

DATA AND RETRIEVAL

Retrieval over the chart: assembling clinical context from standards that were not built for it

ABSTRACT

Before a model drafts anything, software decides what it may read. This paper specifies that read path. It states the interoperability standards exactly — including the three-version gap between the published USCDI and the version adopted at 45 C.F.R. § 170.213 — shows that those standards were built for document exchange rather than question answering, and sets out the design requirements that follow from measured record quality. It also states what better context assembly cannot repair.

Type Architecture and design rationale **References** 28 **Reading time** 14 min

Last reviewed September 2026

1 Context assembly precedes generation

A model asked to draft a clinical note does not read the chart. It reads whatever a retrieval step placed in front of it. That step, deciding which record entries enter the context, in what order and with what metadata, is software somebody wrote. Its failures are not model failures, and a better model does not correct them.

The clearest evidence separating the two comes from an evaluation on real records. Hager and colleagues ran 2,400 real patient cases drawn from MIMIC-IV across four abdominal pathologies. Hospitalists reached 87.5–92.5% diagnostic accuracy; the open models tested reached 58.8% (Llama 2 Chat), 67.8% (OASST) and 65.1% (WizardLM). The second condition matters most: when the models had to gather information autonomously rather than being handed it, accuracy fell further still, from 58.8% to 45.5% in one case.¹

That result is usually read as a claim about agentic tool use. It is better read as a claim about context assembly: performance degrades when the selection of what to read is left to the model. Making that selection well is an architectural problem, not a prompting one.

The evaluation literature has largely not examined it. Of 519 studies of health care applications of large language models, only 5% used real patient care data; 44.5% evaluated on medical examination questions.² An exam question arrives pre-assembled. A chart does not.

Three claims follow. The interoperability standards clinical data reaches a reader through were built for document exchange rather than question answering, and the version binding US exchange is older than most writing assumes. The data they carry is of measured unreliability, making ground-truth treatment of the record a design error. A read path accounting for both has statable requirements, worth demanding of any vendor, this one included.

58.8% → 45.5%

accuracy when a model gathered information itself rather than receiving it

79%

US hospitals able to integrate received records, 2025; 74% in 2021

30–50%

of relevant chronic conditions absent from problem lists

2 The standards, stated exactly

The common errors here run one direction: they credit the installed base with a newer standard than it has.

FHIR

The most recent full release is R5, version 5.0.0, published 26 March 2023, status Trial Use; no R5 content is presently marked normative.³ R6 is not a released standard, sitting at 6.0.0-ballot4 as of September 2026.⁴ Neither is what US exchange runs on. That is R4, version 4.0.1, published 27 December 2018, the base for every US Core release from 3.1.0 onward.⁵ Regulated US exchange runs on a 2018 specification.

US Core and USCDI

US Core stands at STU 9, version 9.0.0, published 31 May 2026 on FHIR 4.0.1; it maps to USCDI v6 and never moved to R5.⁶ USCDI is where the gap is largest and least often stated correctly. Published versions run from v1 (July 2020 errata) to v6 (July 2025, current), with a v7 draft in 2026; publication is annual and decoupled from regulatory adoption.⁷ The Code of Federal Regulations, at 45 C.F.R. § 170.213, adopts exactly two: the July 2020 Errata Version 1, which expired 1 January 2026, and Version 3, which carries no expiration.⁸ USCDI v3 is therefore the only version currently adopted in regulation, three versions behind the published v6. A claimed v5 or v6 compliance obligation describes something regulation does not require. The same trap catches US Core: certification does not require 9.0.0.

SMART, Bulk Data and C-CDA

SMART App Launch has stood at 2.2.0 (STU 2.2) since 30 April 2024, with no new release in over two years.⁹ FHIR Bulk Data Access reached 3.0.0 (STU 3) on 11 December 2025; its asynchronous \$export operation covers Patient, Group and System scope for population-level extraction: a bulk egress

mechanism, not a retrieval mechanism for one encounter.¹⁰ And document exchange is not deprecated: Consolidated CDA reached 5.0.0 (STU 5), US Realm, on 13 June 2026 and remains embedded in certification for transitions of care at § 170.315(b)(1).¹¹ A read path treating C-CDA as legacy is mis-scoped for its substrate.

TABLE 1

Current published version against the version that binds US exchange, September 2026. A dash means the sources cited here establish no adopted version.

Standard	Current published version	What binds in the US
HL7 FHIR	R5, 5.0.0, 26 March 2023, Trial Use; R6 in ballot, not published	R4, 4.0.1, 27 December 2018
US Core IG	STU 9, 9.0.0, 31 May 2026, on FHIR 4.0.1, mapped to USCDI v6	Certification does not require 9.0.0
USCDI	v6, July 2025; v7 in draft	v3, at 45 C.F.R. § 170.213(b); the v1 adoption expired 1 January 2026
SMART App Launch	2.2.0 (STU 2.2), 30 April 2024	—
FHIR Bulk Data Access	3.0.0 (STU 3), 11 December 2025	—
C-CDA	5.0.0 (STU 5), 13 June 2026	Certification, transitions of care, § 170.315(b)(1)
TEFCA Common Agreement	Version 2.1, November 2024	Voluntary; 11 Designated QHINs

3 What the standards were built to do

Read the specifications for what they claim. FHIR defines resources and an interaction model for exchanging them; US Core profiles those resources for the US realm; SMART App Launch defines how an application obtains authorization and launch context; Bulk Data defines asynchronous population extraction; C-CDA defines a document. None defines a query language for clinical questions.

A conformant API can answer “return this patient’s Condition resources.” It cannot answer “what has been tried for this patient’s back pain, and what happened.” The second requires joining conditions to medications to procedures to encounter narrative across time, weighing which entries are current, and recognizing that part of the answer may sit in a scanned document nobody coded. Every vendor putting a model in front of a chart builds that layer on top of the standards, whether or not it says so.

The resource-usage evidence fits: across 112 real-world FHIR applications the most-used resources were Patient (96), Observation (83), Condition (78) and Medication (71).¹² That is the roster of a record viewer, not of a system reasoning across a longitudinal history.

Two published attempts to put FHIR in front of a language model illustrate the bridge without crossing it. FHIR-Former serialized FHIR resources into text for a transformer over 408,267 patients and roughly 1.1 million data points across ten FHIR R4 resource types; ICD-10 prediction accuracy was 94%, but macro F1 only 48%, a signal of class imbalance and of how the headline figure misleads.¹³ LLMonFHIR queried a patient's own FHIR records through a model, physicians rating 210 responses at a median 5/5 for accuracy, understandability and relevance; but the authors named specific weakness in summarizing conditions and retrieving laboratory results, on synthetic data rather than a real-PHI deployment.¹⁴ Those two weaknesses are the two operations a read path exists to perform.

4 What is deployed, as opposed to specified

Federal survey data separates capability from specification. ASTP/ONC's Quick Stat #66, on American Hospital Association Health IT Survey data, measures four exchange domains. In 2025: send 96%, receive 93%, find 94%, integrate 79%, with 76% of hospitals engaging in all four against a 2014 baseline of 23%.¹⁵

Integrate is the domain that governs a read path, and it has not moved: 74% in 2021, 79% in 2022, 78% in 2023, 79% in 2025, flat for four years, while find climbed from 84% to 94% and receive from 87% to 93%.¹⁵ Hospitals are steadily more able to obtain outside data and no more able to incorporate it, so a model reading the local record reads one into which outside data has largely not been integrated.

Fax is not receding

Quick Stat #70 measures methods. For sending summary-of-care records by mail or fax in 2025, 40% of hospitals reported doing so often and a further 34% sometimes: 74% in total. For receiving, 35% often and 46% sometimes: 81%. The direction is not the one usually assumed: often-use for sending rose from 30% in 2023 to 40% in 2025.¹⁶ Most fax figures circulating in health IT are vendor marketing with no traceable source. These are the federal numbers, and they show no channel in retreat.

Connection is not use

TEFCA is the framework built to change this. The Common Agreement stands at version 2.1, released November 2024, its v2 predecessor having introduced FHIR-based exchange.¹⁷ Eleven QHINs are designated as of September 2026.¹⁸ That is a working framework, not yet a substrate a read path can assume.

FHIR API adoption is real and uneven. An ONC data brief reported more than two-thirds of hospitals using an HL7 FHIR API for patient access in 2022, up 12 percentage points from 2021, with adoption lagging among small, independent hospitals and those not on a market-leading EHR.¹⁹ The application side is thinner than the narrative implies: of those 112 FHIR applications, only about 49% used SMART

on FHIR and 44% appeared in no EHR app gallery.¹² And where the pipe is connected it goes unused: EHR audit-log analysis found each additional hour of documentation associated with a 7.1% decrease in the likelihood a primary care physician accesses outside records.²⁰

A read path built for a clean FHIR world, in which every fact is a coded resource, every outside record is integrated and every partner is reachable through a QHIN, will not survive contact with a real practice. The substrate is mixed: coded resources where they exist, documents where they do not, faxes where neither does.

WHAT THIS SURVEY EVIDENCE DOES AND DOES NOT ESTABLISH

Both Quick Stats report hospital self-reports of capability, not measured exchange volume, and describe hospitals rather than ambulatory practices. One caution on the trend: the 2023-to-2025 jump in find and receive is large enough that a survey or question change should be ruled out before reading it as pure behavior change.¹⁵

5 The record is evidence, not ground truth

Standards govern the envelope, not the contents. A Condition resource carrying a stale, duplicated or absent problem is still a conformant Condition resource, and its conformance says nothing about whether the problem is true today.

The measurements are not marginal. A rapid scoping review synthesizing 103 peer-reviewed studies of electronic problem lists in primary care, covering March 2016 to May 2026, found 30–50% of clinically relevant chronic conditions omitted from problem lists and roughly 10% of patients holding duplicate entries. It also found persistent unclear ownership of list maintenance and unresolved disagreement about what constitutes a “problem.”²¹

Demographic fields are no better. Across 569 patients at 13 Massachusetts primary care clinics, EHR sensitivity was 70.9% for identifying African American patients and 83.8% for Hispanic patients; more than one-third recorded as Spanish-preferring completed the study survey in English.²² A more recent analysis reports a 2023 review of 56 datasets finding EHRs often held missing or inaccurate race and ethnicity data, most acute for non-white populations, and a health system in which 66.5% of patients selected different categories than the EHR held.²³

Two consequences follow. The first is that a read path must treat record entries as evidence of varying reliability. A problem entered and reconciled at today’s visit and a problem carried forward untouched for six years are the same resource type and are not the same evidence. Flattening both into a line

reading “Problem list:” destroys the distinction a reviewing clinician needs most.

The second is skipped more often: a read path must be able to state that something is unavailable. Given no medication data, a model will not reliably report having none. Omar and colleagues seeded 300 physician-validated vignettes each with one fabricated element and found six models elaborating on the fabrication rather than rejecting it, from 50–53.3% for the best-performing model to 80–82.7% for the worst, 66% overall. A mitigation prompt reduced that to 44%; setting temperature to zero produced no significant improvement.²⁴ A system that builds on a fabrication it was handed will build on an absence it was not told about: silent omission in the context window is a fabrication generator.

6 Requirements for a read path

What follows is stated as requirements. It describes no index, embedding scheme or ranking function: ChartVoyant publishes none, and a specification should not invent one.

Every assertion is traceable to the entry it came from. Anything placed in front of a model, and anything the model asserts in a draft, must resolve back to a specific record entry in one step. Traceability constrains the retrieval representation rather than following from it: paraphrasing several entries into one paragraph destroys the mapping before the model sees it.

Recency and provenance travel with the content. When and by whom an entry was recorded, and whether it was reconciled or carried forward, are part of its evidential weight; serializing a chart into narrative discards them silently, the failure the problem-list literature predicts.²¹ The W3C’s PROV family gives the right shape, modeling provenance on Entity, Activity and Agent; note that PROV-DM is the Recommendation while the overview document is a Working Group Note, so a design citing the overview as its normative standard has cited the wrong one.²⁵ In FHIR, the Provenance resource is Maturity Level 4, Trial Use, and narrower than assumed: it covers the generation of an entity while AuditEvent covers usage and all other activity.²⁶ A design recording authorship in Provenance and expecting it to answer access questions has a gap.

Unavailability is represented, never omitted. The read path must distinguish three states prose collapses into one: a value recorded, a value recorded as absent, and a value never obtained. “No known drug allergies” and “no allergy data available” are different facts, and only the first belongs in a draft read as a clinical assertion.²⁴

The encounter transcript anchors the draft. ChartVoyant listens to the visit, the transcript fills live, and drafts are produced alongside it; the published audit event for a drafted note, *ai.note.drafted*, is recorded as written by AI and linked to the transcript, immediately after the *visit.transcript.appended* event that produced it. The rationale: the transcript of the encounter now underway is the one source whose recency is not in question. Historical entries inform the draft; the transcript binds it. Suggested code sets in the drafted output, such as the ICD-10-CM M54.16 and G89.29 in ChartVoyant’s public example, stay suggestions for a clinician to accept or reject.

Retrieval never spans practices. ChartVoyant states that every practice's data is walled off from every other one: records, files and audit logs alike. That is a correctness property, not only a privacy one: the set a retrieval may span is the unit at which a wrong answer becomes a disclosure. Isolation belongs below the retrieval layer, enforced where the data lives rather than checked in the query.

Nothing retrieved is a decision. Nothing signs itself; clinical suggestions are drafts for a licensed clinician to review; and no patient information reaches an AI vendor without a signed business associate agreement. The read path inherits all three. The last is a retrieval-design constraint in particular: it settles what may cross a boundary at all.

7 Retrieval provenance as an oversight requirement

A reviewer who cannot see what the model read cannot check what it wrote. That follows from the error distribution, not from a preference about interface design.

In a study of 450 clinical note–transcript pairs across 18 configurations, clinicians annotated 12,999 sentences for hallucination and 49,590 for omission. The hallucination rate was 1.47% (191 of 12,999), 44% of those major; the omission rate was 3.45% (1,712 of 49,590), 16.7% major, with major hallucinations clustered in the Plan section at 21%.²⁷ Omissions ran at more than twice the rate of hallucinations, and an omission is invisible in the output: a reviewer reading only the draft has no signal that anything is missing. Detecting one requires comparison against the source, which requires the source to be reachable.

The second argument is about cost. A systematic review of 40 studies of automation bias found it appearing in single tasks, typically diagnosis rather than monitoring, and specifically where verification complexity was high; the design implication the authors draw is to reduce the cognitive cost of verifying a suggestion rather than instructing people to verify harder.²⁸ Showing the entry an assertion came from, next to the assertion, converts verification from a chart search into a glance. Whether reviewers use it is a separate question, treated in a sibling paper.

8 What better context assembly does not fix

RESIDUAL FAILURE MODES

Better context assembly does not repair a wrong problem list. The 30–50% omission rate for clinically relevant chronic conditions is a property of the record, and retrieving that list perfectly retrieves the omissions with it.²¹ Worse, a stale entry surfaced with clean provenance acquires an authority it has not earned: traceability establishes that a claim came from the record, not that the record was right.

It does not make an absent record present. With integrate flat at roughly 79% for four years, outside data that was never incorporated cannot be retrieved from a chart that does not hold it.¹⁵ Nor can it recover information that only ever existed on a fax nobody indexed, and 81% of hospitals still receive summary-of-care records by mail or fax at least sometimes.¹⁶ To a retrieval step, an unindexed document is identical to one that does not exist.

Two residuals deserve naming. The transcript is itself a derived artifact produced by a recognition system that can err: anchoring a draft to it improves traceability without making the anchor infallible. And nothing here guarantees that a reviewer shown the sources consults them; lowering the cost of verification is a different claim from establishing that it occurs. ChartVoyant publishes no measurement of its own retrieval quality, coding performance or error rates, and none is asserted here.

9 What to ask about a read path

The read path is the least-inspected part of most clinical AI products and among the most decisive for output quality. Six questions separate a specified one from an unspecified one.

1. Which FHIR release and US Core version does the product implement, and when it claims a USCDI version, is that the published one or the one adopted at 45 C.F.R. § 170.213?
2. For any assertion in a generated draft, can you reach the record entry it came from in one step?
3. How does the output distinguish “no allergies” from “no allergy data”?

4. What is the unit of isolation for a retrieval, and what enforces it: the storage layer, or the query?
5. What happens when the relevant document arrived by fax and was never indexed?
6. What reaches the model vendor, under what agreement, before any of this runs?

None asks for a benchmark score. They ask whether the system can account for what it read, the property that makes everything downstream reviewable.

REFERENCES

Entries 3 to 11, 17, 18, 25 and 26 are standards specifications, regulations or governance documents rather than research, and are cited for their text and version status. Entries 15, 16 and 19 are federal survey reports; they record self-reported capability rather than measured exchange volume. Entries 1, 2, 12, 13, 14, 20 to 24, 27 and 28 are peer-reviewed studies. Statements about ChartVoyant's own design are drawn from its public product documentation and carry no reference number. Every figure is quoted at the precision its source reports.

1. Hager P, Jungmann F, Holland R, et al. *Evaluation and mitigation of the limitations of large language models in clinical decision-making*. Nature Medicine. 2024;30(9):2613–2622. doi:10.1038/s41591-024-03097-1 **BENCHMARK**
2. Bedi S, Liu Y, Orr-Ewing L, et al. *Testing and Evaluation of Health Care Applications of Large Language Models: A Systematic Review*. JAMA. 2025;333(4):319–328. doi:10.1001/jama.2024.21700 **SYSTEMATIC REVIEW**
3. Health Level Seven International. *HL7 FHIR Release 5 (Fast Healthcare Interoperability Resources Specification)*. Version 5.0.0, published 26 March 2023. hl7.org/fhir/ **STANDARD**
4. Health Level Seven International. *FHIR Release 6 (continuous integration build / ballot)*. Version 6.0.0-ballot4; in ballot as of September 2026, not published. build.fhir.org/versions.html **STANDARD**
5. Health Level Seven International. *HL7 FHIR Release 4*. Version 4.0.1, published 27 December 2018. hl7.org/fhir/R4/ **STANDARD**
6. HL7 International, Cross-Group Projects Work Group. *US Core Implementation Guide*. Version 9.0.0 (STU 9), built on FHIR 4.0.1, published 31 May 2026. hl7.org/fhir/us/core/ **STANDARD**
7. Assistant Secretary for Technology Policy / Office of the National Coordinator for Health IT. *United States Core Data for Interoperability (USCDI)*. v1 (July 2020 errata) through v6 (July 2025); v7 draft, 2026. isp.healthit.gov/united-states-core-data-interoperability-uscdi **STANDARD**

8. Department of Health and Human Services, ASTP/ONC. *45 CFR § 170.213 — United States Core Data for Interoperability*. Last amended 9 January 2024 (89 FR 1428). ecfr.gov/current/title-45/.../section-170.213 REGULATION
9. HL7 International. *SMART App Launch Implementation Guide*. Version 2.2.0 (STU 2.2), published 30 April 2024. hl7.org/fhir/smart-app-launch/ STANDARD
10. HL7 International. *FHIR Bulk Data Access Implementation Guide*. Version 3.0.0 (STU 3), published 11 December 2025; 4.0.0 ballot 7 August 2026. hl7.org/fhir/uv/bulkdata/ STANDARD
11. HL7 International, Structured Documents Work Group. *Consolidated CDA (C-CDA) Implementation Guide*. Version 5.0.0 (STU 5), US Realm, on CDA Release 2.0, published 13 June 2026. hl7.org/cda/us/ccda/ STANDARD
12. Griffin AC, He L, Sunjaya AP, et al. *Clinical, technical, and implementation characteristics of real-world health applications using FHIR*. JAMIA Open. 2022;5(4):ooac077. [doi:10.1093/jamiaopen/ooac077](https://doi.org/10.1093/jamiaopen/ooac077) CROSS-SECTIONAL
13. Engelke M, Baldini G, Kleesiek J, et al. *FHIR-Former: enhancing clinical predictions through Fast Healthcare Interoperability Resources and large language models*. Journal of the American Medical Informatics Association. 2025;32(12):1793–1801. [doi:10.1093/jamia/ocaf165](https://doi.org/10.1093/jamia/ocaf165) COHORT
14. Schmiedmayer P, Rao A, Zagar P, et al. *LLMonFHIR: A Physician-Validated, Large Language Model-Based Mobile Application for Querying Patient Electronic Health Data*. JACC: Advances. 2025;4(6, Part 1):101780. [doi:10.1016/j.jacadv.2025.101780](https://doi.org/10.1016/j.jacadv.2025.101780) BENCHMARK
15. ASTP/ONC. *Electronic Health Information Exchange by Hospitals*. Health IT Quick Stat #66; source data American Hospital Association Health Information Technology Survey, 2014–2025. healthit.gov/data/quickstats/electronic-health-information-exchange-hospitals/ GOVERNMENT REPORT
16. ASTP/ONC. *Methods Used by Hospitals to Engage in Interoperable Exchange*. Health IT Quick Stat #70; source data American Hospital Association Information Technology Survey, 2018–2025. healthit.gov/data/quickstats/methods-used-hospitals-engage-interoperable-exchange/ GOVERNMENT REPORT
17. The Sequoia Project (Recognized Coordinating Entity) under ASTP/ONC. *Trusted Exchange Framework and Common Agreement (TEFCA) — Common Agreement*. Version 2.1, released November 2024. rce.sequoiaproject.org/common-agreement/ STANDARD
18. The Sequoia Project (Recognized Coordinating Entity). *Designated QHINs*. As of September 2026. rce.sequoiaproject.org/participate/designated-qhins/ OFFICIAL REGISTRY

19. Strawley C, Everson J, Barker W. *Hospital Use of APIs to Enable Data Sharing Between EHRs and Apps*. *ONC Data Brief No. 68*, Office of the National Coordinator for Health IT. September 2023. healthit.gov/data/data-briefs/hospital-use-apis-enable-data-sharing-between-ehrs-and-apps/
GOVERNMENT REPORT
20. Holmgren AJ, Adler-Milstein J, Apathy NC. *Electronic Health Record Documentation Burden Crowds Out Health Information Exchange Use By Primary Care Physicians*. *Health Affairs*. 2024;43(11):1538–1545. [doi:10.1377/hlthaff.2024.00398](https://doi.org/10.1377/hlthaff.2024.00398) CROSS-SECTIONAL
21. Nair R, Bekker I, Singer A, Lau F. *Current state of electronic problem lists in primary care: a rapid scoping review*. *Family Practice*. 2026;43(4):cmag036. [doi:10.1093/fampra/cmag036](https://doi.org/10.1093/fampra/cmag036)
SCOPING REVIEW
22. Klinger EV, Carlini SV, Gonzalez I, et al. *Accuracy of Race, Ethnicity, and Language Preference in an Electronic Health Record*. *Journal of General Internal Medicine*. 2015;30(6):719–723. [doi:10.1007/s11606-014-3102-8](https://doi.org/10.1007/s11606-014-3102-8) CROSS-SECTIONAL
23. Tsalidis A, Bharadwaj L, Shen FX. *Standardization and accuracy of race and ethnicity data: Equity implications for medical AI*. *PLOS Digital Health*. 2025;4:e0000807. [doi:10.1371/journal.pdig.0000807](https://doi.org/10.1371/journal.pdig.0000807) REVIEW
24. Omar M, Sorin V, Collins JD, et al. *Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support*. *Communications Medicine*. 2025;5:330. [doi:10.1038/s43856-025-01021-3](https://doi.org/10.1038/s43856-025-01021-3) BENCHMARK
25. World Wide Web Consortium, Provenance Working Group. *PROV-Overview: An Overview of the PROV Family of Documents*. W3C Working Group Note, 30 April 2013. [w3.org/TR/prov-overview/](https://www.w3.org/TR/prov-overview/)
STANDARD
26. HL7 International, Security and Privacy Work Group. *Provenance (FHIR resource), FHIR R5*. Maturity Level 4, Trial Use. hl7.org/fhir/R5/provenance.html STANDARD
27. Asgari E, Montaña-Brown N, Dubois M, et al. *A framework to assess clinical safety and hallucination rates of LLMs for medical text summarisation*. *npj Digital Medicine*. 2025;8:274. [doi:10.1038/s41746-025-01670-7](https://doi.org/10.1038/s41746-025-01670-7) BENCHMARK
28. Lyell D, Coiera E. *Automation bias and verification complexity: a systematic review*. *Journal of the American Medical Informatics Association*. 2017;24(2):423–431. [doi:10.1093/jamia/ocw105](https://doi.org/10.1093/jamia/ocw105)
SYSTEMATIC REVIEW

RELATED PAPERS

A write-path architecture for an AI-native EMR

The write path this read path feeds.

Evaluating clinical language models

How models behave on real records rather than exam questions.

Regulating AI in certified EHR technology

The API criterion by which outside software reaches the data.