

DATA SHARING AND INTEROPERABILITY OPTIONS

DigiCap explainer on architecture options, trade-offs and practical examples

This explainer compares five common arrangements for sharing operational (beneficiary or non-beneficiary) data between humanitarian organisations. It assesses each model against setup, inclusivity, scalability, reuse, live interaction, data protection, agency control, standards, governance and lifecycle cost. The explainer also illustrates the differences between centralised and federated workflows in deduplication and referral use-cases.

Manual / ad hoc No shared architecture	Centralised One shared platform	Bilateral APIs Pair-by-pair links	Federated One exchange layer	Hybrid Federation + centralised node
--	---	---	--	---

GLOSSARY

Core concepts on interoperability and data sharing

INTEROPERABILITY The ability of independent systems and organisations to exchange data and use it. This requires agreement on technical connections, data structure and meaning, permissions, responsibilities and operating procedures.	DATA SHARING Passing data from one organisation or system to another, manually or automatically. Data can be shared without the systems being fully interoperable.
API – APPLICATION PROGRAMMING INTERFACE A documented digital doorway through which one software system can send a request to another and receive a response automatically.	USE CASE The practical task supported by data exchange, such as deduplication, referral, service mapping, reporting, or payment verification.
LAYERS OF INTEROPERABILITY The areas that must work together: technical connectivity; syntactic data structure; semantic meaning; organisational and legal authority and responsibility; and operational processes for using and maintaining the exchange.	LIVE QUERY An automated check made at the moment information is needed. For example, a registration system can check for relevant prior assistance while a person is being registered.
SYSTEM OF RECORD The authoritative system in which an organisation maintains the original and current version of particular data.	SCHEMA The agreed structure of exchanged data: field names, formats, required fields and permitted values.
DEDUPLICATION Checking whether the same person, household or assistance record may appear more than once. A match identifies a possible overlap for review; it does not automatically mean that assistance is ineligible or inappropriate.	DATA MINIMISATION / MINIMUM DISCLOSURE Collecting and sharing only the information necessary for a defined purpose. In some cases, a yes/no, match/no-match or status response is sufficient.
ACTIVE PARTICIPANT An organisation whose system can initiate requests and respond to requests from other systems through the exchange layer.	PASSIVE PARTICIPANT / SYSTEM A connected system that can receive and respond to authorised requests but does not initiate exchanges itself.

WHICH MODELS ARE ASSESSED?

Manual / ad hoc No shared architecture	Centralised One shared platform	Bilateral APIs Pair-by-pair links	Federated One exchange layer	Hybrid Federation + centralised node
--	---	---	--	---

HOW THE MODELS ARE ASSESSED?

SETUP & TIME-TO-VALUE How quickly can the model become operational, including technical, legal and governance work?	PARTICIPATION & INCLUSIVITY Can organisations without mature systems or API capacity take part?
SCALABILITY What happens when additional organisations join?	REUSE ACROSS USE CASES Can the same infrastructure support referrals, deduplication, reporting, service mapping and other workflows?
LIVE INTERACTION Can systems query, verify, acknowledge and update status in real time or near real time?	DATA PROTECTION & SECURITY How many copies are created, where are they stored, and what is the impact of compromise?
AGENCY CONTROL Does each organisation remain in control of its data and system of record?	STANDARDS & DATA QUALITY How easily can common definitions, schemas and identifiers be maintained?
AUDITABILITY & GOVERNANCE Can exchanges be traced, governed and reviewed consistently?	RESILIENCE, COST & LOCK-IN What are the lifecycle costs, failure points and dependencies on a host or vendor?

ASSESSMENT MATRIX

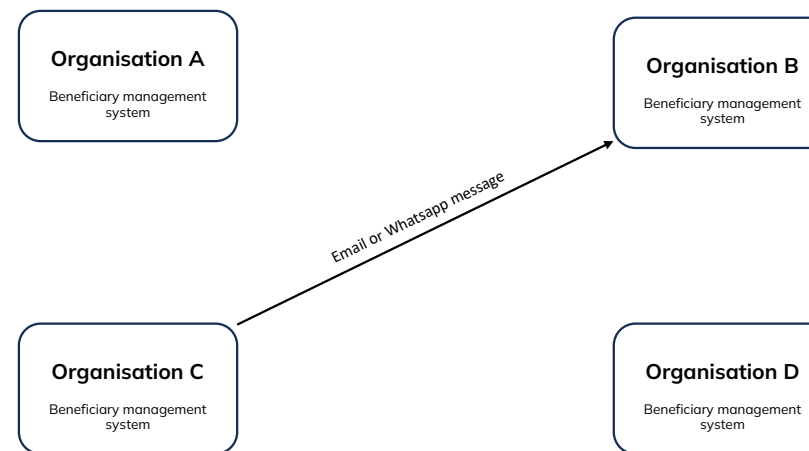
STRONG FIT	CONDITIONAL / MIXED	SIGNIFICANT LIMITATION
-------------------	----------------------------	-------------------------------

OPTION 1: MANUAL / AD HOC EXCHANGE

Manual exchange between otherwise separate agency systems

Each organisation uses their own beneficiary management system. Data, if any, is shared ad hoc via an agreed method such as email, messaging applications, or another manual delivery method.

Best suited for: Temporary, low-volume exchange during an initial emergency phase. It is not a credible target model for a mature multi-agency response.



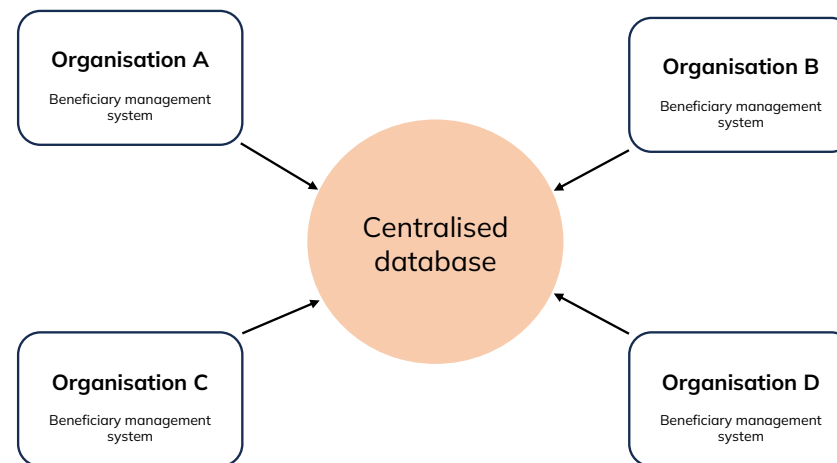
SETUP & TIME-TO-VALUE Can start immediately with familiar tools and no collective infrastructure.	PARTICIPATION & INCLUSIVITY No technical entry barrier, but manual data handling, reporting, and data cleaning are cumbersome for smaller participants.
SCALABILITY Coordination effort grows rapidly with each additional actor, dataset and use case.	REUSE ACROSS USE CASES Each new purpose typically creates another file, workflow or stand-alone tool.
LIVE INTERACTION Manual exchange cannot support point-of-action verification or reliable live status updates.	DATA PROTECTION & SECURITY Copies spread across email, drives and laptops with weak logging, retention and deletion control.
AGENCY CONTROL Each organisation retains full control of its own system and operational decisions.	STANDARDS & DATA QUALITY Definitions and identifiers require repeated manual mapping, and copies can diverge.
AUDITABILITY & GOVERNANCE There is no consistent end-to-end transaction record or shared governance mechanism.	RESILIENCE, COST & LOCK-IN Visible infrastructure cost is low, but hidden staff cost and dependence on individuals are high.

OPTION 2: CENTRALISED PLATFORM

One shared platform stores and manages data for a defined workflow

A shared platform stores data required for a defined function, such as deduplication, referrals, reporting or service mapping. This solution creates a parallel system-of-record of shared data.

Best suited for: A bounded consortium, a defined single workflow, or an environment where many participating organisations have limited technical capacity.



SETUP & TIME-TO-VALUE

A bounded workflow can launch quickly once hosting, roles, retention and the data model are agreed.

SCALABILITY

New actors can be added through accounts or templates, while support and governance concentrate in the host.

LIVE INTERACTION

Live workflows are possible with integration, although periodic file uploads remain common.

AGENCY CONTROL

Participants depend on the platform operator for access, correction, retention and continuity.

AUDITABILITY & GOVERNANCE

One environment can provide consistent access rules, logs and transaction history.

PARTICIPATION & INCLUSIVITY

A portal or file-upload route enables participation without mature systems or API capacity.

REUSE ACROSS USE CASES

A modular platform can support several workflows; many implementations remain single-purpose.

DATA PROTECTION & SECURITY

Controls and audit can be consistent, but concentrated personal data increases breach impact.

STANDARDS & DATA QUALITY

A single schema can enforce common fields, codes, validation rules and statuses.

RESILIENCE, COST & LOCK-IN

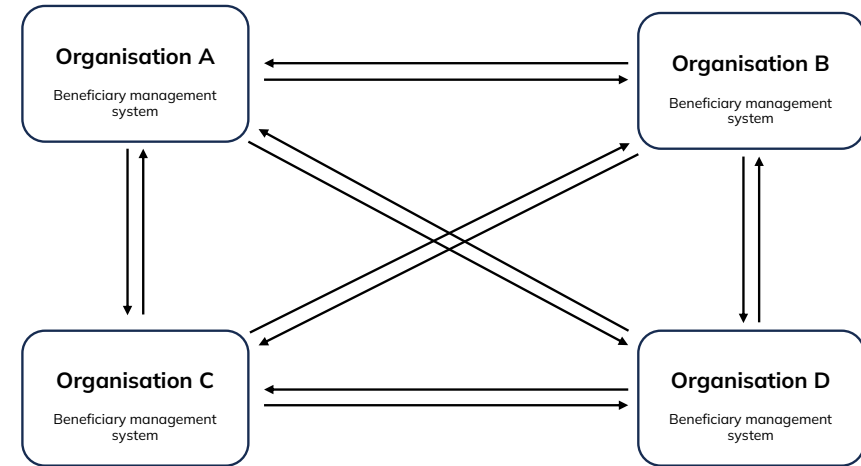
The platform creates a single point of failure and substantial host or vendor dependency.

OPTION 3: BILATERAL API CONNECTIONS

Each pair of organisations connects its systems directly

Each pair of participating organisations connects its systems directly and agrees the technical and legal terms of exchange.

Best suited for: A small number of stable organisations with a specific, high-volume exchange requirement.



SETUP & TIME-TO-VALUE

A direct connection can be quick for one pair, but each new pair repeats technical and legal setup.

SCALABILITY

Connections grow quadratically: 5 organisations need up to 10 links, 10 need 45, and 20 need 190 bilateral linkages.

LIVE INTERACTION

Direct APIs support real-time queries, acknowledgements, status updates and notifications.

AGENCY CONTROL

Each organisation decides what it sends, receives and accepts.

AUDITABILITY & GOVERNANCE

Each pair can log and govern its exchange, but records and agreements remain fragmented.

PARTICIPATION & INCLUSIVITY

Participants need mature systems and technical capacity, or a separate portal, adapter or manual route.

REUSE ACROSS USE CASES

A pair can add several exchanges, but mappings and functions are rebuilt for other pairs.

DATA PROTECTION & SECURITY

Exchange can be direct and purpose-limited, while many interfaces and credentials enlarge the attack surface.

STANDARDS & DATA QUALITY

Two parties can map precisely, but pair-specific mappings create ecosystem-wide semantic fragmentation.

RESILIENCE, COST & LOCK-IN

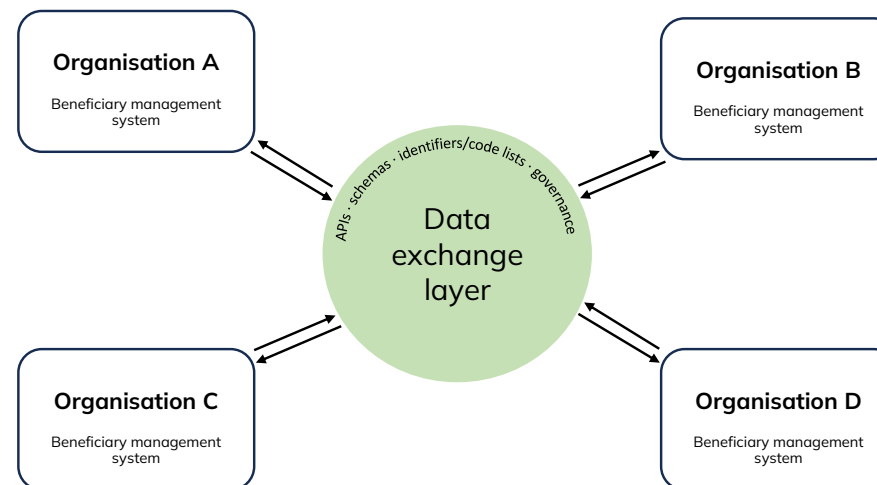
Failures are isolated to individual links, while maintenance cost rises sharply as links multiply.

OPTION 4: FEDERATED DATA EXCHANGE LAYER

Independent systems connect once through a shared exchange layer

Each system connects once to a shared exchange layer; source systems remain systems-of-record, while the layer routes authorised queries, messages, acknowledgements and status updates.

Best suited for: A long-term, multi-agency, multi-use-case ecosystem intended to operate beyond one project, cluster or funding cycle.



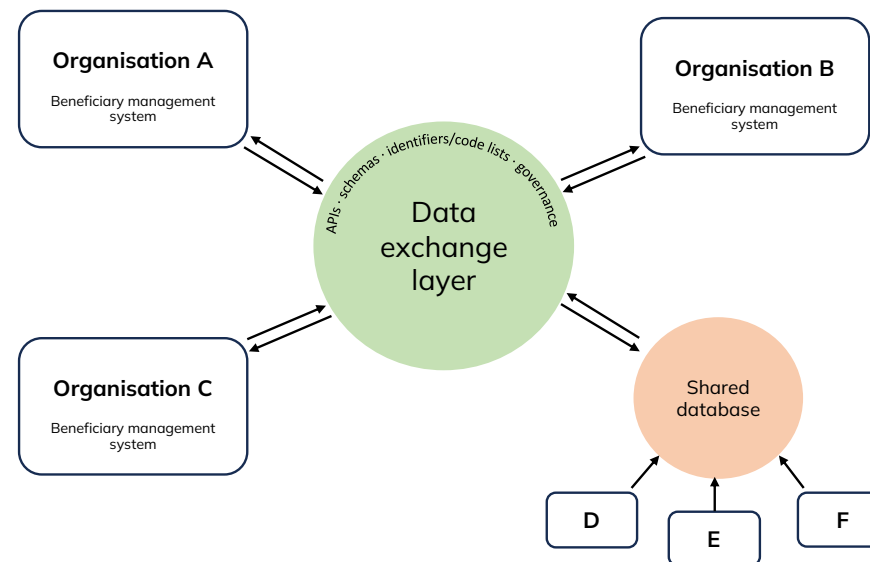
SETUP & TIME-TO-VALUE The first service requires substantial governance, standards, security and integration work.	PARTICIPATION & INCLUSIVITY Integrated participants need one connection; low-capacity actors still require listeners, adapters or a shared node.
SCALABILITY Each organisation normally creates one connection rather than separate links to every participant.	REUSE ACROSS USE CASES The same infrastructure can support referrals, deduplication, reporting, verification and other services.
LIVE INTERACTION Designed for live queries, match responses, acknowledgements and status updates.	DATA PROTECTION & SECURITY Data can remain with the system of record and exchanges can disclose only the minimum required data relevant to the use-case.
AGENCY CONTROL Each organisation retains its source data and chooses the services and purposes in which it participates.	STANDARDS & DATA QUALITY Shared schemas, identifiers, service definitions and versioning can be reused across participants.
AUDITABILITY & GOVERNANCE Transaction logs and correlation identifiers enable end-to-end traceability without centralising beneficiary data.	RESILIENCE, COST & LOCK-IN Open standards reduce lock-in and marginal cost falls with reuse; the exchange layer still needs sustained stewardship.

OPTION 5: HYBRID MODEL: FEDERATED EXCHANGE + SHARED NODE

Federation combined with a shared node for non-integrated participants

A federated exchange layer connects agency systems, while a shared database or portal participates as one node for organisations that cannot yet integrate directly. Combines options 2 and 4.

Best suited to: A response with heterogeneous technical capacity, provided the shared node remains an inclusion route rather than the default repository for all data.



SETUP & TIME-TO-VALUE Operating both an exchange layer and a shared node makes this the most complex model to establish.	PARTICIPATION & INCLUSIVITY API-ready, passive and low-tech participants can use different routes into the same ecosystem.
SCALABILITY Direct integrations scale well, but the shared node can become a capacity and support bottleneck.	REUSE ACROSS USE CASES The same services can be offered through direct APIs and the shared participation node.
LIVE INTERACTION Integrated systems can exchange live; portal and file-based participants may operate with delay.	DATA PROTECTION & SECURITY Direct systems can minimise disclosure, while the shared node introduces a second, centralised risk profile.
AGENCY CONTROL Integrated organisations retain control; shared-node users depend more heavily on the operator.	STANDARDS & DATA QUALITY Common standards can span both channels, but inconsistent translation can create a two-tier data environment.
AUDITABILITY & GOVERNANCE A common transaction layer can provide traceability, while operator and controller roles are more complex.	RESILIENCE, COST & LOCK-IN The model supports gradual transition and fault isolation, but has the highest ongoing operating cost.

EXAMPLE 1: DEDUPLICATION BEFORE ASSISTANCE

Centralised batch checking compared with a federated live query

Centralised model	Federated model
1. Field staff identifies potential beneficiary household. 2. Household is registered and assessed. 3. After returning from the field, staff export identifying data from their registration system, save the file locally and upload it to the shared platform. 4. The platform returns possible matches, which the organisation must review and resolve. 5. Staff finalise the assistance list and inform people whose registration will not proceed because the same household has already received, or is already scheduled to receive, the relevant assistance.	1. Field staff identifies potential beneficiary household. 2. Before starting full registration, the field staff sends a live query from the organisation's own system. The query asks whether the household is already registered for the relevant programme or has received the same type of assistance within the set period. 3. Participating organisations respond through the data exchange layer, ideally with only the minimum result required: no match, possible match, or limited information about relevant prior assistance. 4. Staff interpret the results and either proceed or do not proceed with registration.
Result: The platform prevents a duplicate payment, but it detects the overlap late. Outreach, registration, assessment, data entry and follow-up have already taken place, and another copy of personal data has been created and stored.	Result: Live federated query prevents not only the eventual duplicate payment but also the duplicate work of the staff and expectation management of the potential beneficiary.

MAIN DIFFERENCES

When check happens	Data movement	Main effect
After registration vs. before registration	File export/upload vs. live query	Prevents duplicate payment vs. prevents duplicate work and payment

EXAMPLE 2: INTER-ORGANISATIONAL REFERRAL

A shared referral inbox compared with connected organisational workflows

Centralised model	Federated model
1. Field staff identifies a case in need of referral. 2. The referring organisation enters the referral and agreed case information into a shared platform. 3. Potential receiving organisations must log in to that platform, check whether new referrals are available, and determine whether they match their services, eligibility criteria and geographic coverage. 4. A receiving organisation then accepts or rejects the referral and updates its status. 5. The receiving organisation enters the case data manually into their own system. 6. Referral status (accepted, rejected, etc.) can be checked by the referring organisation on the same platform.	1. Each participating organisation publishes, through an API, an up-to-date catalogue of services it provides, including the locations where they are available and any relevant referral conditions. 2. Field staff identifies a case in need of referral. 3. The referring organisation's own application uses the service catalogue to display suitable receiving services, already filtered by location and other agreed criteria. 4. The agreed minimum case data is sent directly from the referring organisation's system to the receiving organisation's database. 5. The receiving organisation reviews, assigns, accepts, or rejects the referral within its normal system. 6. The referring organisation can query the referral status through the API or receive an automated notification whenever the status changes

MAIN DIFFERENCES

Finding services	Referral submission	Status tracking
Log into a shared platform vs. live service catalogue	Entered into a shared platform vs. sent directly system-to-system	Platform checking/manual updates vs. API-facilitated status updates

CHOOSING THE ARCHITECTURE

The appropriate model depends on scale, duration, use cases and the technical capacity of participating organisations

The architecture should follow the intended scale, duration, number of participants, use cases and technical capacity. A centralised platform may be entirely appropriate for a bounded task; it becomes problematic when each new task creates another platform. Federation is the strongest option where the same exchange infrastructure is expected to support several organisations and use cases over time.

The main questions to ask when making the architecture choice:

- How many organisations will participate and how often will that group change?
- Is the requirement limited to one data workflow or expected to support several use cases?
- Are live queries and status updates necessary or is periodic batch exchange sufficient?
- What proportion of participants can integrate through APIs and what route will others use?
- Who will operate, fund, govern, secure and audit the arrangement over time?

Based on the answers, the following choices are recommended:

Immediate, low-volume, temporary exchange	→ Manual exchange
One bounded consortium workflow	→ Centralised model
Two or three stable, technically capable partners	→ Bilateral APIs
Long-term, multi-agency, multi-use-case ecosystem	→ Federated model
Mixed technical capacity and gradual adoption	→ Hybrid model

DigiCap is a neutral advisory mechanism that helps humanitarian actors address shared digital challenges. It provides independent expertise to tackle interagency challenges around data sharing, interoperability, data governance and responsible digitalisation.