

From Chaos to Coherence

**Private AI Agents, Anonymous Collective Intelligence, and the
Inversion of Consumer Information Asymmetry**

Dragonfly Labs · Algorythmics · April 2026

Abstract

Every company you deal with has a data warehouse, a pricing algorithm, and a team of people professionally optimised to extract the maximum from you. You have a shoebox of paperwork and a vague sense that you're overpaying for something.

This paper describes an architecture that inverts this asymmetry. **Layer 1** is a private AI agent — encrypted, persistent, working entirely in the user's interest — that transforms the chaos of personal finances and administration into coherent, actionable knowledge. **Layer 2** is a collective intelligence network where these agents anonymously pool patterns and return insights that no individual could discover alone.

The result is a system where the act of organising your own life — the personal, mundane, selfish use case — generates, as a structural by-product, the collective bargaining power that the consumer side of the market has never had.

This is not a new idea. Doc Searls articulated it in 2006. The technology to build it didn't exist until now.

1. The Asymmetry

Every company you interact with — your bank, your insurer, your energy provider — operates a Customer Relationship Management system. CRM is a disciplined, data-rich, algorithmically optimised practice for understanding what each customer will tolerate and pricing accordingly. Churn prediction models estimate when you'll leave. Loyalty pricing algorithms calculate how much they can raise your renewal before you notice. Retention teams are deployed when the model flags you as a flight risk.

The customer side has no equivalent. No tool aggregates your own data, remembers your own history, or advocates for your own interests with anything resembling the sophistication that the vendor side deploys against you. You act alone, with incomplete information, against organisations professionally optimised to extract surplus from your inattention.

Akerlof described this information asymmetry in 1970 [1]. Stiglitz and Rothschild formalised screening and signalling in 1976 [2]. Zuboff named the modern variant surveillance capitalism in 2019 [3]. But the structural fact is older than any of them: the vendor side has tools, data, and algorithms. The consumer side has nothing.

2. The VRM Lineage

Doc Searls saw this clearly. In 2006, as a fellow at Harvard’s Berkman Center for Internet & Society, he launched ProjectVRM — Vendor Relationship Management — as the deliberate inverse of CRM [4]. Where CRM gives vendors tools to manage customers, VRM would give customers tools to manage vendors. Searls and his co-authors had already written in *The Cluetrain Manifesto* (1999): “We are not seats or eyeballs or end users or consumers. We are human beings — and our reach exceeds your grasp” [5].

In 2012, Searls published *The Intention Economy*, arguing that markets should grow around buyers, not sellers: “The Intention Economy is about markets, not marketing. You don’t need marketing to make Intention Markets” [6].

The vision was right. The technology was not ready.

ProjectVRM spawned nineteen development efforts — Azigo, Mydex, EmanciPay, SwitchBook, The Mine Project, and others. None survived as consumer products. They shared a common failure mode: they asked the user to do the work. Manually input your bills. Manually categorise your spending. Manually manage your vendor relationships using a clunky web form. The value proposition was sound but the interaction cost was ruinous. Nobody wants to be their own CRM administrator.

Three things that existed in theory in 2006 now exist in practice:

1. **Large language models** that can read an unstructured bank statement, an email from an insurance company, or a PDF utility bill and extract structured, actionable knowledge — without the user lifting a finger.
2. **Commodity encryption** (LUKS2, Argon2id, AES-XTS-plain64) that provides military-grade per-user encrypted vaults at near-zero marginal cost.
3. **Proven federation models** (Mastodon, ActivityPub, Matrix) demonstrating that decentralised networks can operate at scale without a central platform, central authority, or single point of capture.

Searls was twenty years early. The idea was waiting for its technology. The technology is here.

3. The Optimisation Problem

The consumer information asymmetry described above is an instance of a deeper structural problem: single-variable optimisation at scale.

Every system that collapses to a single metric — engagement, yield, click-through, quarterly earnings — works exactly as designed while eroding the context it depends on. The attention economy optimises for time-on-screen and produces anxiety. Loyalty pricing optimises for retention revenue and produces consumer inertia. Comparison sites optimise for referral fees and produce rankings that serve providers, not users.

The consumer sees none of these objective functions. They experience the outputs — rising premiums, buried cancellation flows, auto-renewed contracts, algorithmically personalised pricing — without seeing the machinery. The asymmetry is not just informational. It is computational. One side has optimisation systems. The other side has human memory and a pile of unopened emails.

The architecture described in this paper does not fight optimisation. It gives the consumer side an optimisation system of its own — one whose objective function is declared, auditable, and aligned entirely with the user's interests. Not a single metric, but a multi-dimensional one: cost, time, complexity, risk, renewal timing, provider quality. The vendor side already optimises across multiple dimensions. The consumer side should too.

4. From Chaos to Coherence

Most people’s administrative lives exist in a state of low-grade chaos. Bills arrive by email and are forgotten. Bank statements are downloaded once and never opened again. Insurance renewal dates pass unnoticed. Subscriptions accumulate through inattention. Direct debits continue for services no longer used.

This chaos is not laziness. It is the entirely rational response to an environment that generates more administrative complexity than any individual has the bandwidth to manage. The vendor side understands this perfectly — loyalty pricing, auto-renewal defaults, and buried cancellation processes are all designed to exploit the gap between what you should do and what you have time to do.

The architecture described here addresses this with two layers that serve different purposes but share a single mechanism.

Layer 1 is a private AI agent — called Bob — that transforms personal chaos into coherent knowledge. It reads your email, imports your bank statements, extracts renewal dates, classifies spending, tracks subscriptions, and sends a configurable daily briefing with what matters. The interaction model is not a dashboard you must visit or a spreadsheet you must maintain. It is a conversation. “What am I paying for streaming?” “When does my car insurance renew?” “Show me all transactions over £100 last month.” The agent knows, because it has all your data in one place, cross-referenced and persistent.

The key design decision: the user does no administrative work. The agent reads the email, parses the bank statement, extracts the deadline. The user’s job is to read a ten-second briefing and make decisions. This is what killed every VRM tool before it — they asked the user to be the data entry clerk. Bob is the data entry clerk.

Layer 2 is the Swarm — a collective intelligence network that emerges as a structural by-product of Layer 1. As Bob organises your data, it extracts anonymised, bucketed traits — category, provider, cost band, tenure band — and emits them to a shared intelligence layer. No names. No addresses.

No account numbers. No exact amounts. When enough agents contribute to a cohort, the Swarm synthesises insights and returns them to each agent.

The Swarm creates something that has never existed: an addressable, organised consumer population — anonymous but reachable through their agents. This changes the incentives for every player in the market. A dominant energy provider with high churn in Swarm cohort data cannot ignore the signal. Their competitors cannot ignore it either — because the Swarm makes it trivially easy to surface alternatives to hundreds of thousands of agents simultaneously. The collective leverage is irresistible to both sides. The incumbent must respond to retain. The challenger must respond to acquire. The consumer wins either way, without lifting a finger.

The dual proposition is this: you sign up because you want a personal assistant that handles the boring stuff. That's the selfish use case — entirely sufficient on its own. But the same process that organises your personal chaos generates, without additional effort, the anonymous data that powers collective intelligence for everyone. **From chaos to coherence, and from coherence to collective power.** The first is the reason you sign up. The second is what changes the game.

5. Architecture

5.1 The Private Agent

Each user runs their own isolated instance. There is no shared database, no multi-tenant backend, no cross-user data access. Each instance comprises:

- **An encrypted volume** (LUKS2, AES-XTS-plain64, 512-bit key, Argon2id key derivation with 1 GB memory cost). The user's passphrase is the only key. It is never stored, transmitted, or logged. If the user forgets it, the data is gone. This is the point.

- **A tool-calling AI agent** with persistent memory, connected to the user's email (Gmail, with others planned) and capable of ingesting bank statements (CSV), PDFs, and documents. The agent calls any LLM via API — it is model-agnostic by design, not locked to a single provider. It uses structured tool invocation — not prompt hacks — to read, write, classify, and query the user's data.
- **A configurable daily briefing system** that emails the user with key actions, upcoming costs, calendar events, renewal alerts, and Swarm insights.

The agent currently provides 21 tools spanning bank statement import, spending analysis, subscription tracking, renewal monitoring, memory management, document filing, email processing, and Swarm interaction. It is a production system, not a prototype.

5.2 Privacy by Construction

Privacy in this architecture is not a policy. It is a structural property.

Property	Implementation
Encryption rest	at LUKS2, AES-XTS-plain64, 512-bit, Argon2id (1 GB memory, 4 threads)
Passphrase storage	Never stored. Zero-knowledge at rest.
Isolation	One encrypted container per user. No shared database.
Network egress	Whitelisted: AI model, email provider, Swarm relay. Nothing else.
Volume locking	On shutdown: unmount + cryptsetup luksClose. Data becomes ciphertext.
Auditability	Open source (AGPL-3.0). Network egress verifiable by inspection.
Deletion	Cryptographic erasure. Volume destroyed. Verifiably gone.

When the user's instance is stopped, their data is indistinguishable from random noise. Nobody — not the operator, not a system administrator, not a court order — can read it without the passphrase. While the instance is running, data is decrypted in memory to serve requests. This is the same trust model as Signal, ProtonMail, and every other server-side service that handles encrypted data. The honest statement is: we cannot access your data at rest. While running, the trust boundary is the server operator, exactly as with any hosted service. Self-hosting eliminates this trust requirement entirely.

5.3 The Swarm

The Swarm is an anonymous collective intelligence layer that operates on traits, not identities.

Trait emission. The user's agent extracts bucketed, anonymised traits from their data. A trait might be: { category: "energy", provider: "octopus", cost_band: "£100-150", tenure: "2-5y" }. Categories, providers, cost bands. Never exact amounts, never dates, never names, never addresses, never postcodes, never account numbers. The bucketing is not a policy choice — it is enforced in code before any data leaves the user's instance.

Rotating identity. Each agent's Swarm identity is an HMAC-SHA256 hash of a shared secret and a monthly nonce. Bot IDs rotate every month. The Swarm cannot link this month's traits to last month's. The operator cannot link a bot ID to a user account.

k-Anonymity. Cohort insights are only returned when the cohort contains at least k agents (currently k = 5). Small cohorts are suppressed entirely to prevent re-identification through unique trait combinations [7].

AI synthesis. When cohort density is sufficient, the Swarm runs AI synthesis across the cohort and returns structured insights: provider comparisons, switching statistics, cost anomalies, emerging patterns. The human never sees another human. There is no feed, no profile, no social graph, no messaging, no likes, no comments. Your bot works the crowd on your behalf and brings you intelligence.

Federation. The Swarm follows a hub-optional federation model. The public hub at `swarm.algorythmics.life` is one instance. Anyone can run their own hub — `uvicorn swarm.app:app --port 8803` — and point their agents at it. Hubs can federate: swarms connecting to swarms, sharing aggregated (never individual) trait data across instances. If the public hub disappears, every agent still works. This is the Mastodon model applied to collective intelligence rather than social networking.

Consent. Swarm participation is opt-in. Disabled by default. The user enables it explicitly. When they do, the agent proffers a formal data-sharing agreement on the user’s behalf following the IEEE 7012 standard for machine-readable personal privacy terms [8]. The user controls the terms. They can revoke at any time.

6. The Objective Function

Every AI system has an objective function. Most are hidden. Many are adversarial.

Social media optimises for engagement — time-on-screen, clicks, reactions — because that is what the advertising model rewards. The user’s well-being is not in the loss function. Comparison sites optimise for referral revenue — the provider that pays the highest commission appears first. The user’s cost reduction is a side effect, not the objective. Financial advisers optimise for assets under management. Loyalty programmes optimise for retention through switching costs.

In each case the declared purpose (“connect with friends,” “find the best deal,” “grow your wealth”) diverges from the actual objective function (“maximise ad impressions,” “maximise referral fees,” “maximise management fees”). The user senses this divergence but cannot see the function.

Bob’s objective function is declared, singular, and auditable:

Reduce the cost, complexity, and wasted time of being a consumer.

Bob evaluates “moves” — like a chess engine scoring positions — weighted by confidence and impact. Switch energy provider. Cancel a forgotten subscription. Lock in a renewal before a price rise. Flag an anomalous charge. Bob does the analysis, using its knowledge of the user’s data and, where available, Swarm intelligence from the cohort. The Swarm’s role is distinct: it aggregates anonymous patterns and synthesises collective insights. Bob acts on them. The agent is the decision engine. The Swarm is the intelligence layer.

What the system will never do: optimise for engagement, attention, or advertising revenue. There is no feed to scroll. There is no algorithm nudging you to stay online. There is no ad layer. The ideal interaction is a briefing that takes ten seconds to read and saves £480. Less time spent is better. Less friction is the metric. Reducing waste — of money, of time, of attention — is the only objective.

This is not social media. It is the opposite of social media. Social media captures your attention and sells it. Bob returns your attention and tells you what to do with it.

7. Network Effects Without Surveillance

Every successful platform has a network effect: more users → more value per user. In social media, the network effect depends on identity. You join because your friends are there. The value is in the social graph. This creates lock-in, because leaving means losing your connections.

The Swarm’s network effect depends on data density in cohorts, not identity. More agents contributing energy traits → sharper energy insights for all. More agents contributing insurance traits → better renewal intelligence. The value scales with participation, but participation is anonymous and the user has no identity to lose. There is no switching cost. There is no lock-in. If a better hub exists, point your agent at it.

This is a different kind of scale. One person overpaying for energy is invisible to the market. A million people overpaying for energy — organised, anonymous, addressable through their own agents — is a signal that companies must respond to. Not because we sell the data. Not because we broker the deals. But because organised demand, even when anonymous, is impossible to ignore.

The comparison is to collective bargaining, but without the coordination cost. A trade union requires meetings, votes, representation structures, and human organisers. A buying cooperative requires membership fees and governance. A price comparison site requires you to visit it, enter your details, and trust its rankings. The Swarm requires nothing from the user. Their agent handles it. The collective forms automatically from the by-products of individual coherence.

8. Resilience

The architecture described here is designed to be structurally resistant to capture — by the operator, by a platform, by a government, or by market pressure.

Bob works offline. If the Swarm, the public hub, the email provider, and the internet all disappear, your agent still works. Every tool, every briefing, every query — all operate on local data in a local database on an encrypted volume. The Swarm is an overlay, not a dependency.

The Swarm is federated. Anyone can run a hub. Anyone can point their agents at a different hub. Hubs can exchange aggregated data. The topology is a mesh, not a star. There is no central server that, if seized or shut down, kills the network.

The code is open. AGPL-3.0. If the operator disappears, the code exists. Anyone can fork it, run it, improve it. The licence requires that modifications are shared, preventing proprietary capture.

The privacy model is structural, not promissory. Encryption is not “we promise not to read your data.” It is “we mathematically cannot read your data when your instance is stopped.” Swarm anonymity is not “we promise not to identify you.” It is “the system does not collect the information required to identify you.” This distinction matters. Promises can be broken. Structures cannot — short of rewriting the code, which is open and auditable.

There is no data moat. The conventional platform strategy is to accumulate proprietary data that creates a competitive moat. The Swarm accumulates anonymous, bucketed, rotating-identity traits. This data has collective value but no individual commercial value. It cannot be sold, because it contains no PII. It cannot create lock-in, because the user’s real data stays on their encrypted volume. The moat, if one exists, is the network itself — and the network is decentralised.

9. What This Is Not

It is worth being explicit about what this system is not, because the space is cluttered with things that sound similar but are structurally different.

It is not social media. There is no human interaction. No profiles. No feeds. No likes. No followers. No attention economy. The system is not about connecting people to each other about arbitrary things. It is about building value together — anonymously, automatically, as a by-product of individual utility.

It is not a comparison site. Comparison sites are paid by providers. The provider that pays the highest referral fee gets the top ranking. The user’s interests and the site’s revenue model are structurally misaligned. This system takes no referral fees. It works for the user.

It is not a financial adviser. It surfaces what people in similar positions did. It does not tell you what to do. The distinction is between intelligence and advice — the former informs, the latter directs.

It is not a data broker. Personally identifiable information never leaves the user’s instance. The Swarm sees anonymous, bucketed traits. There is nothing to sell, even in principle.

It is not an advertising platform. There is no ad layer. This is structural, not a policy that could change at a future funding round. The revenue model is subscription, not attention.

10. Related Work

The ideas in this paper draw on a long intellectual lineage.

Information asymmetry. Akerlof’s “Market for Lemons” [1] established that quality uncertainty in markets leads to adverse selection. Stiglitz and Rothschild [2] formalised screening and signalling. The consumer information asymmetry described here is a specific, pervasive instance of these dynamics.

Vendor Relationship Management. Searls’ ProjectVRM at Harvard Berkman Center [4], The Cluetrain Manifesto [5], and The Intention Economy [6] are the direct intellectual ancestors of this work. Searls articulated, with precision and force, the need for customer-side tools as powerful as vendor-side CRM. Customer Commons, born from ProjectVRM, continues this advocacy [9].

Surveillance capitalism. Zuboff [3] named the practice of extracting behavioural surplus from users and selling prediction products to third parties. This system is designed as a structural response to that practice: data that cannot be extracted because it is encrypted, identities that cannot be surveilled because they rotate, and a business model that has no advertising layer.

Privacy-preserving computation. Differential privacy [10] and k-anonymity [7] provide the formal foundations for the Swarm’s trait bucketing and cohort suppression. Federated learning [11] shares the goal of extract-

ing collective value from distributed private data, though the mechanism here (anonymous trait emission) is simpler and more transparent than gradient-based approaches.

Federation. The Mastodon/ActivityPub model [12] proves that decentralised social infrastructure can operate at meaningful scale without a central platform. The Swarm applies the same federation principle to collective intelligence rather than social networking.

11. Deployment

The system described in this paper is not theoretical. It is instantiated and running. Users are onboarded. Traits flow to the Swarm. Daily briefings are sent. Bank statements are imported and classified. The architecture has been security-audited and the privacy model verified against the implementation.

The codebase is open source under AGPL-3.0. The encryption parameters, the Swarm protocol, the trait bucketing logic, and the network egress whitelist are all inspectable. The commitment is: do not trust us — inspect the code.

12. Conclusion

Doc Searls was right, and he was twenty years early. The consumer side of the market has never had tools as sophisticated as the vendor side. The information asymmetry is structural, pervasive, and — until recently — technologically intractable from the consumer's position.

What changed is not ideology. What changed is capability. Large language models can now do what no previous software could: read an unstructured email from an insurance company, a CSV bank statement, a PDF utility bill,

and turn it into structured, actionable knowledge without the user doing anything. Encryption that was expensive is now commodity. Federation that was theoretical is now proven.

The architecture described here exploits this convergence. Layer 1 transforms personal administrative chaos into coherence — a simple, selfish, valuable use case that justifies adoption on its own. Layer 2 harvests the anonymous by-products of that coherence and builds collective intelligence that no individual could generate alone.

The mechanism is a happy dual: you get a personal assistant that handles the boring stuff. The Swarm gets smarter. You get better insights. The market gets a signal it has never had before — organised consumer demand, anonymous but unmistakable. The more people who join, the sharper the intelligence, the stronger the signal. And the network effect doesn't depend on identity, attention, or surveillance. It depends on data density in cohorts. A different kind of scale for a different kind of value.

Nothing about this system fights the market. It simply gives the other side the same tools. That's all. And, perhaps, that's enough.

References

- [1] Akerlof, G. A. (1970). "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, 84(3), 488-500.
- [2] Rothschild, M. & Stiglitz, J. E. (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics*, 90(4), 629-649.
- [3] Zuboff, S. (2019). *The Age of Surveillance Capitalism*. Profile Books.
- [4] Searls, D. (2006). *ProjectVRM*. Berkman Klein Center for Internet & Society, Harvard University. projectvrn.org.
- [5] Levine, R., Locke, C., Searls, D. & Weinberger, D. (1999). *The Cluetrain Manifesto*. cluetrain.com.

[6] Searls, D. (2012). *The Intention Economy: When Customers Take Charge*. Harvard Business Review Press.

[7] Sweeney, L. (2002). "k-Anonymity: A Model for Protecting Privacy." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5), 557-570.

[8] IEEE 7012-2024. Standard for Machine-Readable Personal Privacy Terms.

[9] Customer Commons. customercommons.org. Founded 2012, born from ProjectVRM at Harvard.

[10] Dwork, C. (2006). "Differential Privacy." *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*, 1-12.

[11] McMahan, H. B. et al. (2017). "Communication-Efficient Learning of Deep Networks from Decentralized Data." *Proceedings of AISTATS*.

[12] Mastodon / ActivityPub. W3C Recommendation (2018). activitypub.rocks.

© 2026 Algorithemics. This paper is published under CC BY-NC-ND 4.0. Bob is open source under AGPL-3.0. The Swarm protocol is open. Inspect the code.

bob.algorithemics.life · swarm.algorithemics.life · algorithemics.life