

CITHORUM

Cithorum Knowledge Graph Explainer

One sentence

Jam makes the data cheaper to store, move, and restore. The Cithorum Knowledge Graph turns the same data plane into an operating map: customers, workloads, evidence, incidents, costs, decisions, and outcomes connected in one queryable system.

Why the KG exists

Most AI products sit on top of fragmented tools. They can answer local questions, but they do not know the whole operating picture: which customer owns which workload, which restore was verified, which dataset drove which invoice, which incident changed the SLA, or which evidence supports the answer.

The Cithorum KG is the memory layer above Jam. It keeps the relationships live while Jam handles the economics underneath.

The first KG is not speculative

The first version is **Data-Centre Ops KG** for Cithorum's own Jam Proof Cloud.

Every customer workload Jam touches creates structured telemetry:

- customer
- bucket
- dataset
- raw size
- Jam size
- ratio
- restore status
- CPU and throughput
- invoice basis
- node health
- incident history
- evidence artefacts

The KG turns that telemetry into an operating graph. The first buyer is internal: Cithorum uses it to run the proof environment. The second buyer is external: data-centre operators and regulated infrastructure teams that need the same visibility.

Core architecture

```
Customer data
-> Jam data plane
-> telemetry and metering
-> KG Core
-> products and vertical workflows
```

Jam data plane

Compression, indexing, cryptographic envelope, deterministic restore, and per-customer storage isolation.

Telemetry and metering

Events emitted by Jam: size, ratio, restore result, throughput, CPU, workload type, customer, bucket, and billing basis.

KG Core

Typed entities, typed relationships, provenance, permission checks, query layer, and AI-accessible summaries.

Product surfaces

Dashboards, reports, billing exports, customer portals, vertical workflows, and AI assistants grounded in graph evidence.

What it can answer

Infrastructure questions

- Which customers are consuming the most raw capacity?
- Which workloads produce the strongest Jam savings?
- Which datasets have not had a recent restore proof?
- Which node, disk, queue, or customer bucket caused an incident?
- Which pilot should convert to paid based on measured savings?

Commercial questions

- Which benchmark report supports this invoice?

- Which customer proof can be shown publicly, privately, or under NDA?
- Which partner route created which lead?
- Which workloads map to the next infrastructure raise?

Vertical questions

- Workspace KG: which customer, project, invoice, document, and decision are connected?
- Law KG: which matter, party, precedent, clause, and outcome are connected?
- Sequencing KG: which sample, FASTQ, variant, evidence, and audit trail are connected?

The four-product map

Cithorum should be presented as four clear product surfaces:

1. **Jam Engine** — the deployable data plane.
2. **Data Centre Service** — managed Jam Proof Cloud / hosted storage economics.
3. **Workspace KG** — business memory across tools, documents, decisions, and customers.
4. **Law KG** — private matter memory, precedent graph, and audit-trailed AI over legal work.

Data-Centre Ops KG is the internal foundation that proves the KG before the external verticals scale.

Why this belongs with Jam

Jam creates the telemetry that most KG platforms lack: byte-level economics, restore proof, evidence trails, and customer workload structure. The KG makes that telemetry useful. Together they create the wedge:

- Jam proves measurable savings.
- The Ops KG proves operational intelligence.
- Workspace and Law KG prove vertical workflow value.
- The same core can expand into regulated domains only when controls and paid demand justify it.

Build order

Phase 1 — Jam telemetry

Ship the Jam path first: upload, process, compress, restore, meter, report.

Phase 2 — Data-Centre Ops KG

Graph the proof environment: customers, buckets, workloads, nodes, restores, incidents, costs, and evidence.

Phase 3 — Customer-facing dashboards

Expose benchmark reports, restore proofs, billing exports, and workload history to customers.

Phase 4 — First external KG vertical

Choose Workspace KG or Law KG based on paid design-partner demand. Do not build every vertical at once.

Investor-safe positioning

The KG is not a separate science project. It is how Jam becomes a platform. Jam lowers the cost of the data plane; the KG captures the memory, evidence, and workflows that compound on top of it.