



Ethical AI in the Social Sector



Smita Jain
Director of Data & Analytics
Mastercard Center for
Inclusive Growth



Sarah Di Troia
Managing Director,
OutcomesAI
Project Evident



Rachel Wells
Director of Data Science,
Education
DataKind



Pahniti (Tom) Tosuksri
Director of Product
Management
Grounded Solutions



Data for Social Impact Initiative

OBJECTIVE: To build connection and capacity among social sector organizations working to increase impact *for* and *with* the communities they serve; to cultivate collaborative and equitable data practices across government, philanthropy, nonprofits, and impacted communities.





Chat Question:

How would you describe your current attitude toward AI?

Anything from mostly excited to cautiously interested to apocalyptic, we want to hear it! Please put in the chat.



Refrain: Social sector organizations lag behind the private sector in their capacity to leverage AI to increase impact.

Reframe: Social sector organizations are positioned to lead in developing ethical and equitable approaches to AI.



- Overview of AI in the social sector
- Ethical AI considerations/ frameworks
- One organization's ethical AI journey
- AI tool demonstration/ activity
- How does AI show up in your work?

How are you approaching the ethical implications of these tools?



How familiar are you with using AI?

1. I use AI regularly in my work or daily life.
2. I have a general understanding of AI, but I don't use it often.
3. I've heard of AI, but I don't use it
4. I'm not familiar with AI.





Ethical AI in the Social Sector



Smita Jain
Director of Data & Analytics
Mastercard Center for
Inclusive Growth



Sarah Di Troia
Managing Director,
OutcomesAI
Project Evident



Rachel Wells
Director of Data Science,
Education
DataKind



Pahniti (Tom) Tosuksri
Director of Product
Management
Grounded Solutions Network

Ethical AI Washington University in St. Louis

SARAH DI TROIA

DECEMBER 2024



Who We Are

Project Evident harnesses the power of evidence for greater impact.

We were founded eight years ago, fueled by constructive dissatisfaction with the way evaluation was conducted in the social sector and the ambition to make data and evidence work in favor of practitioners and those they serve.

What is Artificial Intelligence?

AI

Capability of a computer system to mimic human cognitive functions such as learning and problem-solving functions

GENERATIVE AI

Capability of a computer system to create convincing or pleasing imagery, compelling text, or coherent audio outputs

What is the AI Opportunity?

- AI has the ability to accelerate equitable outcomes by:
 - Supporting automation and reduction of repetitive tasks to free up time for higher value activities
 - Reducing decision making complexity
 - Powering R&D to support learning, innovation and improvement

What are Common Applications?

AI

GENERATIVE AI

MACHINE LEARNING

PREDICTIVE ANALYTICS

Predict trends and behavioral patterns by discovering cause-and-effect relationships in data



RECOMMENDATION ENGINES

Use data analysis to recommend products or a course of action that someone might be interested in

NETFLIX

NATURAL LANGUAGE PROCESSING

Understand and recognize meaning in written or spoken language



LARGE LANGUAGE MODEL

Uses large amounts of language to create new sentences based on predicting the next word.



How are practitioners using AI to fuel outcomes?

PREDICTIVE ANALYTICS

Lower the risk a child has after discharge from their long-term psychiatric care, 92% patients had a reduced acuity score.



RECOMMENDATION ENGINES

Increase family income to meet basic needs, connected 1.2 million people to a public benefit.



NATURAL LANGUAGE PROCESSING

Increase connection between families and teachers across language barriers, 15% decline in absenteeism



GENERATIVE AI

How to efficiently and effectively train ~9.5K volunteers crisis counselor and build their confidence to reduce churn

CRISIS TEXT LINE |

How can social sector organizations incorporate AI?

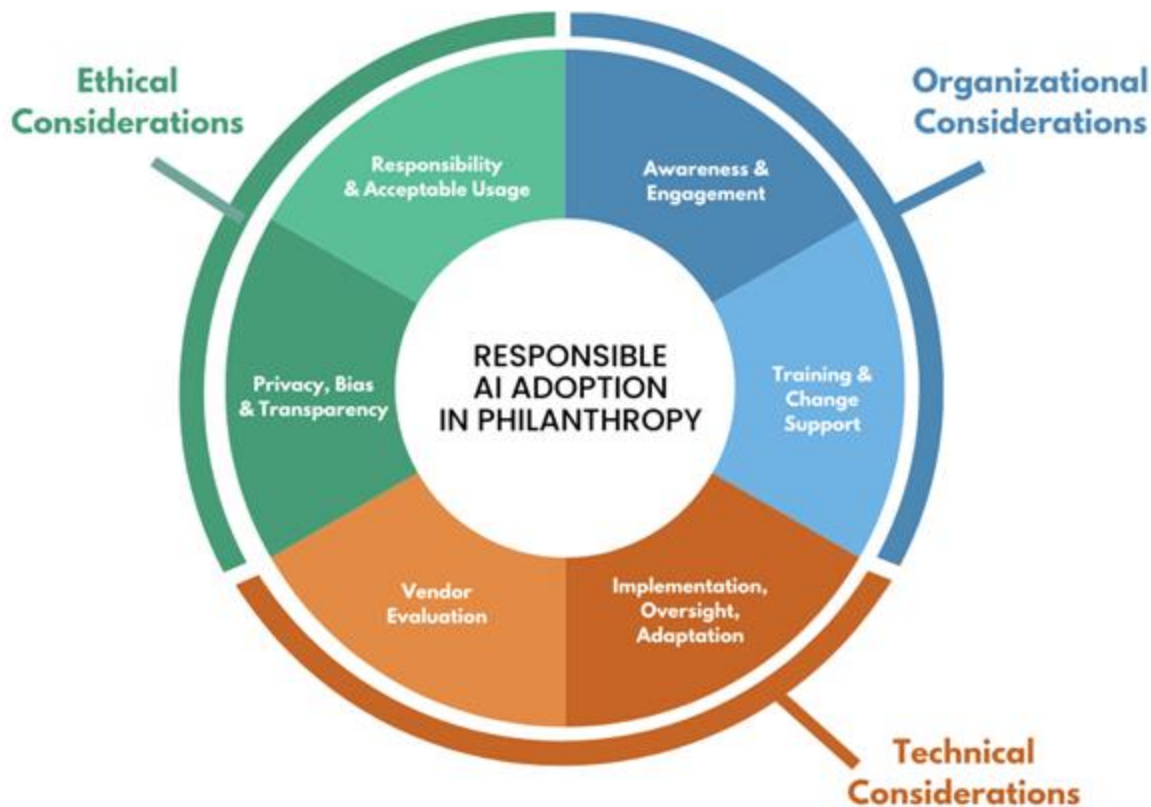


Tools provisioned by the organization for discovery, insight, fundraising, impact assessment, predictive analysis.

Tools provisioned by the organization to save time, predict accurately, and improve operational efficiency.

Staff/stakeholder use of publicly available AI tools that are not provisioned by the organization.

Responsible AI Adoption



Download the Framework



Or visit projectevident.org/resources



Ethical AI

Rachel Wells

Dec 10, 2024

DSI @ WUSTL

DataKind[®]

datakind.org | [@DataKind](https://twitter.com/DataKind)



DataKind®

Harnessing the
power of data
science + AI in
the service of
humanity



© 2020 Medic photo archives



Hallucination

Plausible-looking responses which are factually incorrect

Bias

Model surfacing bias and prejudice as found on the internet; tools being inaccessible or not inclusive across languages and other human differences, increasing inequity

Transparency

Commercial models are usually closed access

Misuse

Data leakage (private data shared by LLM exposed), deep fakes (modification of real content to deceive), or prompt injection (model exploitation/ jailbreak)

Development

Environmental impact, worker exploitation, and intellectual infringement (resulting in plagiarism) involved in model training

Targeting Safer Solutions

Using an LLM to predict technical metadata on data provided

Using an LLM as part of search and summarizing ONLY data provided

Asking an LLM questions where the answers are from its training data

Low Risk

High Risk

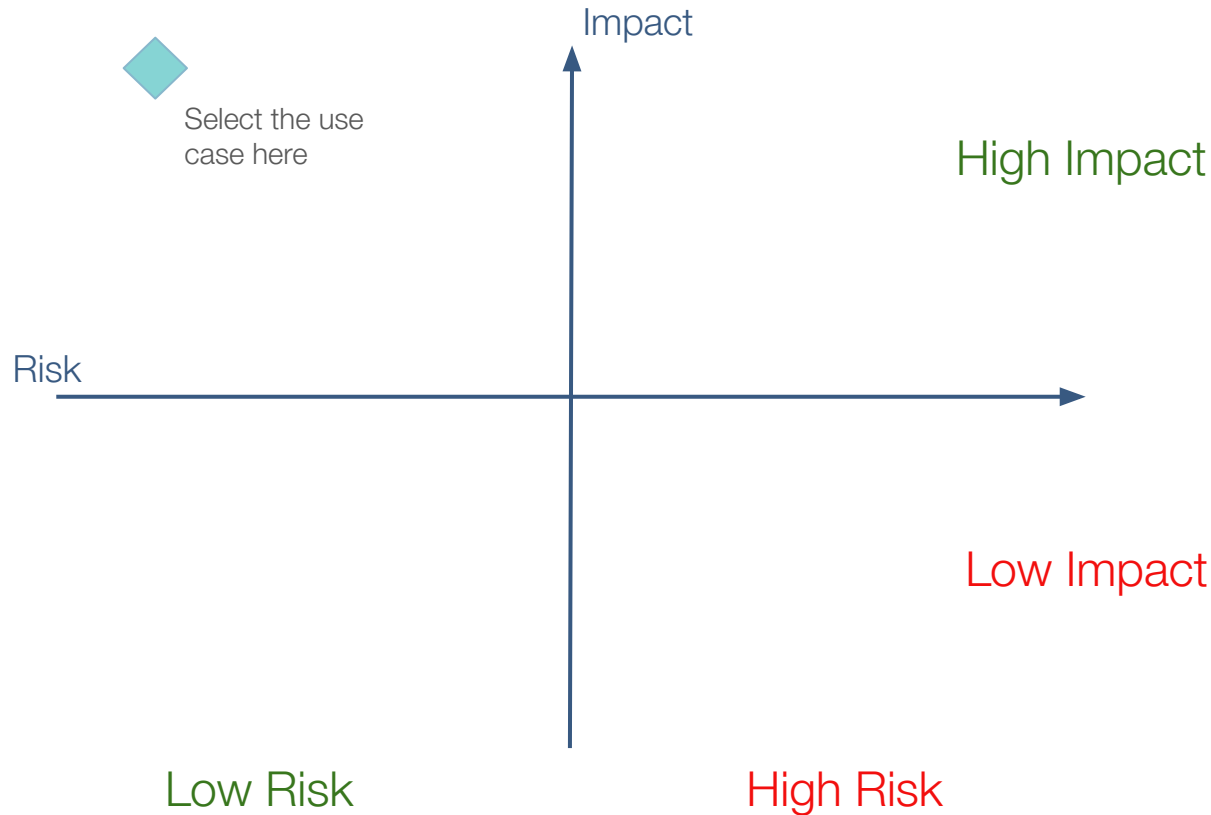
¹ 'LLM' is Large Language Model for example GPT-4, ChatGPT, Llama 2

10 Rules of Thumb for Risk Mitigation

1. **Purpose:** Clearly define the goals and intended use of AI-generated media, whether for internal or external use.
2. **Standardization:** Prefer standardized AI tools to safeguard data and privacy
3. **Account Setup:** Avoid using organizational credentials with public AI tools to prevent privacy issues
4. **Prompt Engineering:** Employ effective prompting techniques for clear, precise communication of expected outcomes.
5. **Confidentiality:** Never share sensitive information with AI tools without authorization, regardless of sensitivity.
6. **Synthetic Media Guidelines:** Follow established synthetic media guidelines, such as those by the Partnership on AI
7. **Contextual Sensitivity:** Ensure AI media is contextually appropriate and non-manipulative
8. **Bias Avoidance:** Ensure no biases (gender, racial, ethnic) are present
9. **Transparency:** Disclose the use of AI tools, the processes involved, and cite sources appropriately
10. **Fact Checking:** Verify the accuracy of AI-generated content

Area(s) of risk	Sample mitigation approach(es)
Hallucination , information loss, errors	Limit data sources, always include references, automated validation techniques, human-in-the-loop fact checking reviews, etc.
Data leakage	Data security systems, select a tool that protects privacy
Misuse , model exploitation, prompt injection (Jailbreak)	Setup model safety guardrails and build in accountability with observability, monitoring, red teaming, etc.
Intellectual infringement	Only present content which adheres to usage rights
Bias	Test for equity and inclusivity for historically disadvantaged communities
Transparency	Select a constrained use case for clarity into how it's working
Inaccessibility (language, internet access, etc.)	Inclusion of translation checks, create accommodations to increase accessibility, etc.
Carbon footprint /climate change	Minimize computing power used & conduct cost benefit analysis
Worker exploitation	Ethical working standards

Define your AI problem statement



Community informed priority data pain points

Actors across the humanitarian community raised similar data challenges hindering effective response & coordination - for which a taxonomy was developed

Pain point	% of pain points in interviews, survey & human lit review	% of pain points in AI research synthesis
Data Harmonization - Interoperability	20%	8%
Data Skills and Capacity - Data Literacy and awareness	18%	13%
Data Discovery - Sharing	6%	3%
Data Availability - Disaggregation	6%	2%
Data Availability - Timeliness	5%	3%
Security and Privacy - Security and permissions	3%	0%
Security and Privacy - Data damage	3%	2%
Data Availability - Completeness	3%	2%
Quality and Accuracy - Consistency	2%	1%
Ethical Challenges - Misuse and distortion	2%	1%
Data Discovery - Search	2%	0%
Data Discovery - Overwhelming amount of data	2%	2%

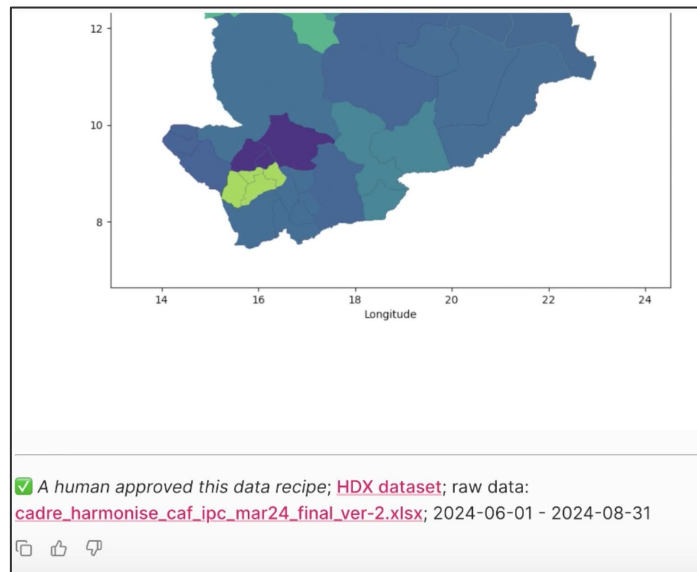
Interoperability - linking datasets together and building data models

Data literacy and awareness - finding and knowing how to use data effectively

Humanitarian AI Assistant *(and Data Recipes AI)*

Goal

Making usable and useful data analysis possible, with lower barriers to access for non-technical users to discover, explore and ask questions of humanitarian datasets, so that humanitarians and front-line responders can act with data-informed responses and organizations can steward data resources more effectively to producing timely and actionable information.



- [hapi_metadata_admin2](#)
- [hapi_metadata_wfp_market](#)
- [hapi_population_social_population](#)
- [hdx_shape_files](#)

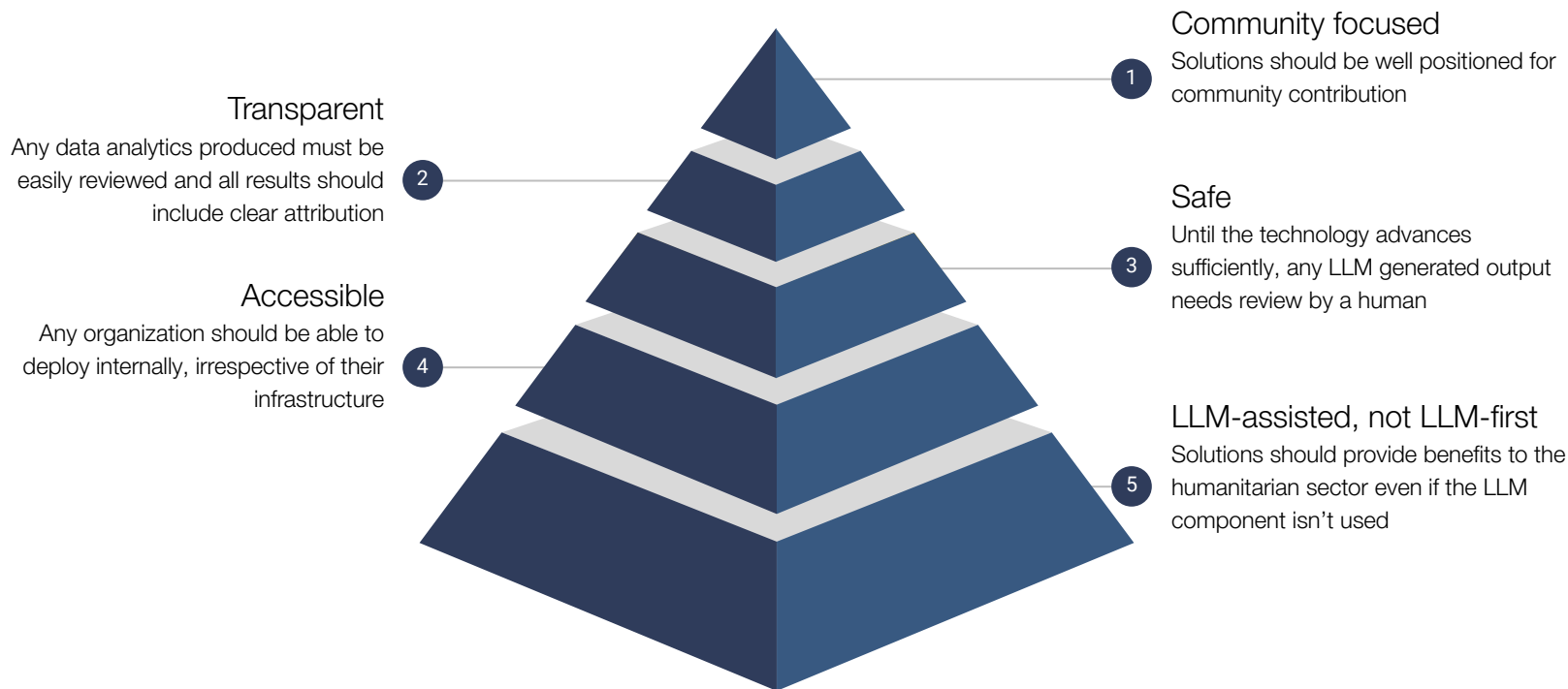
These tables provide a comprehensive dataset for analyzing different dimensions of the Wadi Fira region.

📄 👍 🗨

Plot a map showing food security population in IPC Phase 3 in Chad

several tables. These tables
presence, food prices, food
statistics, and shape files.

Design principles



Student advisors are assigned more students than they can proactively support

*Data-assisted advising
helps advisors use their
limited time to more
efficiently identify and
reach out to those most in
need of help*



Student advisors have many interventions they can implement

Data assisted advising helps advisors identify and build plans for the interventions that fit best for the student

Financial aid support



Academic support and program planning



Building connections to trusted resources



Targeted planning conversations



Outreach to students in need of support

Generate insights on potential areas for support



Adopting data-assisted academic advising

- DataKind has implemented student success tools since 2017 to enable data-assisted advising
- Using our tools, John Jay College has reported a **30% increase** in senior graduation rates in two years, as featured in the [NYTimes](#)
- **We are philanthropy supported to bring this solution to implementation at additional institutions, at no institutional cost**



DataKind's product principles

- **Transparent:** Our features and models will be openly shared, so you can know exactly what variables are leading to identifying those student most at risk of non graduation.
- **Dedicated to bias reduction:** We'll use bias-reducing techniques and regularly review our implementations for fairness and equity.
- **Humans in the loop by design:** Our interventions are designed to be additive to the student experience, and all algorithms are implemented through human actors (advisors).



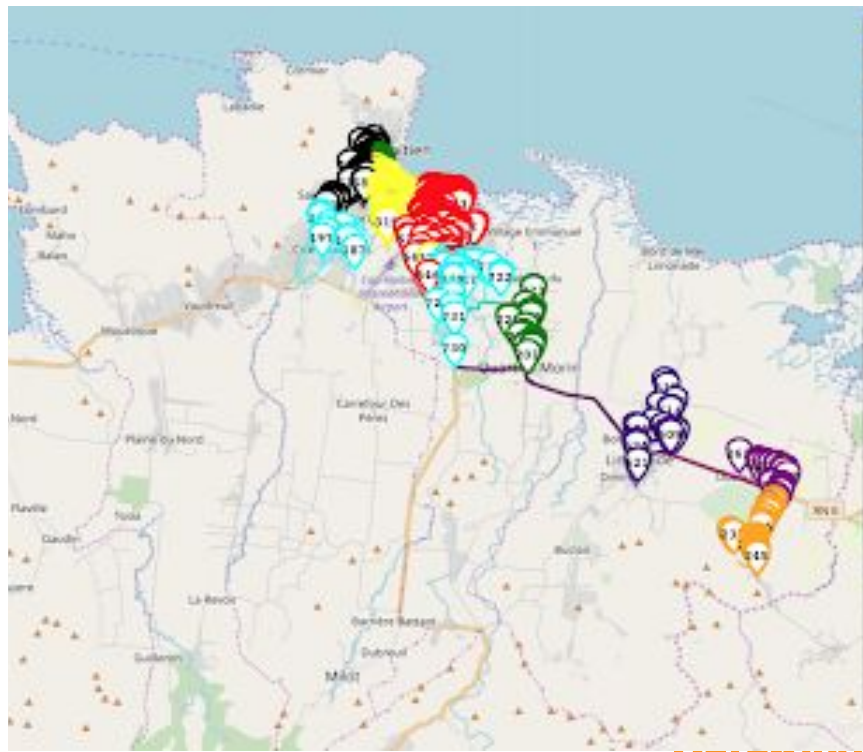
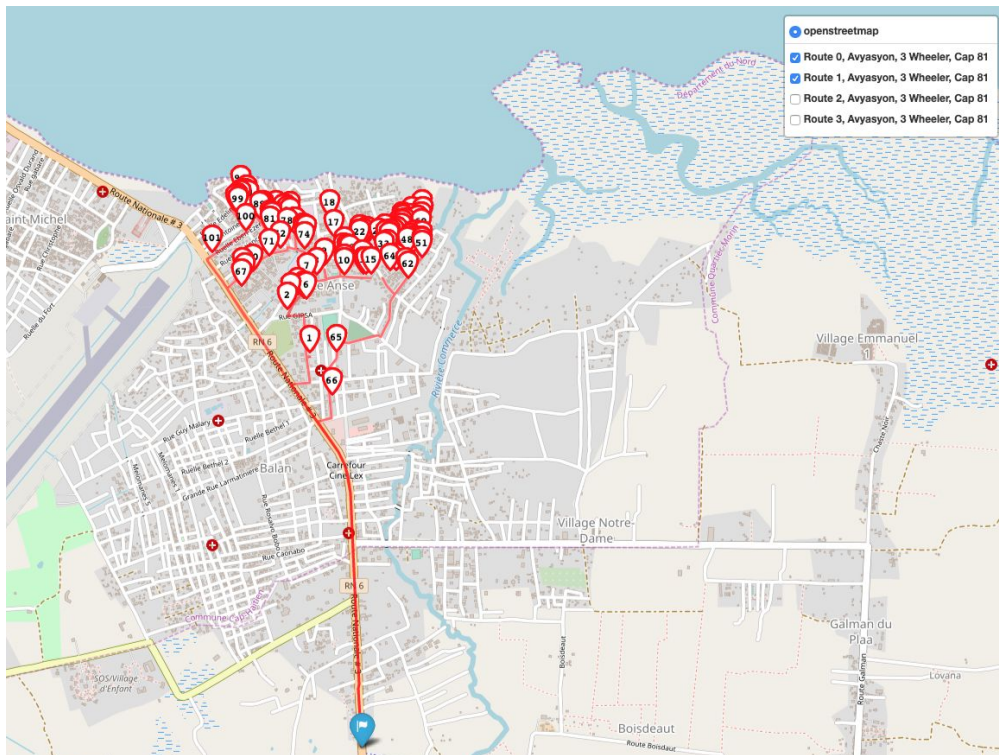


Optimizing service delivery to promote dignity, health and sustainable livelihoods in Haiti



DataKind

DataKind & SOIL's route optimization



How can my organization get started with using AI for a positive social impact?

Baseline learning

Assess the current knowledge of GenAI, data maturity, and technical literacy for different application approaches.

Define learning paths and risks based on the current situation.

Turn values to policy

Build your foundation with standard operating procedures and/or policy.

Identify clear guardrails for your organization in alignment with your values.

Evaluate opportunities

Consider challenges, constraints, and guardrails for risk mitigation for any possible use cases.

Identify and prioritize relevant use cases to define a “vision” and to find focus.

Set up tools & tech to integrate

Use tools developed by others or integrated into your existing systems (ChatGPT, Copilot, etc.)

Develop and integrate into your own tools, if your organization has the capacity and literacy for ethical application.

DataKind[®] Playbook



Nonprofit
Data 101

DataKind[®]

Read the best tips & lessons learned in the [DataKind Playbook](https://datakind.github.io) for social impact professionals (datakind.github.io)

- + Getting started with [Data Engineering](#)
- + [GenAI](#): Getting started, use case selection & risk assessment, setting up tools, and more
- + [Data project stages](#) and best practices and processes for each
- + [Data project scoping](#)

Watch the recordings from our Nonprofit Data 101 webinar series introducing various data science topics to nonprofit practitioners in our [youtube playlist](#)

- + Data project scoping
- + GenAI: the tech, opportunities, and risks
- + A year of GenAI learning
- + Non-data stakeholder engagement
- + ML, AI, Oh My!

More ways to get involved with **DataKind®**

Subscribe to our newsletter to stay up-to-date:

www.datakind.org/subscribe

Connect with us directly at

community@datakind.org



WE



DATA



GROUND SOLUTIONS NETWORK

strong communities
from the ground up

Through its member network and partnerships, GSN exponentially transforms the lasting affordability housing sector in the US.

It does this by preserving, producing and sustaining housing with a specific focus on BIPOC, under-invested in and marginalized communities.



Our Generative AI Journey

A simple ask (Sep '23)



I need some help cutting down
time between meetings and
capturing notes and action items.

Can I use this AI thing?



Research




ai notetaking privacy

privacy policy

privacy concerns

Feeling Lucky

←  r/privacy • 21 days ago
UnitedSorbet127

Warning: AI Company [redacted] violates privacy rights

discussion

I need to share a concerning experience with [redacted]. During a video conference, they recorded and transcribed our conversation without my consent. Despite my requests, they refuse to delete it.

Sometimes, responsible AI means no AI



Aleksandr Tiulkanov

AI, Data & Digital Policy Counsel, LL.M., CIPP/E | I help organisations navigate Digital Policy and operationalise AI and Data governance

January 24, 2023

Open Immersive Reader

Mentioning "facial recognition", in abstract, doesn't always bring positive associations. And when you add "reading emotions", you're almost guaranteed to evoke concern. It looks.

This journalist's [redacted] scare is a reminder that cloud transcription isn't completely private



If you upload data to a service, you can't count on it staying private.
Illustration by Alex Castro / The Verge

/ A reminder of the tradeoffs for ease and simplicity

By Mitchell Clark

Feb 16, 2022, 9:04 PM EST



Research



ALL-Purpose Text Generation

Marketing & Sales Copy

Coding & Web Development

Sound & Music Generation

Other

Knowledge Management & Search

Customer Service Automation

Chatbots

Erotica

Market Intelligence

7,000 AI applications released through Oct 2023

7,000 AI applications projected through end of 2023

IT'S TOO MUCH!

#SCHITTS CREEK

Pop TV

Consulted an expert



Draft me a policy for our non-profit organization around the ethical use of AI tools.

Include restrictions on using any free AI tools to only generic prompts, protecting any proprietary information or data.

Include ethical considerations around racial bias.

Ensure that there is a human element to each process workflow and make a suggestion on how we can do that.

Also include guidance on how we should lean more towards hiring talented creative contractors over using subscription based services for things like video creation, voice overs, and others in this category.

Brainstorm edge cases

for a function with birthdate as in

Explain airplane turbulence

to someone who has never flown b

Message ChatGPT...



Drafted AI Guidelines (Nov '23)



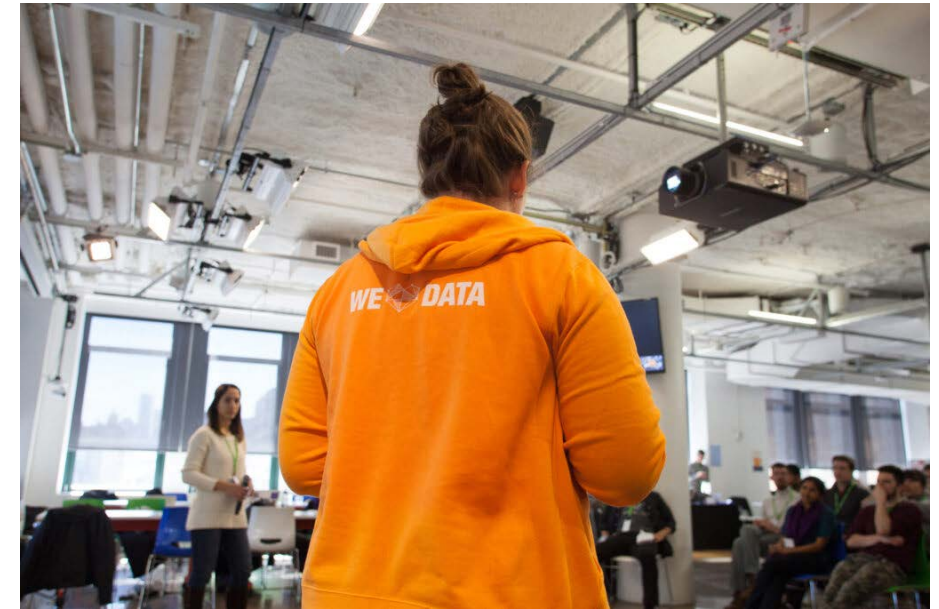
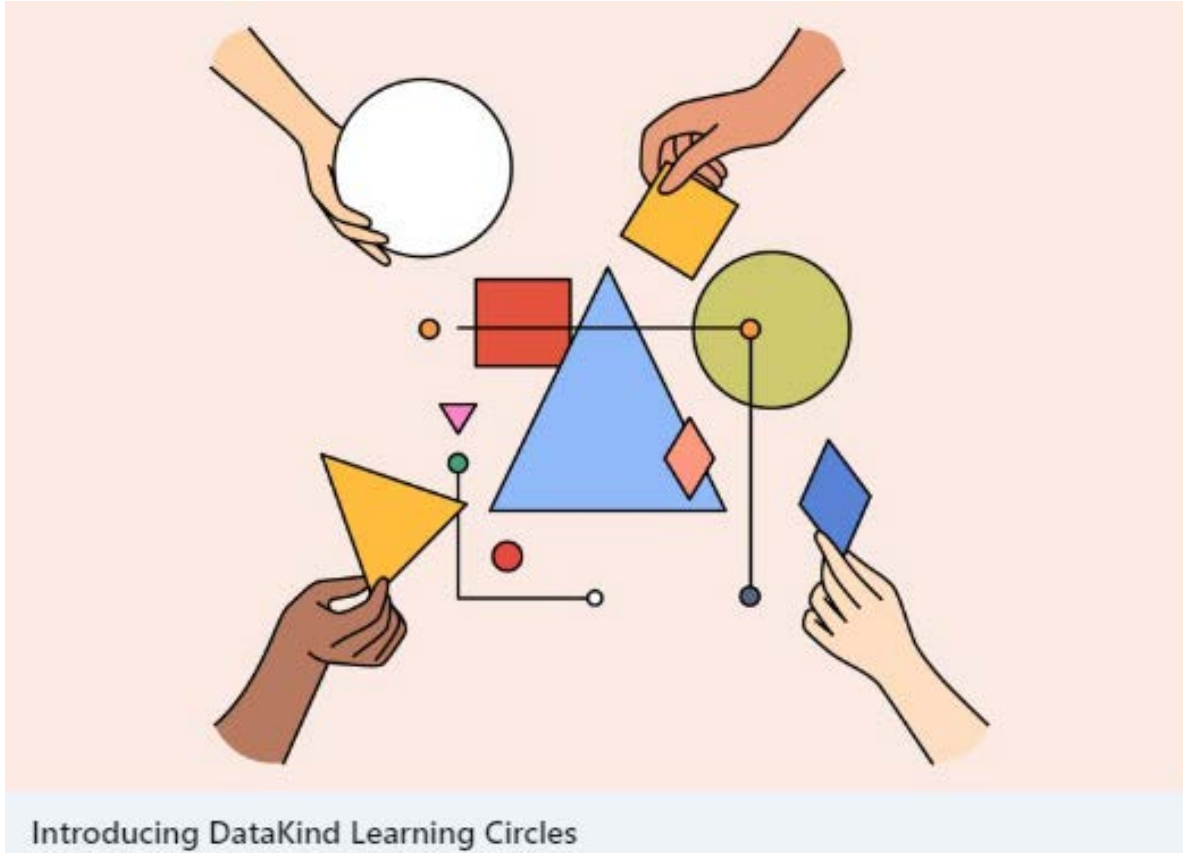
- ★ Ethical Considerations - Racial and other Biases
- ★ Restrictions and Allowances of AI Tools
 - Note-taking Tool
 - Proofreading Tool
- ★ Human Element
- ★ Talent Over Subscription Services
- ★ Augmentation, Not Replacement
- ★ Ongoing Training
- ★ Data Privacy
- ★ Continuous Improvement





Planning a Systematic Approach to Onboarding

Connecting with Real Experts (Jan '24)



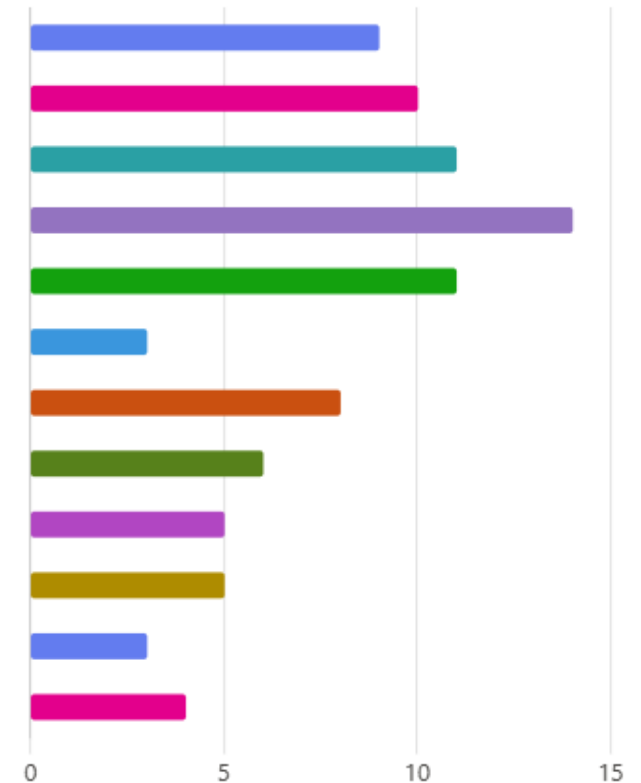
Crossing back over to work



Calling all use cases

3. In what ways would you like to explore using Chat GPT and/or other AI Tools at work?

- Information Search - better and more contextual way to find information online 9
- Reading a paper - used chatpdf or something similar to read a PDF and ask it... 10
- Write Something - asked it to write a paragraph or longer in a specific tone/voic... 11
- Proofreading - analyzed text I prepared for proofreading and grammatical errors 14
- Note Taking - used a "bot" to take notes during a meeting 11
- Speech Coaching - used MS Teams or others to analyze my body language during... 3
- Excel/Data Assistant - used an external tool to analyze a data sheet for insights or creat... 8
- Excel/Coding Assistant - used an external tool to analyze my actions and develop... 6
- Audio/Video Editing - used an external tool to edit videos - identify/eliminate pauses,... 5
- Audio/Video Creation - used an external tool to create videos from prompts 5
- Photo Editing/Creation - used an external tool like Dall-E to create/edit photos 3
- Other 4

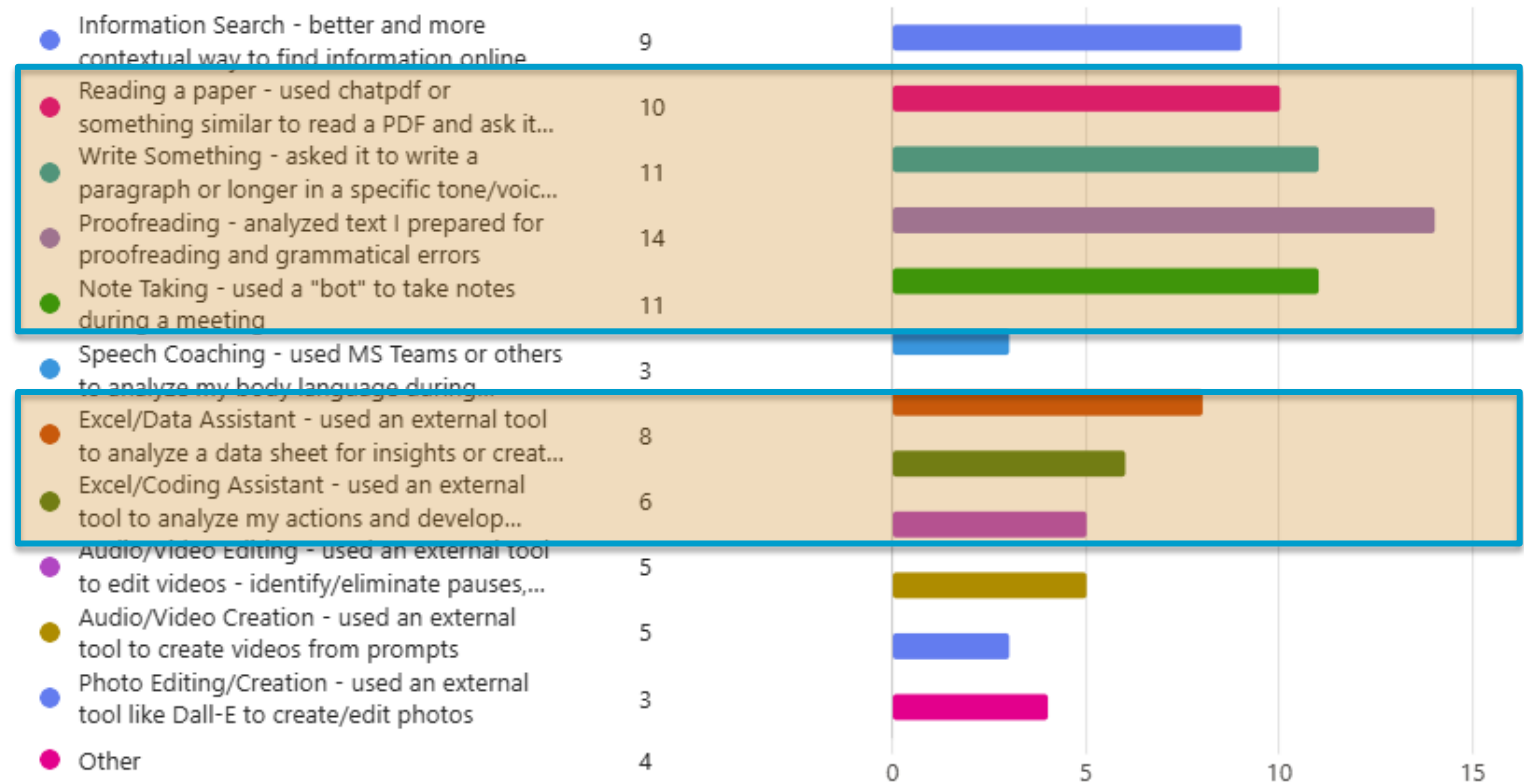


Crossing back over to work



Calling all use cases

3. In what ways would you like to explore using Chat GPT and/or other AI Tools at work?



Taking the Next Step



Tier 1: GenAI Products: Pre-defined Augmentation

**Meeting Notes
Taking**
Summaries
Action Steps

**Video /
Graphics**
Content Design
Editing

Proofreading
Summarizing
Commonizing
Language and
Writing

Cyber Security
Defense
vulnerabilities
Ransomware
protection

Tier 2: GenAI Chat GPT Assistance: Co-Pilots

AI Assistants
Excel
Coding

Coding / Formula
Python
Excel
Salesforce

Tier 3-A: Predictive Modeling & Machine Learning

Machine Learning
Statistical Regressions / Analysis
Trend Identification
Predictions

Tier 3-B: Chatbots and Conversational Knowledge Search

Chatbots / Search
Research & Knowledge
repositories
Help Desk Support
Employee Manual & Policies

Humans in the Approach



Novice

Prompter

Model Trainer

Assess

Any staff with little to no experience with prompting

Staff with higher level awareness and more technical use cases

Organization-wide projects with Lead

Prescribe

Internal training to use pre-defined tools with specific use cases

Basic prompt training alongside coaching, mentoring, and workgroup participation

Project scope technical infrastructure and business analysis of process changes

Tier 1: GenAI Products: Pre-defined Augmentation

Tier 2: GenAI Chat GPT Assistance: Co-Pilots

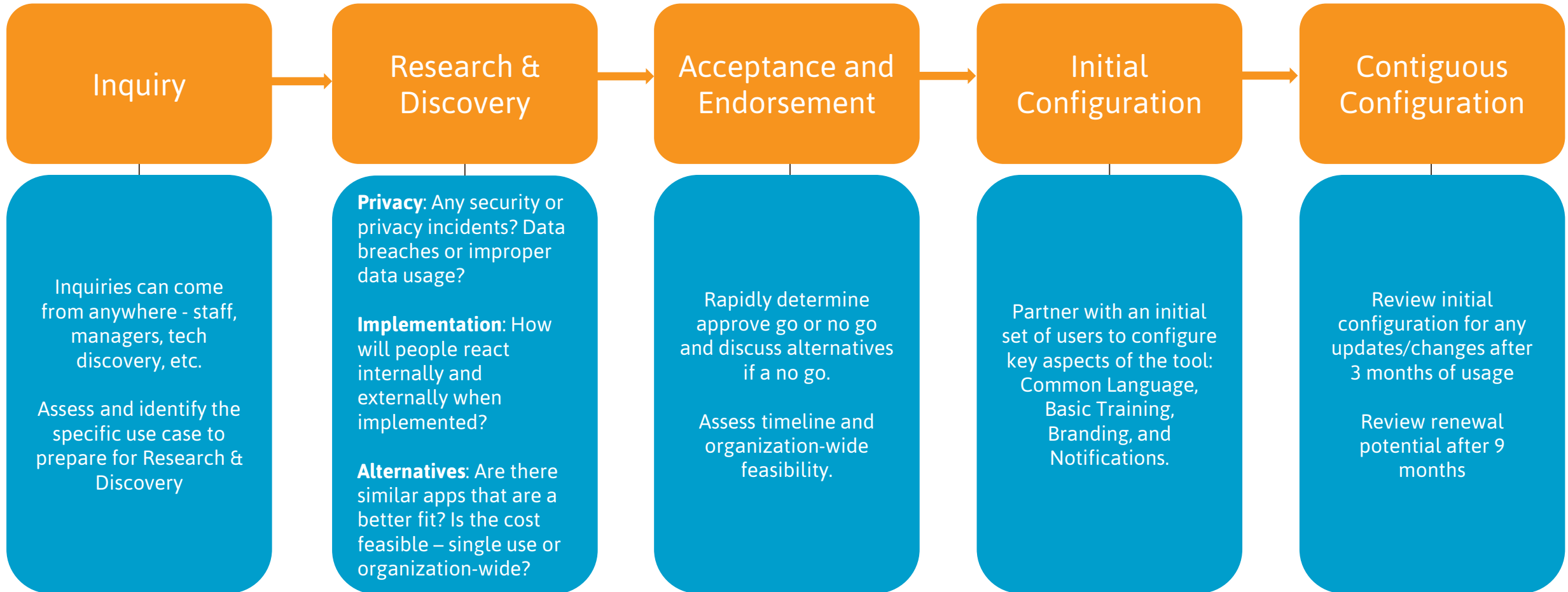
Tier 3-A: Predictive Modeling & Machine Learning

Tier 3-B: Chatbots and Conversational Knowledge Search

Implementation Steps



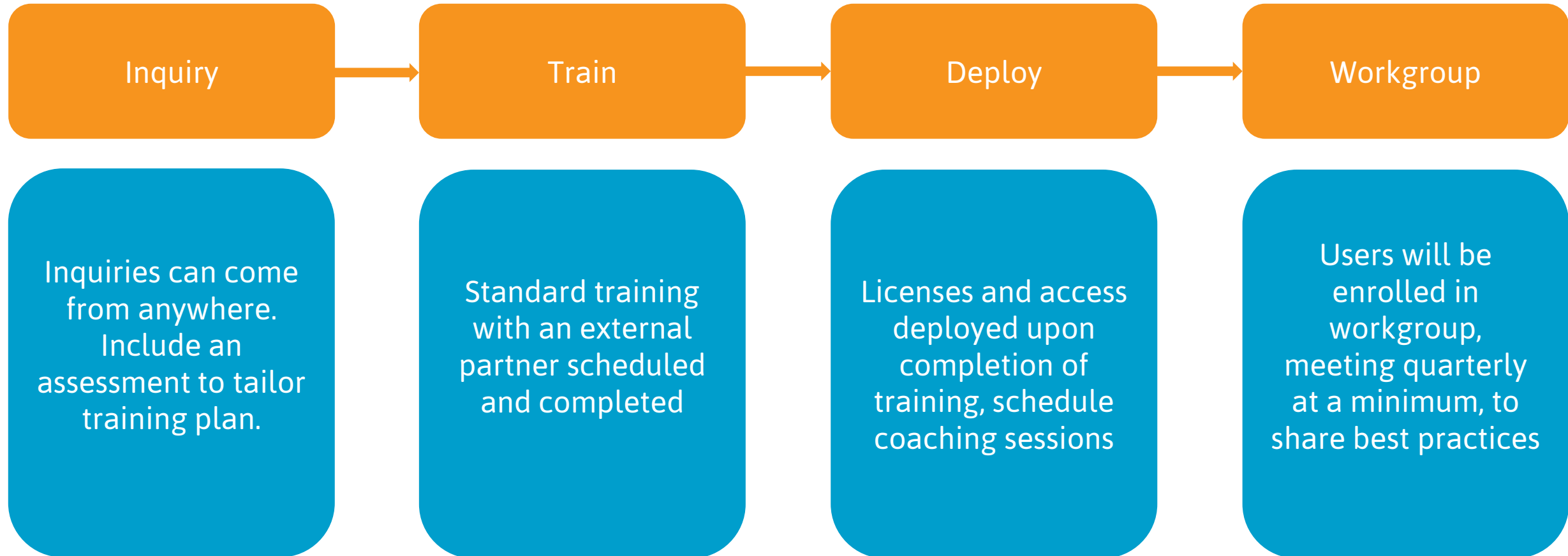
Tier 1: GenAI Products: Pre-defined Augmentation



Implementation Steps



Tier 2: GenAI Chat GPT Assistance: Co-Pilots



Implementation Steps



Machine Learning

Statistical Regressions /
Analysis
Trend Identification
Predictions

Chatbots / Search

Research & Knowledge
repositories
Help Desk Support
Employee Manual & Policies

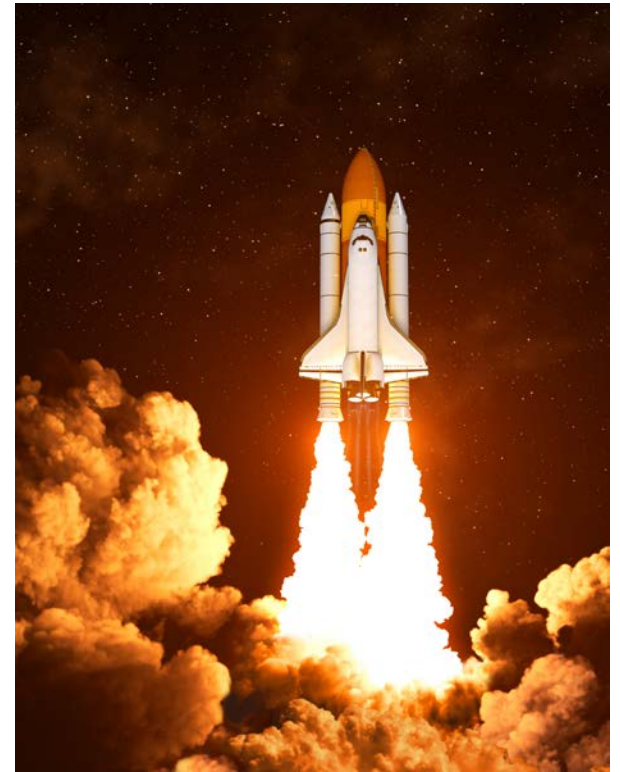
Objectives

- ★ Develop talent (increase LLM usage including MS Co-Pilot and industry-specific LLM's on ChatGPT)
- ★ Develop infrastructure (transition document storage to Sharepoint / OneDrive)
- ★ Plan Internal Roll-outs as MVPs in 2025 with funding partners

Revisiting Guidelines After a Year



- ★ Ethical Considerations - Racial and other Biases
- ★ **Restrictions and Allowances of AI Tools**
 - **Note-taking Tool**
 - **Proofreading Tool**
- ★ **Human Element**
- ★ Talent Over Subscription Services
- ★ Augmentation, Not Replacement
- ★ **Ongoing Training (use case exploration)**
- ★ **Data Privacy**
- ★ **Continuous Improvement**



2024 Uses of AI Tools



Ready Permit AI: Permit Planning & Zoning

By community builder &

Freemium zoning and permit platform for ADUs, modular homes, rentals, and professional projects.



Salesforce Sidekick

By Mitch Lynch

Personal assistant for Salesforce configuration, coding, troubleshooting, solutioning, proposal writing, and more. This is not an official Salesforce product or service.

Prompt Training using pre-trained LLM's on ChatGPT for specific business sectors

Note Taking “Bots” with Qualitative data analysis software

Seeking opportunities to partner with Non-profit AI Accelerators

Salesforce Accelerator

Agents for Impact

As AI agents reshape our world, we champion a future where humans and agents drive social impact together. The Salesforce Accelerator – Agents for Impact equips nonprofits with funding, cutting edge technology, and pro bono expertise to confidently create customized AI agent solutions.

[Submit Interest Form](#)



Prompting – Chat GPT vs Trained LLM



Hi, what can you help me with

< 2/2 >



Hi there! I'm here to streamline your zoning and permit needs, whether you're working on an ADU, modular home, rental property, or larger development project. Here's how I can assist you:

Core Services:

1. **Zoning Information:** Get insights about zoning laws and property use based on your project location.
2. **Site Plans:** Generate instant site plans tailored to your project specifications.
3. **Permit Applications:** Submit permits quickly with automated compliance checks.
4. **Predictive Analysis:** Assess the likelihood of permit approval and identify potential risks.
5. **Compliance Alerts:** Stay updated on local regulations to avoid delays or rejections.
6. **Real-Time Updates:** Track your project and permit status at every stage.

Additional Features:

- **Express Processing:** Speed up your applications.
- **Plan Formatting:** Ensure your documents meet submission standards.
- **Expert Consulting:** Get professional advice tailored to your project.
- **Bundles for Developers and Investors:** Save time and money with curated service packages.

Prompting Basics:

- First question: What can you do for me?
- Ask your question and refine in conversation to get your information
- Mask data if you submit a data set

Prompting – Chat GPT vs Trained LLM



I would like to build a manufactured home in Cleveland, OH in my backyard. What zoning do requirements I need to worry about?

 Memory updated

For building a manufactured home in your backyard in Cleveland, OH, you'll need to check local zoning regulations, which may have specific rules regarding secondary or accessory dwelling units (ADUs). Here are some key requirements to consider:

1. **Zoning District Compatibility:** Verify whether your property's zoning district permits ADUs. Not all residential zones allow additional structures like manufactured homes, and restrictions often vary.
2. **Permit Requirements:** Cleveland requires permits for most building projects, including manufactured homes. You'll likely need to apply for both a zoning permit and a building permit.
3. **Lot Size and Setbacks:** Many areas require a minimum lot size and specific setbacks (distance from property lines). You'll need to check these setback requirements to see if your backyard can accommodate the structure.

I would like to build a manufactured home in Cleveland, OH in my backyard. What zoning do requirements I need to worry about?

Building a manufactured home in your backyard in Cleveland, OH, involves understanding specific zoning requirements, as the city regulates where manufactured homes (also known as mobile or modular homes) can be placed and how they should meet local standards. Here's a general breakdown of the key requirements and considerations:

1. Zoning Districts

- First, confirm your property's zoning district by checking with Cleveland's Planning Department or accessing zoning maps through the city's online resources.
- In Cleveland, manufactured homes are typically allowed in certain residential districts but may not be permitted in all residential zones, especially for accessory use (like placing it in a backyard).
- Common residential zones that might allow manufactured housing include **R-1** (Single-Family Residential) or **R-2** (Two-Family Residential) with additional restrictions.

2. Accessory Dwelling Unit (ADU) Regulations

- Cleveland recently began permitting Accessory Dwelling Units (ADUs) in some areas, which allows secondary living structures on a property. However, a manufactured home may not qualify as an ADU if it does not meet the city's building standards for ADUs.

Drilling Down



what are your sources?

I based this information on common zoning requirements and regulations for ADUs and manufactured homes, as well as general guidance from zoning and city planning practices in Ohio and similar municipalities. For specific details and confirmation of the requirements in Cleveland, I recommend contacting the Cleveland City Planning Commission or checking Cleveland's official zoning codes on their website.



what are your sources for this information?

For zoning and manufactured housing regulations in Cleveland, OH, I used general zoning principles applicable in many U.S. cities and tailored these to Cleveland's known zoning practices. Here are the specific sources and typical resources I would recommend for precise and official information:

1. **City of Cleveland's Zoning Code:** Cleveland's Zoning Code is accessible through the city's planning or building department website. It provides detailed information on district classifications, permitted uses, and specific requirements for accessory structures, ADUs, and manufactured housing. The zoning map is also available online to help identify property zoning classifications.
2. **Ohio Building Codes and HUD Standards:** Ohio has statewide building codes for modular and manufactured homes. Manufactured homes in residential zones must meet federal HUD standards, while modular homes are often subject to local building codes if considered "site-built."

Masking Data



Data Set: Affordable properties we own

- Remove PII – Addresses
- Coded Property IDs
- Include only relevant information

i'm just going to copy/paste it here



Got it! Please paste the data here, and I'll process it for you. Make sure to include the relevant columns (property IDs, lot sizes, and home square footage). I'll keep everything organized and return the results promptly. 😊

```
Subject Lot Size    Sq Ft
14813 Sq.Ft.  A1  1233
12197 Sq.Ft.  A2  1340
9240 Sq.Ft.   A3  1344
6999 Sq.Ft.   A4  1935
```

Thank you for providing the data! I will process this and calculate the **maximum ADU size** for each property based on Minneapolis regulations, preserving your property IDs. Here's how I'll calculate it:

1. Minneapolis ADU Rules:

- Maximum size: **800 sq. ft.** or **50% of the primary dwelling size**, whichever is smaller.
- I'll ensure that no ADU exceeds these limits.

2. Output:

- A table showing the Property ID, Lot Size, Home Size, and the **Maximum ADU Size**.

Let me quickly compute the results and organize them for you. Hold on a moment!

Here is the processed data with the **Maximum ADU Size** calculated based on Minneapolis regulations:

Property ID	Lot Size (Sq. Ft.)	Home Size (Sq. Ft.)	Max ADU Size (Sq. Ft.)
A1	14813	1233	616
A2	12197	1340	670
A3	9240	1344	672
A4	6999	1935	800



Activity: NotebookLM

Explore NotebookLM by uploading your website link(s) and giving it prompts to arrange the information in a variety of ways. <https://notebooklm.google/>

How might this tool be useful in your work?

What kinds of guidelines would you consider in how you approach the tool and the results it provides?



Think about AI in relation to your work.

What AI tools are you using/
considering?

What kinds of guidelines have you set
up or considered in how you approach
them?





We want to hear from you!

Please see link in the chat or use this QR code to access so that you can respond after the event.





Ways to Stay Engaged

- Take the conversation back to your organization!
- Complete our post-event survey
- Sign up for our free online DSI course

