Advancing data sharing and reuse in evidence synthesis

Despite the global growth of evidence synthesis, nearly all evidence projects start from scratch, resulting in extreme levels of duplication and wasted effort. The systems that aim to counter this by enabling the work of one group to contribute to the work of others remain fragmented, under-resourced, and often inequitable. Most data extracted during synthesis projects is never shared, or if shared, is locked in organisation-specific repositories or poorly curated, generic data-sharing platforms. These data are difficult to find and rarely designed for reuse, especially for the questions of most importance in non-English or resource-constrained contexts. Working Group 2 (WG2) has developed a vision and design architecture for a shared infrastructure that enables more efficient, equitable, and trustworthy data reuse across the evidence ecosystem, with the aim of complementing advances in AI to create the conditions for the ESIC vision of comprehensive useful evidence synthesis across all major social questions for the benefit of all.

The proposed solutions are grounded in current needs and existing models, but recognise that implementation will require substantial collaboration, phased development, and global coordination. If successful, these approaches will transform evidence synthesis from a one-time product into a reusable, distributed public good.

The landscape: who needs data and why?

Synthesis data is essential for decision-makers, researchers, evidence teams, and intermediaries across sectors who need timely access to structured, reliable, and contextually relevant evidence. However, current systems are inefficient and duplicative. Many users struggle to locate data, assess its quality, or know whether it can be reused ethically and legally. These challenges are particularly acute in the most resource-constrained settings and for non-English users, who face additional barriers to access, participation, and recognition. Meeting this demand requires systems that enable discoverability, multilingualism, cross-platform interoperability, optimisation for AI systems, and ethical reuse at scale.

Capability gaps and maturity: where are we now?

While many evidence repositories and tools exist, they operate in silos, lack shared standards, and exclude much of the evidence produced in low-resource or non-English settings, limiting trust across the system. Open science approaches are essential for data reuse at scale, however, only a minority of the world's research output is free and open to use by people and automated systems. The lack of shared data standards also makes it difficult for AI systems to achieve the high accuracy required for evidence synthesis, and few repositories provide application programming interfaces (APIs) for machine processing.

Key issues: what's holding us back?

Key barriers include the absence of interoperable formats, fragmented metadata practices, a lack of incentives or funding to share, and unclear licensing and quality standards. Even where technical solutions exist, their uptake is hindered by short-term funding, institutional inertia, and misaligned incentives. Many users are unable to determine whether shared data is up to date, understand the quality of the data, or the processes used to create it. Without intentional coordination and standardisation, the risk is a proliferation of uncoordinated repositories that reduces the value and impact of the included data, leading to a poor experience and confusion for evidence users as they navigate between repositories. Without concerted effort, new systems risk reinforcing existing biases and perpetuating existing inequities. To be effective and equitable, shared infrastructure must be carefully coordinated, user-informed, and sensitive to regional and sectoral variations.

Equity gaps persist in access, participation, and recognition, especially for individuals and organisations in non-English and resource-constrained settings, as well as for underrepresented content such as grey literature or program evaluations.

The ESIC planning process is funded by the Wellcome Trust.

Solutions for progress: what can we do next?

WG2 recommend five complementary solutions, underpinned by a cross-cutting incentive and funding model:

2.1 Federated repository of synthesis data: A connected system of interoperable repositories that enable discovery, access, and reuse of structured synthesis data while respecting varying control requirements and legal constraints. This would result in reduced wasteful duplication of synthesis production, a reduction in time, effort, and resources required to find and produce contextually relevant evidence, and an increase in access to evidence from diverse geographical regions (WG3 3.3).

2.2 Interoperable data standards: Common data formats and record structures for evidence synthesis inputs and outputs to enable automation, platform interoperability, and data reuse across systems. This would facilitate integration of evidence across systems and lower barriers in accessing evidence curated by geographically diverse institutions (WG4 4.2).

2.3 Metadata standards to facilitate data identification and discoverability: Multilingual, modular metadata schemas and taxonomies that support and use AI-assisted tagging to improve evidence discoverability, track provenance and improve visibility of underrepresented evidence. This would enable the incorporation of multiple forms of evidence, increase access to non-English language evidence, facilitate the process of finding evidence, and increase trust and transparency for evidence users (WG4 4.2).

2.4 Open access standards for equitable data sharing and reuse: Open access policies and standards for licensing, attribution, and access rights that promote ethical reuse while accommodating institutional, legal, and policy constraints across settings. This would increase the re-use of evidence from around the world and strengthen transparency and accountability (WG4 4.2).

2.5 Quality assurance of data systems: Efficient and scalable validation and quality assurance of data and data curation practices, based on shared standards, to build the trust required for the reuse of shared data. This would enhance equity monitoring, improve trust and reliability of shared data, and build systemic accountability.

Cross-cutting strategy: sustainable funding and incentive alignment
A shared funding and incentive architecture is essential to implement all five strategies. It must support long-term infrastructure maintenance, enable institutional participation, and align recognition and resourcing models with collaborative, reuse-oriented behaviours.

These strategies are designed for phased implementation, with early development of open science partnerships, metadata and record formats, and data repository and quality assurance models, followed by the longer-term rollout of the global repository system.

Outcomes: What is likely to change?

The five solutions proposed by WG2 will enable a substantial increase in the efficiency of producing evidence syntheses by effectively reusing data. A more connected, equitable, and efficient global infrastructure for data sharing and reuse in evidence synthesis will be created. A connected system of living evidence repositories will enhance discoverability and access to synthesis data across regions and platforms. Standardised record formats and metadata schemas will enhance evidence discoverability, particularly of diverse types of evidence, including grey literature and non-English content and facilitate contributions by AI systems. Open-access policies and standards, along with scalable quality assurance mechanisms, will enhance trust in shared data and ensure the ethical, transparent, and reliable reuse of data.

These outcomes will support broad and diverse global participation and work closely with AI systems to enable the scale, reach, trust and efficiency required to improve responsiveness to decision-maker needs, foster data and AI innovation, and work closely with AI systems to enable the scale, reach, trust and efficiency required to achieve the ESIC vision.

The ESIC planning process is funded by the Wellcome Trust.