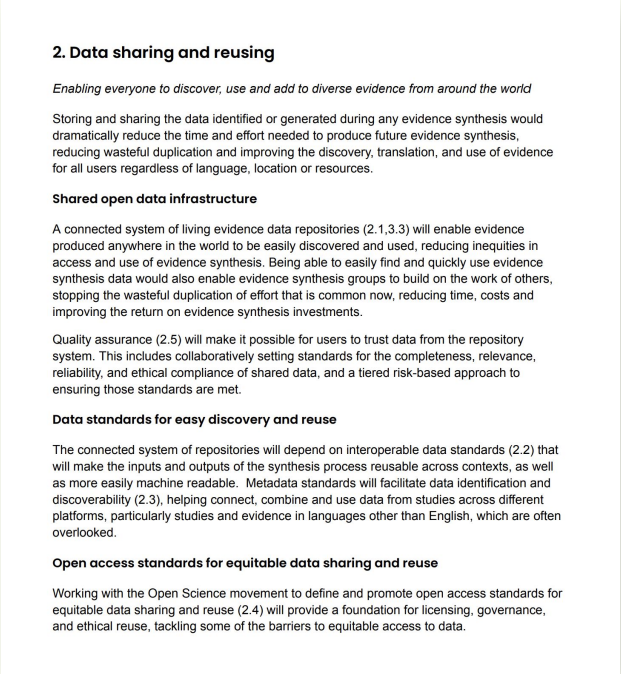
Early thoughts on
Evidence Data Systems

Enabling everyone to discover, use and add to diverse evidence from around the world

- A connected system of living evidence data repositories with data standards, interoperability & quality assurance
    - Deliver open, harmonised, synthesis-ready data
- Ask what society considers to be public good outcomes
- Empower those traditionally excluded from power
- Be driven by and accountable to diverse evidence users
- Harness the incredible capability of AI, cognisant of limitations & uncertainty
- Consider what markets can solve & not solve
- Engage deeply with context & systems beyond ‘evidence’



2. Data sharing and reusing

Enabling everyone to discover, use and add to diverse evidence from around the world

Storing and sharing the data identified or generated during any evidence synthesis would dramatically reduce the time and effort needed to produce future evidence synthesis, reducing wasteful duplication and improving the discovery, translation, and use of evidence for all users regardless of language, location or resources.

Shared open data infrastructure

A connected system of living evidence data repositories (2.1,3.3) will enable evidence produced anywhere in the world to be easily discovered and used, reducing inequities in access and use of evidence synthesis. Being able to easily find and quickly use evidence synthesis data would also enable evidence synthesis groups to build on the work of others, stopping the wasteful duplication of effort that is common now, reducing time, costs and improving the return on evidence synthesis investments.

Quality assurance (2.5) will make it possible for users to trust data from the repository system. This includes collaboratively setting standards for the completeness, relevance, reliability, and ethical compliance of shared data, and a tiered risk-based approach to ensuring those standards are met.

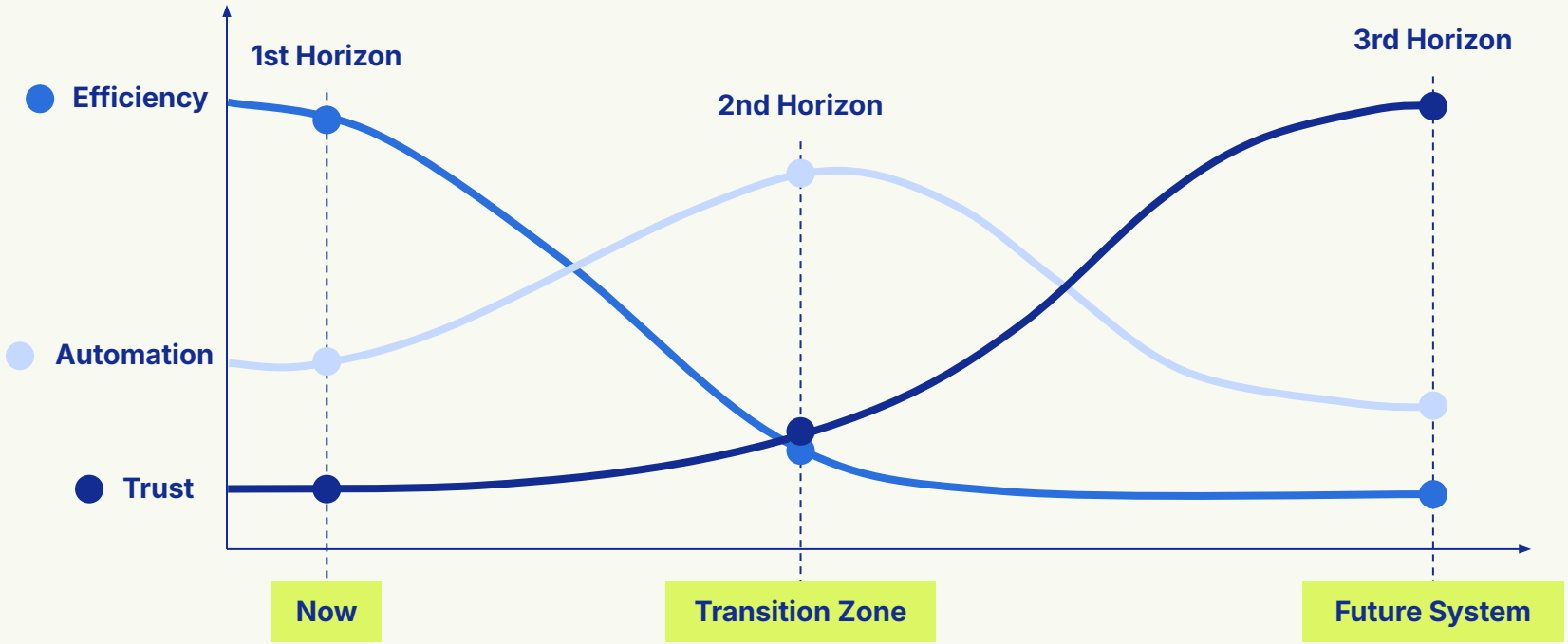
Data standards for easy discovery and reuse

The connected system of repositories will depend on interoperable data standards (2.2) that will make the inputs and outputs of the synthesis process reusable across contexts, as well as more easily machine readable. Metadata standards will facilitate data identification and discoverability (2.3), helping connect, combine and use data from studies across different platforms, particularly studies and evidence in languages other than English, which are often overlooked.

Open access standards for equitable data sharing and reuse

Working with the Open Science movement to define and promote open access standards for equitable data sharing and reuse (2.4) will provide a foundation for licensing, governance, and ethical reuse, tackling some of the barriers to equitable access to data.

Three horizons: efficiency, automation & trust



Sharpe B, et al. Ecology and Society. 2016;21(2):47

Global trends suggest increasing fragmentation & declining trust

- Our future is being shaped by **social fragmentation, geopolitical conflict & AI**
- These forces will shape knowledge systems and understandings of the world
    - **AI** will accelerate our ability to know, but also vulnerability to misinformation, fragmentation & polarisation
    - **Geopolitical conflict** will increase distortion & disruption of information systems by nations states & non-state actors
    - **Science & experts** will be increasingly contested, distorted and co-opted - and likely diminished in the public space
- The forces that create these challenges will make it harder to build the coalitions & global frameworks to solve them

In this context **trust & shared understanding will be hard to achieve**

Edelman Trust Barometer 2025
[33,000 respondents across 28 countries]

"The world is failing to meet the unprecedented challenges of our time because it is ensnared in a vicious cycle of distrust."

Richard Edelman, CEO Edelman

LLM-powered evidence products will drive variable & fragmented understandings, undermining trust

- Society is poorly served by current evidence synthesis offerings
- LLMs have the potential to transform this offering, but effectiveness of LLMs is limited in high-stakes settings where accuracy is critical[1] (quality of the result is a binary, not a percentage[2])
- Low accuracy LLM-powered science summary/decision products will continue to disrupt the market for decision-support product & services
- LLMs, LLM products & associated markets will drive **a highly variable representation of 'what is known'**, even as accuracy improves
- **Evidence cultures are too weak** to generate sufficient countervailing market forces **to overcome these trends**

LLM-powered products drive high variance representations
Models
Stochastic outputs, context sensitivity, linguistic variation, opaque model changes, emergent behaviours
Products
Variable product goals, divergent design iterations, variable response to subjective model outputs
Markets
Domain-specific preferences & biases, market-specific model tuning, variable regulatory signals

[1]Stanford Artificial Intelligence Index Report 2025; [2]Benedict Evans. Are better models better? 22 Jan 2025

Our Vision

A living, open, equitable & extensible evidence data system that enables shared understanding and collective impact.

Planned interventions are necessary but insufficient to mitigate the risk of a low trust future

No Intervention

LLMs, LLM products & associated markets will drive **a highly variable representation of 'what is known'**, even as accuracy improves.

✅**Efficiency**

✅**Automation**

❌Trust

Planned interventions

Global governance & **coordinated action** by aligned, independent organisations. **Independent validation of evidence synthesis AI** against a comprehensive set of benchmarks.

✅**Efficiency**

✅**Automation**

❓Trust

Proposed interventions

A **living, open, equitable & extensible** evidence data system that enables **shared understanding and collective impact.**

✅**Efficiency**

✅**Automation**

✅**Trust**

Core outcomes of key actors create an opportunity to develop a data flywheel

Organisations with Evidence Docs & Data	Evidence Synthesis Tool Developers	Evidence Synthesis AI Model Builders	Evidence Users, Advisors & Intermediaries	Evidence Use Product Developers
Improve discovery & reuse of org data	Reduce end user workload	High model use	Validated, high quality, up to date evidence data	Validated, high quality, up to date evidence data
Enhance org data value	Build user trust	High model performance at low effort & cost	Low effort & cost	Low effort & cost
Low host costs	Integrate AI at low effort & cost	Data for training & testing		
		Low hosting costs		

Core outcomes of key actors create an opportunity to develop a data flywheel

Organisations with Evidence Docs & Data

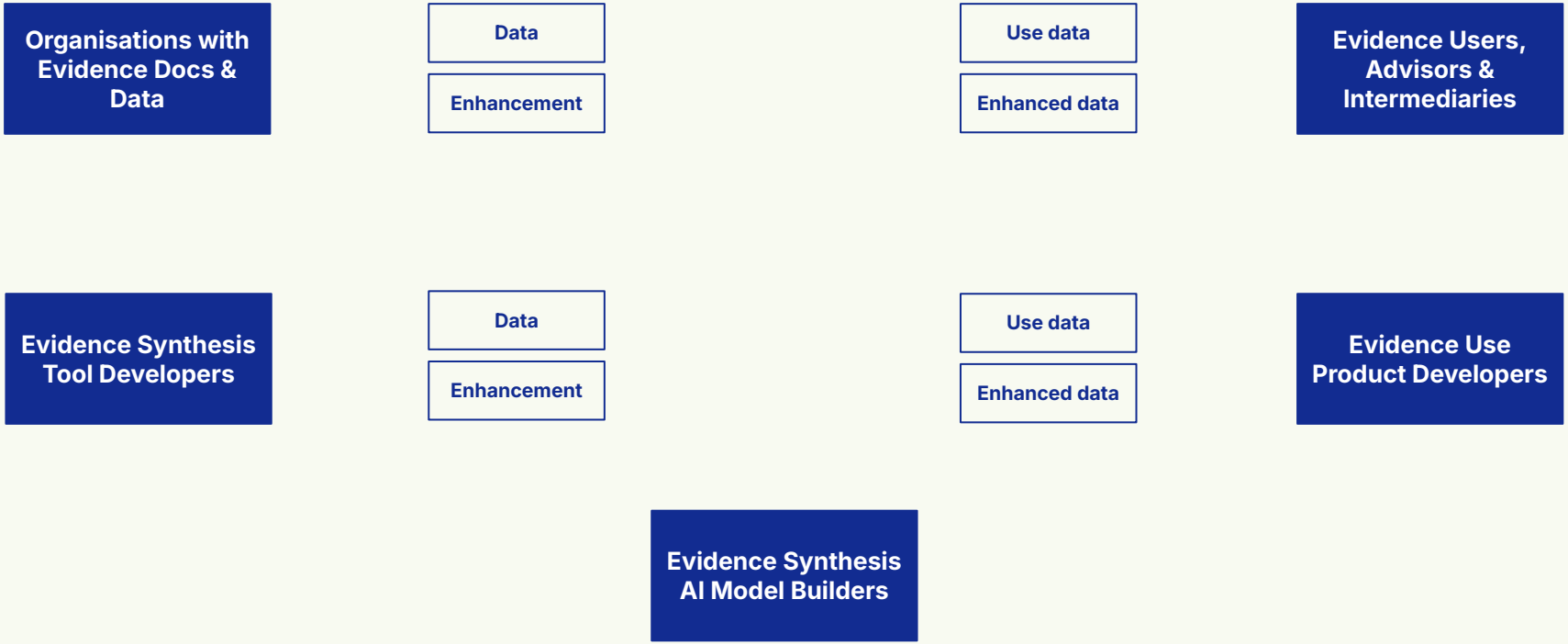
Evidence Users, Advisors & Intermediaries

Evidence Synthesis Tool Developers

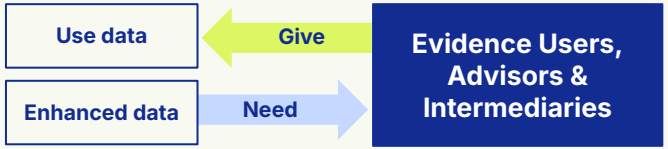
Evidence Use Product Developers

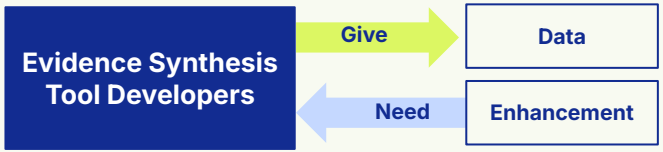
Evidence Synthesis AI Model Builders

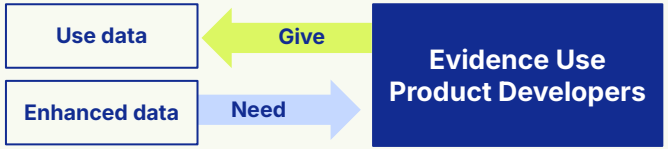
Core outcomes of key actors create an opportunity to develop a data flywheel



Core outcomes of key actors create an opportunity to develop a data flywheel

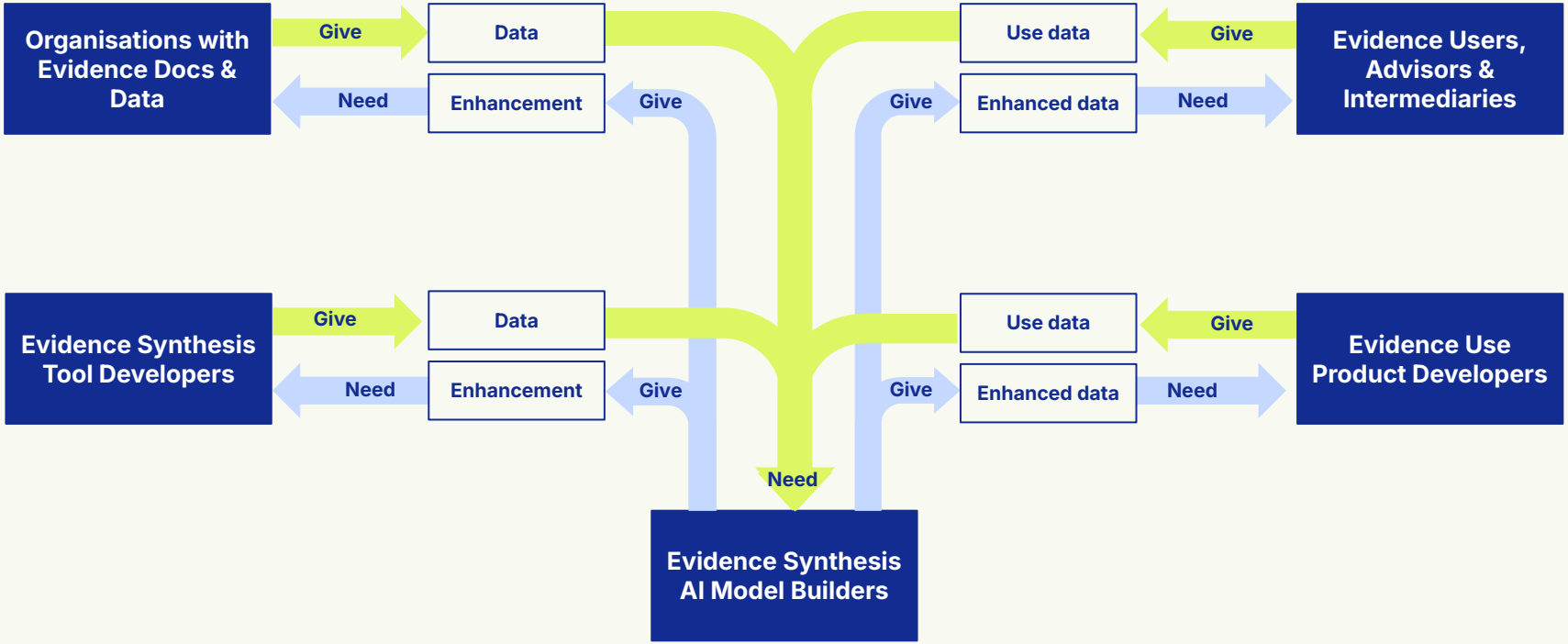




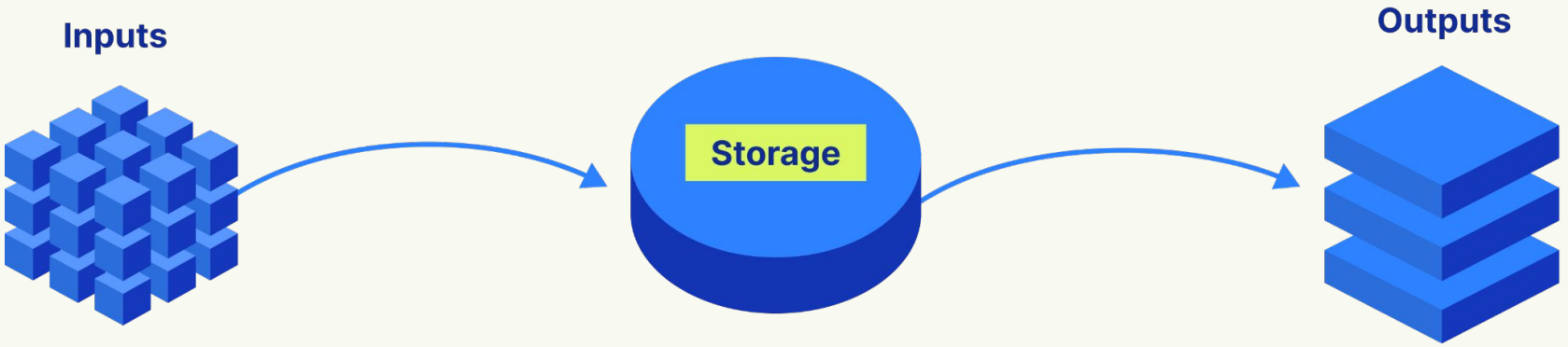


Evidence Synthesis AI Model Builders

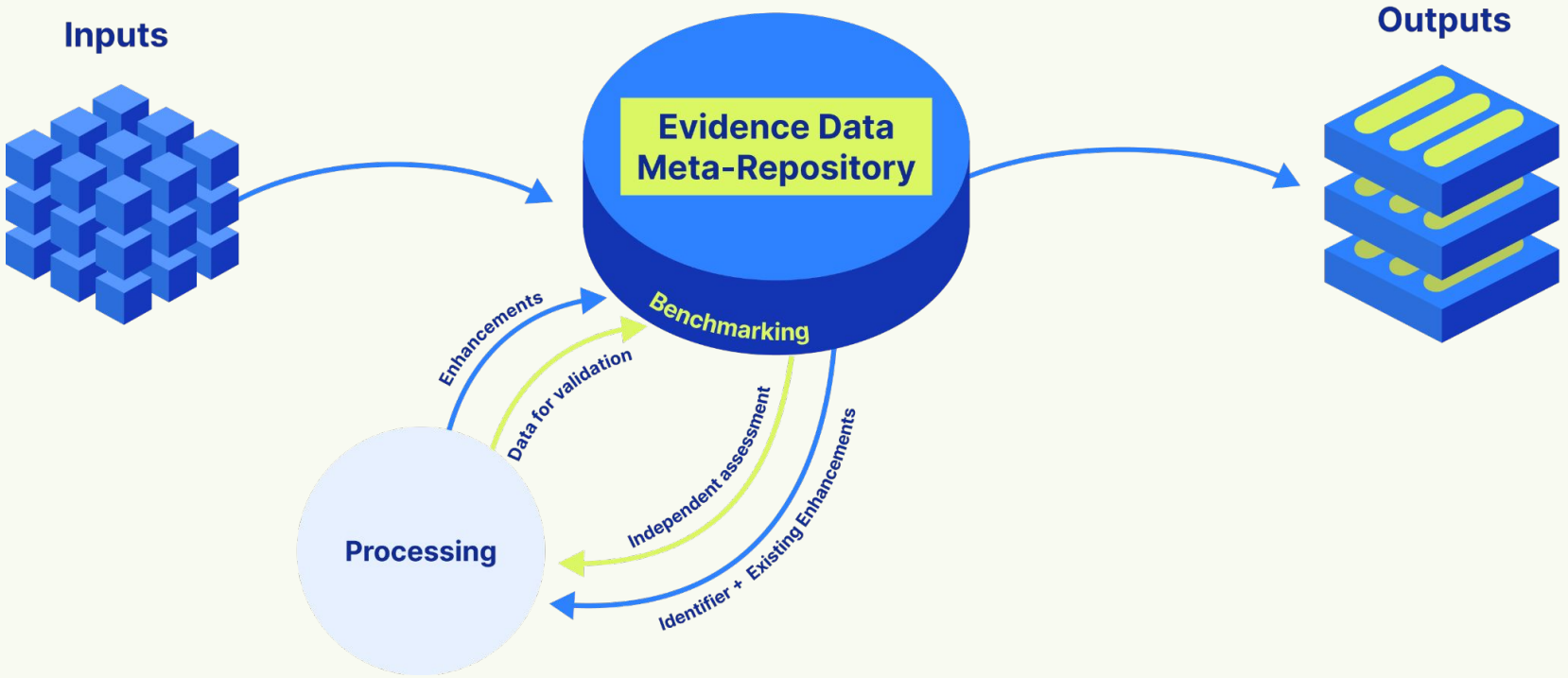
Core outcomes of key actors create an opportunity to develop a data flywheel



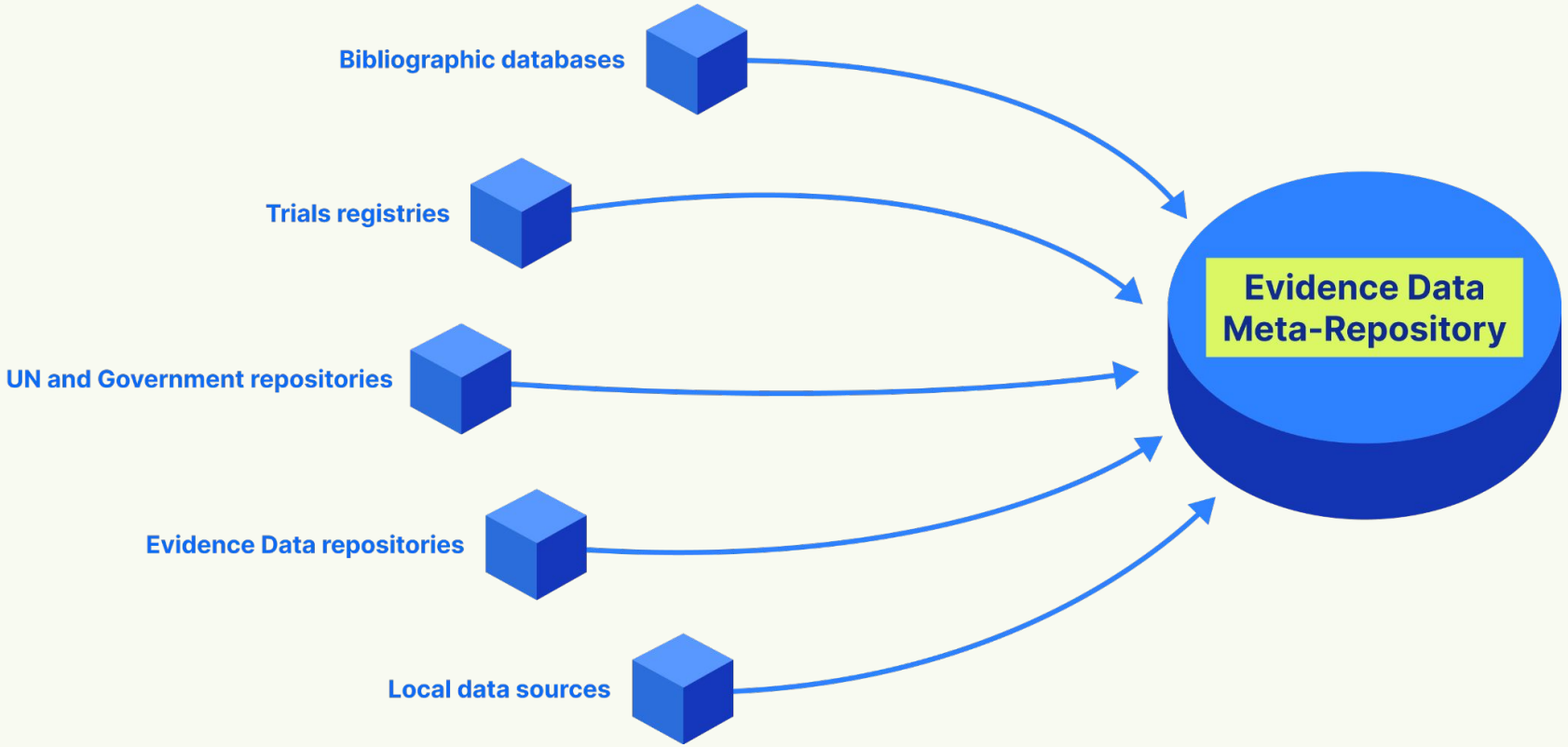
Traditional repository



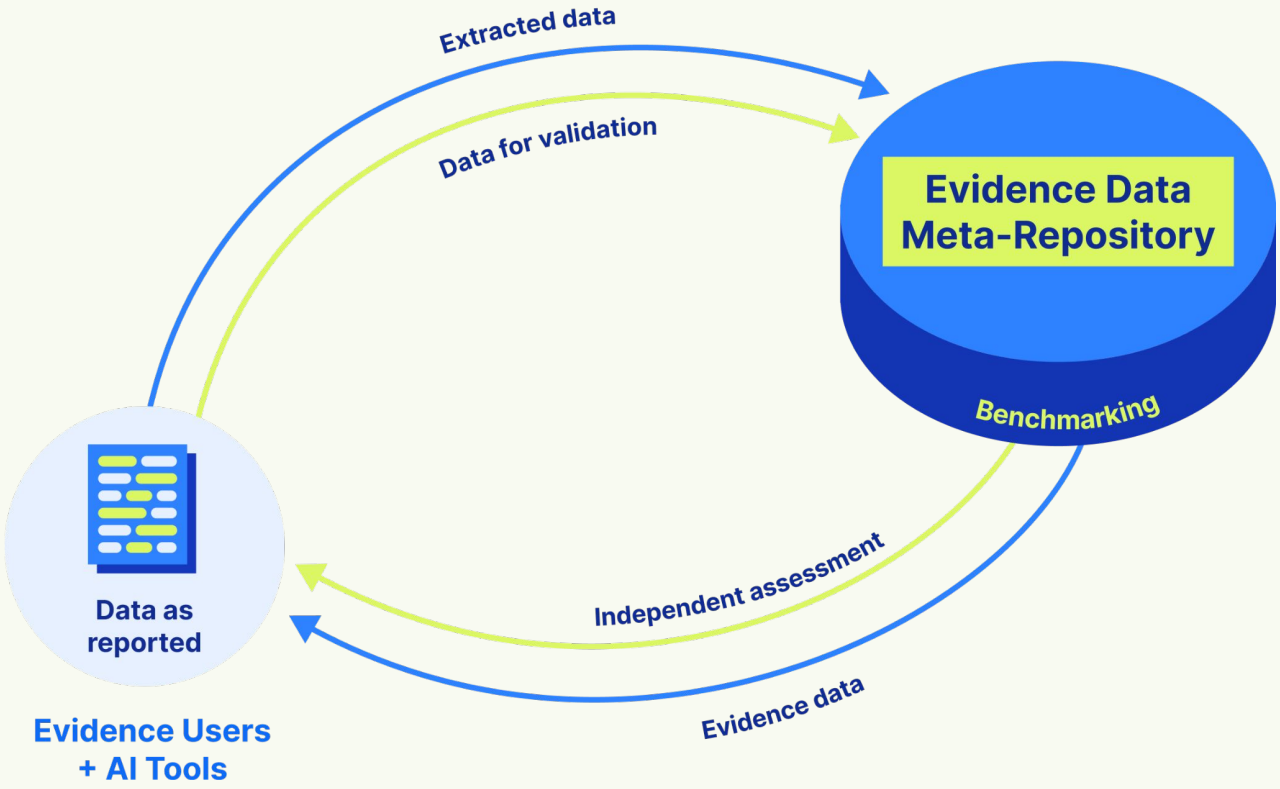
Enhancing repository



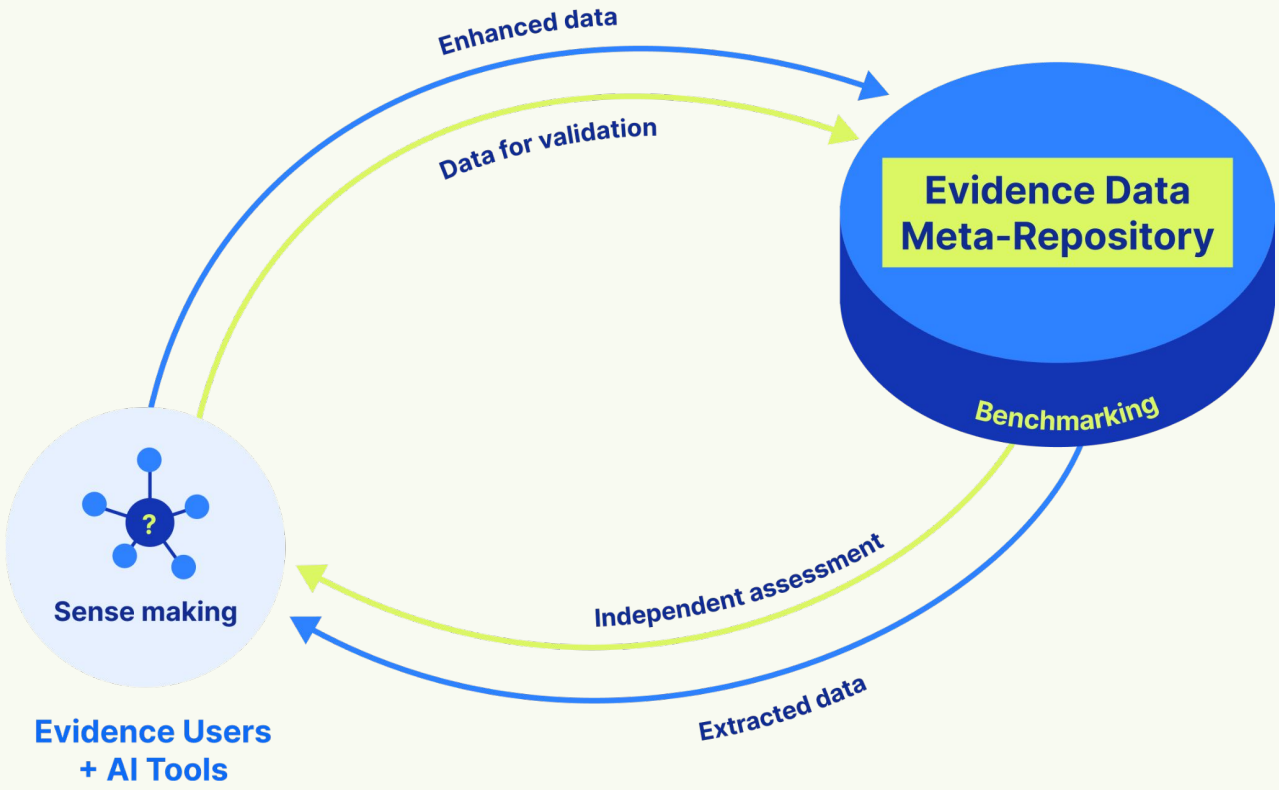
Evidence Data Sources



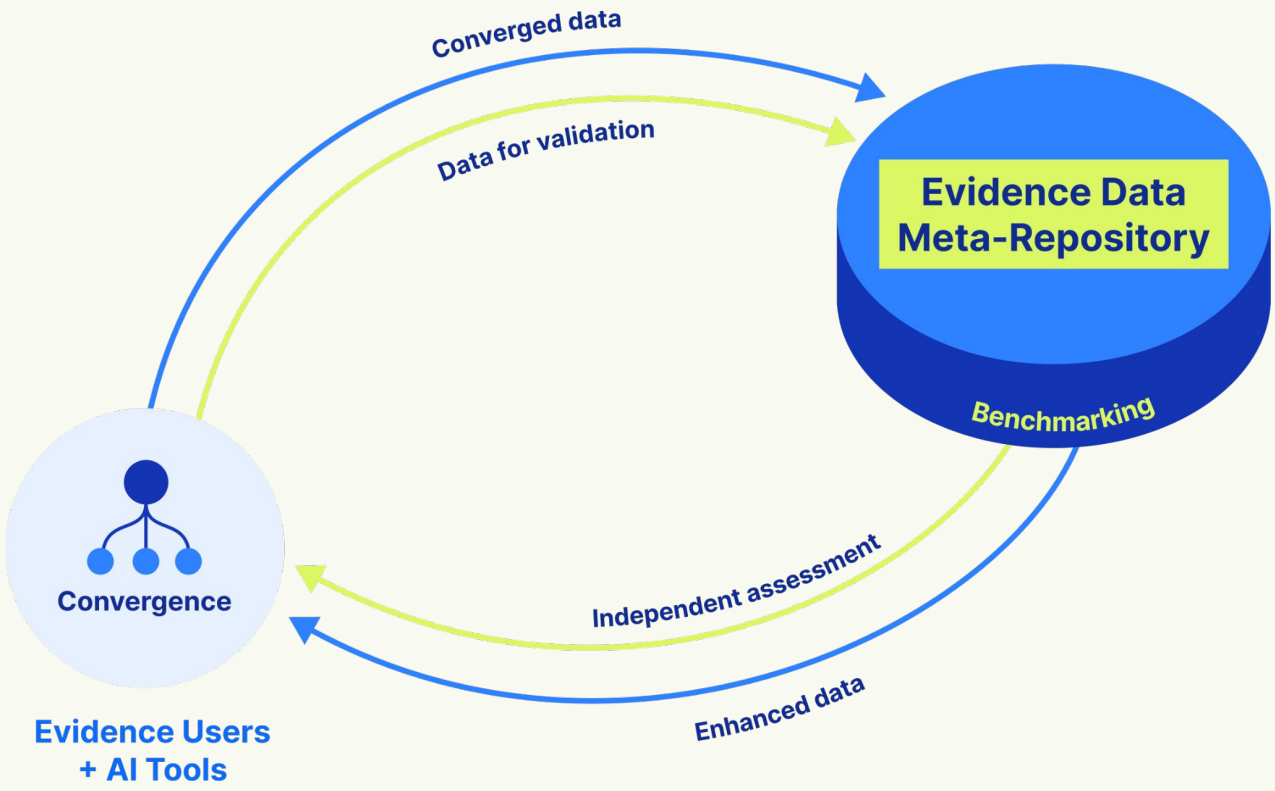
Data as reported



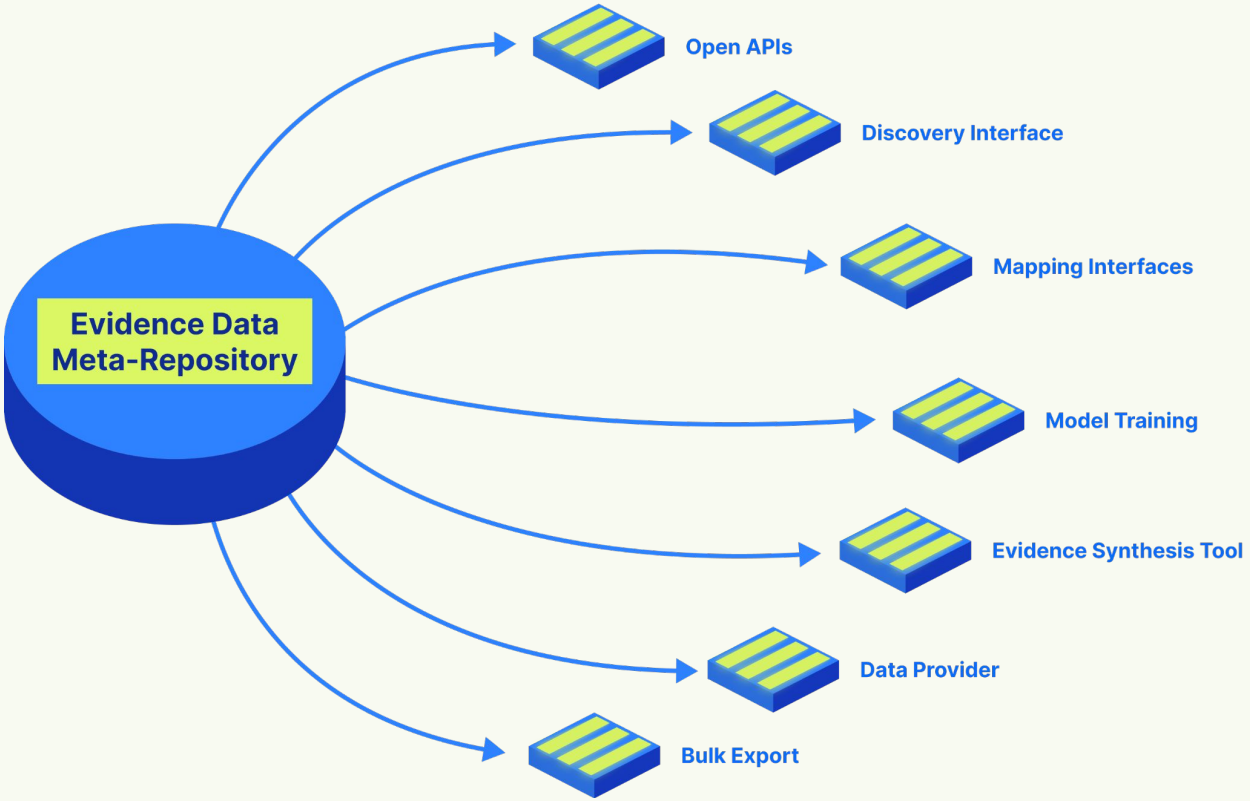
Sense making

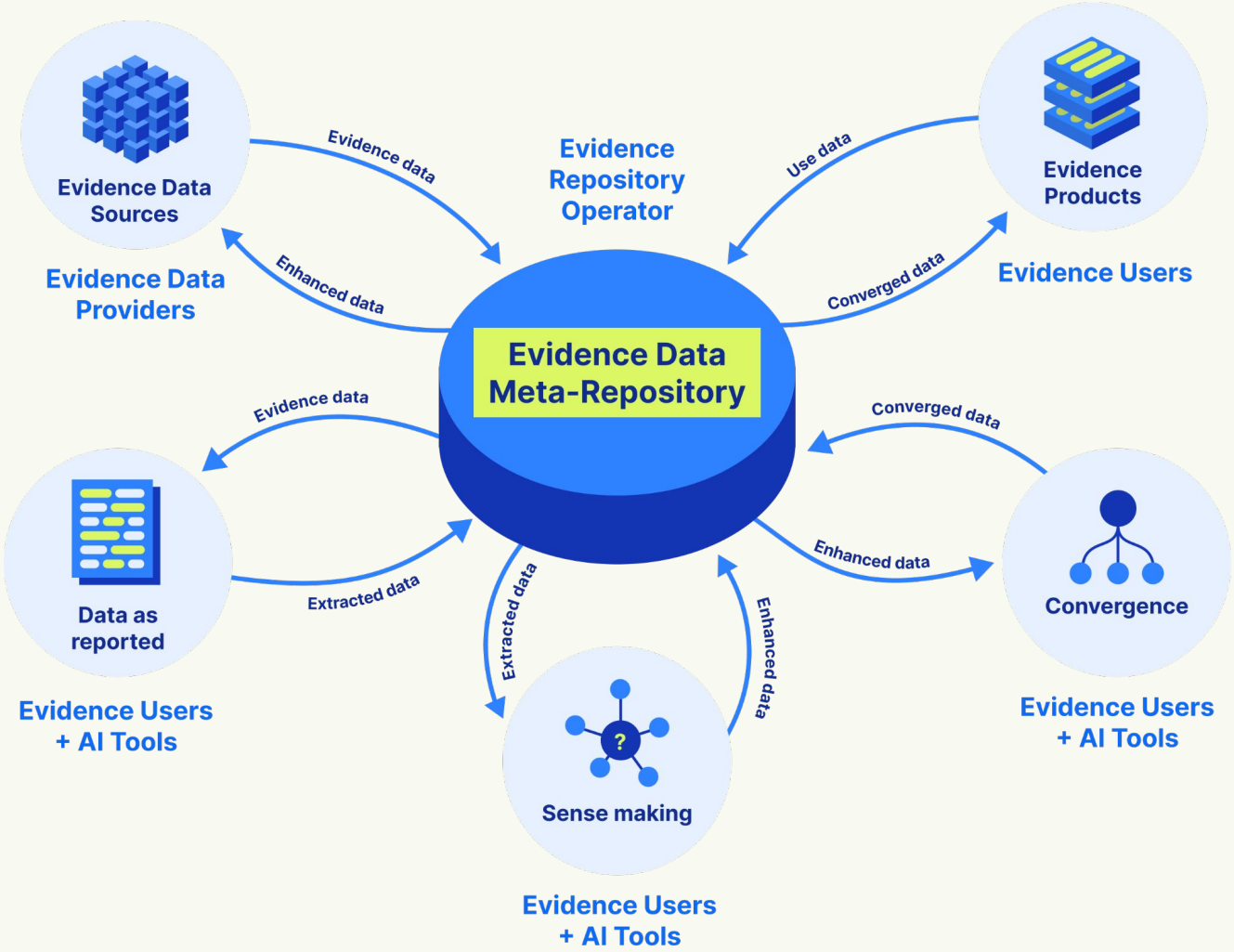


Convergence



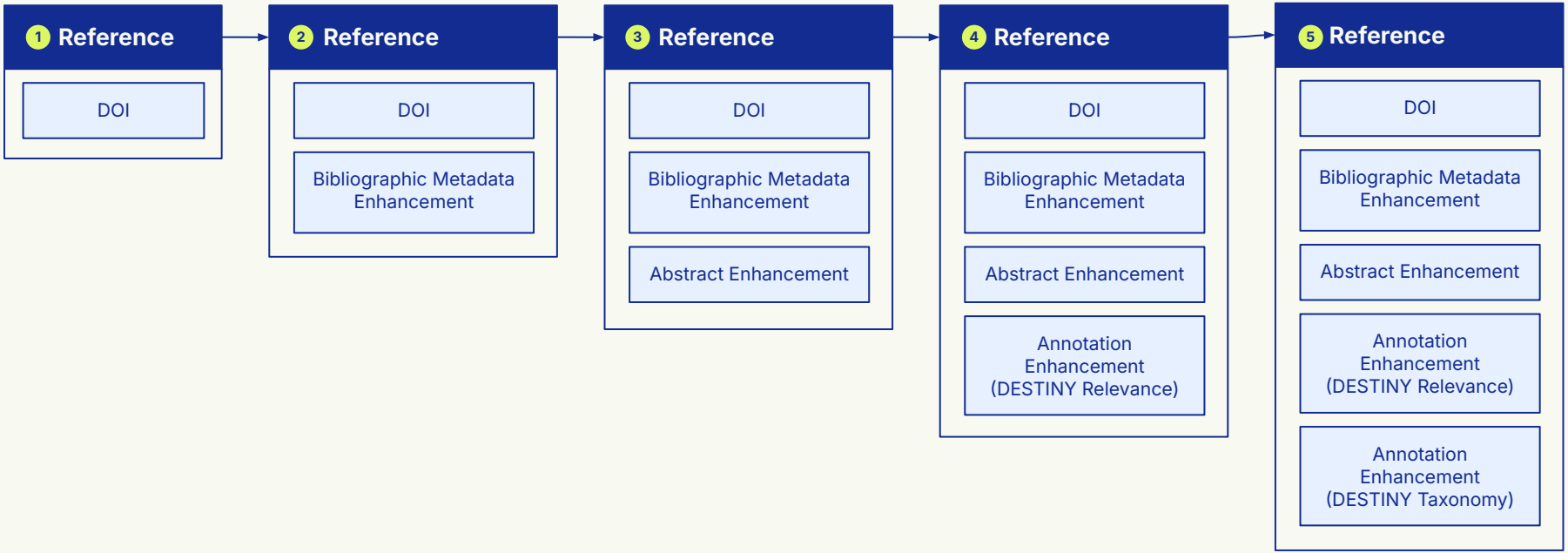
Evidence Products



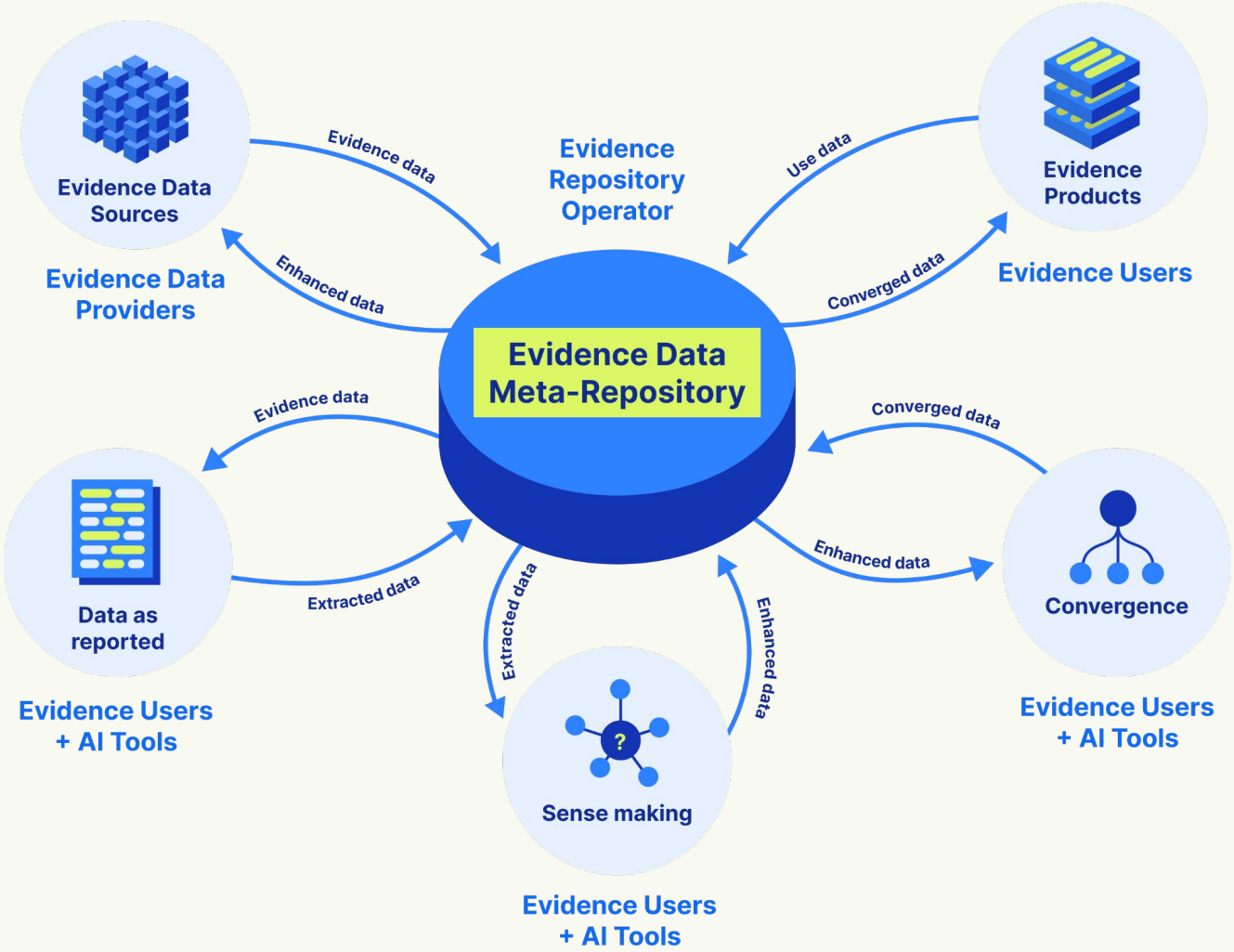

Evidence Data Sources
Evidence Data Providers
Evidence data
Enhanced data
Evidence Repository Operator
Use data
Converged data
Evidence Products
Evidence Users
Evidence Data Meta-Repository
Evidence data
Extracted data
Data as reported
Evidence Users + AI Tools
Extracted data
Enhanced data
Sense making
Evidence Users + AI Tools
Converged data
Enhanced data
Convergence
Evidence Users + AI Tools

Compounding Enhancements

DESTINY Example



We start with only a DOI (❶). We use that DOI to get bibliographic metadata from OpenAlex (❷), but they don't have an abstract for that record so we go to another source for the abstract (❸). Using the title and abstract we now have, we run a relevance classifier (❹). If the record is relevant, we then run a taxonomy classifier (❺).


Evidence Data Sources
Evidence Data Providers
Evidence data
Enhanced data
Evidence Repository Operator
Use data
Converged data
Evidence Products
Evidence Users
Evidence Data Meta-Repository
Evidence data
Extracted data
Data as reported
Evidence Users + AI Tools
Extracted data
Enhanced data
Sense making
Evidence Users + AI Tools
Converged data
Enhanced data
Convergence
Evidence Users + AI Tools