ESIC EVIDENCE SYNTHESIS INFRASTRUCTURE COLLABORATIVE

Memo:
**'Open data system' topics**

Last updated:
**7 January 2026**

- **Purpose of this document**
  1) To inform the drafting of a proposal that will be brought to the ESIC 'community of communities' – about how to pursue the ESIC open data system while building the trust necessary to make it sustainable – in late March
  2) To inform the broader transition from research project(s) led by a few to a co-owned, distributed 'open data system' 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]

- **Nine topics to work through**
  1) Deciding what is being governed and owned and/or the focus of revenue generation and sharing
     a) Data (e.g., synthesis-ready data such as enhancements to OpenAlex records, processed synthesis data such as risk-of-bias assessments and actionable insights)
     b) Open data system (which is and will continue to be 'open source')
     c) ESIC (to support the other parts of the infrastructure needed to create, share and use ESIC synthesis data, although this could also be done through a separate revenue-sharing mechanism)
  2) Selecting an ownership model
     a) Not-for-profit company with share ownership among partners
     b) Entity or initiative housed within another organization but governed separately
     c) Others?
  3) Selecting a governance model
     a) Model that operates independently alongside ESIC's Steering group, Communities council, and Funders interest group (FIG) executive, which (respectively) govern, coordinate and fund for collective impact
     b) Model that leverages these three ESIC groups while still providing an independent forum where issues are debated and proposals crafted, with the proposals then brought to the appropriate ESIC group for further advice or for decision
     c) Model that is fully subsumed within these three ESIC groups (i.e., decisions about the open data system are made in the same way as decisions about any other parts of the infrastructure)
  4) Selecting a set of data-sharing options
     a) e.g., no restrictions on reuse
     b) e.g., reuse by all public-benefit organizations (with appropriate citation), but no reuse by for-profit companies
     c) e.g., reuse by all organizations (with appropriate citation), including with reuse by for-profit companies that participate in a revenue-generation model
  5) Selecting a revenue-generation model
     a) Collect revenue for APIs for AI / tech companies
     b) Collect revenue for APIs for organizations seeking direct connections to their workflows and with the ability to pay (e.g., multilateral development banks for their diagnostics and project analysis; research-funding agencies for their peer-review processes and impact-measurement exercises; organization seeking to use the data to create and sell 'best buys' lists)
     c) Others?
  6) Selecting a revenue-sharing model
     a) Payment proportional to share of data contributions, with a multiplier for the degree of human-in-the-loop contributions (and with a multiplier for maintaining up-to-date data using 'living evidence synthesis' approaches)

   b) Payment proportional to share of data usage (like the Spotify model), with a multiplier for the degree of human-in-the-loop contributions (with the same multiplier as 5a)
7) Selecting additional incentives beyond revenue sharing
   a) Academic incentives (e.g., data-citation model; data-usage-as-impact metrics model)
   b) Others?
8) Selecting appropriate next steps for each of the key evidence-synthesis producing networks
   a) Sustainable business model
   b) Change-management support (e.g., if moving from journal revenue to data-sharing revenue)
9) Selecting appropriate next steps for other organizations or initiatives that can contribute to and benefit from this open data system (e.g., those with applications that could connect to the open data system (e.g., protocol registration with PROSPERO, electronic health records with EBMonFHIR)

- **Assumptions**
  1) Partners are motivated by public benefit, not profit
  2) Partners are operating under a collective-impact model, while still needing to fulfill organizational and professional mandates
  3) Partners require sustained funding flows to support their contributions to the open data system 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
  4) 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
  5) Partners are committed to the FAIR principles (findable, accessible, interoperable and reusable), as well as (where applicable) the CARE principles (collective benefit, authority to control, responsibility, and ethics) for Indigenous data governance

- **Running list of possible resources to leverage**
  1) Open Data Institute: https://theodi.org/
  2) World Bank: https://data.worldbank.org/
  3) Country open-data portals, such as the UK's evaluation registry (https://evaluation-registry.cabinetoffice.gov.uk/) and Singapore's data portal (https://data.gov.sg/)
  4) Lawyers who tried to transition UN entities to open data, such as UNICEF