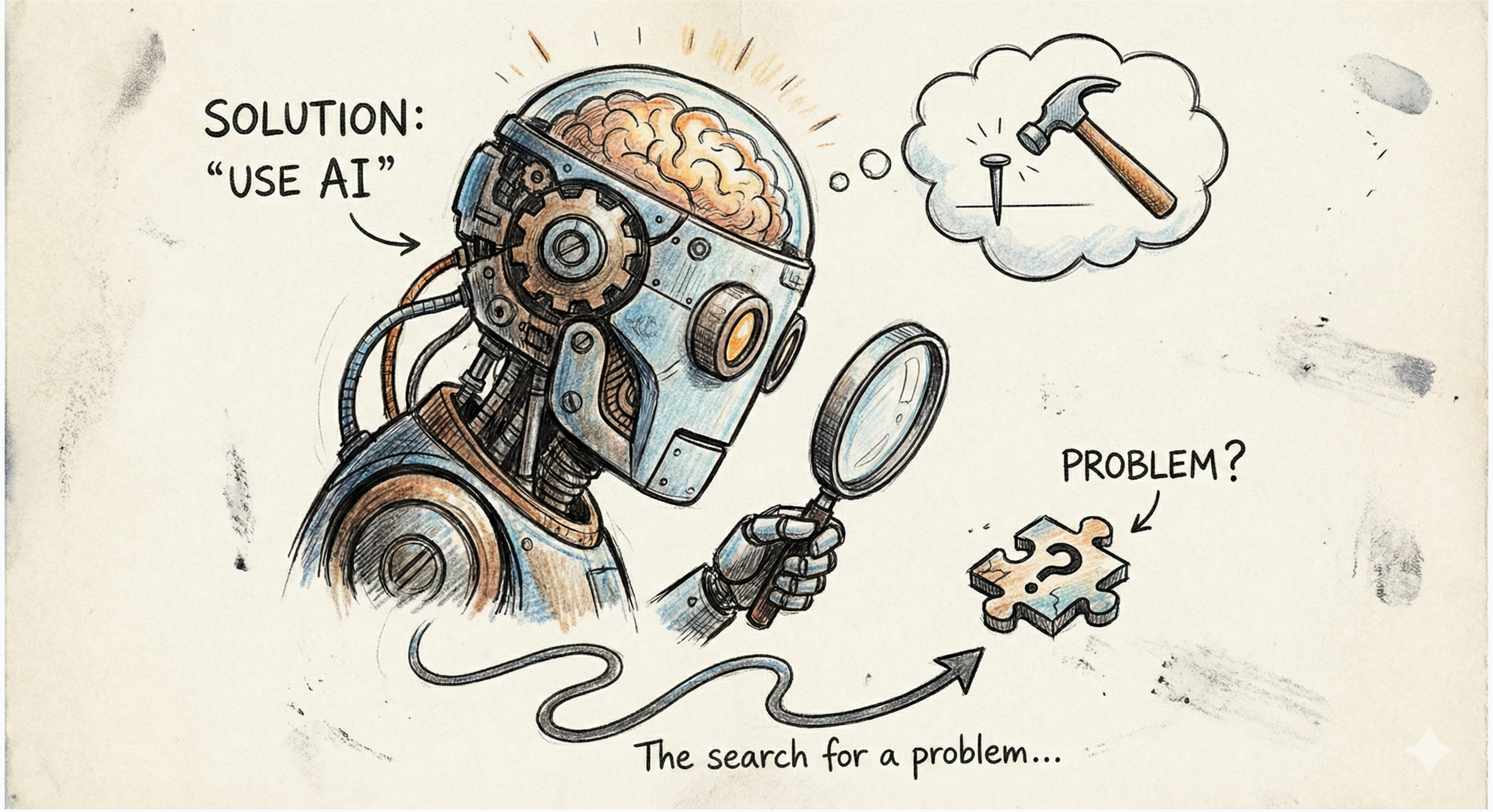
Data Science, AI, and Machine Learning in Public Health using R

Part 2

December 2025

Presented By: Wronski Associates

Applying AI: From Directive to Problem Statement



"Use AI" is not a problem statement—it's a solution looking for a problem. That is a clarifying statement, not a criticism.

Before we can evaluate whether AI is appropriate, we need to articulate what we're actually trying to accomplish. This topic builds that skill.

Common Ways Directives Arrive (and How to Unpack Them)

Key point: Directives come in predictable flavors; each has a different underlying question to surface.

Directive Type	Example	Underlying Question to Surface
Vague area	"We should use AI to improve disease surveillance."	What specific outcome? Measured how? What’s not working now?
Task-based	"Can AI help us with our AI project backlog?"	What’s causing the backlog? What decisions are being delayed? What would "helped" look like?
Peer pressure	"Other states are using AI for fraud detection—we should look into that."	Is our context similar? Do we have the same problem? The same data?
Buzzword-driven	"Let’s use AI to be more data-driven."	What decisions aren’t currently informed by data? What would change if they were?
Solution-embedded	"We need a chatbot for routine public inquiries."	What problem is the chatbot solving? High call volume? Long wait times? Repetitive questions?

The Anatomy of a Good Problem Statement

Key point: A good problem statement is specific, outcome-focused, and solution-agnostic.

A complete problem statement answers:

1. **What decision or action will this inform?**
   - Not "improve outcomes" but "identify which patients to prioritize for outreach"
   - Not "be more efficient" but "reduce time spent manually categorizing incoming requests"
2. **Who needs the output and when?**
   - Is this for frontline staff making daily decisions? For leadership reviewing quarterly trends? For an annual report?
   - Real-time? Daily batch? One-time analysis?
3. **What does success look like?**
   - How will you know it worked?
   - Is there a measurable target? (e.g., "reduce no-show rate from 28% to 20%")
   - What's the counterfactual—what happens if this isn't done?
4. **What's the current state?**
   - How is this handled today?
   - What's the baseline performance?
   - Why isn't the current approach working?

Example Transformation

Before	After
"Use AI to reduce no-shows at WIC clinics."	"Identify WIC participants likely to miss their upcoming appointment so clinic staff can send targeted reminders 48 hours in advance. Success = reduce no-show rate from 28% to 20% within 6 months. Currently, reminders are mailed to everyone 3 days before."

Red Flags in Problem Framing

Key point: Some problem framings are warning signs that more work is needed before proceeding.

Red Flag	**Example**	**Why It’s a Problem**
No clear outcome	"Make things better" / "Improve quality"	You can’t measure success; you can’t define what AI would produce.
Solution embedded	"We need a chatbot" / "Build a predictive model"	You’ve skipped problem definition; you may be solving the wrong thing.
Too broad	"Improve public health" / "Reduce costs"	There’s no specific decision or action; this is a goal, not a problem.
Too narrow (premature)	"Predict which of our 50 clients will default"	The volume may not justify AI; you’ve scoped before understanding.
Unmeasurable	"Make staff feel more supported"	You can’t define or detect success.
No baseline	"We want to be faster"	Faster than what? How fast is it now?

The Reframing Conversation

Key point: Reframing isn’t just an internal exercise. It often requires a conversation with the person who issued the directive.

- Pushing back on framing can feel risky, but it’s essential.
- The goal isn’t to say "no". It’s to say "let me make sure I understand what we’re trying to accomplish."
- Reframing is part of the process.

"I want to make sure we’re solving the right problem. Can you tell me more about what’s not working today?"
"When you say 'improve X,' what would that look like if it worked? How would we know?"
"Other states are doing this—do we know if their situation is similar to ours?"
"If we had a perfect solution, what would you do differently tomorrow?"

From Directive to Problem Statement Summary

Component	Key Takeaway
Why "use AI" isn’t a starting point	AI is a tool; surface the underlying pain point
How directives arrive	Recognize the type; have questions ready
Anatomy of a problem statement	Four elements: decision, audience/timing, success, current state
Red flags	Spot vague, embedded, or unmeasurable framings
The reframing conversation	Have language to push back diplomatically

Applying AI: What Makes a Problem AI-Suitable?

- Is there a clear outcome?
- Does data exist?
- Is there a pattern to learn?
- Is the problem repeatable?
- Does the value justify the effort?
- Can action be taken on the output?

There’s a Clear Outcome or Decision

Key point: Articulate what we’re trying to predict, classify, find, or generate. If we can’t define done, AI can’t help.

- AI needs a target. What is the model supposed to produce?
- The outcome should be specific and observable—not abstract.
- Ask: "If this worked perfectly, what would the output be?"

Clear Outcome	**Unclear Outcome**
Predict whether a patient will be readmitted within 30 days (yes/no)	Improve patient outcomes
Classify incoming emails as complaint, question, or compliment	Make our inbox more manageable
Estimate the number of WIC enrollees next quarter	Plan better for the future
Flag inspection reports that mention structural issues	Find important things in reports

There’s Data—and You Can Access It

Key point: AI learns from examples. No data means no learning. But "data exists" isn’t enough. It needs to be accessed.

- Some AI requires training data: historical examples of inputs and outcomes.
- The data must be accessible. Not locked in another agency’s system, protected by policy, or in someone’s filing cabinet.
- Quality matters: Is the data complete? Consistent? Representative?

Question	Why It Matters
Does the data exist at all?	Can’t train on data you don’t have
Is it digitized and structured?	Handwritten forms or PDFs may require extraction first
Can you legally/ethically access it?	Privacy rules, data sharing agreements, IRB
Is it linkable to the outcome?	You need to connect inputs to what happened
Is there enough of it?	Small samples limit what AI can learn
Is it representative?	Data from one clinic may not generalize to others

Questions to ask:

- Where does this data live? Who controls it?
- How can we get it? What approvals are needed?
- How far back does it go? How many examples do we have?
- Does the data include the outcome we’re trying to predict?

We Have Data… Maybe

- "We have data" is not the same as "we have usable data."
- Access and quality are as important as existence.

There’s a Pattern to Learn

Key point: The phenomenon isn’t random. There’s signal in the data that relates inputs to outcomes. Historical patterns exist and are likely to hold.

- AI finds patterns in data. If the outcome is essentially random, or driven by factors not captured in your data, AI won’t help.
- The pattern needs to be learnable from the inputs you have.
- The pattern needs to be stable. What was true in the past should be reasonably true in the future.

Situation	**Problem**
Outcome is driven by factors you can’t measure	e.g., predicting which patients will comply with treatment when compliance depends on family support you don’t have data on
Outcome is essentially random	e.g., predicting lottery numbers
Outcome is rare	e.g., predicting which of 10,000 facilities will have a catastrophic failure when it’s happened twice in 20 years
Pattern has shifted	e.g., predicting demand based on pre-pandemic data

Questions to ask:

- Is there reason to believe the inputs relate to the outcome?
- Has anyone studied this relationship before? What did they find?
- Is the outcome common enough to learn from?
- Has the underlying process changed recently?

Other Considerations

- AI needs signal, not noise.
- Rare events and shifting contexts are warning signs.

The Problem is Repeatable. Wait…. What?

Key point: AI shines when you need to make similar decisions many times. One-off analyses rarely justify the investment.

One-off analyses rarely justify the investment → Is this really true?

- Building, validating, and deploying AI takes effort. That effort pays off when making the same type of decision repeatedly.
- One-time questions ("Why did the rate of rural hospital closings increase so dramatically last year?") are often better answered with traditional analysis.
- Repeatability also means the problem will keep occurring. There is ongoing value, not just a one-time fix.

High Repeatability (Good for AI)	Low Repeatability (Probably Not AI)
Triage 500 incoming applications per week	Decide whether to open a new regional office
Flag high-risk patients daily	Investigate why one clinic has high turnover
Classify 10,000 public comments on a rule	Write the agency’s strategic plan
Predict monthly caseload	Explain a one-time budget shortfall

Questions to ask:

- How often does this decision get made?
- How many instances are we talking about per week/month/year?
- Will this problem still exist in two years?

The Value Justifies the Effort

Key point: Building, validating, and maintaining AI takes resources. Is the problem high-volume, high-stakes, or high-cost enough to warrant it?

- AI projects aren’t free. They require data preparation, model development, validation, deployment, and ongoing monitoring.
- The value should exceed the cost—either in efficiency gains, better outcomes, or risk reduction.
- "Cool" or "innovative" isn’t a business case.

Value Drivers

Driver	**Example**
High volume	50,000 applications to review per year
High stakes	Wrong decision causes patient harm or legal liability
High cost of current process	3 FTEs spend 80% of their time on manual triage
High cost of errors	Each missed fraud case costs $100,000

Questions to ask:

- What’s the cost of the current process (time, money, errors)?
- What’s the cost of getting it wrong?
- How much improvement would justify the investment?
- Could a simpler solution (rule, checklist, additional staff) get 80% of the benefit at 20% of the cost?

You Can Act on the Output

Key point: There’s an intervention pathway. Knowing the answer changes what gets done.

- A prediction is useless if there’s no intervention pathway.
- "Actionable" means: someone will do something different because of this output.

Actionability Failures

Situation	Problem
No capacity to intervene	Model flags 500 high-risk patients, but only 2 care coordinators are available
No authority to act	Model identifies fraud, but only another agency can investigate
No clear intervention	Model predicts turnover, but there’s no retention program
Output ignored	Model recommends priority, but staff override it every time

Questions to ask:

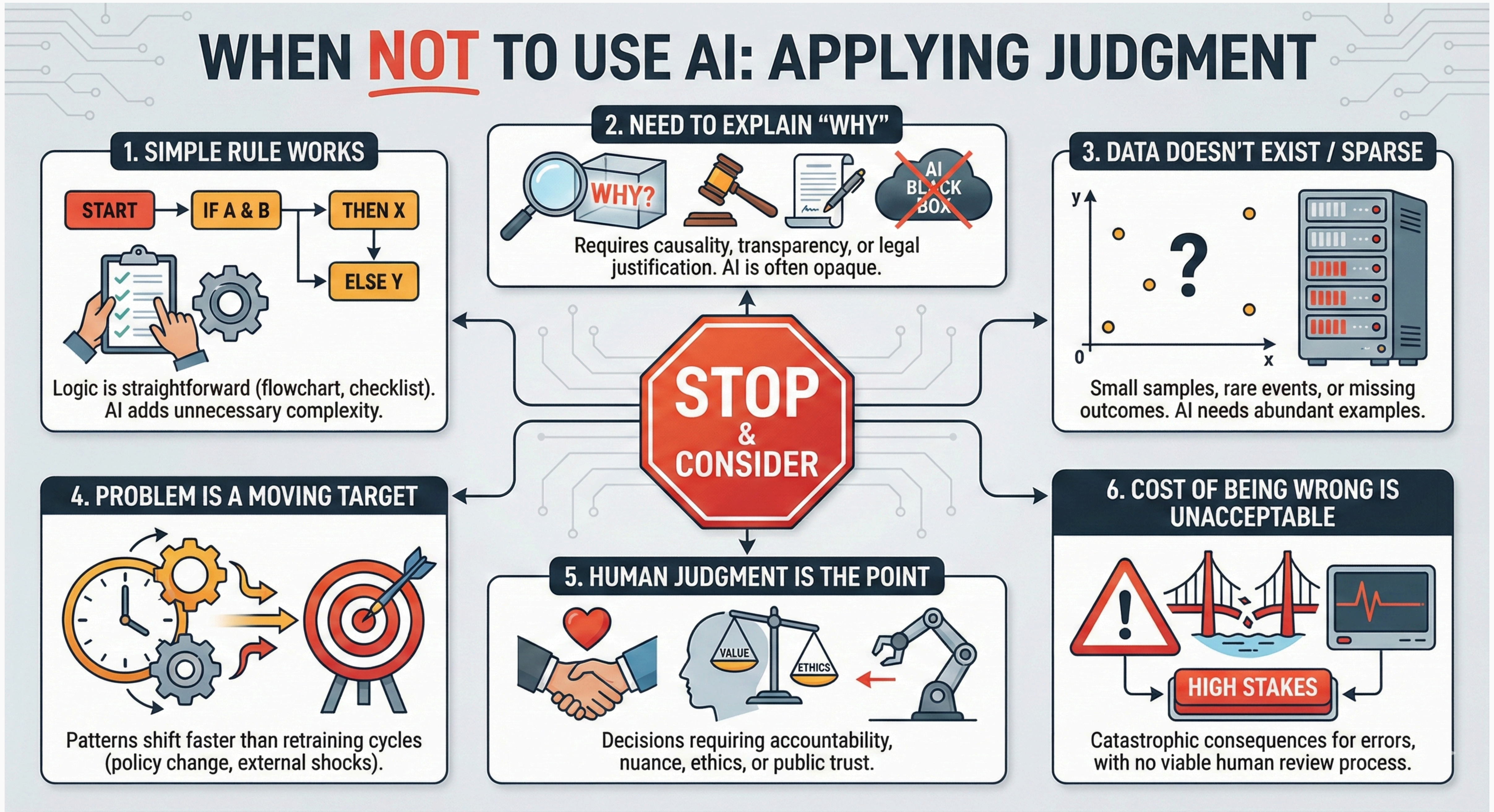
- If the model gives us the answer, what will we do with it?
- Who will act on the output? Do they have capacity?
- Is there an intervention that’s been shown to work?
- What happens if we identify a problem but can’t fix it?

What Makes a Problem AI-Suitable? Summary

Component	**Key Takeaway**
Clear outcome or decision	If you can’t define the output precisely, you’re not ready
Data exists and is accessible	"We have data" ≠ "we have usable data"
Pattern to learn	AI needs signal, not noise; beware rare events and shifting contexts
Problem is repeatable	AI is for repeated decisions at scale, not one-off questions
Value justifies effort	Always ask: "Is there a simpler way?"
Can act on output	Prediction without action is just trivia

Applying AI: When AI is Overkill or Wrong

Knowing when **not** to use AI is as valuable as knowing when to use it.



When a Simple Rule Works

Key point: If the logic can be written as a flowchart or a few IF-THEN statements, AI is not needed. Explainability and maintainability favor simple rules.

- AI is powerful when patterns are complex and hard to articulate. But many decisions follow straightforward logic.
- Rules are transparent, easy to audit, and easy to change.
- AI adds complexity, maintenance burden, and opacity—only justified when rules can’t capture the pattern.

Signs a Rule Might Be Enough

Situation	Example
Logic is already documented	Eligibility criteria for a benefit program
Decisions follow a checklist	Inspection pass/fail based on defined criteria
Experts can articulate the logic	"If A and B, then X; otherwise Y"
Exceptions are rare and known	A few edge cases that can be handled manually

Questions to ask:

- Can a subject matter expert write down the decision rules?
- How often do exceptions occur that the rules can’t handle?
- Would a decision tree or flowchart capture 90% of cases?

When You Need to Explain the "Why"

Key point: If the primary goal is understanding causal mechanisms or defending a decision in a legal or policy context, classical statistics may be more appropriate than AI.

- AI optimizes for prediction, not explanation. Many models (especially deep learning) are "black boxes."
- Some contexts require explainability: legal proceedings, regulatory compliance, public accountability, clinical decision-making.
- Classical statistics—regression, hypothesis testing—is designed to quantify relationships and attribute effects.

When Explanation Matters More Than Prediction

Context	Why Explanation Matters
Policy evaluation	Need to know **if** an intervention caused an outcome, not just predict outcomes
Legal or regulatory decisions	May need to justify why a specific decision was made
Clinical care	Physicians need to explain recommendations to patients
Public accountability	Citizens may demand to know why they were flagged or denied
Root cause analysis	Need to understand **why** something happened, not just predict it will

Questions to ask:

- Will someone need to explain **why** this decision was made?
- Is the goal to understand a relationship or to predict an outcome?
- Are there legal, regulatory, or ethical requirements for transparency?

When Data Doesn’t Exist or Is Too Sparse

Key point: AI can’t conjure patterns from nothing. Small samples, rare events, or missing outcome data are showstoppers.

- AI learns from examples. If there aren’t enough examples, especially of the outcome being predicted, the model won’t learn well.
- "Enough" depends on the complexity of the problem, but dozens or low hundreds is usually too few for most ML approaches.
- Missing outcome data is particularly problematic: A model can’t predict something it’s never observed.

Data Insufficiency Scenarios

Scenario	Problem
Rare outcome	Predicting which of 10,000 facilities will have a catastrophic failure when it’s happened 3 times in 20 years
New program	No historical data exists because the program just launched
Outcome not recorded	You want to predict client success, but "success" was never systematically tracked
Biased sample	Data only exists for people who completed the process, not those who dropped out
Data exists but is inaccessible	Records are in paper files, another agency’s system, or legally restricted

Questions to ask:

- How many examples of the outcome do we have?
- Is the outcome recorded consistently and reliably?
- Are there systematic gaps in who or what is represented?

When the Problem Is a Moving Target

Key point: If the underlying pattern changes faster than the model can be retrained, AI may produce stale or misleading outputs.

- AI models learn patterns from historical data. They assume the future will resemble the past.
- When the underlying process shifts, due to policy changes, external shocks, or evolving behavior, models degrade.
- "Model drift" is manageable with monitoring and retraining, but some contexts change too fast or too unpredictably.

When Patterns Are Unstable

Situation	Problem
Policy or regulation change	Eligibility rules changed; historical patterns no longer apply
External shock	Pandemic, economic crisis, natural disaster disrupts normal patterns
Adversarial behavior	Fraudsters adapt to detection methods; the pattern shifts as you act on it
Rapid evolution	Technology, markets, or behaviors changing faster than retraining cycles
One-time event	Patterns from a unique historical period won’t generalize

Questions to ask:

- Has the underlying process changed recently?
- How quickly does the environment change?
- Can we retrain frequently enough to keep up?
- Are we trying to predict something that actors will game once they know we’re predicting it?

When Human Judgment Is the Point

Key point: Some decisions should remain human because of accountability, nuance, or public trust—even if AI could technically make them.

- Not every decision **should** be automated, even if it **could** be.
- Some decisions require human accountability: someone needs to be responsible for the outcome.
- Some decisions require nuance that AI can’t capture: context, relationships, values.
- Some decisions affect public trust: automating them may be legal but unwise.

When to Keep Humans in the Loop

Situation	Why Human Judgment Matters
High-stakes individual decisions	Denying benefits, removing children, criminal sentencing
Value-laden tradeoffs	Balancing efficiency vs. equity, cost vs. access
Relationship-dependent contexts	Case management, counseling, community engagement
Novel or ambiguous situations	Edge cases that don’t fit historical patterns
Public-facing decisions	Where automation could erode trust or appear dehumanizing

Questions to ask:

- Who is accountable if this decision is wrong?
- Does this decision require context that isn’t in the data?
- Is there value in the human interaction itself?

When the Cost of Being Wrong Is Unacceptable

Key point: If a wrong prediction has catastrophic consequences and a human review cannot be built in, the risk may outweigh the benefit.

- All models make errors. The question is whether the cost of those errors is acceptable.
- False positives and false negatives have different consequences; both need to be understood.
- High-stakes contexts may require human review of all AI recommendations, which may negate efficiency gains.

Error Cost Considerations

Error Type	Low-Cost Example	High-Cost Example
False positive	Sending an unnecessary appointment reminder	Falsely flagging a family for child abuse investigation
False negative	Missing a low-priority inspection	Failing to detect a disease outbreak

Risk mitigation options:

- Human review of all AI-flagged cases (but this adds cost and delay)
- Human review of high-confidence flags only (but low-confidence cases may be wrong)
- AI as one input among many (but staff may over-rely on it anyway)
- Threshold tuning to favor one error type over another (but tradeoffs remain)

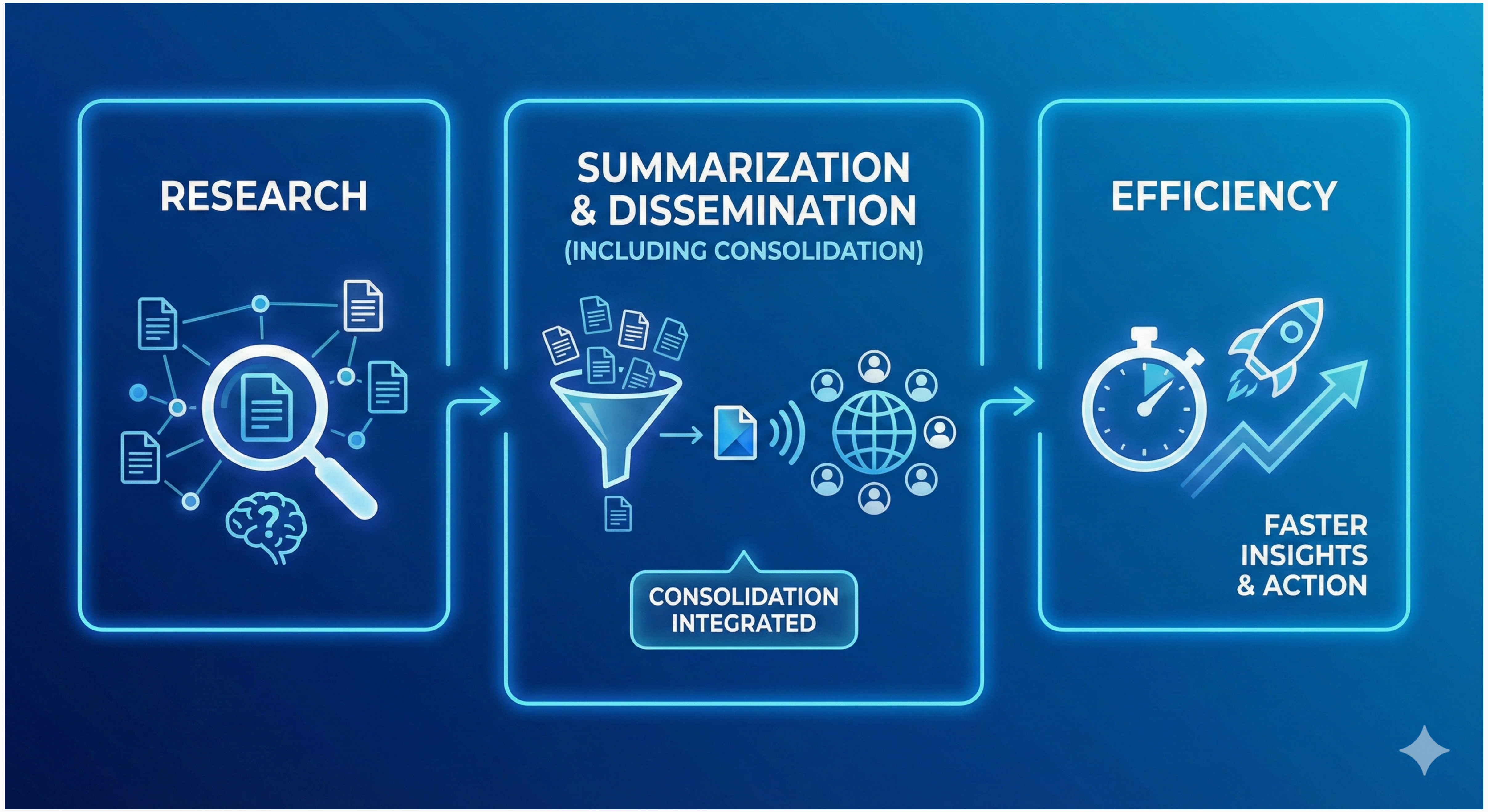
Questions to ask:

- What happens if the model is wrong? Who is harmed?
- Can we build in human review? At what cost?
- Is the error rate acceptable given the stakes?
- Are we comfortable with the tradeoff between false positives and false negatives?

When AI is Overkill or Wrong Summary

Component	Key Takeaway
Simple rule works	AI should earn its complexity; rules are transparent and maintainable
Need to explain the "why"	Prediction and explanation are different goals; choose accordingly
Data doesn’t exist or is too sparse	No data, no AI; verify before proposing
Problem is a moving target	AI assumes the future looks like the past; unstable contexts mislead
Human judgment is the point	"Can we?" is different from "Should we?"
Cost of being wrong is unacceptable	Every model errs; high stakes demand oversight

Applying AI: AI for Research, Summarization, and Efficiency (Opportunity)



AI for Research

Key point: AI can help discover patterns, generate hypotheses, and synthesize evidence at scales that would be impractical for humans alone.

- Research in public health often involves finding patterns in large datasets—disease clusters, risk factors, emerging trends.
- AI excels at pattern discovery: identifying relationships that humans might miss or couldn't process manually.
- AI can also accelerate literature review and evidence synthesis by processing thousands of papers.
- Important distinction: AI is good at **finding** patterns and **generating** hypotheses, but testing hypotheses and establishing causation still requires traditional methods.

Application	What AI Does	Example
Pattern discovery	Finds unexpected relationships in data	Identifying geographic clusters of opioid overdoses
Risk factor identification	Surfaces variables associated with outcomes	Finding predictors of preterm birth in claims data
Anomaly detection	Flags unusual cases or trends	Detecting a spike in ER visits that might indicate an outbreak
Literature synthesis	Summarizes and categorizes research at scale	Reviewing 5,000 papers on vaccine hesitancy to identify themes
Hypothesis generation	Suggests relationships worth investigating	Proposing candidate variables for a formal study

Key question to ask:

Are you trying to find something new in data, or test a specific hypothesis?

- If **finding something new**: AI may help surface patterns worth investigating.
- If **testing a hypothesis**: Classical statistics is likely more appropriate—AI finds correlations, not causes.

AI for Summarization and Dissemination

Key point: AI can condense, translate, and repackage information when the volume or complexity exceeds what humans can efficiently process.

- AI (especially NLP and generative AI) can summarize, translate, and reformat this content.
- Summarization isn’t just "making things shorter". It’s extracting what matters for a specific audience or purpose.

Summarization and Dissemination Applications

Application	What AI Does	Example
Document summarization	Condenses long documents into key points	Summarizing a 200-page regulatory impact analysis
Public comment analysis	Categorizes and summarizes thousands of comments	Identifying themes in 10,000 comments on a proposed rule
Case note synthesis	Extracts key information from unstructured notes	Pulling relevant history from years of case management notes
Plain language translation	Rewrites technical content for general audiences	Converting an epidemiology report into a public fact sheet
Content generation	Creates first drafts for human review	Drafting routine data reports/press releases

Summarization and Dissemination Key question to ask:

Is the source material too voluminous or complex for humans to process efficiently?

- If **yes**: AI can help condense, categorize, or translate.
- If **no**: A human may be faster and more accurate, especially for nuanced content.

Cautions:

- AI summarization can miss nuance or introduce errors—human review is essential.
- Generated content should be reviewed before publication, especially for public-facing materials.
- Sensitive content (legal, clinical) requires extra scrutiny.

AI for Efficiency

Key point: AI can automate repetitive, high-volume tasks that consume staff time—especially classification, triage, and flagging.

- Many agency workflows involve repetitive decisions: categorizing incoming requests, routing cases, flagging exceptions, pre-populating forms.
- These tasks are often low-complexity but high-volume—exactly where AI shines.
- Efficiency gains free up staff for higher-value work that requires human judgment.
- The goal isn't to replace staff but to reduce time spent on routine tasks.

Efficiency Applications

Application	**What AI Does**	**Example**
Classification/routing	Assigns incoming items to categories	Routing constituent emails
Triage/prioritization	Ranks or flags cases by urgency or risk	Flagging enrollees at high risk of hospitalization
Data entry/extraction	Pulls structured data from unstructured sources	Extracting fields from scanned forms
Form pre-population	Fills in known information automatically	Pre-populating inspection forms with history
Anomaly flagging	Identifies outliers for human review	Flagging unusual claims patterns for fraud investigation
Scheduling/matching	Optimizes assignments or appointments	Matching home visitors to families based on needs and availability

AI for Efficiency Key question to ask:

Is there a high-volume, low-complexity task that consumes staff time?

- If **yes**: AI may be able to automate or assist.
- If **no**: The task may not justify the investment, or may require human judgment.

Efficiency vs. replacement:

- AI for efficiency works best when it **assists** humans, not replaces them.
- Staff should review AI outputs, especially for consequential decisions.
- Efficiency gains are real but require upfront investment in building and validating the system.

Overlaps and Combinations

Key point: Many real opportunities span multiple buckets—and that’s fine.

- A single project might involve research (finding patterns), summarization (reporting results), and efficiency (automating ongoing monitoring).
- The buckets are a lens for **finding** opportunities, not a rigid classification system.
- When pitching or scoping a project, it can help to name which bucket(s) it falls into—this clarifies the value proposition.

Examples of Overlap

Project	**Research**	**Summarization**	**Efficiency**
Predict which families will miss SNAP recertification	✓ (pattern discovery)		✓ (triage/flagging)
Summarize public comments and identify themes		✓ (summarization)	✓ (processing volume)
Detect disease outbreaks from syndromic surveillance	✓ (anomaly detection)		✓ (ongoing monitoring)
Generate plain-language summaries of inspection findings		✓ (translation)	✓ (automation)

Where to Look in Your Own Work

Research:

- Do we have large datasets that haven’t been fully explored?
- Are there patterns we suspect but haven’t had time to investigate?
- Is there a body of literature or evidence we need to synthesize?

Summarization/Dissemination:

- Are there documents or reports that take too long to read or summarize?
- Do we struggle to translate technical findings for public audiences?
- Are staff spending time writing routine content that could be drafted automatically?

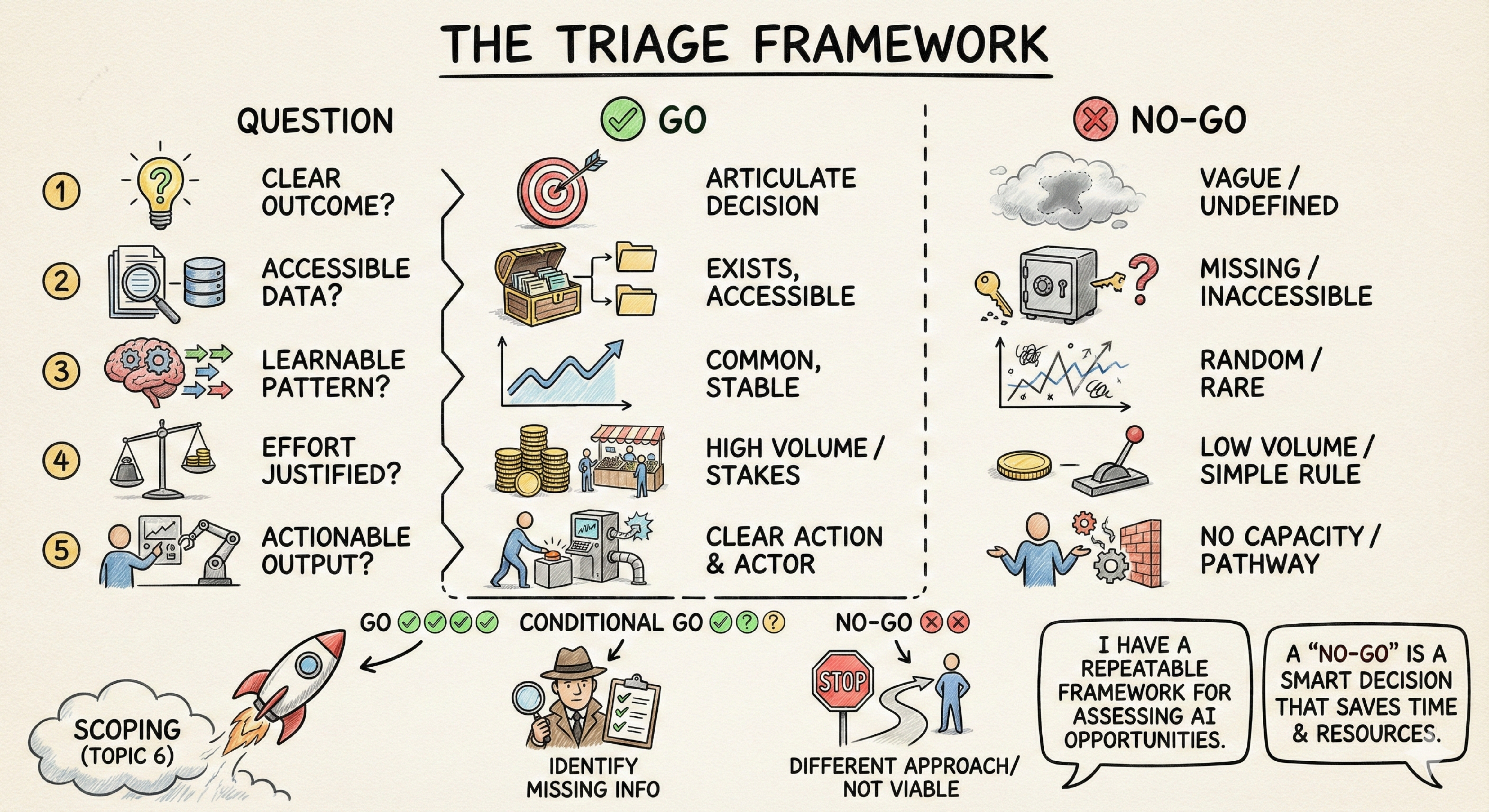
Efficiency:

- What tasks are repetitive and consume significant staff time?
- Are there backlogs in classification, routing, or triage?
- Are staff doing manual data entry that could be automated?

AI for Research, Summarization, and Efficiency Summary

Component	Key Takeaway
AI for Research	AI finds patterns and generates hypotheses; statistics tests them
AI for Summarization	AI handles volume; humans ensure nuance and accuracy
AI for Efficiency	Look for high-volume, low-complexity tasks; AI assists, humans review
Overlaps	Projects often span buckets; use them to find and communicate value
Where to look	Scan your own work using prompts for each bucket

Applying AI: The Problem Triage Framework (Triage)



Is There a Clear Outcome or Decision?

Key point: If you can’t define what success looks like or what decision the AI would inform, stop here.

- This is the first gate. Without a clear outcome, everything else is speculation.
- The outcome should be specific: what exactly will the model produce?
- The outcome should be tied to a decision or action: who will use this output, and for what?
- Vague outcomes ("improve things," "be more efficient") are red flags.

Diagnostic Questions

Question	What You’re Looking For
What would the model output?	A specific prediction, classification, score, or generated content
What decision does this inform?	A concrete action someone would take based on the output
Who uses the output?	A named role or team with capacity to act
How would you know it worked?	A measurable change in outcomes, efficiency, or accuracy

Clear vs. Unclear Outcomes

Clear Outcome	**Unclear Outcome**
Flag patients with >50% probability of 30-day readmission	Improve patient outcomes
Classify emails as complaint, question, or compliment	Make the inbox more manageable
Generate a first draft of the weekly surveillance summary	Help with reporting

Go/No-Go: * **Go:** You can write a sentence: "The model will produce [X] so that [person/team] can [decision/action]." * **No-Go:** You cannot complete that sentence, or it’s vague.

Is There Data—and Is It Accessible?

Key point: No data, no ML. But "data exists" isn't enough—you need data you can actually get your hands on.

- AI learns from examples. You need historical data that includes both the inputs (features) and the outcomes (labels).
- Existence isn't enough: the data must be accessible, usable, and of sufficient quality.
- Common blockers: data is in another system, requires legal agreements, is on paper, or has quality issues.

Diagnostic Questions

Question	What You're Looking For
Does the data exist?	Confirmed, not assumed
Where does it live?	Specific system, database, or files
Who controls access?	Named owner or governance process
What approvals are needed?	Data use agreements, IRB, legal review
Does it include the outcome?	The thing you're trying to predict must be recorded
How much data is there?	Thousands of examples is usually a minimum
What's the quality?	Completeness, consistency, accuracy

Common Data Blockers

Blocker	Example
Data doesn’t exist	"Success" was never tracked; outcome is undefined
Data exists but is inaccessible	Records are in another agency’s system with no data-sharing agreement
Data is on paper or in PDFs	Would require manual extraction or OCR
Data is too sparse	Only 50 examples of the outcome in 5 years
Data quality is poor	Inconsistent coding, missing fields, duplicates
Data doesn’t include the outcome	You have inputs but not what happened next

Go/No-Go:

- **Go:** Data exists, is accessible, includes the outcome, and has reasonable quality and volume.
- **No-Go:** Any of the above blockers are present and can’t be resolved.

Is There a Pattern to Learn?

Key point: AI finds patterns. If the phenomenon is random, rare, or driven by factors not in your data, AI won’t help.

- AI needs signal—a relationship between inputs and outcomes that holds across examples.
- If the outcome is essentially random, or driven by factors you can’t measure, there’s no pattern to learn.
- If the outcome is extremely rare, there may not be enough examples to learn from.
- If the underlying process has changed, historical patterns may not apply.

Diagnostic Questions

Question	**What You’re Looking For**
Is there reason to believe inputs relate to the outcome?	Prior research, domain knowledge, or exploratory analysis
Is the outcome common enough?	Enough positive and negative examples to learn from
Is the process stable?	Patterns from the past are likely to hold in the future
Are the drivers in your data?	The factors that influence the outcome are captured

Warning Signs

Warning Sign	Example
Outcome is essentially random	Predicting lottery numbers
Outcome is extremely rare	3 events in 20 years
Key drivers aren’t in the data	Compliance depends on family support, which isn’t recorded
Process has shifted	Policy change makes historical patterns obsolete
Adversarial context	Fraudsters adapt once they know you’re predicting

Go/No-Go:

- **Go:** There’s reason to believe a learnable pattern exists, the outcome is common enough, and the process is stable.
- **No-Go:** The outcome is random, too rare, or the drivers aren’t in your data.

Is the Effort Justified?

Key point: Could a simpler rule, checklist, or existing process accomplish the same thing? Is the problem high-volume or high-stakes enough to warrant the investment?

- AI projects aren't free. They require data preparation, model development, validation, deployment, and ongoing maintenance.
- The investment should be proportional to the value.
- Sometimes a simple rule or manual process is good enough—or even better.

Diagnostic Questions

Question	**What You're Looking For**
What's the volume?	How many decisions per week/month/year?
What's the cost of the current process?	Staff time, errors, delays
What's the cost of errors?	Consequences of getting it wrong
Could a simple rule work?	Can an expert write down the logic?
Is the improvement worth the investment?	Realistic estimate of benefit vs. cost

When AI May Not Be Justified

Situation	Better Alternative
Low volume (dozens per year)	Manual review
Logic is simple and documented	Rule-based system or checklist
Current process works well	Don’t fix what isn’t broken
Marginal improvement expected	Invest elsewhere
One-time analysis	Traditional statistics or manual review

Go/No-Go:

- **Go:** The problem is high-volume, high-stakes, or high-cost, and simpler alternatives won’t suffice.
- **No-Go:** A simpler approach would achieve most of the benefit at a fraction of the cost.

Can You Act on the Output?

Key point: A prediction is useless if there’s no intervention pathway. Knowing the answer must change what you do.

- This is often the forgotten question. People get excited about prediction without asking: "Then what?"
- Actionability requires: someone who will act, capacity to act, an intervention that works, and a process to deliver the output.
- If you can’t act—due to capacity, authority, or policy—the AI output is wasted.

Diagnostic Questions

Question	What You’re Looking For
What would you do with the output?	Specific intervention or decision
Who would act on it?	Named role or team
Do they have capacity?	Staff time, resources, bandwidth
Is there an intervention that works?	Evidence that the action improves outcomes
Is there a process to deliver the output?	Integration into workflow

Actionability Failures

Failure	**Example**
No capacity	Model flags 500 patients, but only 2 care coordinators available
No authority	Model identifies fraud, but only another agency can investigate
No intervention	Model predicts turnover, but there’s no retention program
No integration	Output sits in a report no one reads
Staff override	Recommendations are routinely ignored

Go/No-Go:

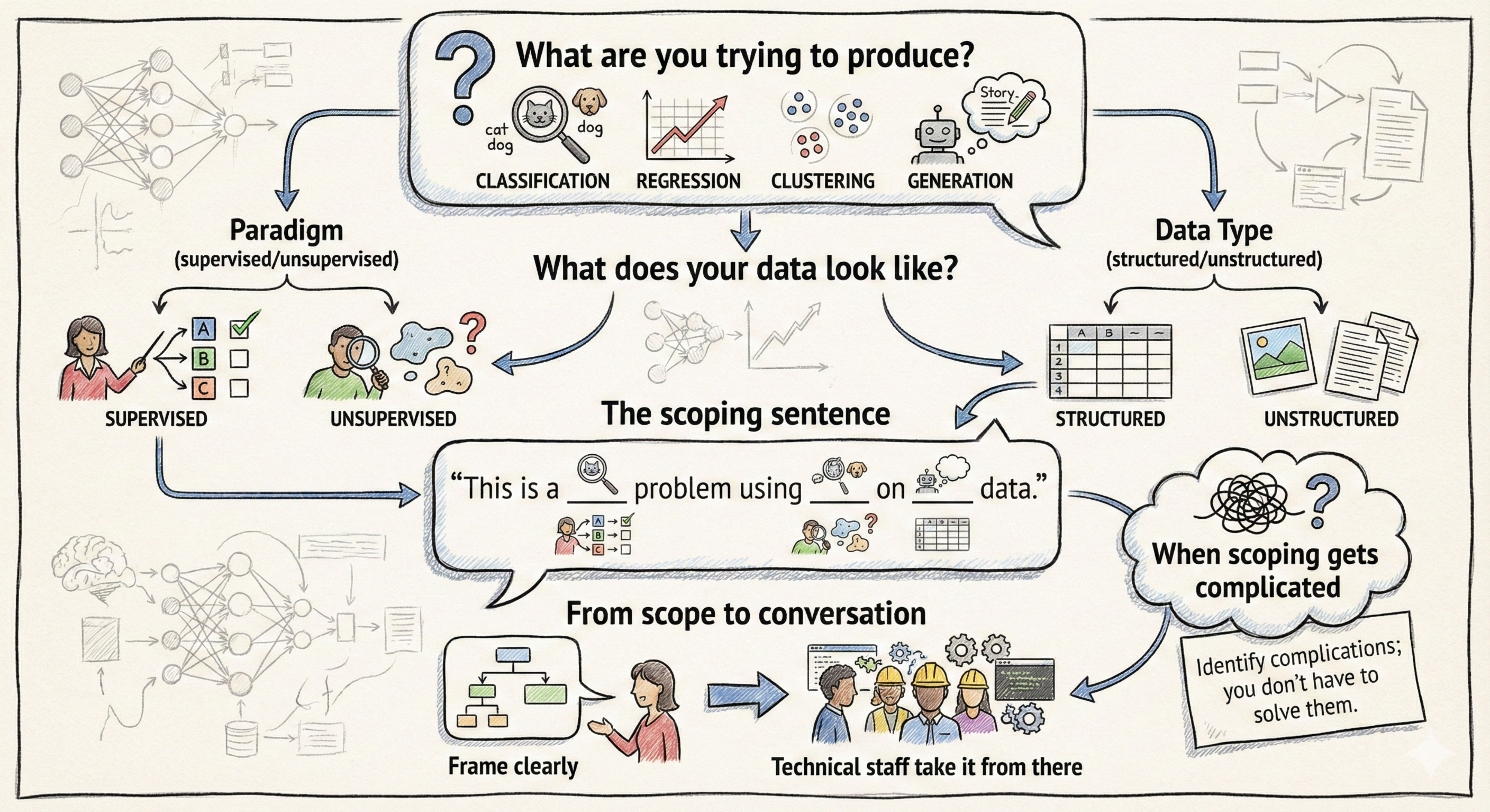
- **Go:** There’s a clear action, a capable actor, and a process to deliver the output.
- **No-Go:** Any link in the action chain is broken.

Putting It Together: The Triage Decision

Key point: The five questions form a sequential filter. Failing any one is a reason to pause—not necessarily to stop forever, but to address the gap.

#	Question	Go	No-Go
1	Is there a clear outcome or decision?	Can articulate output and decision	Vague or undefined
2	Is there data—and is it accessible?	Exists, accessible, includes outcome	Missing, inaccessible, or poor quality
3	Is there a pattern to learn?	Learnable, common enough, stable	Random, rare, or shifted
4	Is the effort justified?	High-volume, high-stakes, no simpler way	Low volume or simple rule works
5	Can you act on the output?	Clear action, capable actor, integrated	No capacity, authority, or pathway

Applying AI: Matching Problems to AI Task Types (Scope)



What Are You Trying to Produce?

Key point: The first scoping question is: what kind of output does the model produce? This determines the task type.

- AI tasks are categorized by what they produce.
- Getting the task type right is essential—it determines what algorithms are appropriate, what data you need, and how you evaluate success.
- The four main task types cover most agency use cases.

The Four Task Types

Task Type	What It Produces	Example
Classification	A category or label	Is this patient high-risk or low-risk? Is this email a complaint, question, or compliment?
Regression	A number	How many WIC participants will enroll next month? What's the predicted length of stay?
Clustering	Groups or segments	What natural groupings exist among overdose cases? Which facilities have similar inspection profiles?
Generation	New content	A summary of this report. A first draft of this communication.

How to Identify the Task Type

If the output is…	Then the task type is…
Yes/no, or one of several categories	Classification
A continuous number (count, amount, duration)	Regression
Groups discovered from the data (no predefined categories)	Clustering
Text, images, or other content created by the model	Generation

Common confusion:

- **Classification vs. Regression:** If predicting "will they or won't they," it's classification. If predicting "how much" or "how many," it's regression.
- **Classification vs. Clustering:** Classification assigns to **predefined** categories (supervised). Clustering **discovers** categories from the data (unsupervised).

What Does Your Data Look Like?

Key point: The structure and labeling of data determines the learning paradigm—supervised or unsupervised.

- **Supervised learning:** You have labeled data—examples where you know the outcome. The model learns to predict the label from the inputs.
- **Unsupervised learning:** You don’t have labels. The model finds structure or patterns in the data without being told what to look for.
- The distinction matters because it determines what’s possible and what data you need.

Learning Paradigms

Paradigm	What You Have	What the Model Does	Example
Supervised	Labeled data (inputs + known outcomes)	Learns to predict the outcome from inputs	Predict readmission (yes/no) from patient data
Unsupervised	Unlabeled data (inputs only)	Finds patterns, groups, or structure	Group outbreak reports by similarity

How Task Type and Paradigm Connect

Task Type	Typical Paradigm
Classification	Supervised (you have labels for each category)
Regression	Supervised (you have the number you’re predicting)
Clustering	Unsupervised (you’re discovering groups, not predicting them)
Generation	Supervised (deep learning trained on examples)

Data Type	Description	Examples
Structured	Organized in rows and columns; each column is a defined field	Spreadsheets, databases, claims data
Unstructured	No predefined format; requires processing to extract information	Free text, images, audio, PDFs

Why it matters:

- Structured data works with most ML algorithms.
- Unstructured data typically requires NLP (for text) or deep learning (for images, audio).
- Many real problems involve a mix—structured fields plus free-text notes.

Putting It Together: The Scoping Sentence

"This is a

[task type]

problem using

[learning paradigm]

on

[data type]

data."

Scoping Sentence Examples

Problem	Task Type	Paradigm	Data Type	Scoping Sentence
Predict which patients will be readmitted within 30 days	Classification	Supervised	Structured	"This is a classification problem using supervised learning on structured data."
Estimate the number of WIC enrollees next quarter	Regression	Supervised	Structured	"This is a regression problem using supervised learning on structured data."
Group disease outbreak reports by similarity	Clustering	Unsupervised	Unstructured	"This is a clustering problem using unsupervised learning on unstructured data."
Generate plain-language summaries of inspection findings	Generation	Supervised (deep learning)	Unstructured	"This is a generation problem using supervised learning (deep learning) on unstructured data."

Scoping Sentence Examples

Problem	Task Type	Paradigm	Data Type	Scoping Sentence
Classify incoming emails by topic	Classification	Supervised	Unstructured	"This is a classification problem using supervised learning on unstructured data."
Predict length of hospital stay	Regression	Supervised	Structured	"This is a regression problem using supervised learning on structured data."
Identify natural groupings among overdose cases	Clustering	Unsupervised	Structured	"This is a clustering problem using unsupervised learning on structured data."
Flag public comments that contain misinformation	Classification	Supervised	Unstructured	"This is a classification problem using supervised learning on unstructured data."

Why this matters:

- The scoping sentence gives technical staff a starting point.
- It ensures you and the technical team are talking about the same thing.
- It helps identify what data preparation is needed (e.g., labeling, text processing).

When Scoping Gets Complicated

Key point: Not every problem fits neatly into one box. Recognize common complications and know when to consult.

- Some problems involve multiple task types (e.g., classify, then summarize).
- Some problems are borderline (e.g., classification with many categories vs. regression with discretized output).
- Some problems require hybrid approaches (e.g., structured data plus free text).
- It’s okay not to have all the answers—the goal is to get close enough to have a productive conversation with technical staff.

Common Complications

Complication	Example	What to Do
Multiple steps	Flag high-risk patients (classification), then generate outreach letter (generation)	Scope each step separately
Borderline task type	Predict "low / medium / high" risk—is this classification or regression?	Either can work; discuss with technical team
Mixed data	Structured claims data plus free-text case notes	May need separate processing pipelines; flag for technical team
Unclear labels	You want to classify, but outcomes aren’t consistently recorded	May need to create labels first (labeling project)
Rare outcome	Classification where only 2% of cases are positive	May need special techniques (class imbalance); flag for technical team

From Scope to Conversation

Key point: The output of scoping is not a technical specification. Usually, it's a starting point for conversation with technical staff or vendors.

- Participants are not expected to design the solution.
- The goal is to communicate clearly: what's the problem, what's the task type, what data is available, and what complications exist.
- Technical staff will take it from there—selecting algorithms, designing features, setting up evaluation.

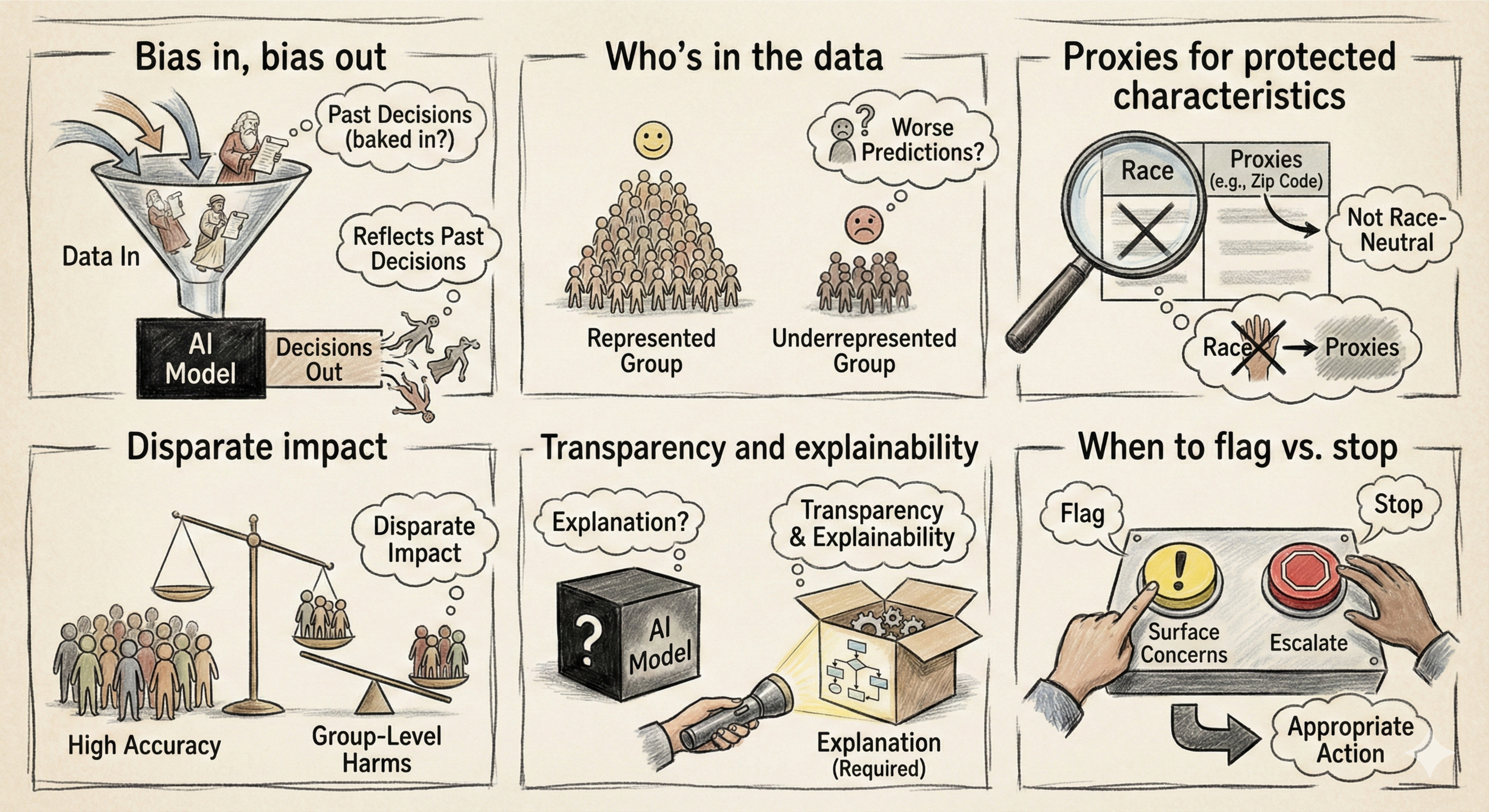
What to Bring to the Conversation

Element	What to Communicate
Problem statement	From Topic 1—what decision or action does this inform?
Triage results	From Topic 5—the five questions and your answers
Scoping sentence	Task type, paradigm, data type
Data description	What data exists, where it lives, volume, quality, access
Complications	Anything that doesn't fit neatly—mixed data, unclear labels, rare outcomes
Concerns	Ethical flags, political sensitivities, timeline constraints

Matching Problems to AI Task Types (Scope) Summary

Component	Key Takeaway
What are you trying to produce?	Task type: classification, regression, clustering, or generation
What does your data look like?	Paradigm (supervised/unsupervised) and data type (structured/unstructured)
The scoping sentence	"This is a [task type] problem using [paradigm] on [data type] data."
When scoping gets complicated	Identify complications; you don’t have to solve them
From scope to conversation	Frame clearly; technical staff take it from there

Applying AI: Surfacing Bias and Ethical Flags



Bias In, Bias Out

Key point: If historical data reflects past inequities, the model will learn and perpetuate those patterns.

- This isn’t a bug—it’s how the technology works. The model finds patterns; it doesn’t know if those patterns are fair.
- The question isn’t "is there bias?" (there almost always is) but "what kind, how much, and does it matter for this use case?"

How Bias Gets Into Data

Source of Bias	Example
Biased decisions	Child welfare investigations historically initiated more often in Black and low-income communities
Biased access	Healthcare utilization data reflects who had insurance and access, not who needed care
Biased measurement	Performance reviews that systematically rate certain groups lower
Biased labels	"Fraud" labels based on who got investigated, not who actually committed fraud
Selection bias	Data only includes people who completed a program, not those who dropped out

Questions to ask:

- What human decisions produced this data?
- Were those decisions made equitably?
- Who was more likely to be in this dataset? Who was less likely?
- If we train on this data, whose patterns are we learning?

Who’s in the Data and Who’s Missing?

Key point: If certain populations are underrepresented in training data, the model may perform poorly for them. Participants should ask: does this data reflect the population we’ll apply this to?

Common Representation Problems

Problem	Example
Geographic gaps	Model trained on urban clinic data applied to rural populations
Demographic gaps	Model trained mostly on white patients applied to diverse population
Temporal gaps	Model trained on pre-pandemic data applied post-pandemic
Institutional gaps	Model trained on one hospital’s data applied across a health system
Outcome gaps	Rare outcomes for certain groups mean less data to learn from

Questions to ask:

- Who is in this dataset? Who is missing or underrepresented?
- Does the training data reflect the population where we’ll deploy this?
- Are there groups for whom we have very few examples?
- Have we checked whether the model performs differently across groups?

Transparency and Explainability Expectations

Key point: Some decisions require explanation to the affected person or the public.

- Some models (especially deep learning) are "black boxes"—they produce outputs, but it's hard to say why.
- Some contexts require explainability: legal proceedings, clinical decisions, public accountability.
- Even if not legally required, explainability builds trust and enables oversight.
- There's often a tradeoff: more complex models may be more accurate but less explainable.

When Explainability Matters

Context	**Why Explainability Matters**
Benefit denials	Applicants may have a right to know why they were denied
Clinical recommendations	Physicians need to explain decisions to patients
Child welfare	Families and courts may demand to know why a flag was raised
Public-facing decisions	Citizens expect transparency in government decisions
Audits and oversight	Regulators or legislators may require explanation

Questions to ask:

- Will anyone need to explain why this decision was made?
- Can we articulate what factors drove the model's recommendation?
- Is there a legal or regulatory requirement for explainability?
- Would a "black box" decision erode public trust?

When to Flag vs. When to Stop

Key point: Not every concern is a dealbreaker.

- Bias and ethical concerns exist on a spectrum. Some are serious enough to stop a project; others can be noted and managed.

Escalation Framework

Level	Description	Example	Action
Note and monitor	Minor concern; manageable with design choices	Model uses zip code; could proxy for race	Document; test for disparate impact; monitor
Pause and consult	Significant concern; needs expert input	Training data reflects known historical bias	Consult with equity office or ethics board before proceeding
Stop and escalate	Serious concern; potential for significant harm	Model would be used for high-stakes decisions on vulnerable population with no human review	Escalate to leadership; do not proceed without explicit approval

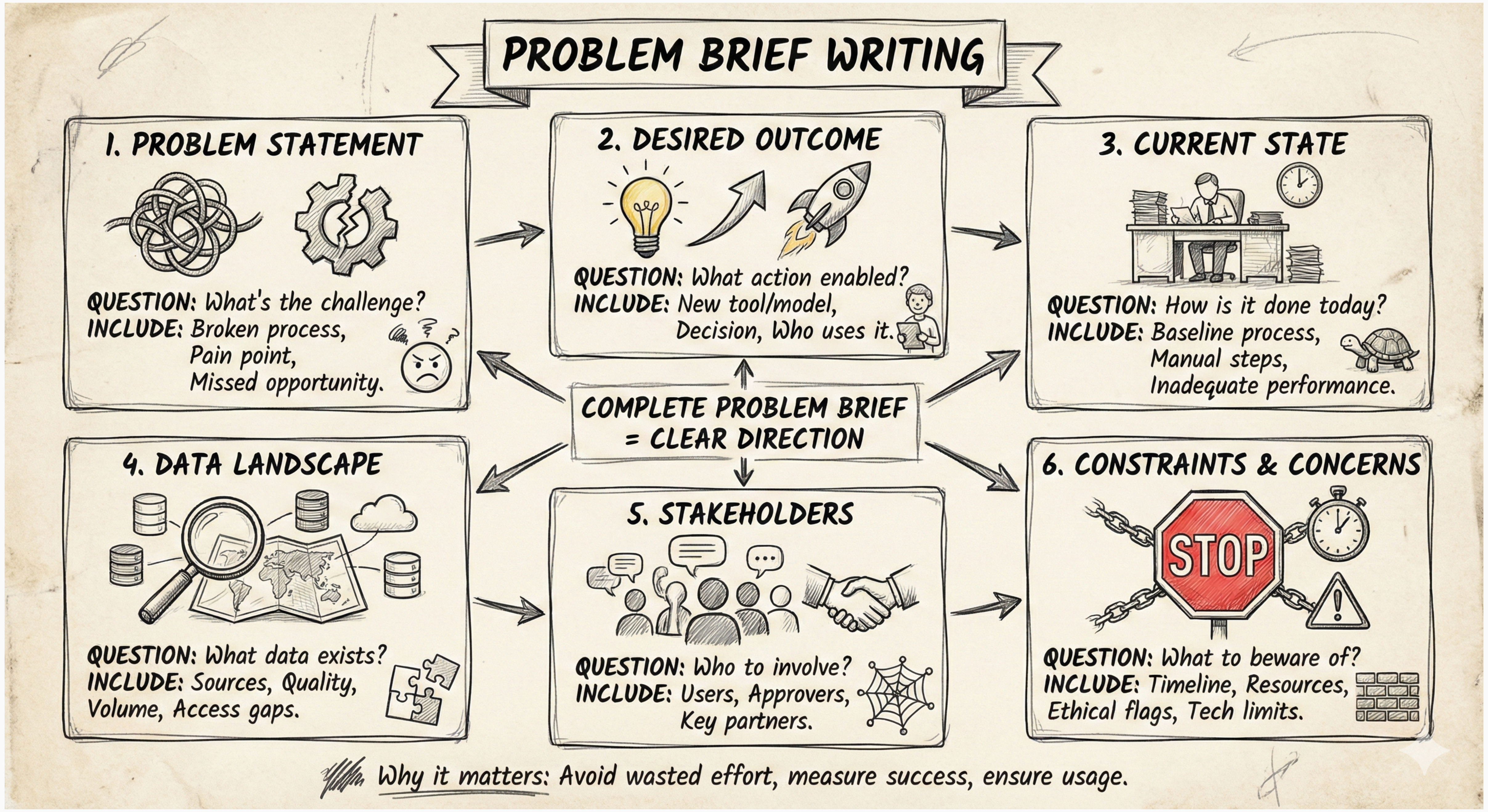
Questions to ask:

- Is this a concern I can address with design choices, or does it need outside input?
- Who in my organization should know about this?
- What's the worst-case scenario if this concern isn't addressed?
- Is there a formal review process this should go through?

Surfacing Bias and Ethical Flags Summary

Component	**Key Takeaway**
Bias in, bias out	Historical data reflects past decisions—ask what biases are baked in
Who’s in the data	Underrepresented groups may get worse predictions
Proxies for protected characteristics	Removing race doesn’t make a model race-neutral
Disparate impact	Overall accuracy can hide group-level harms
Transparency and explainability	Some decisions require explanation
When to flag vs. stop	Surface concerns; escalate appropriately

Applying AI: Writing a Problem Brief



What a Problem Brief Is (and Isn’t)

Key point: A problem brief is a clear, concise articulation of the problem, the desired outcome, and the context—written for someone who doesn’t know the program area. Articulating a problem clearly is the bridge between "we should use AI" and actually getting something built.

What a Problem Brief Does and Doesn’t Do

It Does	It Doesn’t
Articulate the problem clearly	Prescribe a technical solution
Define what success looks like	Specify algorithms or tools
Describe the current state	Include project timelines or budgets
Identify stakeholders and constraints	Assign tasks or responsibilities
Flag concerns and unknowns	Resolve all questions

Core Elements of a Problem Brief

Key point: A complete problem brief has six elements. Each one answers a question that technical staff or decision-makers will need answered.

The Six Elements

Element	**Question It Answers**	**What to Include**
Problem statement	What's the challenge?	What's not working, what's the pain point, or what opportunity exists. Be specific.
Desired outcome	What decision or action would this enable?	What the model would produce, who would use it, and what they would do with it.
Current state	How is this handled today?	The baseline: current process, performance, volume, and why it's inadequate.
Data landscape	What data exists?	Where the data lives, what it contains, volume, quality, access, and gaps.
Stakeholders	Who needs to be involved?	Who would use the output, who needs to approve, who needs to be consulted.
Constraints and concerns	What should we be aware of?	Timeline, resources, ethical flags, political sensitivities, technical limitations.

Writing a Good Problem Statement

Key point: The problem statement is the foundation. It should be specific, outcome-focused, and solution-agnostic.

- A good problem statement describes the pain point without embedding a solution.
- It’s specific enough that someone outside the program area can understand it.
- It focuses on outcomes (what’s not working) rather than activities (what you want to do).

Problem statement formula:

[Who] is experiencing [what problem], which results in [what consequence]. Currently, [how it’s handled], but [why that’s inadequate].

Weak vs. Strong Problem Statements

Weak Problem Statement	Strong Problem Statement
"We need a chatbot for our call center."	"Constituents calling our benefits hotline wait an average of 18 minutes to reach a representative. 40% of calls are simple status checks that don’t require human assistance. This leads to frustration and ties up staff who could handle complex cases."
"We want to use AI to improve maternal health."	"Women in rural counties are 2.3x more likely to experience severe maternal morbidity than those in urban areas. We lack early warning indicators to identify high-risk pregnancies before complications occur. Currently, risk assessment is done at the first prenatal visit using a paper checklist."
"We should predict fraud."	"Our program integrity unit reviews 200 provider claims per month for potential fraud, selected based on manual referrals. We estimate we’re missing 80% of fraudulent claims. Each undetected fraud case costs an average of $45,000."

Describing the Data Landscape

Key point: Data landscape includes: what data exists, where it lives, what format it's in, how much there is, what's known about quality, and what approvals are needed to access it. It's okay to have gaps but they need to be identified.

What to Include in Data Landscape

Question	What to Document
What data exists?	Tables, fields, systems, sources
Where does it live?	System names, databases, file locations
What format is it in?	Structured (database, spreadsheet) or unstructured (text, PDFs)
How much is there?	Number of records, time range covered
What's the quality?	Completeness, consistency, known issues
Does it include the outcome?	Is the thing you're trying to predict recorded?
What approvals are needed?	Data use agreements, IRB, legal review
What's missing or unknown?	Gaps, uncertainties, things you need to find out

Example data landscape description:

"Data sources include Medicaid claims (5 years, 2M+ records, stored in our data warehouse), vital records (birth certificates linked by patient ID), and WIC enrollment data (separate system, requires data sharing agreement). Claims data is generally complete but has known issues with diagnosis coding consistency. We do not currently have social determinants of health data. Outcome variable (hospital readmission within 30 days) is recorded in claims. Access requires DUA with Medicaid agency (in progress)."

Common Mistakes in Problem Briefs

Key point: Certain mistakes are predictable and avoidable. Recognizing them helps participants write better briefs and critique others' briefs.

Mistake	Example	Why It’s a Problem
Embedding a solution	"We need a machine learning model that…"	Skips problem definition; may be solving the wrong thing
Being too vague	"Improve outcomes for clients"	Can’t evaluate success; technical staff don’t know what to build
Being too broad	"Use AI to transform our department"	No specific problem; no way to scope or prioritize
Overcomplicating	"A system that predicts X, flags Y, generates Z, and integrates with A, B, C"	Too many goals; likely to fail or take forever
Ignoring the "so what"	"Predict which clients will miss appointments" (but no intervention)	Prediction without action is useless
Skipping constraints	No mention of timeline, resources, or sensitivities	Sets up for surprise blockers later
Assuming data exists	"We’ll use our data to…" without verification	May discover data doesn’t exist or isn’t accessible

How to avoid these mistakes:

- Write the problem statement before mentioning any technology.
- Ask: "If someone outside my agency read this, would they understand the problem?"
- Ask: "What would we do differently if we had the answer?"
- Ask: "What could go wrong, and have I mentioned it?"

The Review Process: Giving and Receiving Feedback

Key point: Problem briefs improve with feedback.

- Writing is iterative. A first draft is rarely final.
- Peer review catches blind spots—things that are obvious to you but not to others.
- A good critique focuses on clarity, completeness, and feasibility—not style preferences.

Review Rubric

Criterion	Questions to Ask
Problem is specific and outcome-focused	Can I explain the problem to someone outside this program? Is there a clear pain point?
Success is defined	Do I know what "working" looks like? Is it measurable?
Current state is described	Do I understand how this is handled today and why that’s inadequate?
Data is addressed	Do I know what data exists, where it is, and what’s uncertain?
Stakeholders are identified	Do I know who will use the output and who needs to be consulted?
Concerns are flagged	Are constraints, risks, and sensitivities mentioned?
No solution is embedded	Does the brief describe a problem, or does it prescribe a solution?

Giving and Receiving Feedback

Giving feedback:

- Start with what’s clear and strong.
- Ask questions rather than making assertions: "I wasn’t sure what you meant by…" or "What would happen if…?"
- Focus on gaps, not style: "I didn’t see anything about data access—is that a concern?"
- Be specific: "The problem statement is clear, but the success criteria are vague" is more useful than "needs work."

Receiving feedback:

- Assume good intent—the goal is a better brief.
- Listen for confusion—if the reviewer is confused, others will be too.
- Ask clarifying questions: "What would make this clearer?"
- Revise based on patterns—if multiple reviewers flag the same issue, it’s real.

Writing a Problem Brief Summary

Component	**Key Takeaway**
What a brief is	Frames the problem; doesn’t prescribe the solution
Core elements	Six elements: problem, outcome, current state, data, stakeholders, constraints
Problem statement	Specific, outcome-focused, solution-agnostic
Data landscape	Name what exists, where it is, and what’s unknown
Common mistakes	Embedding solutions, being vague, ignoring "so what"
Review process	Use the rubric; give and receive feedback constructively