

Getting Started with AI: A Practical Guide for Flood and Stormwater Agencies

Prepared by the NAFSMA Innovation Work Group | March 2026

Artificial intelligence is no longer a concept on the horizon for public agencies. It is already embedded in federal flood forecasting systems, being deployed by peer agencies for permit review and asset management, and available through off-the-shelf tools that require no specialized technical staff to operate. The question for NAFSMA member agencies is not whether to engage with AI, but how to do so deliberately and effectively.

This guide is designed to help agency leaders and program managers answer three practical questions: What is AI actually good for in our context? Where does it fall short and what risks do we need to manage? And how do we get started without overcommitting resources or making avoidable mistakes?

The guidance draws on examples from peer agencies, NAFSMA's Innovation Work Group discussions, and the growing body of documented AI applications in flood control, stormwater management, and public infrastructure.

SECTION 1

Where AI Delivers Real Value for Public Agencies

AI is not a single technology. It is a family of tools, ranging from simple pattern-matching algorithms to large language models that can read and draft documents. The applications most relevant to flood and stormwater agencies fall into five categories.

1. Pattern Recognition in Large Datasets

AI excels at finding patterns in datasets too large or complex for manual analysis. For water resource agencies, this translates directly to practical problems: identifying which storm drains are most likely to fail based on age, material, and surrounding conditions; extracting finished floor elevations for thousands of structures from lidar point cloud data; detecting defects in CCTV pipe inspection footage; and flagging anomalies in sensor data that might indicate a gauge malfunction or an emerging flood event.

Fort Worth's AI-assisted storm drain prioritization model and San Antonio River Authority's lidar-based structure elevation extraction are both examples of pattern recognition delivering faster, more consistent results than manual methods at a fraction of the cost.

2. Speed and Scalability in Forecasting and Modeling

AI models, once trained, can run flood simulations orders of magnitude faster than traditional physics-based models. LSU's compound flood prediction system generates a 72-hour simulation in four seconds on a standard laptop, compared to 15 minutes on a supercomputer using conventional methods. NOAA's AI-enhanced weather models deliver equivalent forecast accuracy at 9% of the computing cost of legacy systems.

For agencies that need to run multiple scenarios quickly during an active weather event, or that want to expand flood modeling capacity without expanding their computing infrastructure budget, AI-assisted modeling is a meaningful capability upgrade.

3. Administrative Workflow Automation

Generative AI tools, including large language models like the one that produced this document, are well suited to the high-volume, repetitive writing and document review tasks that consume staff time without requiring specialized expertise. Applications already in use at peer agencies and governments include:

- Permit application pre-screening for completeness and code compliance, with AI flagging missing documents or deficiencies before the application reaches a staff reviewer
- Grant narrative drafting based on project descriptions, with AI producing a structured first draft that staff then review and refine
- Summarizing lengthy regulatory documents, environmental reports, or federal guidance into actionable staff briefings
- Drafting routine correspondence, public notices, agenda items, and board reports from structured inputs
- Monitoring federal regulatory dockets and flagging relevant proposed rules, notices of funding opportunity, and policy changes

Honolulu's permit pre-screening AI reduced the wait time to reach a plan reviewer from six months to two to three days. Hamilton, Ontario reported a 60% reduction in overall permit processing time and a 70% reduction for residential permits.

4. Improving Accuracy Through Error Correction and Calibration

One of the most practical near-term applications of AI for agencies that rely on federal data is using it to improve the accuracy of that data for local conditions. University of Michigan's Errorcastnet system, for example, identifies and corrects systematic errors in NOAA's National Water Model predictions, improving forecast accuracy four to six times for lead times of one to ten days.

Agencies that have invested in local gauge networks and historical event data are well positioned to apply similar approaches: training AI models on the gap between federal forecasts and local observations, then using those models to produce locally calibrated predictions.

5. Prioritization and Decision Support

AI tools can help agencies answer prioritization questions that are technically tractable but practically difficult at scale: Which of our 50,000 catch basins is most overdue for cleaning? Which properties in our jurisdiction have the highest expected annual flood damage? Which permit applications in our backlog are most at risk of becoming problematic?

When AI is applied to these questions with good underlying data, it allows programs managers to direct limited staff time and capital budgets toward the highest-value work, and to document the basis for those decisions in a way that is defensible to boards, auditors, and the public.

SECTION 2

Where AI Falls Short: Honest Limits and Real Risks

Understanding where AI fails is as important as understanding where it succeeds. The agencies that have had the worst experiences with AI tools are typically those that expected too much, skipped the validation step, or deployed AI in contexts where it was not well suited. The following categories represent the clearest limitations for public agency applications.

AI IS NOT WELL SUITED FOR:

- Decisions that require legal accountability or public due process
- Novel situations with no relevant historical data to learn from
- Contexts where errors have serious safety or legal consequences and cannot be caught by human review
- Replacing professional judgment on complex technical or policy questions
- Generating precise numerical outputs where accuracy must be independently verified
- Tasks requiring deep local context that is not captured in data
- Replacing community engagement, public hearings, or stakeholder input

AI TENDS TO WORK WELL FOR:

- High-volume, repetitive tasks with clear right/wrong answers
- Pattern recognition across large datasets with good historical examples
- First-draft generation where human review follows
- Speeding up processes where the bottleneck is administrative, not technical
- Synthesizing information from many sources into a structured summary
- Flagging potential issues for human review rather than making final decisions
- Augmenting, not replacing, physics-based models and professional expertise

Key Risks to Understand and Manage

- **Hallucination and factual errors.** Large language models can generate plausible-sounding text that is factually wrong. They may cite regulations that do not exist, invent statistics, or misstate program requirements. Any AI-drafted document that will be acted

on, submitted to a funder, or made public must be reviewed by staff with subject matter expertise before it goes out.

- **Bias in training data.** AI models learn from historical data. If that data reflects past inequities, the AI will reproduce them. An asset management model trained on inspection history from wealthier neighborhoods may underpredict failure risk in areas that were historically under-inspected. Agencies should test AI outputs for geographic and demographic consistency before relying on them for capital allocation.
- **Overconfidence in model outputs.** AI systems, particularly commercial products, often present outputs with a false sense of precision. A flood model that produces inundation maps to the inch is not necessarily more accurate than one that produces outputs to the foot. Agencies should ask vendors to demonstrate model performance against actual events before deployment and should treat AI outputs as one input among several rather than as ground truth.
- **Data privacy and security.** Many AI tools, particularly cloud-based services, process and may retain the data you submit. Agencies should review vendor data handling policies carefully before submitting sensitive information, including property records, infrastructure system data, or personnel information. Check whether your jurisdiction has data governance policies that apply to AI vendor contracts.
- **Vendor lock-in.** AI tools that require your data to be in a proprietary format, that do not provide API access, or that do not allow you to export your models can create significant switching costs. Agencies should negotiate for data portability and model transparency from the outset, and should be cautious about allowing core business processes to become dependent on a single vendor's platform.
- **Staff displacement concerns.** AI tools that automate tasks previously done by staff can generate internal resistance if the change management is handled poorly. Be transparent about what AI is being used for, involve affected staff in the evaluation and implementation process, and be clear that the goal is to reduce low-value administrative burden, freeing staff for higher-value work.

SECTION 3

Assess Your Agency's Readiness

Before selecting or procuring any AI tool, it is worth taking an honest inventory of your agency's current capabilities and constraints. The following questions are not a formal scoring rubric. They are prompts for an internal conversation that will help you identify where to start and where to invest first.

DATA READINESS

- Do we have a current, accurate GIS database of our flood control and stormwater infrastructure?
- Do we maintain consistent CCTV or inspection records for our storm drains, culverts, or levees?

- Do we have historical event data (gauge readings, flood extents, damage records) that is digitized and accessible?
- Are our records in a format that can be exported and used by external tools (standard formats, not locked in legacy systems)?
- Do we have a data governance policy that addresses cloud services and third-party data processing?

STAFF AND CAPACITY READINESS

- Do we have at least one staff member who can evaluate AI tool performance and is not dependent on a vendor to interpret results?
- Do we have staff time to pilot a new tool, review outputs, and refine the process before full deployment?
- Is leadership willing to invest in staff training alongside tool procurement?
- Have we identified a staff champion who can drive adoption and troubleshoot early implementation challenges?

PROBLEM READINESS

- Have we identified a specific problem that AI is likely to help with, rather than looking for a general 'AI solution'?
- Do we have a way to measure success, so we can tell whether the tool is working?
- Is the problem high-volume, repetitive, or data-intensive enough that AI offers real efficiency gains?
- Are the stakes low enough that errors during a pilot phase can be caught and corrected before causing harm?

SECTION 4

A Phased Approach to AI Adoption

NAFSMA recommends a phased approach that starts with lower-risk administrative applications, builds staff confidence and organizational learning, and then moves toward more technically complex deployments as capacity develops. This mirrors the adoption path followed by the peer agencies highlighted in NAFSMA's AI case studies.

PHASE 1: Start with Administrative Tools

Months 1–6: Low technical risk, immediate time savings

- Identify one or two high-volume administrative tasks where AI can provide a useful first draft or pre-screen: permit intake review, grant narrative drafting, regulatory summary preparation, or board report production
- Select a commercial AI writing tool (most agencies are already paying for Microsoft 365 Copilot or Google Workspace, both of which include AI writing features)
- Train two to three staff members on effective use, including how to review and correct AI outputs
- Establish a simple review protocol: AI produces draft, staff reviews for accuracy and local context, supervisor approves before any external use
- Document time savings and error rates to build the internal case for expanded adoption

PHASE 2: Apply AI to Asset Management and Prioritization

Months 6–18: Moderate technical investment, measurable operational value

- Audit your infrastructure data to confirm it is current, consistent, and accessible
- Identify a specific capital prioritization question where AI-assisted analysis would improve on current methods: which pipes to rehabilitate, which catch basins to clean first, which properties to include in a buyout program
- Engage an engineering firm or technology partner with demonstrated experience in the specific application
- Run a pilot on a subset of your system and validate outputs against what staff would have decided using traditional methods
- If results are consistent with professional judgment and the tool adds analytical depth, expand to full system

PHASE 3: Integrate Federal AI Data Products

Ongoing: Dependent on GIS infrastructure

- Confirm your agency has current GIS infrastructure and staff capable of consuming web service data feeds
- Register as a FEMA Cooperating Technical Partner if not already, which positions your agency to receive and contribute to Risk MAP and AI-enhanced flood mapping products
- Connect NOAA's Flood Inundation Mapping data to your emergency operations platform so that real-time AI-generated inundation forecasts are available to your operations staff during flood events
- Identify gaps between federal model outputs and local conditions, and document those gaps as a basis for local model calibration or supplemental data collection

PHASE 4: Pursue Technical AI Applications

12–36 months: Requires data foundation and technical capacity

- Develop or procure AI-assisted flood modeling or warning system capabilities once the data and staff foundations are in place
- Consider partnerships with peer agencies, universities, or USACE ERDC to share the cost and learning associated with more complex AI deployments
- Evaluate compound flood modeling tools (particularly relevant for coastal and urban agencies dealing with combined storm surge and rainfall-driven flooding)
- Assess the feasibility of AI-assisted inspection tools for levee, dam, or conduit safety work, drawing on USACE ERDC's published work on automated defect detection

SECTION 5

Procuring AI Tools: Questions to Ask Before You Sign

AI procurement is different from traditional software procurement in ways that can create significant long-term risk if not addressed upfront. The following questions should be part of any RFP, vendor evaluation, or contract negotiation involving AI tools.

About the Model and Its Performance

- What data was this model trained on, and how recent is it?
- How has performance been validated against real-world outcomes? Can you provide documented results from peer agencies?
- What is the model's known failure rate, and in what types of situations does it perform worst?
- How does the model handle situations that fall outside its training data?
- Is the model explainable? Can it tell us why it made a particular prediction or recommendation?

About Data Ownership and Security

- Who owns the data we submit to your system?
- Is our data used to train or improve your model? Can we opt out?
- Where is our data stored, and what security certifications does your platform hold?
- What happens to our data if we terminate the contract?
- Does your platform comply with applicable state and local data governance requirements?

About Vendor Relationships and Continuity

- Can we export our data and any custom model configurations in a standard format?
- What happens to our workflows if your company is acquired or discontinues this product?
- What is your update and versioning policy? Will model updates change outputs in ways that affect our processes?
- What training and support are included, and what is the ongoing support cost?

SECTION 6

Staff Training for Responsible AI Use

AI tools are only as reliable as the people using them. No matter how capable a generative AI platform is, it produces plausible-sounding outputs — not guaranteed-correct ones. Flood and stormwater agencies deal in technical, regulatory, and financial information where errors have real consequences. That makes staff training not a nice-to-have, but a core implementation requirement.

The foundational expectation for all NAFSMA member agencies: Every AI-generated document must be verified by a qualified staff member before use. Every research output produced with AI assistance must be independently confirmed against primary sources before it is relied upon for decisions, submittals, or public communications.

Why Verification Is Non-Negotiable

Large language models predict the most statistically probable next word — not the factually correct one. In practice, this means they can produce grant citations that don't exist, regulatory citations to outdated rules, flood depth estimates that sound authoritative but are fabricated, and project cost figures drawn from irrelevant comparisons. The Georgetown Beeck Center frames this well: just because AI *can* help with a task doesn't mean it *should*, or that it will do a good job. Only the staff member who owns the work can catch those errors before they cause harm.

Core Training Components

Effective staff training should cover four areas: what AI tools actually do (and how they fail), how to write prompts that produce useful outputs, how to verify what the tool produces, and what not to put into AI systems. Training does not need to be lengthy. A two-hour workshop using real agency examples is more valuable than a comprehensive course that staff never complete.

RECOMMENDED TRAINING CURRICULUM FOR AGENCY STAFF

Module 1: How AI Tools Work (30 minutes)

- What large language models do: pattern matching and probability, not knowledge or reasoning
- Common failure modes: hallucination, outdated training data, false precision
- Live demonstration: ask the tool a question you already know the answer to, then spot where it goes wrong

Module 2: Effective Prompting (30 minutes)

- How to give context, role, and format instructions to get better outputs

- Practice with agency-specific tasks: grant paragraph, public notice, permit checklist

Module 3: Verification Protocols (30 minutes)

- The mandatory review checklist: every fact, citation, and number must be confirmed
- How to use the AI tool itself to check its own work (with caution — see verification prompts below)
- When to escalate: outputs that require legal, engineering, or regulatory review before use

Module 4: Data Sensitivity and What Not to Share (30 minutes)

- Enterprise tools vs. consumer tools: different rules for what you can submit
- Categories that must never go into a public AI tool: personnel records, enforcement case details, unpublished infrastructure vulnerability data, attorney-client communications
- The rule of thumb from the Georgetown Beek Center: don't put anything into a public AI tool that you wouldn't be comfortable telling an AI company employee over coffee

Sample Agency AI Use Guidelines

The following guidelines are designed to be adapted by member agencies. They are intentionally written in plain language so they can be shared directly with staff, included in onboarding materials, or posted near shared workstations.

[AGENCY NAME] SAMPLE AI USE POLICY FOR STAFF — ADAPT AS NEEDED

What AI tools can help you with

- Drafting first versions of documents you will then review and edit (grant narratives, board reports, public notices, permit checklists, meeting minutes)
- Summarizing long documents you have already read and can fact-check against
- Translating technical content into plain language for public communications
- Generating tabletop exercise scenarios and training materials for internal review

What you must always do

- Review every AI-produced document for factual accuracy before it leaves the agency or is used to make a decision
- Verify every statistic, regulation citation, program name, funding amount, and deadline the tool produces by checking the original source
- Disclose AI assistance on documents where your agency's policy or the receiving party requires it
- Get supervisor approval before any AI-assisted document is submitted externally (to funders, regulators, or the public)

What you must never do

- Submit AI-generated content to a funder, regulator, or the public without review and approval
- Enter personnel records, enforcement case details, attorney-client communications, or unpublished vulnerability data into any AI tool
- Use AI output as a substitute for professional engineering, legal, or regulatory judgment
- Rely on AI-produced research without verifying the underlying sources independently

Verification Prompts: Using AI to Check AI Work

One useful technique is to ask the AI tool to critique its own output. This does not replace human verification — the tool can hallucinate its self-critique just as it can hallucinate the

original output — but it often surfaces obvious errors and strengthens the document before the staff reviewer takes over. The following prompts have been tested for agency use cases.

VERIFICATION PROMPTS FOR STAFF USE

General factual check — run after any draft document

“Review the document above. List every specific claim, statistic, regulation citation, program name, funding amount, and deadline. For each one, tell me how confident you are it is accurate and flag any that I should verify with an outside source before submitting this document.”

Research verification — run after any AI-assisted research summary

“For each finding in the summary above, provide the specific source (publication, agency, date, and URL if available) where that information can be confirmed. If you do not have a specific source for a claim, say so explicitly rather than generating a plausible-sounding citation.”

Grant narrative check — before submitting to a funder

“Read this grant narrative and identify: (1) any eligibility claims that should be verified against the Notice of Funding Opportunity, (2) any cost estimates or match percentages that appear to be estimates rather than confirmed figures, (3) any program requirements referenced that I should confirm are still current, and (4) any statements that could be seen as overclaiming project outcomes.”

Regulatory document check — for compliance summaries and technical reports

“This document references federal and state regulations. For each regulatory citation, tell me: the specific rule or code section, the date it was last amended (if you know it), and whether any portions of this regulation were subject to proposed changes or litigation as of your training cutoff. Flag any citation where I should independently verify that the current rule text matches what is described here.”

Responsible use self-assessment — before deploying any new AI application

“Before we use AI for [specific task], help me think through: (1) What decisions will this output influence? (2) What is the potential harm if the output is wrong? (3) What data sensitivity is involved? (4) What human oversight will be in place? (5) What changes if this is scaled up? (6) Who will own ongoing monitoring and quality checks?”

SECTION 7

Everyday AI Applications and Responsible Use Framework

The Georgetown Beeck Center’s June 2026 webinar for NAFSMA members identified ten specific everyday uses for large language models in flood and stormwater agencies, along with a practical framework for deciding whether a given application is appropriate. This section draws directly from that guidance.

What LLMs Do Well

Large language models are specifically well suited for five categories of work: summarizing long documents into key points; extracting structured data from unstructured text; comparing document versions side by side; rewriting content for clarity or a different audience; and converting narrative descriptions into checklists, tables, or structured formats. These capabilities map directly to the administrative bottlenecks most agencies face.

Top 10 Everyday AI Uses for Your Agency

FROM THE BEECK CENTER: TOP 10 EVERYDAY AI USES (ALL REQUIRE HUMAN REVIEW)

- Help draft grant narratives and compliance checks
- Translate technical reports into plain-language public notices
- Summarize public comments, workshop notes, and stakeholder feedback
- Check permit submittals for completeness against a known checklist
- Turn operations manuals into SOPs, checklists, and training guides
- Draft RFP language and compare vendor proposals side by side
- Generate tabletop exercise scenarios for storm response training
- Query and synthesize historical maintenance records and work orders
- Perform first-pass review and categorization of inspection photos
- Draft meeting minutes, action items, and board memos from notes or recordings

Source: Georgetown Beeck Center, Practical AI in Flood & Stormwater Management, NAFSMA Webinar, June 16, 2026

What Is Ready vs. What Is Not

Not every AI application is appropriate for government use today. The Beeck Center draws a clear line between applications that are likely ready for agency deployment and those that carry unacceptable risk without significantly more maturity and oversight infrastructure.

LIKELY READY FOR AGENCY USE

- Smart sewer real-time control with human oversight
- AI-assisted pipe inspection image review
- LLM support for grants, permits, and public communications (with review)
- Drone imagery analysis for post-storm damage assessment

PROBABLY NOT READY

- Autonomous “agentic” control of interlocking public infrastructure with no human in the loop
- AI-driven crisis decision-making without human authority and accountability in the loop

Six Questions to Ask Before Deploying Any AI Application

The Beeck Center recommends agencies work through six questions before deploying any AI tool — whether it is a new procurement or an expanded use of an existing platform. These questions are designed to surface risks that are easy to miss when a tool is generating impressive outputs.

- **What decisions does this output influence?** Trace the output downstream to understand what actions it will drive and who will be affected.
- **What is the harm if the output is wrong?** Scale your oversight investment to the severity of potential consequences. Low-stakes drafting needs light review. High-stakes decisions need independent verification.
- **What data sensitivity is involved?** Apply the enterprise vs. consumer tool distinction. Sensitive operational, legal, or personnel data must stay out of consumer-facing platforms.

- **What human oversight exists?** Name the specific person responsible for reviewing outputs before they are acted upon. “Someone will check it” is not a sufficient answer.
- **What changes if this scales?** A tool that works acceptably on ten documents per month may produce systematic errors if used on a thousand. Test at scale before committing.
- **Who owns monitoring and evaluation?** Assign a staff member to track error rates, staff complaints, and changes in tool behavior after vendor updates. AI systems are not static and need ongoing attention.

SECTION 8

Getting Connected Through NAFSMA

NAFSMA's Innovation Work Group was established to help member agencies navigate exactly these kinds of emerging technology questions. The Work Group provides a peer network for sharing what is working, a forum for raising questions with federal agency counterparts, and a channel for developing collective positions on AI-related policy issues as they emerge.

If your agency is actively evaluating or deploying AI tools, we want to hear from you. Member experiences are the most valuable input the Work Group has, and sharing what you have learned benefits the full NAFSMA community.

HOW TO ENGAGE

- Contact Sunny Simpkins to be connected with the Innovation Work Group
- Share your agency's AI projects or questions for inclusion in future member resources
- Volunteer for a peer interview to contribute to NAFSMA's ongoing AI case study research
- Attend Innovation Work Group sessions at the 2026 Annual Meeting in Annapolis, Maryland (July 13-16)

Contact: Sunny Simpkins, Executive Director | sunnys@nafsma.org | (202) 289-8625 | www.nafsma.org

National Association of Flood & Stormwater Management Agencies | P.O. Box 4336, Silver Spring, MD 20914