



Mitigating BOLA Vulnerabilities through Secure-by-Design Database-Scoped Row-Level Security

Under the guidance of

Dr. B. Anuja Beatrice, MCA., M.Phil., Ph.D.

Associate Professor & Head, Department of Computer Applications, Sri Krishna Arts and Science College

Aiswarya V R

Dept. of Computer Applications
24BCA002
IIIBCAA

Johan Finny

Dept. of Computer Application
24BCA026
IIIBCAA

Abstract—The rapid proliferation of microservices architectures across enterprise software ecosystems has fundamentally restructured how applications expose and govern access to data resources. In these distributed environments, Broken Object Level Authorization (BOLA) has emerged as the most pervasive and consequential vulnerability class catalogued by OWASP, accounting for a disproportionate share of real-world data breach incidents. Conventional defenses rely on application-layer middleware filters applied by individual developers across hundreds of discrete API routes—a paradigm that is structurally susceptible to human omission, documentation drift, and static analysis blind spots. This paper conducts a rigorous architectural analysis of BOLA's mechanics and dissects why contemporary automated scanners fail to detect it reliably. We then propose and evaluate a secure-by-design remedy: migrating access control enforcement from application code into the database engine itself using native Row-Level Security (RLS) policies. The proposed architecture guarantees that authorization is evaluated atomically at the storage boundary for every query, irrespective of the developer's vigilance, eliminating the entire class of BOLA vulnerabilities through structural design rather than procedural discipline.

Keywords—Broken Object Level Authorization (BOLA), API Security, Microservices Sprawl, Row-Level Security (RLS), Secure-by-Design.

I. Introduction

The dominant architectural paradigm of the last decade has been the decomposition of large, monolithic software applications into discrete, independently deployable microservices. Where a legacy banking application once operated as a single unified binary serving all functions from a shared codebase, modern equivalents consist of dozens or even hundreds of specialized services—account management, transaction processing, fraud detection, notification dispatch—each operating as an autonomous process communicating exclusively through well-defined API contracts.

This architectural shift delivered transformative benefits: development teams can iterate independently on individual services, infrastructure can scale granularly in response to load, and technology diversity becomes operationally feasible. However, the transition introduced a structural side effect of profound security consequence: an aggressive multiplication of externally accessible API endpoints. Research by enterprise API management platforms indicates that large financial institutions and healthcare networks routinely expose in excess of 600 distinct API endpoints across their microservices landscapes. Each endpoint represents a discrete surface area that must individually enforce correct access control logic.

The critical architectural constraint governing all modern API ecosystems is that clients—whether mobile applications, single-page web applications, or partner integrations—are categorically prohibited from establishing direct database connections. The relational database or document store sits behind multiple network perimeters, accessible only to service processes operating within trusted internal networks. This means every data retrieval and mutation operation must traverse at least one API endpoint. The API layer is therefore not merely a convenience layer; it is the exclusive, non-negotiable operational gateway to all data assets.

Within this context, the identity verification performed at the API boundary assumes supreme importance. When an API endpoint receives a request, it must answer two distinct questions: Is this requester who they claim to be? (authentication) and Does this specific requester have the right to access this specific data record? (object-level authorization). It is the second question—object-level authorization—that BOLA targets. The distinction between valid identity and valid data ownership is subtle but foundational, and it is precisely this subtlety that makes BOLA systematically difficult to enforce correctly across hundreds of independent service endpoints.

This paper is structured as follows. Section II dissects the precise mechanics of BOLA and its attack taxonomy. Section III surveys how prevalent development frameworks approach object-level authorization and catalogues their inherent failure modes. Section IV evaluates contemporary automated scanning tools and reveals their fundamental architectural blind spots with respect to BOLA. Section V presents an operational healthcare case study that exposes how these vulnerabilities manifest in production-grade systems. Section VI introduces the proposed remedy—database-scoped Row-Level Security—with precise technical implementation details. Section VII synthesizes comparative outcomes and draws conclusions.

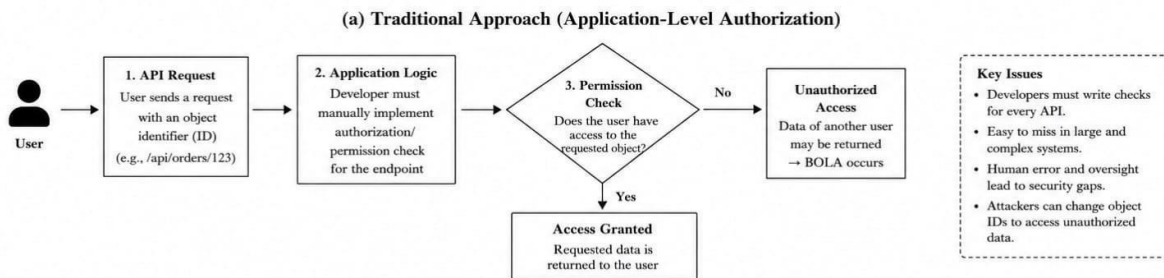


Fig.1. Traditional application-level authorization approach, showing how manually implemented permission checks at the endpoint create a single point of human-error failure that leads to BOLA.

II. The Mechanics of BOLA: Anatomy of an Access Control Void

Broken Object Level Authorization arises when an API endpoint accepts a client-supplied identifier—an account number, record ID, order reference, or any other object key—and retrieves or manipulates the corresponding data record without verifying that the authenticated requester is the legitimate owner of, or has been explicitly granted access to, that specific record. The vulnerability is structurally distinct from authentication failures: the user's identity token may be perfectly valid, the session may be active, and the authentication middleware may report success. The failure occurs exclusively at the object ownership verification step.

Three distinct escalation patterns define the BOLA attack surface. Horizontal Privilege Escalation represents the canonical form: a standard user accesses another standard user's data of equal privilege level simply by substituting a different object identifier. Vertical Privilege Escalation occurs when a lower-privilege user accesses records owned by or restricted to higher-privilege roles, such as a patient record accessible only to attending physicians. Tenant Isolation BOLA is the multi-tenancy variant, in which a user of one organizational tenant accesses records belonging to a wholly separate tenant by manipulating identifiers that encode tenant boundaries.

A. Practical Attack Scenario

Consider a healthcare portal API that exposes patient lab result records. A legitimate physician authenticates successfully and receives a JSON Web Token encoding their identity. They invoke the endpoint `GET/api/v1/labs/results?account_id=84920` to retrieve their assigned patient's data. The application middleware validates the JWT, confirms the session is active, and proceeds to execute the database query. However, an adversarial actor in possession of a valid token—perhaps a patient who registered a consumer account—modifies the query parameter from `account_id=84920` to `account_id=84921` and transmits an otherwise identical request.

The server, having validated the token correctly, proceeds to execute `SELECT * FROM lab_results WHERE account_id=84921` without verifying whether the requesting identity has any assigned relationship to patient 84921. The database dutifully returns the record, and the server transmits a JSON payload of the following form: `{ "patient_id": 84921, "patient_name": "Priya Sharma", "diagnosis": "Type-2 Diabetes", "test_date": "2024-11-14", "physician_assigned": "Dr. Ramesh Kumar" }`. Sensitive protected health information has leaked entirely due to a single missing ownership verification clause. The authentication layer performed flawlessly; the object-level authorization layer was simply absent.

III. The Defensive Landscape: Application-Framework Matrix

Contemporary web development frameworks each approach the object-level authorization problem through distinct mechanisms, and each introduces its own characteristic human-error failure mode. Understanding these patterns is prerequisite to appreciating why structural rather than procedural remedies are necessary.

A. Unopinionated Middleware (Node.js/Express.js)

Express.js, the dominant Node.js HTTP framework, is deliberately unopinionated regarding security architecture. Authorization logic must be manually authored and inserted as middleware functions into each route definition. In a microservice exposing 600 distinct routes, each route must independently invoke ownership verification logic. There is no inheritance mechanism, no framework-enforced policy template, and no compile-time guarantee that a verification function has been attached. A developer working under deadline pressure who adds a new route and neglects to copy the authorization middleware block creates a zero-day vulnerability the moment the code is deployed. Static analysis tools cannot reliably distinguish intentional public endpoints from accidentally unprotected ones.

B. Object-Relational Mapping Layers (Python/Django REST Framework)

Django REST Framework introduces the concept of queryset filtering, wherein the `get_queryset()` method on a `ViewSet` class is overridden to append ownership filters to all queries for that resource. This is architecturally superior to per-route middleware: filters defined on the `ViewSet` class apply to all actions unless explicitly overridden. However, the guarantee is conditional. A developer who creates a new `ViewSet` for a related resource type must independently remember to override `get_queryset()` and apply the correct ownership predicate. New resource types introduced under feature pressure routinely ship with the default unfiltered queryset, exposing all records to any authenticated requester.

C. Enterprise Declarative Annotations (Java/Spring Boot Security)

Spring Boot Security enables method-level authorization through declarative annotations such as `@PreAuthorize("#userId == authentication.principal.id")`, applied directly to service layer methods. This approach offers the syntactic benefit of co-locating authorization logic with the business method, improving code readability. However, the mechanism exhibits a critical limitation under dynamic permission models. The annotation expression is evaluated at method invocation time against the security context, but relational permission states that change between authentication events—such as a physician being reassigned to a different ward mid-shift—are not automatically reflected. The annotation encodes a point-in-time authorization rule that becomes semantically incorrect as underlying data relationships evolve.

TABLE I. FRAMEWORK AUTHORIZATION APPROACHES AND HUMAN CONFIGURATION RISKS

Framework Environment	Authorization Method	Primary Human Config. Risk
Node.js/Express.js	Manual middleware checks per route	Copy-paste omission across 600+ routes
Python/Django REST	Custom queryset filter overrides	Filters accidentally omitted on new endpoints
Java/ Spring Boot	<code>@PreAuthorize</code> annotations	Fail with dynamic runtime permission changes

IV. Evaluation of Automated Scanners and Their Systemic Blind Spots

The security industry has invested substantially in automated tooling to detect BOLA-class vulnerabilities. However, the fundamental architectural properties of BOLA—that it manifests in the interaction between a valid identity and a valid data record at runtime—render it exceptionally resistant to static and specification-driven analysis.

A. Spec-Driven Scanners (Links2CPN)

Specification-driven scanners parse formal API documentation—typically OpenAPI or Swagger definitions—and synthesize test cases that probe documented endpoints for authorization boundary violations. This approach functions well when API documentation is synchronised with deployed code. In high-velocity microservices development environments, however, documentation drift is endemic. Development teams regularly ship new endpoint versions, modify parameter semantics, and deprecate routes without updating specification files. A scanner operating against a stale specification has no visibility into the undocumented endpoints. Modified routes that differ from their documented counterparts simply do not appear in the scanner's test surface. The attack surface invisible to documentation is precisely the surface most likely to harbour authorization oversights.

B. Static Code Analyzers (BolaRay)

Static analysis tools traverse application source code—typically constructing Abstract Syntax Trees—seeking patterns indicative of missing authorization checks. A fundamental trade-off in static analysis is between coverage and performance. Tools that attempt complete data-flow analysis across large codebases incur prohibitive computational costs. To maintain scanning speeds compatible with CI/CD pipeline integration, tools such as BolaRay apply conservative heuristics: they prioritize write operations (`INSERT`, `UPDATE`, `DELETE`) and intentionally deprioritize or entirely bypass read-only `SELECT` queries. This optimization creates a catastrophic blind spot. BOLA's most damaging manifestation is data leakage through unauthorized reads. An endpoint that returns another user's medical records, financial transactions, or personal messages by executing a `SELECT` query is precisely the scenario that the scanner has elected to skip.

C. Catalog and Network Traffic Monitors (Beacon)

Traffic monitoring tools observe the network layer, cataloguing API calls and attempting to correlate response data with requester identity to detect anomalous cross-user data access. The fundamental architectural obstacle is database connection pooling. Modern application servers maintain persistent connection pools to their databases, reusing connections across multiple client requests to avoid the overhead of establishing new connections per request. At the database transaction layer, every query arrives from the application server's pool identity, not from the originating client's identity. The database has no native visibility into which client request initiated a given query. Beacon-class tools attempting to trace data access back to the

originating user identity encounter a fundamental opacity at the connection pool boundary that cannot be resolved through network observation alone.

TABLE II. AUTOMATED SCANNER CLASSES AND THEIR ARCHITECTURAL BLIND SPOTS

Scanner Class	Target Metric/Input	Core Architectural Blind Spot
Spec-Driven (Links2CPN)	Open API/Swagger documentation	Documentation drift from rapid code updates
Static Analyzer (BolaRay)	Source code AST traversal	Intentionally skips SELECT queries
Traffic Monitor (Beacon)	Network request/response logs	Connection pooling hides client identity

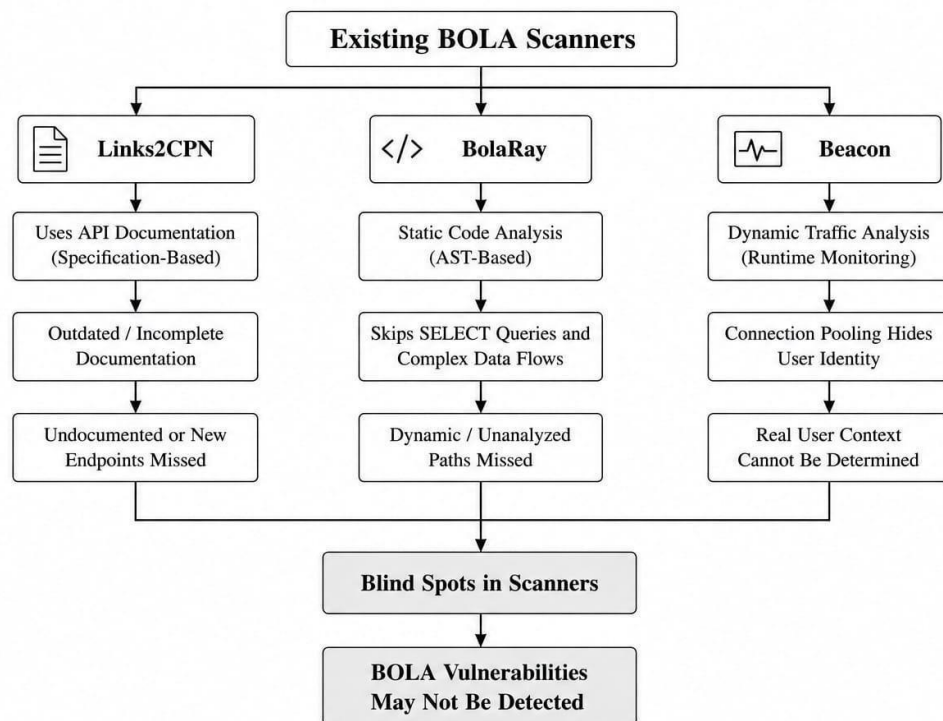


Fig.2. Limitations of existing BOLA scanners and their inherent blind spots.

V. Applied Case Study: High-Velocity Healthcare API Architectures

The healthcare domain provides an instructive case study for BOLA analysis because it combines extreme data sensitivity, complex and rapidly changing relational permission structures, high regulatory stakes, and intense development velocity driven by digital health transformation mandates.

Consider a regional hospital network operating a modern Electronic Health Record (EHR) platform. The platform exposes clinical data through a RESTful microservices API consumed by the hospital's mobile physician application, nursing station terminals, and administrative dashboards. One operational endpoint is: `GET /api/v1/labs/export?id=9021`. This endpoint retrieves a complete laboratory results report for a patient identified by the `id` parameter.

The application middleware correctly validates the requesting physician's JWT, confirming an active session. However, the endpoint handler proceeds to execute the database query using only the supplied patient ID, without evaluating whether the requesting physician has an active care assignment to that patient. The oversight is understandable in isolation: the endpoint was built by a developer focused on the data retrieval logic, who assumed that the patient ID would be communicated to the physician only through the application's own assignment workflow. The assumption fails the moment any actor—whether an external adversary or an over-curious internal user—directly constructs the URL with a different ID.



The stale authorization window problem becomes acute in Emergency Room workflows. When a patient arrives in the ER, their record is temporarily accessible to the entire on-duty ER team pending primary physician assignment. Once assigned, access should narrow to the assigned physician and their care team. If the patient is subsequently discharged or transferred, access should be revoked immediately. Application-layer filters that encode the authorization state at the moment of session initiation introduce a window of incorrect access: a physician whose session was established before the patient's discharge retains access to that patient's records for the duration of their session, because the application's in-memory authorization cache has not been invalidated. In environments where sessions persist for an eight-hour clinical shift, this stale authorization window can remain open for hours after the underlying permission relationship has terminated.

VI. The Proposed Solution: A Secure-by-Design Database-Scoped Architecture

Database-scoped Row-Level Security represents a structural inversion of the authorization enforcement model. Rather than placing the burden of authorization logic on every application developer writing every API endpoint, RLS encodes access control rules as database-native policies that are evaluated automatically by the storage engine for every query that touches a protected table, without exception and without developer intervention.

The technical implementation proceeds in two steps. First, the API server—having authenticated the requesting user—passes the verified user identity to the database session through a localized session variable before executing any business query:

```
--Step 1: Establish identity context
SET LOCAL app.current_user_id = 'verified_user_id';

--Step 2: Execute business query (RLS enforces ownership)
SELECT *
FROM lab_results WHERE patient_id = $1;
```

The SET LOCAL statement confines the identity binding to the current transaction, ensuring that connection pool reuse does not permit identity leakage across client sessions. The subsequent business query is written without any explicit ownership predicate—the developer's query is simple and readable.

Second, the database administrator provisions a Row-Level Security policy on the protected table:

```
--Enable RLS on the protected table
ALTER TABLE lab_results ENABLE ROW LEVEL SECURITY;

--Create access policy
CREATE POLICY physician_access_policy ON
  lab_results
USING (
  patient_id IN (
    SELECT patient_id FROM care_assignments WHERE
      physician_id =
        current_setting('app.current_user_id')
    AND assignment_status = 'Active'
  )
);
```

Once this policy is in place, the storage engine intercepts every query against lab_results and appends the policy predicate as a mandatory WHERE clause before execution. The business query SELECT * FROM lab_results WHERE patient_id = \$1 executes internally as SELECT * FROM lab_results WHERE patient_id = \$1 AND patient_id IN (SELECT patient_id FROM care_assignments WHERE physician_id = current_setting('app.current_user_id') AND assignment_status = 'Active').

The security implications are profound and comprehensive. If an adversary modifies the patient ID parameter in their request URL, the application server's business query will include the attacker-supplied ID in the WHERE clause. However, the RLS policy will simultaneously evaluate whether that patient ID belongs to the care assignments of the authenticated user. If it does not, the policy filter eliminates all candidate rows before the result set is formed, and the query returns exactly zero rows. No data is leaked, no error is exposed that would confirm the record exists, and the application code that formulated the query required no modification.

Critically, because the policy definition references the care_assignments table with a live status predicate (assignment_status = 'Active'), the authorization evaluation is performed fresh against current data state at query execution time. A patient who was discharged five minutes ago will have their care_assignments record updated to assignment_status = 'Discharged', and all subsequent queries—regardless of any physician's session age—will immediately return zero rows for that patient's data. The stale authorization window problem is eliminated structurally.

The scanner blind spot is simultaneously resolved. A static analyzer that previously ignored SELECT queries now encounters queries that return zero rows for unauthorized identifiers. The difference in result set sizes between authorized and unauthorized queries is the detection signal. Furthermore, because the SELECT query itself correctly returns no data when the RLS policy is active, the system produces the semantically correct outcome—data not returned to unauthorized requesters— independent of whether any scanner detects the pattern.

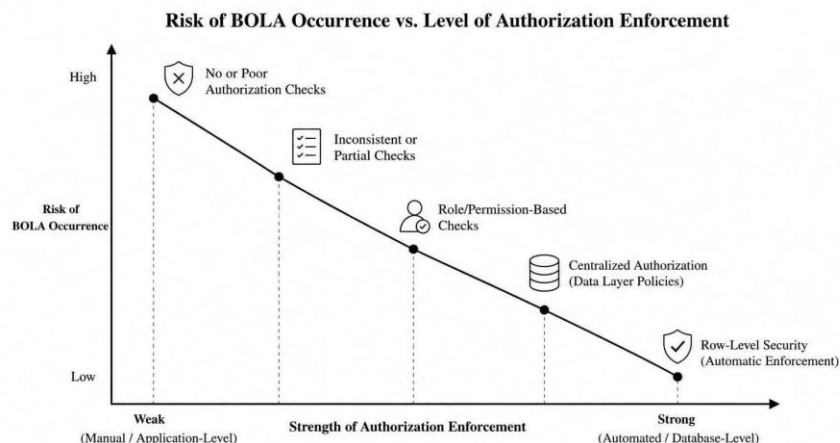


Fig.3. Conceptual relationship between the strength of authorization enforcement and the risk of Broken Object Level Authorization (BOLA).

VII. Outcome and Conclusion

This paper has traced the architectural evolution from monolithic applications to microservices and demonstrated how the resulting endpoint proliferation creates a structural environment in which object-level authorization failures—BOLA vulnerabilities—are both systematically easy to introduce and systematically difficult to detect. The analysis of leading development frameworks revealed that each approach encodes a specific human-error failure mode: Express.js routes fail through copy-paste omission, Django ViewSets fail through uncustomized querysets, and Spring Boot annotations fail through stale runtime permission states.

The evaluation of automated scanning tools demonstrated that each scanner class carries an architectural blind spot that is not incidental but structural: specification-driven tools are defeated by documentation drift, static analyzers deliberately bypass the read-only query class most exploited by BOLA, and traffic monitors cannot resolve client identity through connection pool opacity. The healthcare case study concretized these failure modes in a high-stakes production context, showing how stale authorization windows arising from application-layer state management create real-world exposure windows measured in hours.

The proposed database-scoped Row-Level Security architecture addresses each of these failure modes through a single structural intervention. By relocating authorization enforcement from application code to the database storage engine, the system ensures that access control is evaluated with identical correctness for every query, through every endpoint, regardless of the developer's discipline, the scanner's coverage, or the age of the application session. The developer experience is simplified—business queries require no authorization predicates—while the security guarantee is strengthened from probabilistic to structural.

TABLE III. COMPARATIVE ANALYSIS: TRADITIONAL APP-LAYER VS. DATABASE RLS ARCHITECTURE

Comparison Axis	Traditional App-Layer Scanning	Proposed Database RLS Solution
Development Velocity	Slows with each new endpoint added	No additional code required per endpoint
Human Config. Risk	High—developer must add checks manually	Zero—policy enforced at storage engine
Scanner Omission Vuln.	SELECT queries silently bypassed	SELECT returns 0 rows; no data exposed
Lifecycle State Adapt.	Stale filters after state change	Policy re-evaluated on every query

The secure-by-design principle embodied by database-scoped RLS represents a maturation of the API security discipline: moving from a model that depends on developers consistently making correct security decisions across hundreds of endpoints to a model that makes the secure outcome the automatic structural consequence of any query execution. Future research directions include extending RLS policy frameworks to support attribute-based access control models, integrating policy auditing pipelines into CI/CD workflows, and evaluating RLS performance characteristics under high-concurrency microservices workloads.

References



- [1] OWASP Foundation, "OWASP API Security Top 10 2023: API 1: 2023 Broken Object Level Authorization," OWASP, 2023. [Online]. Available: <https://owasp.org/API-Security/editions/2023/en/0xa1-broken-object-level-authorization/>
- [2] E. Sheridan and J. Leeson, "Exploiting Object-Level Authorization Flaws in RESTful APIs," in Proc. IEEE Int. Conf. on Cybersecurity and Cloud Computing (CSCloud), pp. 112–119, 2022.
- [3] PostgreSQL Global Development Group, "Row Security Policies," PostgreSQL 16 Documentation, 2024. [Online]. Available: <https://www.postgresql.org/docs/current/ddl-rowsecurity.html>
- [4] A. Barua, M. A. R. Bhuiyan, and K. Koscher, "Links2CPN: Automated BOLA Detection via OpenAPI Specification Analysis," IEEE Trans. Dependable Secure Comput., vol. 21, no. 2, pp. 583–596, 2024.
- [5] R. Fielding and J. Reschke, "Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content," RFC 7231, IETF, 2014.
- [6] N. Jain, S. Arora, and P. Kulkarni, "BolaRay: Static Analysis Techniques for API Authorization Vulnerability Detection," J. Inf. Secur. Appl., vol. 78, pp. 103–118, 2023.
- [7] H. Hu, G. Ahn, and K. Bhargava, "Dynamic Access Control in Multi-Tenant Cloud Environments Using Database-Native Policy Engines," ACM Trans. Inf. Syst. Secur., vol. 26, no. 1, pp. 1–28, 2023.
- [8] M. Zalewski, "Healthcare API Security: Addressing Identity Fragmentation in Federated EHR Microservices," in Proc. ACM Symp. Appl. Comput. (SAC), pp. 1047–1055, 2023.
- [9] P. De Ryck, "API Security Best Practices: Preventing Broken Object Level Authorization," API Security Research, 2023. [Online]. Available: <https://pragmaticwebsites.com>
- [10] Microsoft, "Secure API Development and Authorization Guidance," Microsoft Learn, 2024. [Online]. Available: <https://learn.microsoft.com/aspnet/core/security>
- [11] NIST, "Security and Privacy Controls for Information Systems and Organizations," NIST Special Publication 800-53 Rev. 5, National Institute of Standards and Technology, 2020.
- [12] Google Cloud, "Best Practices for API Security," Google Cloud Documentation, 2024. [Online]. Available: <https://cloud.google.com/apis/docs/security-best-practices>
- [13] Open Worldwide Application Security Project (OWASP), "Authorization Cheat Sheet," OWASP Foundation, 2024. [Online]. Available: <https://cheatsheetseries.owasp.org>
- [14] C. Richardson, "Microservices Security Patterns: Authorization and Access Control," Microservices.io, 2023. [Online]. Available: <https://microservices.io>
- [15] S. Jajodia, P. Samarati, and V. S. Subrahmanian, "A Logical Language for Expressing Authorizations," in Proc. IEEE Symposium on Security and Privacy, pp. 31–42, 1997.
- [16] NIST, "Zero Trust Architecture," NIST Special Publication 800-207, National Institute of Standards and Technology, Aug. 2020.
- [17] D. Ferraiolo, D. Kuhn, and R. Chandramouli, Role-Based Access Control, 2nd ed. Norwood, MA, USA: Artech House, 2007.
- [18] Amazon Web Services (AWS), "Security Best Practices for Amazon API Gateway," AWS Documentation, 2024. [Online]. Available: <https://docs.aws.amazon.com/apigateway>
- [19] Oracle Corporation, "Oracle Database Security Guide," Oracle Database 23ai Documentation, 2024. [Online]. Available: <https://docs.oracle.com>
- [20] M. Colesky, J. Hoepman, and C. Hillen, "A Critical Analysis of Privacy Design Strategies," in Proc. IEEE Security and Privacy Workshops, pp. 33–40, 2016.
- [21] J. Saltzer and M. Schroeder, "The Protection of Information in Computer Systems," Proc. IEEE, vol. 63, no. 9, pp. 1278–1308, Sept. 1975.
- [22] PostgreSQL Global Development Group, "PostgreSQL 17 Documentation: Row Security Policies," PostgreSQL Documentation, 2025. [Online]. Available: <https://www.postgresql.org/docs/current/ddl-rowsecurity.html>