or result in harm from inappropriate use of AI in critical situations (e.g., medical emergency). Overreliance on AI systems can compromise autonomy and weaken social ties.

5.2 Loss of human agency and autonomy

Humans delegating key decisions to AI systems, or AI systems making decisions that diminish human control and autonomy, potentially leading to humans feeling disempowered, losing the ability to shape a fulfilling life trajectory or becoming cognitively enfeebled.

Socioeconomic & Environmental

6.1 Power centralization and unfair distribution of benefits

AI-driven concentration of power and resources within certain entities or groups, especially those with access to or ownership of powerful AI systems, leading to inequitable distribution of benefits and increased societal inequality.

6.2 Increased inequality and decline in employment quality

Widespread use of AI increasing social and economic inequalities, such as by automating jobs, reducing the quality of employment, or producing exploitative dependencies between workers and their employers.

6.3 Economic and cultural devaluation of human effort

AI systems capable of creating economic or cultural value, including through reproduction of human innovation or creativity (e.g., art, music, writing, code, invention), can destabilize economic and social systems that rely on human effort. This may lead to reduced appreciation for human skills, disruption of creative and knowledge-based industries, and homogenization of cultural experiences due to the ubiquity of AI-generated content.

6.4 Competitive dynamics

AI developers or state-like actors competing in an AI 'race' by rapidly developing, deploying, and applying AI systems to maximize strategic or economic advantage, increasing the risk they release unsafe and error-prone systems.

6.5 Governance failure

Inadequate regulatory frameworks and oversight mechanisms failing to keep pace with AI development, leading to ineffective governance and the inability to manage AI risks appropriately.

6.6 Environmental harm

The development and operation of AI systems causing environmental harm, such as through energy consumption of data centers, or material and carbon footprints associated with AI hardware.

AI System Safety, Failures, & Limitations

7.3 Lack of capability or robustness

AI systems that fail to perform reliably or effectively under varying conditions, exposing them to errors and failures that can have significant consequences, especially in critical applications or areas that require moral reasoning.

7.4 Lack of transparency or interpretability

Challenges in understanding or explaining the decision-making processes of AI systems, which can lead to mistrust, difficulty in enforcing compliance standards or holding relevant actors accountable for harms, and the inability to identify and correct errors.

Causal Taxonomy of AI Risks

Category	Level	Description of how the risk is presented in evidence
Entity	AI	The risk is caused by a decision or action made by an AI system
	Human	The risk is caused by a decision or action made by humans
	Other	The risk is caused by some other reason or is ambiguous
Intent	Intentional	The risk occurs due to an expected outcome from pursuing a goal
	Unintentional	The risk occurs due to an unexpected outcome from pursuing a goal
	Other	The risk is presented as occurring without clearly specifying the intentionality
Timing	Pre- deployment	The risk occurs before the AI is deployed
	Post- deployment	The risk occurs after the AI model has been trained and deployed
	Other	The risk is presented without a clearly specified time of occurrence

From: <https://airisk.mit.edu>

The **Entity** variable captures which, if any, entity is presented as the main cause of the risk.

- When the risk is attributed to **AI**, it means that the risk arises from decisions or actions made by the AI system itself, such as generating harmful content or disempowering humans.

- Conversely, when **Humans** are seen as the source, the risks are implied to be due to human actions like choosing poor training data, intentional malicious design, or improper use of AI systems.
- The **Other** category captures cases where the focal entity is not a human or AI or is ambiguous.

The **Intent** variable captures whether the risk is presented as occurring as an expected or unexpected outcome from pursuing a goal.

- **Intentional** risks are those that occur as expected outcomes from pursuing a specific goal, such as a case where AI is intentionally programmed to act deceptively or to exhibit bias.
- **Unintentional** risks reflect unintended consequences, such as an AI system inadvertently developing biases due to incomplete training data.
- The **Other** category captures risks where the intent is not clearly specified.

The **Timing** variable captures the stage in the AI lifecycle at which the risk is presented as occurring.

- **Pre-deployment** risks are those that arise before the AI system is fully developed and put into use.
- **Post-deployment** risks arise after the AI has been deployed, including issues like the misuse of AI for harmful purposes.
- The Other category is used for risks that do not have a clearly defined time of occurrence.

Appendix D: Recommended Readings

Across almost nine months of work, Task Force members collaborated in the process of finding relevant resources – articles, research, guidance, etc – that could inform our reflections and recommendations. We have aggregated these resources in [this Knowledge Base](#). A prioritized subset of these resources are shared here in this appendix.

All of the positions put forth in this report are grounded in analysis drawn from these materials.

Principles

[AI Principles](#), Organisation for Economic Co-operation and Development

[2024 Blueprint for an AI Bill of Rights](#), the White House

[“Guidelines for Participatory and Inclusive AI.”](#) Partnership on AI’s Global Task Force for Inclusive AI

Research

[“Gen AI: Too Much Spend, Too Little Benefit?”](#) Goldman-Sachs Global Macro Consulting

[“Promising algorithms to perilous applications: a systematic review of risk stratification tools for predicting healthcare utilisation,”](#) BMJ Health Care and Informatics

[“AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking.”](#) Michael Gerlich, Societies Journal.

[“The more sophisticated AI models get, the more likely they are to lie,”](#) Ars Technica

[Large Language Models, Small Labor Market Effects.](#) Humlum & Verstergaard working paper

[AI-Powered Rules as Code: Experiments with Public Benefits Policy.](#) Beeck Center’s Digital Benefits Network

Risks

MIT’s AI Risk Repository: <https://airisk.mit.edu>

AI Risk Management Framework, National Institute for Standards in Technology
<https://www.nist.gov/itl/ai-risk-management-framework>

[Taxonomy of AI Trustworthiness](#), UC Berkley Center for Long Term Cybersecurity

["A Collaborative, Human-Centred Taxonomy of AI, Algorithmic, and Automation Harms."](#) Many authors, AI and Automation Incidents and Controversies project

[Chatbots Are Primed to Warp Reality](#); Matteo Wong, the Atlantic

Analysis

AI Snake Oil by Sayash Kapoor, Arvind Narayanan
<https://press.princeton.edu/books/hardcover/9780691249131/ai-snake-oil>

"Envisioning Information Access Systems: What Makes for Good Tools and a Healthy Web?"
Emily Bender and Chirag Shah, <https://dl.acm.org/doi/10.1145/3649468>

Inescapable AI – The Ways AI Decides How Low-Income People Work, Live, Learn, and Survive, TechTonic Justice. <https://www.techtonicjustice.org/reports/inescapable-ai>

"Watching the Generative AI Hype Bubble Deflate," Harvard Ash Center for Democratic Governance and Innovation.
<https://ash.harvard.edu/resources/watching-the-generative-ai-hype-bubble-deflate>

"Diversity and Artificial Intelligence: A Social Work Challenge." Sue Coyle
<https://www.socialworktoday.com/archive/ND19p12.shtml>

Guidance

AI 101 – Aspen Digital – <https://www.aspendigital.org/report/ai-101>

Member Data and the Ethics of AI, ASAE – Center for Association Leadership
https://www.asaecenter.org/resources/articles/an_plus/2025/03-march/member-data-and-the-ethics-of-ai

From Principles to Practice, 2024 White House report:

<https://web.archive.org/web/20250116062754/https://www.whitehouse.gov/ostp/ai-bill-of-rights/from-principles-to-practice>

The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed

RAND Corporation, https://www.rand.org/pubs/research_reports/RR2680-1.html

Public AI: Making AI work for everyone, by everyone. Mozilla

https://assets.mofoprod.net/network/documents/Public_AI_Mozilla.pdf

Digital Decision-making Tool Center for Democracy and Technology,

<https://cdt.org/insights/digital-decisions-tool/>