TECHNOLOGY POSITION PAPER

The Honest Broker

Siloed data is killing people.
How can AI fix it?

February 2026 • Part 2 of the Unpaid Contractor series

IN August 2002, two ten-year-old girls disappeared from the village of Soham in Cambridgeshire. Holly Wells and Jessica Chapman had been lured into the home of their school caretaker, Ian Huntley, who murdered them and disposed of their bodies. When the subsequent inquiry by Sir Michael Bichard examined how Huntley had been appointed to a position of trust around children, it uncovered something that would haunt British policing for two decades. Humberside Police held extensive intelligence on Huntley, and had received multiple allegations of sexual offences against minors stretching back to 1995. None of that intelligence was accessible to Cambridgeshire Constabulary when it ran his background check. The systems didn't talk to each other. Two children died in the gap between databases.

Fifteen years later, on 22 May 2017, Salman Abedi detonated a shrapnel-packed bomb in the foyer of the Manchester Arena as crowds left an Ariana Grande concert. Twenty-two people were killed and more than a thousand injured. The public inquiry chaired by Sir John Saunders found that MI5 had received intelligence about Abedi on two separate occasions in the months before the attack but had not shared it with Counter Terrorism Policing. The inquiry concluded there was a realistic possibility that the attack might have been prevented if that intelligence had been acted upon. MI5's Director General said he was profoundly sorry. The families of the dead said MI5 could not be trusted to tell the truth.

These are not edge cases. They are symptoms of a structural condition that pervades the British state and much of the democratic world: critical information locked in institutional silos, inaccessible to the people and agencies that need it, at the moments when it matters most.

the urge to merge...

Every time a data-sharing failure kills someone, the institutional reflex is the same: centralise. Build a bigger database. Connect everything. The logic feels irresistible. If the problem is that systems don't talk to each other, surely the answer is to put all the data in one place?

The UK has tested this theory at extraordinary scale and cost. In 2002, the same year Holly Wells and Jessica Chapman were murdered, the NHS launched the National Programme for IT in an attempt to create a single, integrated electronic patient record for every person in England, and connect 30,000 GPs to 300 hospitals through one centralised system. It was described as the largest civilian IT programme in the world. The original budget was £6.2 billion. By the time it was scrapped in 2011, the bill had exceeded £10 billion and the Public Accounts Committee described it as one of the worst and most expensive contracting fiascos in history. It delivered almost nothing of clinical value.

The failure was predictable and predicted. The programme was top-down, politically motivated, and imposed without adequate consultation with the clinicians who would use it. But the deeper lesson is architectural. Centralisation doesn't solve the governance problem, it just displaces it. You end up with a single catastrophic point of breach, a monolithic system too rigid to adapt to local needs, and no granularity in access control. The very thing that makes a centralised system appealing – everything in one place – is what makes it dangerous. The more data you aggregate, the higher the value of the target and the greater the damage when something goes wrong.

Policing has walked the same path with the same results. The Bichard Inquiry's first recommendation was the urgent creation of a national police intelligence system. The Police National Database was eventually launched in 2011, seven years after Bichard reported, at a cost

of £75 million. It now sits alongside the older Police National Computer, a system dating from 1974, and the two have never been properly integrated. A programme to merge them into a single Law Enforcement Data Service was launched in 2016 with a budget of £671 million. By 2021, costs had ballooned to £1.1 billion, timelines had slipped repeatedly, and the programme had been effectively reset after the police lost confidence in the Home Office's ability to deliver. As of early 2026, the PNC is running on infrastructure so elderly that its replacement is now classified as critical national infrastructure. The ambition to unify police data under one roof has consumed well over a billion pounds and remains, more than two decades after Soham, substantively unfinished.

The pattern is consistent. The state identifies a data-sharing failure. It commissions a centralised solution. The solution takes years longer and costs many times more than projected. Local institutions resist because the system doesn't meet their needs. The resulting architecture is brittle, expensive to maintain, and still doesn't deliver the joined-up intelligence that was promised. Meanwhile, the underlying problem of siloed information that can't be accessed when it's needed persists.

a panopticon, inverted...

There is an alternative. Instead of dragging every database into a single warehouse, leave the data where it is and broker access to it.

Our previous paper argued that agentic AI should be treated as a contractor rather than a tool and given its own identity rather than becoming the user's digital doppelganger. The model scales. Where the individual model has one principal and one or more contractors, a public-sector deployment has a network of contractors mediating between the citizen, the state, and the institutions that hold the data.

The architecture is straightforward. Each data silo – the PNC, the PND, the DVLA, HMRC, the NHS Spine, the Home Office immigration records – retains full control of its own data behind its own gatekeeper agent. The gatekeeper understands the data it protects, the policies governing access, and the legal frameworks within which disclosure is permitted. It does not surrender data on demand. It evaluates requests.

On the other side of the exchange, the requesting party, whether a police officer, a clinician, a caseworker, or eventually a citizen, operates through their own agent: an *honest broker* that formulates queries, presents justifications, and negotiates access on behalf of the principal. The broker doesn't need to know what's in the silo, it just needs to know what its principal is authorised to ask for and why.

Data never moves to a central repository. It is queried in place, filtered by policy, and delivered as the minimum necessary fragment. The gatekeeper logs every request, every justification, and every response. The broker logs every query it sent and every answer it received. Both sides of the transaction are auditable. This is the opposite of a panopticon. It is a system in which every act of data access is visible, attributable, and challengeable.

one query, five silos...

Consider a working example. A detective constable in the West Midlands is investigating a series of linked burglaries. They need to establish whether a suspect has prior convictions, whether any vehicles registered to the suspect have appeared on ANPR cameras near the crime scenes, and whether there is intelligence held by other forces linking the suspect to an organised crime group.

Today, the detective queries the PNC for criminal records and the PND for intelligence, each through separate interfaces with separate access protocols. ANPR data sits in yet another system. If there is relevant intelligence held by another force, it may or may not be discoverable depending on how it was tagged, whether it was uploaded to the PND, and whether the detective knows to look for it. This is the fragmented landscape that has persisted since Soham despite billions of pounds of investment.

Under a brokered model, the detective's agent submits a single structured request: "Suspect X, investigation reference Y, offence category Z. Requesting: criminal history, ANPR hits within these geographic and temporal parameters, cross-force intelligence." The request carries the detective's warrant number, their force, the investigation reference, the legal basis for the query, and the justification for each data category requested.

Each gatekeeper evaluates the request against its own access policies. The PNC gatekeeper confirms that the detective is authorised to access criminal records for the stated purpose and returns the relevant history. The ANPR gatekeeper checks geographic and temporal scope, confirms proportionality, and returns matching hits. The PND gatekeeper evaluates the intelligence request, applies the originator's dissemination restrictions, and returns what the detective is permitted to see. If a gatekeeper denies a request, the denial and its reason are logged on both sides.

The critical difference is not speed, although speed is a benefit. The critical difference is governance. Every data flow is policy-mediated, logged, and attributable. The detective gets a more comprehensive picture than they would today, assembled from disparate sources without any of those sources surrendering control. And crucially, the architecture doesn't require the

systems to be merged. They can remain separate, administered by their respective authorities, each evolving at their own pace. What connects them isn't a monolithic database. It's a protocol.

This is the crucial engineering insight. The brokered model does not require any of those underlying systems to be replaced, merged, or modernised. The PNC can remain on its ageing Fujitsu mainframe. The PND can stay where it is. The ANPR system, the DVLA database, the Home Office immigration records – none of them need to change. The broker layer sits on top, mediating access through a common protocol while each silo retains full sovereignty over its own data and its own disclosure rules. This is not a plan for the next decade. It is deployable on the infrastructure that exists today. It is, in effect, the integration of legacy systems without the integration, achieving through a policy-mediated access layer what a billion pounds of attempted database mergers has so far failed to deliver. The Home Office's own NLEDS programme quietly arrived at a version of this conclusion when it abandoned the monolithic PNC-PND merge in 2020 and pivoted to what it called "federated search capability." The brokered model takes that instinct and gives it a governance framework.

health hazard...

If policing is the most dramatic illustration of data-silo failure, the National Health Service is the most pervasive. GP records are routinely invisible to accident and emergency departments. Mental health records are walled off from primary care. A patient presenting at one hospital with a drug allergy recorded at another may have no way of communicating that fact to the clinician treating them. The tension is identical to policing – privacy demands compartmentalisation, patient safety demands access. The consequences of getting the balance wrong are measured in avoidable harm.

The failed NPfIT attempted to solve this by building one system to rule them all. The brokered model solves it differently. A patient's agent, operating under the patient's authority with permissions the patient has explicitly granted, can broker access to their own records across providers. The GP's gatekeeper shares what the patient has authorised, and the hospital's gatekeeper receives only what is clinically relevant. The mental health trust's gatekeeper applies its own, stricter disclosure rules. No provider surrenders its entire database. The patient's record is assembled dynamically at the point of need, filtered by consent and clinical context, and the audit trail shows exactly what was shared, why, and with whom.

This is not a speculative architecture. It is a direct translation of how human data governance already works when it works well. A patient can already ask their GP to forward specific records to a specialist. They can authorise a hospital to access their prescription history. They can refuse to share their mental health notes. The brokered model automates what is already happening

manually, at the speed and scale that modern healthcare demands, without requiring a decade-long, multi-billion-pound integration programme.

follow the money…

The financial sector presents the same structural tensions under different regulatory pressure. Anti-money laundering obligations require banks, solicitors, accountants, and other regulated entities to report suspicious activity to the National Crime Agency. The UK Financial Intelligence Unit receives over 850,000 Suspicious Activity Reports a year, stored on a central database that now holds over 4.5 million records. Each SAR adds a piece of a picture - a bank sees a pattern of unusual transfers, a solicitor feels a property transaction doesn't quite add up, an accountant flags cash flows that don't match the declared business activity. No single institution sees the whole landscape.

The challenge for law enforcement is to connect those fragments without requiring every financial institution to open its books. A brokered network does exactly this. An investigator's agent submits a query with a legal basis, a justification, and a defined scope. Gatekeeper agents at each institution evaluate the request against their regulatory obligations and return only what is lawfully disclosable. The investigator gets a composite intelligence picture assembled from multiple sources without anyone in the chain having to expose statute-protected data. The architecture mirrors the existing SARs regime but operates in real time, at machine speed, with granular access control and an end-to-end audit trail.

when the state gets it wrong…

Thus far, the argument has focused on how brokered access helps the state do its job better – catch criminals, treat patients, follow dirty money. But there is another side to the data-silo problem that receives far less attention and demands far more.

In 2010, the UK Border Agency destroyed thousands of landing cards documenting the arrival dates of Caribbean-born residents who had settled in Britain from the late 1940s onwards. When the Home Office subsequently introduced its hostile environment policy, demanding documentary proof of legal residence as a condition of employment, housing, healthcare, and banking, many members of the Windrush generation found themselves unable to prove what had always been true: that they were here lawfully. At least 83 people were wrongly deported. An unknown number lost jobs, homes, and access to medical care. The Windrush Lessons Learned Review concluded that what happened was foreseeable and avoidable, the result of ignorance and thoughtlessness compounded by immigration regulations tightened with wholesale disregard for the rights of the people they would affect.

The Post Office Horizon scandal, in which a faulty accounting system was trusted over the testimony of hundreds of sub-postmasters for more than two decades, is another case in point. Data held by one institution, in this case the Post Office's proprietary IT system, was treated as infallible. The people it accused had no effective means of interrogating the data, challenging its accuracy, or presenting a counter-narrative. The system was opaque. The individuals it destroyed had no broker acting on their behalf.

These are not principally technology failures. They are governance failures in which the state's data systems were weaponised against the people they were supposed to serve. The consistent theme is asymmetry: the institution holds the data, controls access to it, and has the resources to defend its version of events. The individual has none of these.

the bedevilled's advocate...

This is where the brokered model does something that no centralised system can. It gives the citizen an agent of their own.

A citizen's broker is not primarily a tool for accessing government services, although it can be that. It is their data advocate. It operates under the citizen's authority to query the systems that hold data about them, to request disclosure, to challenge inaccuracies, and to maintain a verifiable record of every interaction. When the Home Office claims you have no right to reside, your broker queries the immigration database, the HMRC employment records, the NHS registration history, and assembles the evidence that you do. When a faulty system accuses you of theft, your broker requests the underlying data, logs the request and the response, and creates an audit trail that an independent reviewer can examine.

This is not science fiction. It is the logical extension of a principle that already exists in law. The Data Protection Act 2018 and the UK GDPR give every individual the right to access the personal data that organisations hold about them, the right to rectification if that data is inaccurate, and the right to an explanation of automated decisions. These are powerful rights on paper. In practice, exercising them requires navigating labyrinthine bureaucratic processes that are slow, opaque, and often deliberately obstructive. A citizen's broker automates the exercise of rights that already exist, at machine speed, with a complete audit trail.

The implications for accountability are profound. In a brokered architecture, the citizen's agent maintains its own record of every query and every response. If a database confirms your status today and the record is destroyed tomorrow, the broker's log preserves the evidence that it once existed. It cannot rely on a faulty database to convict you of a crime you didn't commit without that database being interrogable. It cannot deny you services on the basis of data you are not

permitted to see because your broker can see it, and can log what it found. The audit trail exists on both sides. The citizen's record of the transaction is as authoritative as the state's.

voices from the coalface…

This paper has been informed by conversations with senior professionals serving in the public sector whose daily work involves navigating the practical consequences of fragmented data governance. Their consistent testimony is that the problem is not a shortage of data but a shortage of legible mechanisms for getting the right data to the right person at the right time. The technology exists. The governance frameworks do not.

The contractor model proposed in our first paper addressed the question of how to integrate agentic AI safely and effectively into team-based workflows. The brokered model extends that principle into the institutional landscape. Instead of principals running contractor teams, we have networks of contractors, gatekeepers and brokers, negotiating data access on behalf of their respective principals within a framework of automated policies, logged transactions, and mutual accountability. The trust architecture is the same. The scale and reach are different.

For organisations, the brokered model offers something that centralisation has repeatedly failed to deliver: interoperability without the friction of integration. Each institution retains sovereignty over its own data, and each defines its own access policies. The gatekeeper agent enforces those policies consistently and at scale, mitigating the inevitable errors and inconsistencies that plague manual data-sharing agreements. New data sources can be added to the network without rebuilding the architecture and existing systems can participate without being replaced. The model is additive, not disruptive.

For governments, the brokered model resolves a tension that has sabotaged public-sector IT initiatives for decades: how to enable data-sharing whilst preserving institutional autonomy and individual rights. Centralisation sacrifices autonomy and legitimacy, and concentrates risk. Fragmentation preserves autonomy but creates the information gaps that gut competence and wreck lives. A brokered network threads the needle. It enables cross-boundary intelligence without dissolving boundaries.

quis custodiet……

The obvious objection is trust. Who watches the brokers? What stops an agent from overstepping its remit, fabricating justifications, or colluding with another agent to circumvent access controls?

The answer is the same as for human contractors: bounded authority, audit trails, and legal

liability. A broker operates under a statutory mandate from its human principal. Every query it issues carries a justification that can be reviewed after the fact. Every response it receives is logged by both the broker and the gatekeeper. The audit trail is bilateral, immutable, and subject to independent oversight. If a broker submits a query that exceeds its mandate, the gatekeeper denies it and the denial is recorded. If a gatekeeper discloses data that shouldn't have been disclosed, the broker's record of the transaction provides the evidence.

This is not a theoretical safeguard. It is how professional accountability already works in regulated sectors. When a solicitor requests disclosure of financial records, the request is documented, the legal basis is stated, and the response is logged. When a detective requests a production order, the application is reviewed by a judge. The brokered model formalises and automates existing accountability mechanisms. It makes them swifter, more consistent, and harder to circumvent than their manual equivalents.

The more subtle risk is mission creep: the gradual expansion of access permissions over time, normalising surveillance under the guise of efficiency. This is a real and serious concern, but it is a concern that applies equally to any data-sharing architecture. The brokered model's advantage is legibility. Because every access request is logged with its justification and every permission boundary is explicitly defined, scope creep is visible. It can be measured, audited, and challenged. In a centralised system, access expands silently because the data is already there. In a brokered system, every expansion of access is a policy decision that leaves a trace.

the quiet inversion...

There is a perception, widely held and not entirely unreasonable, that AI in the hands of the state is a surveillance technology. The history of government data systems, from GCHQ's bulk interception programmes to the Home Office's handling of Windrush, gives people ample cause for scepticism. The default assumption is that more AI means less privacy, that anything which makes it easier for the state to access data will inevitably be used to extend its reach into people's lives.

The brokered model inverts that assumption. It doesn't just make government more efficient, it makes the citizen more powerful. A personal agent that brokers, logs, and challenges data requests on the citizen's behalf is not a surveillance tool but a constitutional safeguard. It is the data equivalent of the right to legal representation – an agent that acts in your interest, understands the rules, and keeps a record. In a post-Windrush, post-Horizon landscape where the state's data systems have repeatedly been wielded against individuals, giving every citizen their own AI broker is more than a quirky technology add-on. It's a long-overdue fundamental democratic corrective.

Consider the asymmetry that currently exists. The state has vast computational resources, access to linked databases across multiple agencies, and the institutional capacity to process millions of records. The individual has a Subject Access Request form and six to eight weeks of waiting. The brokered model doesn't eliminate that asymmetry, but it narrows it dramatically. A citizen's broker armed with the legal rights the citizen already possesses, operating at machine speed, with a complete and independently verifiable audit trail, fundamentally changes the balance of power in the relationship between the individual and the institution.

This is far from being an anti-government argument. Effective governance requires effective data sharing. The police need to connect intelligence across forces. Clinicians need to see patient records across providers. Financial regulators need to trace money across institutions. The question has never been about whether data should be shared, but *how*, and under whose control. The brokered model addresses that question in a way that serves both sides of the equation. The state gets the intelligence it needs and the citizen gets a society that works, based on a framework that honours their statutory rights, and a governance architecture that, probably for the first time in British history, acknowledges them as equals.

At a moment when the public sector is under sustained pressure to do more with less, to modernise creaking infrastructure without ballooning budgets, and to restore public trust in institutions that have been haemorrhaging it for decades, a governance model that simultaneously improves operational capability, reduces integration costs, strengthens civil liberties, and creates a visible audit trail is not just an attractive proposition, it's a rare convergence of institutional self-interest with the public good.

The contractor model makes the state more efficient and the citizen more confident in their citizenship - a confidence rooted, moreover, in the profound conviction that in a democracy, governance should work for the citizen, and not the other way around.

This paper extends the contractor model developed in The Unpaid Contractor. It has been informed and inspired by conversations with serving public-sector professionals. If this is your field, or you have a particular interest in getting this right, we would welcome your feedback.

HACCU.AI