DigiCap.

Digital advisory for the humanitarian sector

SETTING UP A FEDERATED DATA EXCHANGE SYSTEM

Practical roadmap from first agreement to a sustainable multi-agency service

This DigiCap explainer sets forth a generic five-phase roadmap for country contexts to setting up a multi-agency, multi-use-case federated data exchange infrastructure. A federated approach keeps source records in participating systems, while a governed exchange layer connects them for an agreed purpose.

1. DEFINE	2. GOVERN	3. DESIGN	4. PILOT	5. OPERATE
Agree why and for whom data should be exchanged	Set decision rights, lawful sharing and safeguards	Standardise meaning, data and connections	Test limited, real workflows before scale-up	Maintain trust, security and funding over time

GLOSSARY

Core concepts relevant to federated data exchange

FEDERATED EXCHANGE Independent systems exchange approved requests and responses while retaining their source records.	**CONTROLLER / PROCESSOR** Legal roles for deciding data purposes and processing it; definitions vary by applicable law.
SYSTEM OF RECORD The organisation's authoritative system for maintaining and correcting its original information.	**DATA-SHARING AGREEMENT** Binding rules between parties on purpose, access, responsibilities, safeguards and accountability.
USE CASE The practical purpose for an exchange, such as checking overlap or sending a referral.	**MINIMUM DISCLOSURE** Only the information necessary for a specific authorised purpose is exchanged.
EXCHANGE LAYER Shared infrastructure that authenticates, routes and records authorised transactions.	**PRIVACY-PRESERVING MATCHING** Methods for comparing records while limiting disclosure (no method guarantees zero identity risk).
COMMON SCHEMA Agreed field names, structures, types, codes and message versions.	**AUDIT TRAIL** Limited information showing who exchanged what kind of data, when, why and with what result.
SHARED MEANING Agreement on what each field and status actually mean across organisations.	**REDRESS** A usable route to challenge, correct and remedy harmful or incorrect outcomes.

LAYERS WHICH NEED TO WORK TOGETHER

Technical	**Structure**	**Meaning**	**Authority**	**Operations**
Can systems connect securely?	Do fields and messages align?	Do terms and rules mean the same thing?	Who may share, for what purpose?	Who checks, corrects, supports the service?

THE FIVE-PHASE ROADMAP

1. DEFINE	2. GOVERN	3. DESIGN	4. PILOT	5. OPERATE
Agree why and for whom data should be exchanged	Set decision rights, lawful sharing and safeguards	Standardise meaning, data and connections between systems	Test limited, real workflows before scale-up	Maintain trust, security and funding over time
OUTPUT Use-case statement & participant mapping	**OUTPUT** Governance & safeguard rules	**OUTPUT** Standards & security design	**OUTPUT** Test results & go/no-go decision	**OUTPUT** Runbook, funding & review
DECISION GATE Is federation the most appropriate solution?	**DECISION GATE** Are lawful authority and access rules in place?	**DECISION GATE** Are shared meanings and standards in place?	**DECISION GATE** Is it working and are risks mitigated?	**DECISION GATE** Should it continue, expand, change, close?

GENERAL PRINCIPLES & CONSIDERATIONS

Purpose before technology	→	Start with a concrete operational need. Choose technology only after design is clear.
Collective governance, agency control	→	Agree shared rules and oversight, while each organisation retains control of its source data and authorised participation.
Data minimisation and protection	→	Exchange only what is necessary. Queries, hashes, responses and logs can still expose sensitive information if risks are not mitigated.
Human oversight of consequential decisions	→	Possible matches, uncertain results or missing responses must not automatically lead to exclusion. Ensure appropriate review and correction.
Inclusive participation	→	Different technical capacities require different participation routes. API integration should ideally not be a prerequisite for participation.

PHASE 1: USE-CASE DEFINITION

Define the operational problem worth solving and test whether federation is the right response

DECISION	QUESTIONS TO ANSWER	AGREED EVIDENCE OR OUTPUT
Purpose	What decision or workflow improves if organisations can exchange data? Who benefits?	A one-sentence use case; intended result; decision owner.
Boundaries	Which people, programmes, locations and periods are included? Which uses are prohibited?	Scope and exclusions; allowed query types.
Participants	Who holds the relevant information? Which groups or agencies lack voice or capacity?	Actor map, systems inventory and participation needs.
Feasibility	Can source systems answer reliably? Is identity quality sufficient and sharable? Are connectivity and resources adequate?	Data-quality baseline, feasibility issues and likely costs.
Alternative options	Have alternatives such as bilateral links, a bounded shared tool or no exchange?	Recorded rationale for using a federated model.

PHASE OUTPUT	DECISION GATE
• Use-case statement • Participant and system mapping • Baseline measures • Initial risk assessment	Proceed only when the use is necessary, realistic, acceptable to participants and worth the possible cost.

PHASE 2: GOVERNANCE

Separate collective decision-making, legal responsibilities and day-to-day technical operations

DECISION	QUESTIONS TO ANSWER	AGREED EVIDENCE OR OUTPUT
Decisions and voice	Who admits members, approves uses, changes rules, resolves conflicts and represents affected people?	Governance charter, voting and appeal arrangements.
Legal authority	Which laws apply? What is the lawful basis? Who controls or processes each data flow?	Role mapping, legal assessment and signed agreements.
Access limits	Who can ask what, for which purpose? What data stays off-limits, including external authority requests?	Purpose-and-access matrix; rules for exceptional requests.
Protection and remedy	How will people be informed, contest a match, correct records and receive help if harmed?	Risk assessment, complaints and correction procedures.
Host and operator	Who holds contracts and funds? Who actually operates the service? How is independence checked?	Named institutional roles, oversight and resourcing.

PHASE OUTPUT	DECISION GATE
• Governance charter • Defined responsibilities • Data-sharing and operating agreements • Proportionate protection assessment	No live personal-data exchange until lawful authority, access rules, accountability and usable safeguards are in place.

PHASE 3: DATA AND TECHNICAL DESIGN

Make the meaning of each request and response as clear as the technology carrying it

DECISION	QUESTIONS TO ANSWER	AGREED EVIDENCE OR OUTPUT
Shared definitions	Do “household”, “assistance received”, “accepted referral” and time windows mean the same thing?	Agreed data dictionary, terms and codes.
Minimum messages	What is the least each request and response must reveal? Which status and error codes are required?	Versioned schemas and interface contracts.
Record matching	What identifiers exist? How are spelling variants, duplicates, false matches and uncertain results handled?	Documented matching methods, confidence rules and human review.
Security & logging	Who authenticates callers? How are permissions, encryption, keys, rate limits, logs and retention handled?	Trust and threat models; audit/log retention design.
Participation routes	What if a partner has no API or stable internet? When would a shared node introduce extra risk?	Adapter, portal or delayed-exchange design, with fallback.

PHASE OUTPUT	DECISION GATE
• Service blueprint • Schemas and definitions • Access controls • Matching rules • Integration and testing specifications	Each participant can provide and interpret the same minimum response, and the security design fits the context.

PHASE 4: PILOT / PROOF OF CONCEPT

Start with invented data; then test a small, reversible live service with willing participants

DECISION	QUESTIONS TO ANSWER	AGREED EVIDENCE OR OUTPUT
Correctness	Can the service recognise matches, non-matches and uncertain cases across different data quality?	False-positive and false-negative checks; review procedures.
Exchange behaviour	What happens with stale, duplicate, missing, delayed or failed messages?	Idempotency, retries, timeouts, status and error tests.
Protection and security	Are access limits enforced? Are logs minimal? Can leaks, coercion and misuse be detected?	Security checks; incident drill; deletion verification.
Real-world decisions	Can staff interpret warnings, correct mistakes, override errors and help affected people?	Observed case-handling and redress tests.
Pilot boundaries	Which partners, use cases, data, locations and time periods are included? How do we stop safely?	Pilot protocol, manual fallback and rollback plan.

PHASE OUTPUT	DECISION GATE
• Pilot evaluation • Quality and safety metrics • Outstanding issues • Documented go / revise / stop decision	Go live only if benefits are demonstrated, serious risks are mitigated and corrective paths work in practice.

PHASE 5: OPERATE AND EXPAND

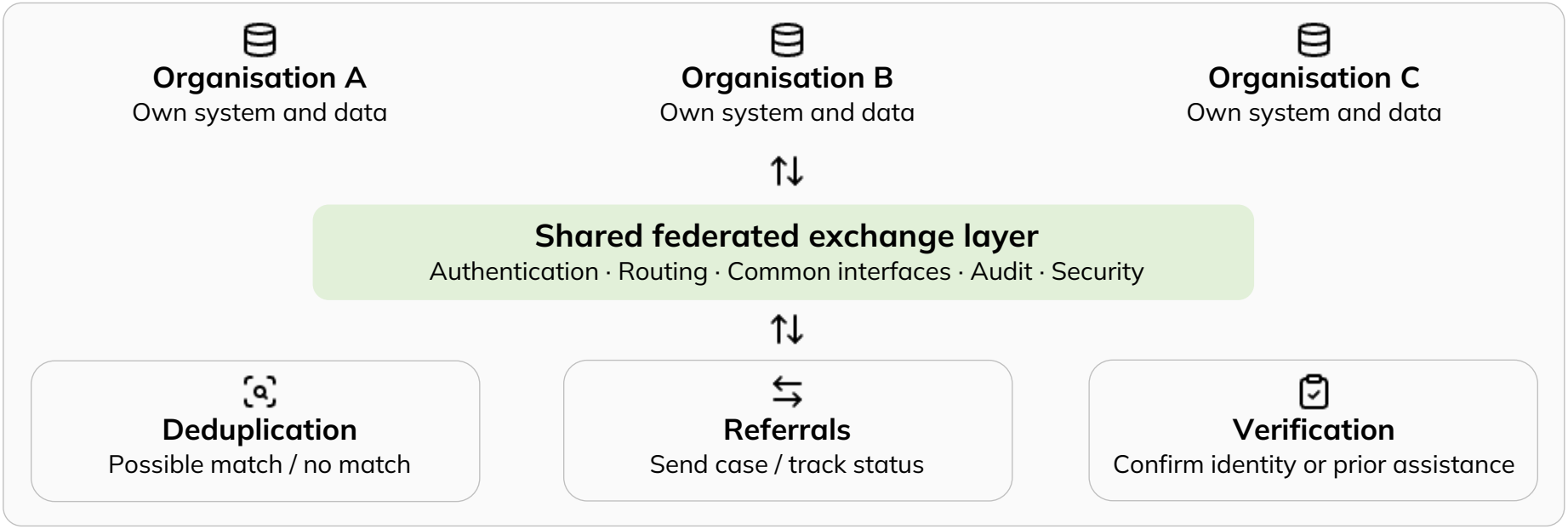
Running the data exchange is a continuing responsibility, not the end of the implementation project

DECISION	QUESTIONS TO ANSWER	AGREED EVIDENCE OR OUTPUT
Service operations	Who handles incidents, outages, support, backups, technical updates and vendor performance?	Runbook, service targets, emergency contacts and contracts.
Data and participant quality	Who fixes stale records, broken mappings, weak security or new versions of connected systems?	Quality reviews, onboarding and version-change plans.
Governance oversight	How are audits, complaints, requests for new uses and exceptional data demands reviewed?	Scheduled governance review, transparent decisions and audits.
Financial sustainability	Who pays operating costs and future integrations? What if the host or supplier withdraws?	Multi-year budget, cost allocation and continuity plan.
Change or exit	When must a new use case return to earlier phases? How are memberships or the whole service ended?	Change control, revocation, data handling and exit procedures.

PHASE OUTPUT	DECISION GATE
• Operational plan • Sustainable budget • Audit and review cycle • Onboarding • Change and exit arrangements	Continue or expand only while the system remains necessary, secure, inclusive, demonstrably useful and funded.

ONE FEDERATION, MULTIPLE USE-CASES

Build the shared infrastructure once but agree on the purpose, permissions and rules for each use case separately



THREE RULES FOR EXPANDING THE FEDERATION

1. **Connected does not mean authorised:** An organisation can join one service without agreeing to participate in every other exchange.
2. **Data minimisation remains essential:** Even when data stays in participating organisations' systems, queries, responses and logs may contain sensitive information. Exchange only what is necessary and appropriate.
3. **New purpose means new approval:** Additional services can reuse the infrastructure, but require their own justification, permissions and technical specifications.

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