

We urgently need an **open data system** (ODS) to streamline how we share and reuse evidence-synthesis data.

This data system will underpin access to trustworthy, actionable **insights** on all the big questions (or problems) of our time – from policy-scale, AI-enabled, living evidence syntheses – including how the insights vary by groups and contexts. Imagine every study (or evaluation) – from anywhere in the world – addressing the same question... all in one place, each with key data (e.g., about effect sizes) extracted and quality assessed, and with an overall synthesis about what was learned and with what equity considerations and confidence or caveats.

This data system will provide access to the synthesis- and AI-ready **data** underpinning these insights in case people want to contextualize the insights for a different group or context, or people want to adjust how studies were identified and selected or how evidence was extracted, appraised and synthesized. Imagine that we extract data once and conduct quality (e.g., risk-of-bias) assessments once, and use the data many times (and thereby avoid the extraordinary duplication happening in evidence synthesis every day).

We hope to come to **agreement** on and to begin implementing an agreement on an open data system sometime between April and June. We need the core elements of a workable system in place by September, by which time 11 [sectoral hubs](#) and two [regional hubs](#) (and one sub-regional hub) will be dependent on it. We then need to continue to add elements quickly and in an initial priority order to be set by the sectoral and regional hubs. In the absence of an open data system, we will continue to waste time and money and will (likely irrevocably) cede evidence synthesis to tech solutions that do not get the balance right between rigour and speed.

We acknowledge the fast-changing context in which these discussions are happening, and the adoption challenges this presents. These changes include AI, of course, but also significant adjustments to funding flows in international assistance, greater focus on 'localization,' new methods and concepts like 'living synthesis,' challenges to journal revenue from new publishing models, etc. In such a context, an open data system that lowers infrastructure and back-end processing costs over the medium to long-term is even more important, although the shorter-term adoption issues are likely more complex.

Four key steps are important to understand in [how we got here](#):

- a working group – operating as part of the broader ESIC planning process from January through June 2025 – articulated many of our collective aspirations for the sharing and reusing of evidence-synthesis data
- the Funder executive announced on 21 September 2025 that an open data system would be one of the five year-1 foundational investments
- the Funder executive shared on 11 October 2025 a memo outlining our proposed approach to technology-company engagement
- we convened a series of three touchpoints about the proposed open data system in October and November 2025.

The key documents prepared for or emerging from these steps can be found on the '[how we got here](#)' webpage, with the most important of these documents including:

- memo from the Funders executive – shared on 12 December 2025 – that describes the [next steps](#), including a decision by the Funders executive to support Global South co-leadership to deliver on the 'remaining' open data system capabilities (with funding from the Wellcome Trust) and to honour the rigour of the past selection processes and support co-leadership by the team **co-led** by James Thomas
- [sample use cases](#)
- existing and proposed [capabilities](#) for the open data system.

In keeping with the decisions communicated in the [Funders executive memo](#) mentioned above, we are proposing to build on the source code developed by a consortium of organisations and projects associated with the Wellcome Trust's 'DESTINY' investment, as well as potentially by others [documented in the 'Data sharing and reusing' working group reports](#) and to be more fulsomely documented [by the proposed consolidating entity](#). Building on what already exists and what is being planned will allow us to go much farther, faster, as well as create lasting legacies for these contributors. We are planning discussions with these contributors about how to get the balance right in how representation – particularly in ownership and governance (what we call 'straw dogs' 2 and 3 later in this report) – is distributed among those who played key roles in getting the source code to where it is today and the several (or many) groups who will be developing source code in future and the many groups that will be contributing to (and using) data in future. We also want to ensure we continue to get the balance right in how representation helps to make good on our commitment to shift power to the Global South.

Four key processes are important to understand about our [planning and engagement](#) process between January and June of this year:

- a [seven-member planning group](#) worked together – and with facilitation and business/legal input provided by the Wellcome Trust – from mid-January to mid-March 2026 to draft a proposal for how the open data system will work – **this 'straw dogs' report** – which they are now sharing with our 'community of communities'
- the sharing of this 'straw dogs' report **kicked** off an engagement process that included a community webinar on Monday 30 March and an offer for bilateral and small-group discussions throughout April **and May** to answer additional questions, respond to additional feedback, or talk through potential implications for particular groups
- the planning group revised the report based on the very helpful feedback it received, including some very detailed and constructive feedback from several individuals, and now is posting this revised version in two formats (one with the additions/edits highlighted and a second without the highlights)
- in parallel to this cross-sectoral planning on an open data system, a minimum-viable product of a synthesis-ready evidence repository is being developed for the education sector – **for launch in early May** – and feeding lessons learned back into the cross-sectoral process.

Additional details can be found on the '[planning and engagement](#)' webpage.

This report contains **ten 'straw dogs'**:

- [Straw dog 1: Consolidating entity and licensing](#)
- [Straw dog 2: Ownership \(including PACE acting on behalf of the 'community of communities'\)](#)
- [Straw dog 3: Governance \(including an ODS leadership group\)](#)
- [Straw dog 4: Data sharing and reusing \(including benefits, protections and quality assurance\)](#)
- [Straw dog 5: Revenue generation \(including tech company engagement\)](#)
- [Straw dog 6: Revenue sharing \(including additional financial modeling and methods work needed\)](#)
- [Straw dog 7: Incentives for participation \(including benefits, one-time transition costs, and lowering risks\)](#)
- [Straw dog 8: Next steps for the key evidence-synthesis producing networks \(including coaching, business models, and change-management support\)](#)
- [Straw dog 9: Next steps for other entities that can contribute to and use the ODS](#)
- [Straw dog 10: Development-related milestones.](#)

The report also contains a [questions and answers \(Q&A\)](#) section.

You can use the above hyperlinks to navigate to any of the straw dogs.

A '**straw dog**' is typically taken to mean a brainstormed simple draft proposal intended to generate discussion of its disadvantages and to spur the generation of new and better proposals. We **framed the initial request for community feedback more positively: what would make the proposal even better?**

We were delighted by the robust engagement with and feedback on the straw dog report, which both improved the report but importantly also affirmed that the general direction of travel is right and that we are ready to consider the 'straw dogs' as forming the basis on which more detailed planning and execution can now progress.

Please **continue to** send your feedback to info@evidencesic.org.

Two related considerations to keep in mind when reviewing the straw dogs are:

- will they help us pursue the ESIC open data system (ODS) while building the **trust** necessary to make it sustainable
- will they help us transition from research project(s) led by a few to a co-owned, distributed ODS that people trust (i.e., feel invested in, contribute insights to & engage in testing of (e.g., for/in different contexts & sectors), & add data to and draw data from).

We documented **five assumptions** underpinning our work:

- partners are motivated by public benefit, not profit (e.g., producers want their work enriched, found, used and appropriately acknowledged)
- partners are operating under a collective-impact model, while still needing to fulfill organizational and professional mandates
- partners require sustained funding flows to support their contributions to the ODS and to ESIC more generally, while recognizing that there are efficiencies to be gained by participating in ESIC's broader efforts to make evidence synthesis radically more timely, relevant and affordable
- partners are committed to open science, which in this context means primarily open-source technology and open data, while still needing to respect intellectual property rights
- partners are committed to the FAIR principles (findable, accessible, interoperable and reusable), as well as (where applicable) the CARE principles for Indigenous data governance (collective benefit, authority to control, responsibility, and ethics).

Straw dog 1: Consolidating entity and licensing

ESIC requires a consolidating entity for all ODS-related:

- data and source code licensing (i.e., which licenses are used for which purposes) and contributor license agreements, as well as oversight on the provenance tracking and visibility and access controls that ensure adherence to agreed licensing (i.e., which protocols are in place for which categories of contributors and users)
- revenue generation and sharing terms and mechanisms (e.g., establishing under what terms AI companies would be allowed to use the data)
- procurements of new capabilities, and more generally priorities for the development of new tech solutions
- policies (e.g., supporting adherence to relevant laws and risk-management protocols), **procedures, protocols and rules that govern people's interactions with the ODS, as well as a community-owned model for developing and managing these policies, procedures, protocols and rules**
- access to good advice and support (e.g., defending – within the limits of our resources – the source code and data if someone uses them outside of the terms of the license, pushing back if a group asserts a right that cannot be defended).

The next two straw dogs – ownership and governance – pertain to this consolidating entity.

Regarding the first bullet point above, ESIC requires appropriate:

- licensing of **data**, with the proposal being to use the same data licensing already in use
 - data from OpenAlex (and arguably from any sources that can currently be retrieved with Google searches)
→ [CC0 1.0](https://creativecommons.org/licenses/by/4.0/) (which is very permissive)

- data from publishers like Elsevier and Wiley, which have their own data licenses
- data from synthesis producers like Cochrane authors, evaluation report producers like UNEG members, etc., who also have their own data licenses
- data created through (parts of) the ODS (e.g., record enhancement) → [CC0 1.0](#) (again very permissive)
- licensing of **source code**, with the proposal being to use the same source-code licensing as being used by DESTINY
 - repository → [GNU Affero General Public License v3.0](#) (which allows people to reuse the source code as long as they respect the same level of 'open source' commitment)
 - robots that perform tasks such as assigning meta data → [Apache License 2](#) (which allows people to reuse the source code without committing to make it 'open source,' which we note creates a market advantage for those with current money-making ventures and which reflects a legacy decision by DESTINY that we may want to return to – to better understand and perhaps re-consider – as a community)
- **contributor** license agreement (CLA), which defines the conditions under which contributions are made and can help with issues such as:
 - re-licensing
 - knowing whose responsibility it is to defend the source code and data if someone uses them outside of the terms of the license.

We note the importance of separating: 1) source code for the open data system (which we originally called the 'data platform,' includes both the repository and the robots, and must be 'open source'); 2) data; 3) AI tools that leverage the platform and data. While our focus is 1 and 2, tech companies need to know the terms of engagement for any AI tools with 1 and 2. We discuss tech company engagement as part of straw dog 5.

We anticipate that ownership of third-party applications that interact with the ODS (e.g., platforms contributing data to and drawing on data from the ODS, robots that are adapted to support workflows specific to a given sectoral or regional hub's needs) will remain with the developers of these third-party applications, albeit with ODS source code adaptations following the appropriate ODS source-code licensing provisions.

Straw dog 2: Ownership (including PACE acting on behalf of the 'community of communities')

In the near term, and in keeping with the commitment to Global South co-leadership expressed in the Funder executive memo shared with the community on 12 December 2025, the Pan-African Collective for Evidence (PACE) will act – on behalf of ESIC's community of communities – as the consolidating entity for the ODS. PACE is a non-profit organization and it will use any budgetary surpluses to re-invest in and hence advance the ODS (or the broader aims of ESIC). PACE will be asked to adopt an ongoing 'test and learn' approach to its ODS role and the Wellcome Trust, as the initial funder for this ODS role, will use a milestones-based approach (or 'start-up mentality') to funding. PACE will be asked to lead an agile, performance-based management of multiple streams of tech development, which we return to in straw dog 10, and it will do so in partnership with the EPPI-Centre, which has independently secured significant funding to develop what is now becoming the ODS, and with a broader Global South 'tech alliance' that it will help to create. This means that PACE will need to respond if we encounter challenges with access to or usefulness of tech solutions (including their applicability beyond domains for which they were originally developed, such as for randomized trials), need to accommodate special scenarios (a government wanting a gated approach to data sharing pending trust building; an opportunity to experiment with a different approach to revenue generation and sharing), encounter new forms of risk that need to be managed, and need to bring delivery 'failures' to the Communities council for discussion, as we increasingly build the trust needed to ensure mutual accountability.

PACE will over time explore **at least two alternative future scenarios**, both of which were considered to introduce complexities that would needlessly delay progress with the ODS if they were pursued right now.

One alternative future scenario is the creation of a **new non-profit company** (NPC) in South Africa, and also voluntarily registered as a non-profit organization (NPO). Some or all the board members of this new entity could be appointed by one or more of the ESIC Steering group, Communities council, Funders executive (with the advice of the ODS leadership group) to ensure strong alignment to existing 'governance' groups. In practice, this creates an NPC that is directly governed by ESIC's structures.

A second alternative future scenario is the creation of a **new non-profit company – with a different organizational form** – that builds on much of the above (e.g., South African registration) and also allows for:

- leveraging the governance and financial-management track records of the contributing entities, which would almost certainly help with meeting funders' due-diligence requirements
- formalizing and broadening a sense of collective ownership via representatives of different categories of interest holders and particularly ones that have robust governance and financial-management track records, such as, in the nearer term:
 - global public good producer networks that have historically been market shapers for, and implementation leaders with, digital solutions (e.g., Cochrane Collaboration and Campbell Collaboration)
 - evidence-intermediary networks that will be key contributors to and users of the ODS, particularly their supporting secretariats (e.g., Instituto Veredas for Hub LAC)
 - hybrids of the above and tech solutions developers (e.g., PACE and the EPPI-Centre).

Under this scenario, PACE will need to **develop a strategy to identify and invite other groups to join such a company** over time, such as when there are designated regional hubs and if/when they meet entry conditions around governance and financial-management track records, while **continuing to ensure that additions ensure a continued transition in power to the Global South. PACE will need to engage the ODS leadership group and ESIC's broader 'governance' groups in such expansions of ownership.**

The Wellcome Trust has drafted **principles** that underpin PACE's proposed role in acting – on behalf of ESIC's community of communities – as the consolidating entity for the ODS and that reflect the commitment of **several key organizations** to work together in supporting this arrangement and in exploring alternative future scenarios. Once agreement is reached on the principles, the Wellcome Trust will formalize these principles through a **memorandum of understanding** between itself and **an initial set of key** partners, which will serve as a pre-condition for funding PACE to play this role.

Straw dog 3: Governance (including an ODS leadership group)

ESIC requires the creation of (or transition of the ODS planning group to become) an **ODS leadership group** (or coordination or management group) that will discuss emerging issues and make recommendations or craft proposals to bring to the appropriate ESIC 'governance' group for decision (or advice). The ODS leadership group will:

- develop overarching policy (e.g., license selection)
- discuss and approve strategically important recommendations from PACE about operational issues (e.g., priorities for coding)
- maintain a running list of issues to work through (e.g., documenting licenses currently in use, documenting technical features that might affect interoperability, and documenting policies and practices that might require a change-management approach)
- act as the 'Intellectual Property Management Group' called for in the draft principles discussed in the preceding paragraph, which includes providing updates to the ODS contributors and receiving and acting on feedback from these contributors.

Straw dog 1 – the roles of the coordinating entity – provides examples of the types of decisions to be made by the ODS leadership group. Straw dog 2 – particularly the ‘test and learn’ approach to these roles, the milestones-based approach (or ‘start-up mentality’) to funding, and the need for an agile, performance-based management of multiple streams of tech development – provides examples of the types of ethos to be maintained by the ODS leadership group. Both straw dogs could help to inform the **recruitment** of (new) people who give the ODS leadership group the right breadth of expertise and skillsets.

The ODS leadership group will honour ESIC's commitment to include Global South, citizen and early-career leaders in all advisory and decision-making bodies.

Straw dog 4: Data sharing and reusing (including benefits, protections and quality assurance)

The ODS can provide five types of **benefits to contributors** (and users), all of which related to their work being enriched, found, used and appropriately acknowledged:

- access to synthesis data so that it can be
 - incorporated into workflows (including via APIs), such as diagnostics and project analysis by multilateral development banks, evaluation planning (e.g., theories of change) and recommendations development by evaluation offices, design of primary research and evidence syntheses, etc.
 - re-analyzed in preparing new syntheses (e.g., contextualize the insights for a different group or context, or adjust how studies and evaluations were identified and selected or how evidence was extracted, appraised and synthesized)
 - used to create ‘best buys’ lists and other decision-ready formats
 - used to create feedback loops for next-generation research and evaluation (improving relevance and rigour, and reducing waste)
 - used to train one’s own small language models
- access to enhancements to contributors’ own reports and meta-data (e.g., structured abstracts, sectoral taxonomy mapping, context mapping, and many others) that can then be added to contributors’ own repositories, making them even more valuable to contributors’ target audiences
- access to more users of contributors’ work (e.g., evaluations being incorporated in evidence syntheses, evidence syntheses being used by multilaterals and government policymakers)
- access to usage data that complements contributors’ existing performance metrics (e.g., ODS-usage data to complement contributing repository’s usage data; data-usage-as-impact metrics to complement peer-reviewed journal articles, which will require advocacy to shift academic incentives to incorporate a data-citation model)
- access to good advice and support (e.g., defending – within the limits of our resources – the source code and data if someone uses them outside of the terms of the license, pushing back if a group asserts a right that cannot be defended).

The ODS can provide three types of **protections** that contributors could rely on (versus the status quo or risk of AI companies mining contributors’ data without constraints):

- seat, or representative with a seat, at the governance table(s) where decisions will be made
- opportunity to make explicit decisions about
 - ✓ allowing contributors’ data to be indexed and how any caveats are noted (e.g., not all reports are included)
 - ✓ data licensing conditions
 - ✓ constraints on or suggestions for how data extracted from contributors’ work are presented in a new context
 - ✓ how ethical issues are addressed (e.g., providing data that reduces the risk of increasing discrimination; linking to a platform like Vivli that has worked through the ethical use of individual participant data)
- capability to operationalize these decisions by, for example, distinguishing public data from restricted data, with the latter requiring authentication (to ensure adherence to data licensing conditions)
 - ✓ control over the terms under which contributors’ data are shared with AI companies

- ✓ opportunity to have contributors' staff provided with authentication.

The ODS can assure **quality** in seven ways so that people have the confidence to use it:

- require adherence to harmonized, transformed methods for policy-scale, AI-enabled living evidence syntheses, which includes:
 - leveraging accepted guidance (e.g., Cochrane handbook for systematic reviews of interventions; UNEG guidance for evaluation synthesis)
 - filling gaps in methods and standards for primary studies (e.g., beyond randomized controlled trials)
 - developing systematic and transparent approaches to identifying, quantifying (where possible) and communicating uncertainty
- require adherence to agreed taxonomies
- require some form of 'accreditation' of the individual or team adding data (e.g., Cochrane reviewer; METIUS team member; evidence intermediary **that is part of a Global South network like the Africa Evidence Network and Hub LAC and** using well-documented, robust processes, such as adhering to many steps covered by AMSTAR criteria; **WHO** synthesis/guideline team member)
- require some form of peer review of the submitted data (e.g., Campbell peer review)
- allow for users to document critiques and/or add a 'second' set of data that they feel is better
- have an adjudication mechanism – administered by the sectoral and/or regional hubs and inspired by the Wikipedia model of decentralized policing of 'rules' – to assess critiques and multiple competing data sets
- conduct random audits (by humans, AI or a combination) and use the insights to prioritize 'clean ups.'

As with any of the lists of potential contributors and beneficiaries in this report, the organizations listed above are provided for illustrative purposes and many more can and will be added.

Straw dog 5: Revenue generation (including tech company engagement)

ESIC requires a **sustainable mix of revenue** that can support both ongoing system development and the data contributions that give the system its unique power, which will likely require a blend of:

- funding from philanthropies, research funders, national funders and international-assistance funders that are committed to a **global public good** that is efficiently maintained and that uses a collective-impact approach
- revenue from multilateral development banks and other multilateral entities, governments, research funders and others that want and can afford to pay for **APIs** that enable direct connections to their workflows
- engineer secondments, collaborative research projects and direct funding from tech companies that are interested in one or both of **global public good** investments and targeted investments that create new **products** (and hence intellectual property and branding opportunities for ESIC as well)
- similar secondments, projects and funding from a **'public AI' model** if one emerges among the 'middle powers.'

This blend is likely to shift over time, and our community of communities is fortunate that fairly significant funding for the open data system as a global public good – point 1 in the list above – has been secured for the next few years.

Because the promise and perils of **tech company engagement** – point 3 in the list above – are both so large and so time-sensitive, ESIC requires community agreement on **nine points**:

- 1) **engagement goals**
- 2) **an 'AI for policy'** problem statement
- 3) counterfactuals to not engaging, and the alternatives for engaging
- 4) safeguards
- 5) partnership parameters
- 6) reasons why long-term partnership is key
- 7) clarity about the synergies in what tech companies bring to the table

- 8) people and organizations from whom we can seek advice
- 9) next steps.

Such agreement would allow us to pursue exploratory meetings with relevant tech companies both now and in future.

First, our proposed **goals** for tech company engagement are:

- to shift tech companies' trajectories towards greater rigour – as our 'community of communities' would define it – in the 'AI for policy' tools that they are developing
- to accelerate our own thinking and our delivery of tools by accessing engineering, frontier AI models, evaluation and other capabilities from tech companies and securing their sponsorship of our applications to their charitable arms
- to build the trust needed to explore the possibility of long-term partnerships that would truly shift tech companies' and our trajectories in the 'AI for policy' space through ongoing demand-side engagement, AI and research methods transformation, and human-in-the-loop (HITL) processes.

Second, the proposed framing of the **problem statement** is that AI cannot yet be safely and responsibly, and with the right balance of human-in-the-loop processes:

- be used in each step in the evidence-synthesis process or identify, quantify (where possible) and communicate uncertainty
- be used to conduct the local applicability assessments needed to determine which evidence (created in particular contexts and with a focus on particular groups) likely applies in a given context and for a particular group
- mix and match different forms of local evidence (e.g., data analytics, evaluation, behavioural science) and global evidence (i.e., what we are focused on with ESIC) in each of the four typical steps in a policy analysis
- combine all the above to answer bigger goal-related questions or cross-sectoral questions in a contextually relevant way (e.g., what are the 'best buys' in education for a country like ours? the best ways in any sector to improve under-5 mortality in a country like ours?)
- combine all the above to support its real-time use in **evidence interrogation that happens every day in policy work and in** policy-engagement processes like the stakeholder dialogues and citizen panels run by many of our evidence intermediary partners, **as well as** the guideline-development processes run by profession-focused colleagues.

The first of the five elements of problem statement is where much of the evidence-synthesis community is focused, however, tech companies are interested in AI for policy (or AI for decision-making), not just AI for evidence synthesis. A more fulsome **exploration** of all five elements of the problem statement **was undertaken** at an ESIC meeting on 30 and 31 March, **and this thinking is now informing our current approach.**

Third, it is important to note the **counterfactuals** to ESIC – armed with the safeguards, partnership parameters and long-term orientation – engaging with (largely) unregulated US tech companies:

- we don't engage with tech companies and they work with management-consulting firms and others to develop solutions that become the 'new normal' but do not incorporate the types of systematic and transparent approaches (existing and future) that many in our 'community of communities' value
- we don't engage with tech companies and they harvest and use our data in unexpected ways
- we don't engage with tech companies and within 6-12 months, many organizations in our community lose opportunities to AI tools
- we don't engage with tech companies and we lose the prospect of a sustainable funding stream for ESIC.

The **alternatives** to engaging with tech companies each have disadvantages:

- **public / sovereign AI institutions** are not moving fast enough

- Chinese tech companies cannot be supported by the Wellcome Trust given it cannot flow funding to mainland China.

Fourth, ESIC needs to stipulate the **safeguards** that are critical to us:

- we focus on creating global public goods and extend the same offer of partnership on tech-solutions development to any relevant tech companies
- we agree on partnership parameters (see below)
- we continue to control access to the ODS **via licensing provisions (and authentication for any restricted data)**
- our tech company partners commit to applicable aspects of a safety pledge, such as they don't collect data on the questions being asked by government policymakers and other users.

Fifth, ESIC needs to articulate the **partnership parameters** that are critical to us:

- rigour, not just speed
- communication of what we know and with what confidence and caveats (includes communication of uncertainty), not just 'black and white' answers
- societal impact, with commercial impact a focus only for sustainability
- Global South leadership opportunities and an equity focus
- non-exclusivity and no unfair advantage, with partnerships offered – on an ongoing basis – to any viable, comparable 'competitor' (e.g., if we are speaking with Google and Anthropic, we should also speak with Microsoft)
- safeguards (see above)
- longer-term partnership (see below), which could include:
 - some IP co-ownership (and, if possible, co-branding) of any product that is developed
 - pathway to funding for sustaining future data flows to the 'walled garden' of the ODS.

Sixth, ESIC needs to make the case why **long-term partnership** is key:

- ongoing demand-side engagement – among decision-makers and evidence intermediaries – is key for stimulating demand and building capacity among users, for product-related learning and improvement, and for 'kitemarking' products that are 'good enough for prime time' and then supporting their roll-out at scale
- ongoing methods transformation is key for ongoing refinements
- ongoing human-in-the-loop processes are key for quality control of data flows.

What ESIC wants to avoid is a 'dine and dash' approach to a partnership where tech companies access data and domain expertise once, a 'story' they can tell, and professional development for their engineers, and then move on. At the same time, ESIC needs to be prepared with an exit strategy that balances the need for building and sustaining trust on the one hand with the need to adapt if we find ourselves going down a path that doesn't look promising or we find our partner becoming comfortable with only getting 70-80% of the way to where we think we need to be.

Seventh, ESIC needs to appreciate the synergies in **what tech companies bring to the table**:

- cloud computing platform (ideally with servers in Europe or another region/country with strict regulation and low-carbon footprint)
- engineering to accelerate the delivery of key functions
- AI research
 - e.g., teaching models to do synthesis properly (as opposed to text summarization as they currently offer)
 - e.g., limiting or digitally certifying AI's creation or manipulation of the data being ingested
- contribution to sustainable funding, including via data access
- 'value' signal to funders and other partners
- 'way in' to their products, which would dramatically increase access to and use.

Eighth, ESIC will need a 'short list' of **people and organizations from whom we can seek advice**, and these could fall into three 'buckets':

- **technical people** who could provide us with lessons learned from bringing together multiple technical teams to work on a common goal
- **tech company/AI-savvy people** who could provide us with the information we need to engage in a more balanced 'negotiation' (e.g., what is the true information value of our different data 'asset classes'), advise us (e.g., on partnership parameters), and/or act as a 'red team' (in simulations to challenge us)
 - list available upon request
- **organizations that could advise us**
 - AI for Public Good: <https://aiforgood.itu.int/about-us/>
 - Digital Public Goods Alliance: <https://www.digitalpublicgoods.net/digital-public-goods>
 - First International Data Union: <https://firstdataunion.org/>
 - Open Knowledge Foundation: <https://okfn.org/en/>
 - Principles of Open Scholarly Infrastructure: <https://openscholarlyinfrastructure.org>, particularly the ideas of using time-limited funds used for time-limited activities and of generating revenue from services, not data
 - Renaissance Philanthropy (which has been involved in support public goods like biobanks): <https://www.renaissancephilanthropy.org/>
 - RESET Digital for Good: <https://en.reset.org/>.

Ninth, ESIC will now pursue a set of **next steps**:

- organize face-to-face meetings with tech companies to explore possible joint projects, safeguards and partnership parameters
- explore in parallel opportunities with public / sovereign AI institutions
- convene a tech-savvy advisory group for ad hoc discussions as engagement issues arise or can be anticipated
- add a requirement for an ethical analysis when procuring the living inventory of AI-DESTs (e.g., ethical considerations that users of a tools should be aware of given the company behind it or the specifics of the tool).

While the ODS planning group did not have time to get to it, the group noted that **additional exploratory** work would be helpful on:

- Freemium model (e.g., how it works and its pros and cons)
- OpenAlex (how it is sustainably funded).

We will also need to return to more detailed sustainability discussions once we have a better sense of where tech company engagement takes us and the proposed exploratory work has been completed.

Straw dog 6: Revenue sharing (including additional financial modeling and methods work needed)

ESIC needs to **share revenue proportional to share of data usage**, with a multiplier for the degree of human-in-the-loop contributions and for maintaining up-to-date data using 'living evidence synthesis' approaches.

The consolidating entity for the ODS will need to undertake the **financial modeling** necessary for potential data contributors to understand what this will mean for the scale and sustainability of their future funding, including how the funding will compare to publishing revenues that currently underpin some organizations' business models.

The consolidating entity will also need to develop the **methods** required to measure 'quality' data usage (e.g., downstream in terms of impact and mid-stream in terms of addressing sectoral hub-prioritized questions and foregrounding equity considerations as part of robust processes) and to incorporate this in the revenue-sharing process.

Straw dog 7: Incentives for participation (including benefits, one-time transition costs, and lowering risks)

ESIC needs to continue to find ways to:

- make the ongoing **benefits** of data sharing for rights-holders (e.g., data visibility/usage, revenue) greater than ongoing costs of data sharing for rights-holders (e.g., loss of usage data, loss of revenue from other sources)
- compensate for the **one-time transition costs** (e.g., sharing usage data, paying one-time development costs, and paying for access to legacy data)
- more generally, continue to find ways to **lower the risk** for different groups (e.g., for legacy data holders for whom this is a big change in their core business models, for small groups in the Global South for whom this is their big chance for leadership, and for all of us with AI).

A more fulsome listing of the benefits, as well as the protections, can be found in **straw dog 4**. As noted in that section, ESIC and others have an important **advocacy** role to play in shifting academic incentives to incorporate a data-citation model (so one receives appropriate academic credit when one's data are used, just as one currently receives academic credit when one's peer-reviewed journal articles are cited). ESIC hopes to explore with the Impact Funders Forum (IFF) a jointly organized session on academic incentives at the IFF meeting being planned in Singapore in November 2026.

Straw dog 8: Next steps for the key evidence-synthesis producing networks (including coaching, business models, and change-management support)

ESIC needs to actively support the key evidence-synthesis producing networks to participate fully in the ODS:

- coaching on how to operate a **design or prototyping lab** to test and adapt new approaches alongside existing operations (e.g., member of the ODS leadership group, role in sectoral hubs, role in the methods transformation, piloting of policy-scale AI-enabled living evidence syntheses, role in the living inventory of AI-digital evidence-synthesis tools and related AI experimentation)
- sustainable **business model**
- **change-management support**, which includes support in three domains
 - cultural (e.g., bringing communities of volunteers along on this journey, including seeing evidence synthesis as a step on a pathway to influencing decision-making but rarely something read as a stand-alone document by a policymaker or citizen)
 - operational (e.g., if moving from journal revenue to data-sharing revenue)
 - incentives (e.g., if moving from authors seeking a citable publication to authors also being granted a citable dataset)
- payment of **one-time transition costs**
- payment for legacy synthesis data
- ongoing payment for flows of new synthesis data
- opportunity to enrich existing data
- policies and procedures (e.g., akin to the on-boarding materials that METIUS is developing for sectoral hubs)
- support for broadening the participation of Global South leaders in their work, including in the design or prototyping lab.

We consider these key networks to be Cochrane, Campbell, Collaboration for Environmental Evidence and JBI, while noting concentric circles of groups involved in producing global public goods:

- synthesis groups working independently
- sectoral hubs
- evidence intermediaries that produce both global public goods and contextualized evidence syntheses.

While the ODS planning group did not have time to get to it, the group noted that additional work could involve undertaking an analysis of the different groups in terms of their objectives, incentive structure, and data, staffing and other parameters.

Type (most with both stocks & flows)	Examples (provided for illustrative purposes only; more can & will be added)	Objectives	Incentive structure	Data, staffing & other parameters
Synthesis networks	Cochrane, Campbell, CEE, JBI			e.g., mostly volunteers
Synthesis groups working independently	IRC, JUNO WHO guideline contributors			e.g., mostly unstructured data
Sectoral hubs	DESTINY, METIUS groups			
Regional hub members	AEN and Hub LAC members			
Hosts of databases of studies / evaluations, syntheses (and synthesis protocols), and synthesis-containing HTAs and guidelines	DEP (3ie), DEREc, UNEG, EEF HE, HSE, SSE PROSPERO			
Synthesis tool developers	HEvKA			

Straw dog 9: Next steps for other entities that can contribute to and leverage the ODS

After or alongside initial efforts to actively support the key evidence-synthesis producing networks, ESIC needs to support the hosts of existing databases that can contribute data to the ODS and leverage the ODS to better meet the needs of their target audiences:

- UNEG engagement is in progress (and engagement with other evaluation database owners needs to be geared up)
- 3ie is being engaged via METIUS and the education-sector minimum viable product, noting that their database includes evaluations and syntheses (and engagement with other synthesis database owners needs to be geared up as well)
- PROSPERO engagement is being planned, noting that they represent both a data contributor (as York) and potential platform contributor (as NIHR)
- Vivli – a platform supporting the sharing of individual participant data from randomized controlled trials and observational studies – is underway, and these discussions will also provide a way in to other groups like the Centre for Open Science and the Generalist Repository Ecosystem Initiative
- WHO norms and standards engagement is in progress, noting that they represent a potential platform division of labour – evidence synthesis versus guidelines – and potential data contributor by WHO-commissioned teams.

ESIC will also need to work through:

- how it prioritizes additional existing open-source databases (e.g., by the anticipated value of their data to the ODS)
- how it prioritizes additional groups that want to begin using and writing to the ODS
- whether and how it prioritizes third-party applications that access the ODS and potentially write new data in.

Straw dog 10: Development-related milestones

The sectoral hubs need to know when key open data system (ODS) capabilities will become available so they can establish their own capabilities in ways that optimally leverage the ODS. The regional hubs will also need to know when to expect what ODS capabilities. Three different sets of capabilities need to be considered here:

- [ODS capabilities](#)
- [sectoral hub capabilities](#)

- [regional hub capabilities](#).

Our current estimate of the timeline for delivering on key ODS capabilities is as follows:

- September 2026
 - sector-specific repository capable of ingesting **scholarly publication records** from OpenAlex and – using robots – providing key enhancements (bibliographic metadata, abstract enhancement, and annotation enhancement for sector relevance and for sector taxonomy → this will help sectoral (and regional) hubs with identifying and selecting studies for inclusion in an evidence synthesis (and would be our first opportunity to test contributor license agreements)
 - sector-specific **'question bank'** platform capable of accepting submissions and supporting clarifications → this will help sectoral (and regional) hubs with prioritizing topics and producing evidence syntheses in timely, demand-driven ways
 - first draft of the **living inventory of AI-DESTs** (AI digital evidence-synthesis tools) → this will help sectoral (and regional) hubs with choosing, and reporting on their use of, AI-DESTs
- December 2026
 - sector-specific repository capable of also accepting uploads of **'grey' literature** that is not included in OpenAlex (e.g., government and UN evaluations) → like with the first bullet in September 2026, this will help with identifying and selecting studies for inclusion in an evidence synthesis
 - sector-specific repository capable of supporting PROSPERO-type **synthesis protocol registration** → this will help sectoral hubs with actively supporting our movement towards full coverage of key questions
 - sector-specific repository capable of accepting uploads of **processed synthesis data from existing databases** (i.e., data both about evidence syntheses and about single studies/evaluations) → this will help with extracting data once and conducting quality (e.g., risk-of-bias) assessments once, and using the data many times (and thereby avoid the extraordinary duplication happening in evidence synthesis every day)
- March 2027
 - sector-specific repository capable of automating **full-text retrieval, screening, and data extraction** → like with the first bullet in each of September and December 2026, this will help with identifying and selecting studies for inclusion in an evidence synthesis
 - sector-specific repository capable of accepting uploads of **extracted data from synthesis teams** → like with the third bullet in December 2026, this will help with extracting data once and using the data many times
 - sector-specific repository with a proof of concept for **translation** of one priority language → this will help with accepting and processing data in multiple languages
- June 2027
 - sector-specific repository capable of automating **critical appraisal** → this will help with risk-of-bias and other quality assessments
 - sector-specific repository capable of accepting uploads of **critical-appraisal data** → same as previous bullet
 - sector-specific repository capable of supporting **cross-registration of synthesis protocols** with other platforms → this will help with reducing duplication in evidence syntheses

Key ODS governance-related capabilities are as follows:

- September 2026
 - ODS leadership group and ODS technical-delivery group established, with majority membership from the Global South
 - sub-contracts for development pipeline concluded.
 - legal opinion and roadmap on new non-profit company concluded
- December 2026

- depending on opinion received, pathway identified for creation of a new non-profit company with majority ownership by the Global South
- June 2027
 - launch of new non-profit company.

We will be creating a visual depicting the roles of the ESIC Steering group, ESIC Communities council, the ODS leadership group, and the ODS technical -delivery group.

Three key ODS capabilities are not yet included in this timeline:

- new capabilities required by a transition to policy-scale, AI-enabled living evidence syntheses
- applications that serve up actionable insights for different groups
- translation into all remaining UN languages plus Portuguese

Suggestions for staffing, guidance and additional functionalities will be further discussed:

- dedicated staff available for onboarding and support to community members (with requests prioritized by their centrality to the infrastructure, such as sectoral and regional hubs, and/or by the anticipated value of adding their data to the ODS)
- guidance for existing tech developers about the data standards, taxonomies, licensing terms and other considerations that will facilitate future onboarding into the ODS
- guidance for patient data contributors about 'accreditation' mechanisms
- additional functionalities suggested by some experienced individuals (e.g., author-level private storage with selective sharing to the 'shared corpus').

A more detailed version of these milestones, including the responsible group as well as complementary milestones about governance and about methods transformation, underpins this summary.

Additional work will also be needed in the months ahead:

- plain-language versions of this report for different target audiences, which we will co-create relevant members of our Communities council (e.g., citizen partners, science advisors)
- alternative revenue-generation models (e.g., Freemium model and the model used by OpenAlex), as proposed in straw dog 5
- financial modeling, as proposed in straw dog 6
- risk-mitigation plan.

Questions and answers

1) Where do I find information on...?

We found that many comments and questions that were submitted during our open consultation had been addressed or could be answered in the existing draft of the straw dogs report, which re-affirmed for us that this densely written report can be difficult to navigate. We have adjusted the 'navigation' list at the bottom of page 2 to include additional details about what can be found in what section. We also encourage readers to use the document search function to find specific details on topics that interest them.

2) How will the ODS serve contributors?

Readers will find details about benefits to data sharing, protections for those sharing data, and quality assurance of shared data in straw dog 4 (Data sharing and reusing), which now has the subtitle 'including benefits, protections and quality assurance.' The benefits to data sharing include:

- access to synthesis data so that it can be incorporated into workflows, re-analyzed in preparing new syntheses and in other ways
- access to enhancements to contributors' own reports and meta-data that can then be added to contributors' own repositories, making them even more valuable to contributors' target audiences
- access to more users of contributors' work
- access to usage data that complements contributors' existing performance metrics
- access to good advice and support.

3) How will the ODS serve the sectoral hubs?

The ODS will serve as the data platform supporting the use of AI-enabled digital evidence-synthesis tools – by sectoral hubs – in:

- ongoing evidence surveillance aligned to sector-specific taxonomies
- production of living evidence products, including living evidence maps and policy-scale AI-enabled living evidence syntheses, both by leveraging existing synthesis-ready data (ingested from OpenAlex) and leveraging synthesis-derived data (from existing evidence syntheses)
- priority setting via sectoral question banks
- moving us to full coverage of the most important questions being asked via a centralized PROSPERO-type function.

A recently updated version of the sectoral hub capabilities can be found [here](#).

4) How will the ODS serve the regional hubs?

The ODS will serve as the data platform supporting the use of AI-enabled digital evidence-synthesis tools – by national and local evidence intermediaries – in:

- ongoing evidence surveillance aligned to regional, national and local priorities
- production of timely, demand-driven, contextualized evidence products, both by leveraging existing synthesis-ready data (ingested from OpenAlex) and leveraging synthesis-derived data (from existing evidence syntheses)
- priority setting via regional question banks
- partnering with sectoral hubs on the agile delivery and local packaging of global public goods, such as living evidence maps and policy-scale AI-enabled living evidence syntheses.

A list of the regional hub capabilities can be found [here](#).

We welcome suggestions for additional questions that you would like to see answered. Please send questions to info@evidencesic.org.