Journal of Secure Digital Technologies
Volume 1 (№2) September 2026

ISSN (E): 3136-9340
Sertificate № KZ19VPY00128988, date of issue: 25.12.2025

The journal is published 4 times a year
(March, June, September, December)

EDITORIAL TEAM
Editor-in-Chief
Dr. Leila Rzayeva, Association prof., PhD, Digital government support center (Kazakhstan)

Deputy Editor-in-Chief
Dr. Ali Myrzatay, PhD, Digital government support center (Kazakhstan)

Executive Secretary
Dr. Abdenova Zhuldyz, PhD, Digital government support center (Kazakhstan)

EDITORIAL BOARD MEMBERS (KAZAKHSTAN)
Seitkulov E.N. – Candidate of Physical and Mathematical Sciences, Professor Computer Science, Computer Engineering and Control
Abdenova G.A. – PhD in Automation and Control, Candidate of Technical Sciences, Professor L.N. Gumilyov Eurasian National University
Atanov S.K. – Doctor of Technical Sciences, Professor.
Sergazin G.K. – PhD, Associate Professor Computer Science, Computer Engineering and Control
Satybaldina D.Zh. – Candidate of Physical and Mathematical Sciences, Professor Director, Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University

INTERNATIONAL EDITORIAL BOARD MEMBERS
Ibraheem Shayea – PhD in Electrical and Electronic Engineering Istanbul Technical University (Turkey)
Yadigar Nasiboglu Imamverdiyev – DSc in Technical Sciences, Associate Professor Information Security, Computer Science and Computer Technology (Azerbaijan)
Erdal Irmak – PhD in Electrical and Electronics Engineering, Professor Gazi University (Turkey)
Korhan Kayisli – PhD in Electrical Electronics Engineering, Associate Professor Gazi University (Turkey)
Yonghao Wang – PhD in Electronic Engineering, Associate Professor Queen Mary, University of London (UK)
Amir Mosavi – PhD in Applied Informatics, Professor of AI Obuda University (Hungary)
Gerald Feldman – PhD in General Computing, Associate Professor Birmingham City University (UK)
Mohamed Ragab – PhD in Computer Science, Assistant Professor Birmingham City University (UK)
Seref Sagiroglu – PhD in System Engineering, Professor Gazi University (Turkey)

MANAGING EDITOR
Dr. Ali Myrzatay, PhD, Digital government support center (Kazakhstan)

TECHNICAL EDITOR
Dr. Ali Myrzatay, PhD, Digital government support center (Kazakhstan)

CONTENTS

Cyber Incident Damage Assessment for Public Digital Systems and Critical Infrastructure 3

Cyber Incident Damage Assessment for Public Digital Systems and Critical Infrastructure 37

Verifiable but Not Correct: Blockchain Trust Boundaries for Autonomous Agents 71

Calibrated Digital Twins and Geometry-based Perception for Automated Driving in Heavy Rain and Snowfall: Evidence From Research Testing in China and Implications for Kazakhstan. 88

Personal Data Protection in Kazakhstan's AI Enabled Public Services 109

Asking Another Jurisdiction What It Holds Without Telling It What You Hold: Private Artifact Matching with a Ledger-Anchored Transcript 131

On-Chain and Off-Chain Data Placement in Hybrid Blockchain Health Systems 143

An Integrated Project Management Maturity Model for Digital Qazaqstan Initiatives: Embedding Cybersecurity, AI Governance and Forward-Looking Cyber-Risk Requirements into ST RK ISO 21502-2022..... 158

Appendix A. Indicators and evidence for the baseline practice groups 181

Privacy Risks and Data Protection Readiness in AI Marketing in Kazakhstan..... 183

Supplementary Materials..... 205

Yerassyl Aikenov*
Master's student, School of Cybersecurity
255374@astanait.edu.kz, ORCID: https://orcid.org/0009-0003-9287-7183
Astana IT University, Kazakhstan

ON-CHAIN AND OFF-CHAIN DATA PLACEMENT IN HYBRID BLOCKCHAIN HEALTH SYSTEMS

Abstract: Which clinical data belongs on the ledger? Hybrid blockchain designs answer class by class, quietly, and the answer usually hardens into habit before anyone writes down why. We formalise it. Placement becomes constrained optimisation over storage tiers — each carrying its own verifiability, erasability, exposure, read latency and ledger footprint — and over clinical data classes carrying regulatory and service-level duties. The problem turns out to be a multiple-choice knapsack instance, hence NP-hard; compliance then prunes it so hard that the hardness stops mattering. Kazakhstan supplies the setting, and its statutory duties supply the constraints. Nine FHIR-aligned classes, six tiers, one worked instance. Regulatory and latency limits alone shrink 10,077,696 candidate assignments to 48; a ledger budget takes it to 12; and every uniform placement policy in common use lands outside that region entirely. The optimiser then buys per-item commitments for seven classes and batched commitments for the two carrying the heaviest instance counts — volume drives that split, not sensitivity. Re-running the model ten thousand times with every estimated parameter perturbed by up to twenty-five per cent left the pruning and the infeasibility of uniform policies intact in every draw, while the batching assignment held in 76.1 per cent of draws. Parameters are illustrative and tabulated in full, and every number reported here comes out of the model rather than off a deployment.

Keywords: *hybrid blockchain; electronic health records; data placement; multiple-choice knapsack; right to erasure; design space analysis; HL7 FHIR*

Размещение Данных в Реестре и Вне Его в Гибридных Медицинских Блокчейн-Системах

Аннотация: Какие клинические данные должны попадать в реестр? Гибридные блокчейн-архитектуры отвечают на этот вопрос класс за классом, негласно, и ответ обычно закрепляется как привычка раньше, чем кто-либо записывает его обоснование. Мы формализуем этот выбор. Размещение становится задачей оптимизации с ограничениями над уровнями хранения, каждый из которых характеризуется своей проверяемостью, возможностью удаления, степенью раскрытия, задержкой чтения и объёмом в реестре, и над классами клинических данных, несущими регуляторные обязательства и требования к уровню обслуживания. Задача оказывается экземпляром задачи о многовариантном рюкзаке и потому NP-трудна; однако требования соответствия отсекают пространство

решений настолько сильно, что эта трудность перестаёт иметь значение. Контекст задаёт Казахстан, а его законодательные требования задают ограничения. Рассмотрены девять классов данных, согласованных с FHIR, шесть уровней хранения и один рабочий пример. Одни только регуляторные ограничения и ограничения по задержке сокращают 10 077 696 возможных назначений до 48; бюджет реестра оставляет 12; и каждая распространённая единообразная политика размещения оказывается целиком за пределами этой области. Оптимизатор затем выбирает поэлементные обязательства для семи классов и пакетные — для двух классов с наибольшим числом экземпляров: это разделение определяет объём данных, а не их чувствительность. Десять тысяч повторных прогонов модели при отклонении каждого оценочного параметра до двадцати пяти процентов сохранили отсечение и недопустимость единообразных политик во всех прогонах, а распределение пакетирования — в 76,1 процента прогонов. Параметры носят иллюстративный характер и приведены полностью, а все числа получены из модели, а не из реального развёртывания.

Ключевые слова: *гибридный блокчейн; электронные медицинские записи; размещение данных; задача о многовариантном рюкзаке; право на удаление; анализ пространства проектных решений; HL7 FHIR*

Медициналық Гибридті Блокчейнде Деректерді Тізілімде және Одан Тыс Орналастыру

Аңдатпа: Қандай клиникалық деректер тізілімге түсуі керек? Гибридті блокчейн архитектуралары бұл сұраққа әр класс бойынша, үнсіз жауап береді, ал бұл жауап әдетте оның негіздемесі жазылмай тұрып-ақ әдетке айналып кетеді. Біз осы таңдауды формалдаймыз. Деректерді орналастыру сақтау деңгейлері бойынша шектеулері бар оңтайландыру есебіне айналады: әр деңгейдің өз тексерілу дәрежесі, жою мүмкіндігі, ашылу дәрежесі, оқу кідірісі және тізілімдегі көлемі бар, ал клиникалық деректер кластары нормативтік міндеттер мен қызмет көрсету деңгейіне қойылатын талаптарды көтереді. Есеп көп нұсқалы рюкзак есебінің дербес жағдайы болып шығады, демек NP-күрделі; алайда сәйкестік талаптары шешімдер кеңістігін соншалықты қатты қысқартатындықтан, бұл күрделіліктің іс жүзінде маңызы қалмайды. Зерттеудің контексін Қазақстан анықтайды, ал оның заңнамалық міндеттері шектеулерді құрайды. FHIR стандартымен үйлестірілген тоғыз деректер класы, алты сақтау деңгейі және бір жұмыс мысалы қарастырылды. Тек нормативтік және кідіріс шектеулерінің өзі 10 077 696 ықтимал тағайындауды 48-ге дейін қысқартады; тізілім бюджеті оларды 12-ге дейін азайтады; ал кең таралған бірыңғай орналастыру саясаттарының әрқайсысы осы аймақтан толықтай тыс қалады. Оңтайландырушы жеті класс үшін жеке міндеттемелерді, ал даналар саны ең көп екі класс үшін топтамалық міндеттемелерді таңдайды: бұл бөліністі деректердің

сезімталдығы емес, олардың көлемі анықтайды. Әрбір бағаланған параметр жиырма бес пайызға дейін өзгертілген он мың қайталама есептеу шешімдер кеңістігінің қысқаруын және бірыңғай саясаттардың жарамсыздығын барлық жағдайда сақтады, ал топтамалау бөлінісі есептеулердің 76,1 пайызында сақталды. Параметрлер иллюстрациялық сипатта және толық келтірілген, ал барлық сандар нақты ендіруден емес, модельден алынған.

Түйін сөздер: *гибридті блокчейн; электрондық медициналық жазбалар; деректерді орналастыру; көп нұсқалы рюкзак есебі; деректерді жою құқығы; жобалау кеңістігін талдау; HL7 FHIR*

1. Introduction (Literature Review)

Placement decides the rest. Once an architect has settled which clinical data classes go on a ledger, which go off it behind a commitment, and which never touch a ledger at all, the compliance posture, the storage bill and the audit story are largely determined — and reversing the choice later means migrating live patient records, which nobody does willingly.

The literature treats this decision unevenly. Payloads off-chain, hashes on-chain — proposals state it, then hurry on to consensus tuning or key management, as though the split were self-evident. It is not. Ask what each option actually proves. A hash on a permissioned ledger, a Merkle commitment anchored in public, and a batched commitment covering sixty-four records differ in what they demonstrate, what they leak, what they cost, and whether anyone can ever undo them. Health data is exactly where those differences bite: statutory retention pulling one way, erasure duties the other [1], [2].

Kazakhstan gives this work its concrete setting. Clinical documentation across the national system already runs electronically; the Code on Public Health defines personal medical data and the aggregator role that handles it [3]; a WHO-supported programme has been reshaping statutory forms toward a digital-friendly information model [4]. Walk a ledger-based service into that environment and it inherits the whole obligation set before a single design choice gets made.

This work converts those obligations into constraints and the architectural preferences into an objective, then asks what remains. The contribution is fourfold. A tier model separates the properties an architect actually trades off, rather than the platform labels usually compared. Regulatory duties, audit duties and service-level targets become hard constraints over that model. The resulting placement problem turns out to be a multiple-choice knapsack instance, which locates it precisely in a well-studied complexity class. And an instantiation on nine FHIR-aligned classes reports a design-space contraction of five orders of magnitude, together with an optimal placement that no uniform policy reproduces.

Three strands bear on the problem, and they have largely developed apart from one another.

1.1 Deciding Whether and Where to Use a Ledger

Wüst and Gervais handed the field its decision flowchart: do you need a blockchain at all, or will a managed database do [5]? Eberhardt and Tai catalogued the off-chaining patterns whose trade-offs we formalise here [6]. Xu and colleagues built a taxonomy meant to carry architectural reasoning [7]. Useful work, all of it. It classifies. Nobody optimises, and nobody carries regulatory duty into the constraint set.

1.2 Ledger Architectures for Health Records

MedRec put permission management on a public ledger [8]. Ancile added proxy re-encryption and interoperable access control [9]; BBDS and MedBlock went after cloud-resident sharing [10], [11]; Healthcare Data Gateways moved the patient to the centre of risk control [12], and later work carried that pattern into mobile-cloud settings [13]. Permissioned substrates followed — Fabric as the usual base [14], PREHEALTH bolting on anonymous credentials [15], latency behaviour by now reasonably well charted [16]. Off-chain custody hardened into practice through IPFS-backed designs [17], [18], [19], [20], [21], and reviews map what resulted [22], [23], [24]. Every one of these systems fixes a placement. Not one derives it.

1.3 Mutability and Erasure

Ateniese and colleagues opened the redactable-blockchain line with chameleon hashing [25]; Derler et al. sharpened rewriting down to attribute-level policies [26]; Deuber et al. threw out trapdoor custody in favour of on-chain voting [27]. Politou's group documents the legal pressure driving all of it [28], [2]. We do not treat redactability as the answer. It is one tier among six, carrying its own price and its own weakened verifiability.

The remainder of the paper is organized as follows. Section 2 sets out the system model, the problem formulation, the solution method and the experimental setup. Section 3 reports the results, Section 4 discusses their implications and the threats to validity, and Section 5 concludes.

2. Methodology

2.1 System Model

Let I index clinical data classes and T index storage tiers. A placement is a map $x : I \to T$ assigning every class to exactly one tier.

2.1.1 Data Classes

Class i carries eight parameters. Payload size s_i, instances per patient-year n_i, reads per instance f_i. Then the duties: a required verifiability level θ_i, a required erasability level e_i, and a ledger-presence flag p_i marking classes whose existence has to be independently attestable. Finally a read-latency ceiling λ_i and a sensitivity weight σ_i, the latter standing in for how much damage a disclosure would do.

Two deserve comment. Erasability and ledger presence fight each other by construction — consent grants and access log entries have to survive for audit, so they carry

$p_i = 1$ and $e_i = 0$, while ordinary clinical records carry the reverse. Uniform placement dies right there. The model puts that tension on the table instead of burying it.

2.1.2 Storage Tiers

Tier t carries a ledger cost rate κ_t, a verifiability level ver_t, an erasability level era_t, a presence flag $pres_t$, an exposure coefficient ρ_t, a read latency lat_t, and a per-instance ledger footprint. Table 2 lists the six tiers used here. They span the practical range: full payload on a public chain, full payload on a permissioned chain, redactable payload via chameleon hashing [25], [26], off-chain custody with a per-item commitment, off-chain custody with commitments batched into a Merkle tree of sixty-four [29], and off-chain custody with no ledger record whatsoever.

Batching is where the interesting economics live. A batched commitment costs one sixty-fourth of the ledger bytes but proves membership at coarser granularity, so an auditor learns that a record belonged to a batch rather than that it stood alone. The model encodes that difference as ver = 0.75, against 0.95 for per-item commitments.

2.1.3 Cost Components

Three costs and one quality loss feed the objective. Storage combines ledger and off-chain bytes at their respective rates. Read latency piles up as $n_i f_i lat_t$, disclosure exposure as $n_i \sigma_i \rho_t$. Verification granularity loss accumulates as $n_i \sigma_i (1 - ver_t)$, which docks the cheap tiers for weakening what an auditor can establish. Each component gets normalised by its maximum over $I \times T$. That puts the four weights on one scale and lets a reader treat them as priorities rather than as unit conversions.

2.2 Problem Formulation

2.2.1 Hard Constraints

Compliance does not belong in an objective function. A regulator does not accept a slightly cheaper design that is slightly less lawful, so duties enter as feasibility constraints, and cost operates only inside the region they leave.

C1 verifiability. $ver_t \geq \theta_i$. Integrity of the class must be demonstrable to an external auditor at the level the class demands.

C2 erasability. $era_t \geq e_i$. Classes subject to withdrawal of consent must sit where the data can be rendered irrecoverable [1], [2], through destruction of the key material that protects it [30].

C3 ledger presence. $pres_t = 1$ whenever $p_i = 1$. Audit-bearing classes must leave a trace no operator can silently drop.

C4 latency. $lat_t \leq \lambda_i$. Interactive classes cannot be parked behind public-chain confirmation.

C5 ledger budget. $\Sigma_i L(i, x(i)) \leq B$, where L is the annual on-ledger footprint of class i under its tier and B is the budget per patient-year. This is the only constraint that couples classes to one another, and it is what makes the problem combinatorial rather than separable.

Residency slips in implicitly. Every tier from the fourth onward keeps payloads inside the national perimeter, so a jurisdiction that forbids foreign replication simply deletes the first tier from every admissible set.

2.2.2 Optimisation Problem

Writing $x_{it} \in \{0,1\}$ for the indicator that class i lands on tier t, the problem becomes

$$\min \Sigma_i \Sigma_t x_{it} (\alpha c^s_{it} + \beta c^l_{it} + \gamma r_{it} + \delta d_{it}) \quad (1)$$

subject to $\Sigma_t x_{it} = 1$ for every i, to $x_{it} = 0$ whenever tier t violates any of C1 through C4 for class i, and to $\Sigma_i \Sigma_t x_{it} L_{it} \leq B$.

The structure is exact, and it deserves a name. Classes partition into disjoint groups; exactly one option gets picked from each; one linear capacity constraint ties the picks together. That is the multiple-choice knapsack problem in minimisation form [31], equivalent by negation to the maximisation form people have studied since the late nineteen-seventies. MCKP generalises 0–1 knapsack and inherits its NP-hardness [32], [33].

Hardness is not the practical obstacle here, though. Pruning is so severe that the surviving space is trivially enumerable, and the useful observation is which constraint does the pruning.

2.3 Solution Method

C1 through C4 apply class by class and hand back admissible sets $A_i \subseteq T$. Enumerate $\Pi_i |A_i|$ candidates exactly, drop whatever breaches the budget. At the scale reported below, that runs in milliseconds.

Real deployments will not stay this small. Several hundred classes across multiple facility types, separate budgets per consortium channel — that instance leaves enumeration far behind. Standard MCKP machinery covers the regime: prune dominated options, order what survives within each class by incremental efficiency, solve the linear relaxation in O(n) by partitioning, then close the integrality gap with branch and bound along the lines of Pisinger's minimal algorithm [34]. The relaxation throws in a bound as well, telling an architect the most a given budget increase could possibly buy. In a procurement conversation that bound is often worth more than the optimum itself.

2.4 Experimental Setup

The instantiation uses nine data classes aligned to FHIR resource types [35], [36] and the six tiers of Section 2.1.2. Parameters appear in Tables 1 and 2. They are illustrative: sizes and instance counts reflect plausible ambulatory workloads, and sensitivity weights encode ordinary clinical intuition about relative disclosure damage. Tables 1 and 2 state them in full, so a reader who disagrees can substitute their own and rerun the model.

Weights are $\alpha = 0.25$ for storage, $\beta = 0.15$ for latency, $\gamma = 0.35$ for exposure and $\delta = 0.25$ for granularity loss. The ledger budget is B = 8 KB per patient-year, chosen to bind: the

minimum feasible budget under C1 through C4 is 5.22 KB, so the constraint has real force without making the instance infeasible.

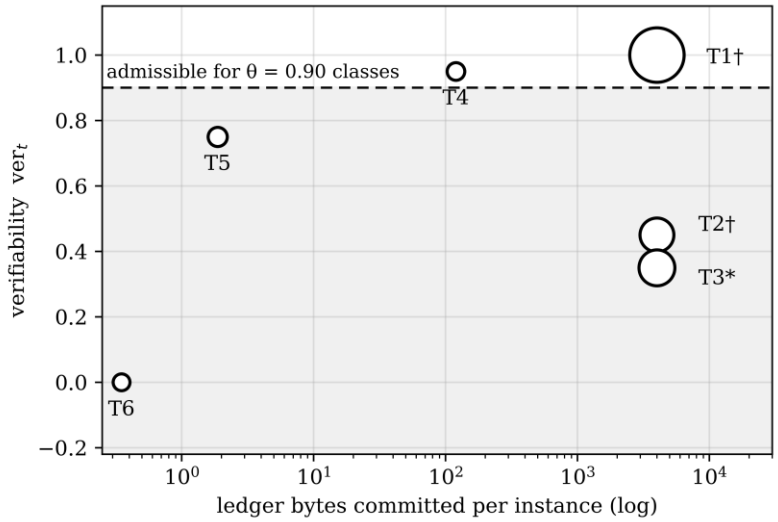


Fig. 1. The six tiers in the footprint–verifiability plane. Marker area scales with exposure ρ_t. † = not erasable, * = redactable only.

Fig. 1 places the six tiers in the plane that matters most for this decision, ledger footprint against verifiability, with marker area scaling to exposure. The θ = 0.90 line cuts through it and eliminates most of the plane for the strictest classes.

Table 1. Data classes (illustrative parameters, per patient-year)

Class	**Size**	**n**	**θ_i**	**e_i**	**λ_i (ms)**	**σ_i**
Patient demographics	2 KB	1	0.00	1.0	400	0.60
Encounter metadata	4 KB	6	0.40	1.0	400	0.40
Observation	1 KB	120	0.75	1.0	250	0.50
DiagnosticReport	30 KB	8	0.90	1.0	500	0.80
ImagingStudy	25 MB	1.5	0.90	1.0	2000	0.90
MedicationRequest	2 KB	14	0.75	1.0	250	0.70
Dispense record *	1 KB	10	0.90	0.0	600	0.55
Consent grant *	0.3 KB	20	0.90	0.0	200	0.15
Access log entry *	0.2 KB	180	0.75	0.0	600	0.20

* audit-bearing: $p_i = 1$, so a ledger trace is mandatory.

Table 2. Storage tiers

Tier	Ledger B / inst.	ver	era	ρ	lat
T1 public, payload	s_i	1.00	0.00	1.00	2500
T2 permissioned, payload	s_i	0.45	0.00	0.35	300
T3 redactable payload	s_i	0.35	0.85	0.40	420
T4 off-chain, per-item	120	0.95	1.00	0.05	180
T5 off-chain, batched/64	1.875	0.75	1.00	0.07	150
T6 off-chain only	0	0.00	1.00	0.04	120

3. Results

3.1 Contribution of the Compliance Envelope

Six tiers, nine classes, $6^9 = 10{,}077{,}696$ unconstrained placements. Verifiability alone cuts that to 10,368. Erasability takes it to 288. Ledger presence removes nothing further — the classes demanding it already sit on tiers that supply it. The latency ceiling leaves 48. The budget leaves 12. Fig. 2 shows the funnel.

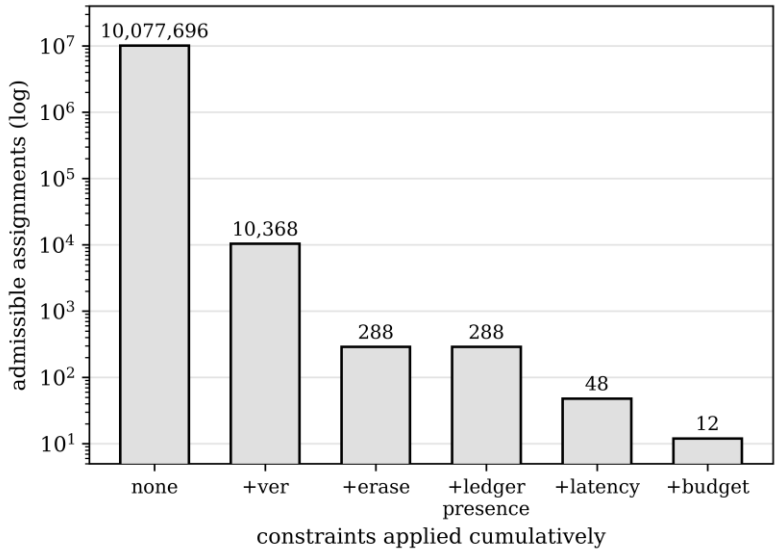


Fig. 2. Compliance constraints remove five orders of magnitude from the design space before any cost term is evaluated.

Five orders of magnitude vanish before anyone consults a single cost coefficient. That is probably the most transferable thing in this paper, and it flips the usual emphasis: architects tune weights and argue storage economics while the compliance envelope has already decided nearly everything. Optimise without writing that envelope down and you are searching a region you never characterised.

3.2 Infeasibility of Uniform Policies

Table 3 evaluates the four uniform policies that recur in the literature against the optimum. Not one of them is feasible. Placing payloads on a permissioned ledger consumes 38,866 KB per patient-year, roughly 4,858 times the budget, and the ImagingStudy class alone accounts for most of it. The budget is not its only problem: a permissioned payload tier cannot clear a 0.75 verifiability floor, cannot erase, and cannot answer inside 250 ms, so C1, C2 and C4 fall with C5. Committing every class per-item needs 42.25 KB, over budget by a factor of five. Batching every class is cheap at 0.66 KB but fails C1 for the four classes demanding $\theta = 0.90$. Pure off-chain storage fails both C1 and C3.

Table 3. Policies compared at B = 8 KB per patient-year

Policy	Ledger KB	J	Gran. loss	Status
All payload, permissioned	38 866	0.584	68.8	C1, C2, C4, C5 fail
All per-item commitment	42.25	0.082	6.25	C5 fails
All batched commitment	0.66	0.063	31.3	C1 fails
All off-chain only	0.00	0.059	125.1	C1, C3 fail
Model optimum	7.64	0.172	25.5	feasible

The unconstrained objective values of the infeasible policies are lower than the optimum's, which is exactly as it should be. Feasibility costs something. A design that ignores C1 through C5 will always look cheaper on paper, and that is precisely how ledger pilots reach procurement before failing legal review.

3.3 Volume Rather Than Sensitivity Governs Batching

The optimum assigns per-item commitments to seven classes and batched commitments to two: Observation and Access log entry. Those are the two classes with the highest instance counts, 120 and 180 per patient-year. They are not the two least sensitive classes — Consent grant carries the lowest sensitivity in Table 1 at $\sigma = 0.15$, yet it receives a per-item commitment, because $\theta = 0.90$ forbids anything coarser.

Granularity is therefore rationed by volume subject to a verifiability floor, not by how much anyone would mind the data leaking. That result is not obvious from the outside, and it contradicts the intuition that cheap proofs should be spent on unimportant data.

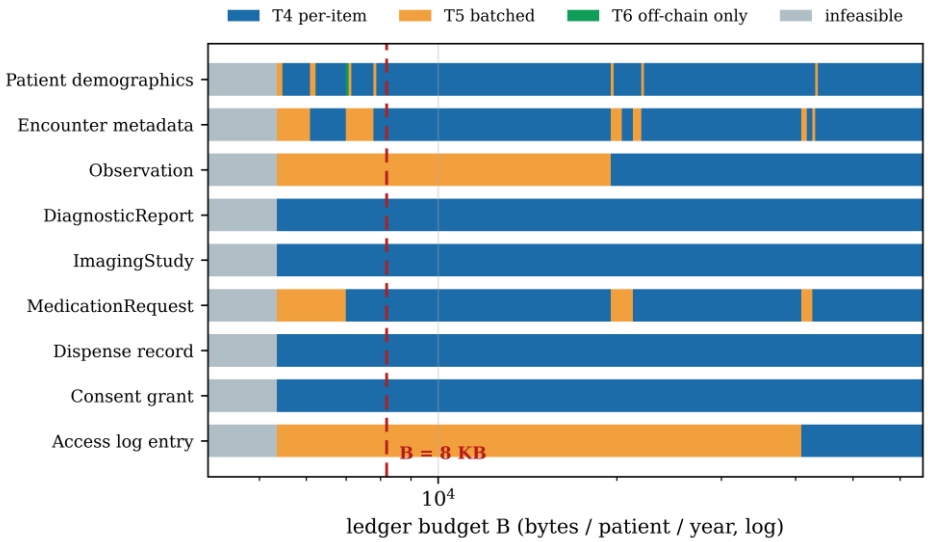


Fig. 3. Optimal placement as the ledger budget grows. Per-item commitments are bought first for low-volume classes; high-volume classes migrate last.

Fig. 3 traces how the placement evolves as B grows from 4 KB to 64 KB. The optimiser buys per-item commitments first for low-volume classes, and the high-volume classes migrate last, at budgets several times the minimum. An architect reading this chart can see directly what an extra kilobyte of ledger budget buys and where it stops mattering.

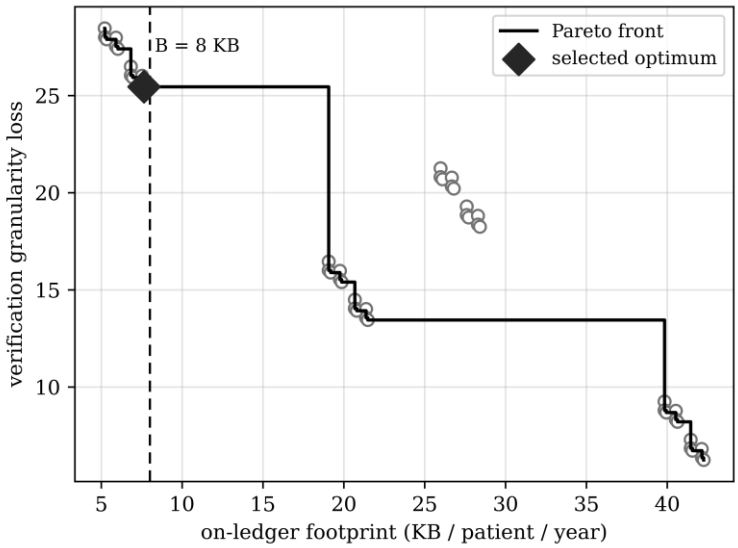


Fig. 4. All 48 admissible assignments. The budget line selects one point on the front; relaxing it buys granularity at a decreasing rate.

Fig. 4 shows all 48 admissible assignments in the footprint–granularity plane together with the Pareto front. The front is short and steep near the minimum feasible budget, then flattens: past roughly 42 KB per patient-year, further ledger spending buys almost no additional verification granularity, because every class that can use a per-item commitment already has one.

3.4 Sensitivity to Parameter Error

Fair question, and the honest answer needs numbers rather than a promise. Tables 1 and 2 hold estimates. An estimate that drifts twenty-five per cent, in a health system nobody has instrumented yet, is ordinary rather than alarming. So the model was rerun ten thousand times with every payload size, instance count and sensitivity weight drawn independently from ±25 per cent of its tabulated value. θ, e, p and λ stayed fixed throughout, because those are duties and service targets rather than measurements — a jurisdiction does not jitter its own retention law.

Two of the three headline results never broke. Pruning stayed above four orders of magnitude in every draw. All four uniform policies stayed infeasible in every draw. Neither result touches the numbers at all: both turn on which tier clears a verifiability floor and which does not, and moving a payload size does not move a floor.

The third result is softer, and saying so costs nothing. Batching landed on the highest-volume classes in 76.1 per cent of draws. In the remaining quarter the instance counts had shuffled far enough that a mid-volume class captured the batched commitment instead. Volume rations granularity robustly; which classes end up batched does not survive the same

abuse. An architect substituting their own counts should expect the assignment to move even where the principle holds.

A reference implementation accompanies this paper. It reads Tables 1 and 2 and recomputes the funnel, Table 3, the optimum and the budget sweep from scratch, so every figure reported above can be checked against the parameters that produced it rather than taken on trust.

4. Discussion and Future Trends

4.1 Implications for Practice

Three implications follow for anyone building a hybrid ledger service in a regulated health system.

First, publish the envelope before the architecture. If constraints eliminate 99.9995 percent of the design space, then a paper reporting a placement without reporting its constraints has withheld the part that did the reasoning. Future system descriptions should tabulate θ, e, p and λ as routinely as they already tabulate throughput.

Second, treat batching as a budgeted resource. The literature usually sells batched commitments as a scalability trick. Here they behave like currency — spent where volume runs highest, withheld where an auditor needs to isolate one record — and the allocation shifts predictably as the budget moves.

Third, be careful comparing objective values across feasibility boundaries. Every infeasible policy in Table 3 scores better than the optimum. Any benchmark that ranks architectures on cost or latency alone, without checking the envelope first, will reliably recommend designs that cannot be deployed.

4.2 Threats to Validity

Parameters are illustrative. Nothing here claims that the specific numbers in Tables 1 and 2 describe any particular health system. The qualitative results follow from the ordering of the parameters rather than their exact values, and Section 3.4 puts numbers on how far that holds: severe pruning and the infeasibility of every uniform policy survived all ten thousand perturbed draws, while the batching assignment survived roughly three in four. A system whose ordering differs will reach different conclusions, and the model provides the instrument for finding that out.

The cost model is linear and additive, which understates interactions. Storing a class off-chain changes the operational burden of key management for every other class sharing a key hierarchy [37], [38], and attribute-based access control shifts the exposure coefficients in ways a single ρ_t cannot capture [39]. Exposure itself collapses to a scalar here; a serious risk treatment would use a distribution.

The model also assumes placement is static per class. Real records change status — a study under active care is not a study under statutory retention twelve years later — and a lifecycle-aware version of this model, in which classes migrate between tiers on defined

events, is the obvious next step. That version is not a straightforward extension: migration costs and backup semantics interact awkwardly with erasure duties [40].

Finally, no deployed system has yet tested the model. Nothing here rests on measured latencies, throughput or storage figures from running infrastructure, and the reader should treat every number here as a consequence of stated assumptions.

5. Conclusion

Placement in hybrid blockchain health systems is a constrained optimisation problem, and naming it as one changes what an architect can say about their design. This paper gave the model, located it as a multiple-choice knapsack instance, and instantiated it on nine FHIR-aligned classes. The compliance envelope removed five orders of magnitude from the design space; all four uniform policies in common use turned out infeasible; and the optimum rationed proof granularity by data volume under a verifiability floor. Perturbing every estimated parameter by a quarter left the first two findings untouched and the third standing in three draws out of four. The next step is the lifecycle-aware formulation, together with a Fabric testbed that will replace the illustrative latency and storage parameters with measured ones.

Acknowledgments

Not applicable.

Author Contributions

Yerassyl Aikenov: Conceptualization, Methodology, Software, Formal analysis, Visualization, Writing – original draft, Writing – review & editing.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

All parameter values used by the model are given in full in Tables 1 and 2. The reference implementation that reproduces the reported results is available at https://github.com/yerassylaikenov/JSDT_article2_Aikenov.git.

List of Supplementary Materials

Supplementary File S1: reference implementation (source code) that reproduces the constraint funnel, Table 3, the optimum and the budget sweep shown in Figures 1–4.

References

[1] Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation). (2016). *Official Journal of the European Union, L119*, 1–88.

[2] Politou, E., Alepis, E., & Patsakis, C. (2018). Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of Cybersecurity, 4*(1), tyy001. https://doi.org/10.1093/cybsec/tyy001

[3] Code of the Republic of Kazakhstan No. 360-VI "On Public Health and the Healthcare System". (2020, July 7). https://adilet.zan.kz

[4] Nurgaliyeva, Z., Spatayev, Y., Syla, S., & Yessenbayev, B. (2024). Paving the way to establishing the digital-friendly health and care information model in Kazakhstan. *International Journal of Medical Informatics, 192*, 105610. https://doi.org/10.1016/j.ijmedinf.2024.105610

[5] Wüst, K., & Gervais, A. (2018). Do you need a blockchain? In *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)* (pp. 45–54). IEEE. https://doi.org/10.1109/CVCBT.2018.00011

[6] Eberhardt, J., & Tai, S. (2017). On or off the blockchain? Insights on off-chaining computation and data. In *Service-Oriented and Cloud Computing (ESOCC 2017)* (Lecture Notes in Computer Science, Vol. 10465, pp. 3–15). Springer.

[7] Xu, X., Weber, I., Staples, M., Zhu, L., Bosch, J., Bass, L., Pautasso, C., & Rimba, P. (2017). A taxonomy of blockchain-based systems for architecture design. In *2017 IEEE International Conference on Software Architecture (ICSA)* (pp. 243–252). IEEE. https://doi.org/10.1109/ICSA.2017.33

[8] Azaria, A., Ekblaw, A., Vieira, T., & Lippman, A. (2016). MedRec: Using blockchain for medical data access and permission management. In *2016 2nd International Conference on Open and Big Data (OBD)* (pp. 25–30). IEEE. https://doi.org/10.1109/OBD.2016.11

[9] Dagher, G. G., Mohler, J., Milojkovic, M., & Marella, P. B. (2018). Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology. *Sustainable Cities and Society, 39*, 283–297. https://doi.org/10.1016/j.scs.2018.02.014

[10] Xia, Q., Sifah, E. B., Smahi, A., Amofa, S., & Zhang, X. (2017). BBDS: Blockchain-based data sharing for electronic medical records in cloud environments. *Information, 8*(2), 44. https://doi.org/10.3390/info8020044

[11] Fan, K., Wang, S., Ren, Y., Li, H., & Yang, Y. (2018). MedBlock: Efficient and secure medical data sharing via blockchain. *Journal of Medical Systems, 42*(8), 136. https://doi.org/10.1007/s10916-018-0993-7

[12] Yue, X., Wang, H., Jin, D., Li, M., & Jiang, W. (2016). Healthcare data gateways: Found healthcare intelligence on blockchain with novel privacy risk control. *Journal of Medical Systems, 40*(10), 218. https://doi.org/10.1007/s10916-016-0574-6

[13] Nguyen, D. C., Pathirana, P. N., Ding, M., & Seneviratne, A. (2019). Blockchain for secure EHRs sharing of mobile cloud based e-health systems. *IEEE Access, 7*, 66792–66806. https://doi.org/10.1109/ACCESS.2019.2917555

[14] Androulaki, E., Barger, A., Bortnikov, V., Cachin, C., Christidis, K., De Caro, A., Enyeart, D., Ferris, C., Laventman, G., Manevich, Y., Muralidharan, S., Murthy, C., Nguyen, B., Sethi, M., Singh, G., Smith, K., Sorniotti, A., Stathakopoulou, C., Vukolić, M., … Yellick, J. (2018). Hyperledger Fabric: A distributed operating system for permissioned blockchains. In *Proceedings of the Thirteenth EuroSys Conference* (Article 30). ACM. https://doi.org/10.1145/3190508.3190538

[15] Stamatellis, C., Papadopoulos, P., Pitropakis, N., Katsikas, S., & Buchanan, W. J. (2020). A privacy-preserving healthcare framework using Hyperledger Fabric. *Sensors, 20*(22), 6587. https://doi.org/10.3390/s20226587

[16] Xu, X., Sun, G., Luo, L., Cao, H., Yu, H., & Vasilakos, A. V. (2021). Latency performance modeling and analysis for Hyperledger Fabric blockchain network. *Information Processing & Management, 58*(1), 102436. https://doi.org/10.1016/j.ipm.2020.102436

[17] Jayabalan, J., & Jeyanthi, N. (2022). Scalable blockchain model using off-chain IPFS storage for healthcare data security and privacy. *Journal of Parallel and Distributed Computing, 164*, 152–167. https://doi.org/10.1016/j.jpdc.2022.03.009

[18] Zhan, W., Chen, C.-L., Weng, W., Tsaur, W.-J., Lim, Z.-Y., & Deng, Y.-Y. (2022). Incentive EMR sharing system based on consortium blockchain and IPFS. *Healthcare, 10*(10), 1840. https://doi.org/10.3390/healthcare10101840

[19] Benet, J. (2014). *IPFS—Content addressed, versioned, P2P file system* (arXiv:1407.3561). arXiv.

[20] Steichen, M., Fiz, B., Norvill, R., Shbair, W., & State, R. (2018). Blockchain-based, decentralized access control for IPFS. In *2018 IEEE International Conference on Blockchain* (pp. 1499–1506). IEEE.

[21] Kumar, R., & Tripathi, R. (2021). Towards design and implementation of security and privacy framework for Internet of Medical Things (IoMT) by leveraging blockchain and IPFS technology. *The Journal of Supercomputing, 77*(8), 7916–7955.

[22] Hasselgren, A., Kralevska, K., Gligoroski, D., Pedersen, S. A., & Faxvaag, A. (2020). Blockchain in healthcare and health sciences—A scoping review. *International Journal of Medical Informatics, 134*, 104040. https://doi.org/10.1016/j.ijmedinf.2019.104040

[23] Agbo, C. C., Mahmoud, Q. H., & Eklund, J. M. (2019). Blockchain technology in healthcare: A systematic review. *Healthcare, 7*(2), 56. https://doi.org/10.3390/healthcare7020056

[24] Lopez, L. J. R., Millan Mayorga, D., Martinez Poveda, L. H., Amaya, A. F. C., & Rojas Reales, W. (2024). Hybrid architectures used in the protection of large healthcare records based on cloud and blockchain integration: A review. *Computers, 13*(6), 152. https://doi.org/10.3390/computers13060152

[25] Ateniese, G., Magri, B., Venturi, D., & Andrade, E. (2017). Redactable blockchain—or—rewriting history in Bitcoin and friends. In *2017 IEEE European Symposium on Security and Privacy (EuroS&P)* (pp. 111–126). IEEE.

[26] Derler, D., Samelin, K., Slamanig, D., & Striecks, C. (2019). Fine-grained and controlled rewriting in blockchains: Chameleon-hashing gone attribute-based. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*. Internet Society.

[27] Deuber, D., Magri, B., & Thyagarajan, S. A. K. (2019). Redactable blockchain in the permissionless setting. In *2019 IEEE Symposium on Security and Privacy (SP)* (pp. 124–138). IEEE. https://doi.org/10.1109/SP.2019.00039

[28] Politou, E., Casino, F., Alepis, E., & Patsakis, C. (2021). Blockchain mutability: Challenges and proposed solutions. *IEEE Transactions on Emerging Topics in Computing, 9*(4), 1972–1986.

[29] Merkle, R. C. (1988). A digital signature based on a conventional encryption function. In *Advances in Cryptology—CRYPTO '87* (Lecture Notes in Computer Science, Vol. 293, pp. 369–378). Springer.

[30] Kissel, R., Regenscheid, A., Scholl, M., & Stine, K. (2014). *Guidelines for media sanitization* (NIST Special Publication 800-88 Rev. 1). National Institute of Standards and Technology.

[31] Sinha, P., & Zoltners, A. A. (1979). The multiple-choice knapsack problem. *Operations Research, 27*(3), 503–515. https://doi.org/10.1287/opre.27.3.503

[32] Garey, M. R., & Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*. W. H. Freeman.

[33] Martello, S., & Toth, P. (1990). *Knapsack problems: Algorithms and computer implementations*. Wiley.

[34] Pisinger, D. (1995). A minimal algorithm for the multiple-choice knapsack problem. *European Journal of Operational Research, 83*(2), 394–410.

[35] HL7 International. (2023). *HL7 FHIR Release 5*. https://hl7.org/fhir/R5/

[36] Zhang, P., White, J., Schmidt, D. C., Lenz, G., & Rosenbloom, S. T. (2018). FHIRChain: Applying blockchain to securely and scalably share clinical data. *Computational and Structural Biotechnology Journal, 16*, 267–278. https://doi.org/10.1016/j.csbj.2018.07.004

[37] Dworkin, M. (2007). *Recommendation for block cipher modes of operation: Galois/Counter Mode (GCM) and GMAC* (NIST Special Publication 800-38D). National Institute of Standards and Technology.

[38] Shamir, A. (1979). How to share a secret. *Communications of the ACM, 22*(11), 612–613.

[39] Goyal, V., Pandey, O., Sahai, A., & Waters, B. (2006). Attribute-based encryption for fine-grained access control of encrypted data. In *Proceedings of the 13th ACM Conference on Computer and Communications Security (CCS)* (pp. 89–98). ACM.

[40] Politou, E., Michota, A., Alepis, E., Pocs, M., & Patsakis, C. (2018). Backups and the right to be forgotten in the GDPR: An uneasy relationship. *Computer Law & Security Review, 34*(6), 1247–1257. https://doi.org/10.1016/j.clsr.2018.08.006