

Semantic Digital Libraries in Public Administration: Optimizing Certificate Request Management

Giovanni Carau¹, Pasquale D’Avino², Donatella Firmani^{3,†}, Elio Gullo² and Luigi Laura¹

¹Università Telematica Internazionale Uninettuno

²Dipartimento della Funzione Pubblica

³Sapienza Università di Roma

Abstract

This work extends the research presented in “Semantic Digital Libraries in Public Administration” [1] evolving the proposed knowledge graph into a distributed semantic digital library. Implemented as Linked Open Data (LOD), this infrastructure defines the structure, entities, attributes, relationships, and rules governing certificate request management in Public Administration (PA) [2].

The primary objective of this study is to integrate the PA distributed library and its ontologies into the National Semantic Catalog SCHEMA.gov.it [3] to enrich and improve the certification delivery process across PAs, reducing both semantic and structural errors.

The previous study introduced a centralized semantic framework based on RDF triples (Subject, Predicate, Object) [4], [5] for representing and interconnecting the entities involved in certificate request processes.

This work transforms the centralized knowledge graph into a distributed Linked Open Data (LOD) ecosystem [6], enabling semantic orchestration and navigation among remote information resources, interoperability across administrations, and enhanced administrative transparency. Building upon this evolving distributed semantic libraries, the study proposes a microservices architecture for managing real data certificates, along with a method to prioritize the implementation of the microservices [7]. The microservices ecosystem is designed to handle different types of citizen certificate requests, leveraging the semantic library as a structural rule engine.

The proposed architecture enables Public Administrations, through dedicated portals, to interact with an ecosystem of microservices by orchestrating their certificate request delivery. These services leverage the semantic library to define, validate, and manage certificate requests, ultimately performing update operations on the underlying citizen datasets. To achieve this goal, an empirical model is introduced to support the planning and prioritization of microservice development for digital certificate issuance and management, contributing to a more efficient process organization and a sustainable evolution of digital public services. The proposed method identifies the centrality measure within the knowledge graph that best represents the actual distribution of request volumes for a subset of certificates—validated through ANPR data [8]—and generalizes this measure to the complete certificate domain to guide the prioritization of microservice implementation.

Keywords

Semantic Digital Libraries, Knowledge Graphs, Distributed Semantic Digital Library, Public Administration Services / Interoperability, Microservices Orchestration, Certificate Management Systems,

1. Introduction

The research presented in “Semantic Digital Libraries in Public Administration” [1] addressed one of the major challenges faced by the Italian Public Administration (PA) [2]: the fragmentation and inefficiency of document and certificate management processes. The previous work proposed a semantic centralized approach based on a Knowledge Graph to represent entities, documents, and their relationships, providing a unified and interoperable framework for certificate request management. In the current work, the Knowledge Graph is explicitly evolved into a distributed semantic digital library: a structured, reusable, and evolvable distributed and federate repository that defines entities, attributes, relationships,

IRCDL 2026: 22nd Conference on Information and Research Science Connecting to Digital and Library Science, February 19-20, 2026, Modena, Italy

[†] Funding of Donatella Firmani.

✉ giovanni.carau@gmail.com (G. Carau); p.davino@governo.it (P. D’Avino); donatella.firmani@uniroma1.it (D. Firmani); e.gullo@governo.it (E. Gullo); luigi.laura@uninettuno.it (L. Laura)

ORCID: 0000-0003-0358-3208 (D. Firmani); 0000-0002-9421-8566 (E. Gullo); 0000-0001-6880-8477 (L. Laura)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

and rules governing certificate request processes. This architecture is designed to support transparency and process automation across multiple administrative levels, including national, regional, and local authorities. Building on these foundations, the present study focuses on advancing the practical implementation of that model. While the previous publication demonstrated the conceptual and technical feasibility of using a Knowledge Graph that serves solely as a semantic backbone, this work establishes the evolution of the Knowledge Graph into a distributed semantic reference library that actively drives definition, validation logic, and decision rules for certificate management microservices. These microservices, in turn, consume real citizen's certificate data in compliance with the library's semantic rules. The motivation for this extension arises from the need to optimize the allocation of resources and development efforts in the digital transformation of public services implemented via a microservices ecosystem [7]. Specifically, by identifying the most representative centrality measures within the Knowledge Graph that most correlate with the actual distribution of certificate requests (based on ANPR - the National Resident Population Registry - data [8]), it becomes possible to determine which certificate management microservices should be implemented first. In summary, this study represents a natural evolution and continuation of the previous research: it moves from conceptual modeling toward methodological and operational validation, contributing to the development of intelligent, graph-driven microservice ecosystems for digital certificate issuance and management within the Public Administration [2].

2. A semantic digital library for certificate request management in public administration

This section builds upon the frameworks and methodologies introduced in “Semantic Digital Libraries in Public Administration” [1]. It details their evolution from a centralized Knowledge Graph into a distributed semantic reference library that leverages Linked Data to enable interoperability across Public Administrations. The study first provides an overview of the methodologies applied in the previous work, followed by a detailed description of the process used to transform the centralized graph—including its ontologies and governing rules—into a Linked Open Data (LOD) ecosystem [6]. This evolution is extended through an architecture of consumer microservices [7], invoked by web client portals to orchestrate complex certificate request scenarios. Finally, a method is introduced to evaluate various centrality measures within the Knowledge Graph, identifying the metric most closely correlated with actual ANPR [8] certificate request volumes. This selected measure is then generalized across the entire certificate domain to prioritize microservice implementation. These services dynamically interact with the semantic library, using SPARQL queries, to retrieve structures, dependencies, and rules, ensuring consistent and error-free behavior across diverse Public Administration portals.

2.1. Methodologies and Frameworks

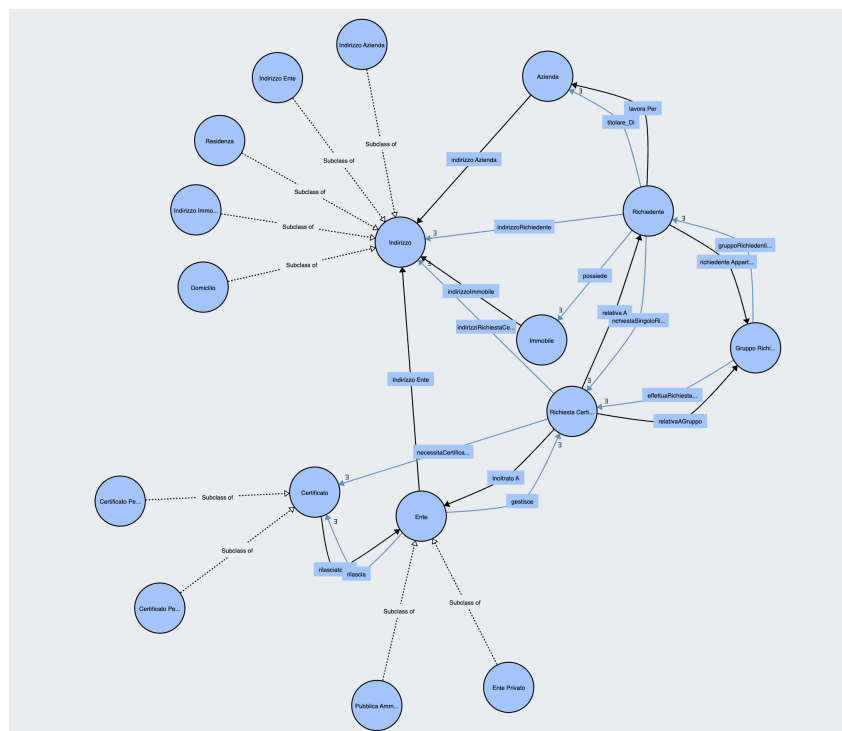
The previous study investigated a Knowledge Graph model designed to streamline and automate certificate request management within Public Administration, with the goal of creating a configurable, maintainable, and widely accessible framework. Various technologies were evaluated, including ontologies and OWL for modeling, RDF for structured data representation, and specialized tools for RDF graph management, visualization, and analysis. Based on criteria such as flexibility, efficiency, and adherence to Semantic Web standards (RDF, RDFS, OWL), the following tools were selected:

1. WebVOWL [9]: For ontology creation.
2. OnOntology [10]: For ontology publication.
3. Excel Template: For manually preparing the dataset.
4. RdfLib (Python) [5]: For loading and orchestrating the Knowledge Graph.
5. NetworkX [11]: For data navigation, application of partitioning criteria, and the computation of centrality measures.
6. Pyvis [12]: For graph rendering and user interaction.

This work presents the evolution of the Knowledge Graph into a distributed semantic reference library based on Linked Open Data (LOD) [13], enabling interoperability and collaborative participation among Public Administrations [2]. Client-side applications—such as portals and certificate management microservices—can query the LOD semantic library via SPARQL endpoints to retrieve semantic knowledge and defined ontology-driven rules and reasoning across entities.

The present study builds upon the conceptual and semantic framework introduced in “Semantic Digital Libraries in Public Administration” [1], which proposed a semantic architecture for managing certificate requests within Public Administration through a knowledge-based digital library. That previous work established the ontology and RDF-based structure [4] for representing entities, relationships, and processes involved in certificate management, providing the semantic foundation for the operational implementation of a knowledge graph.

The ontology includes several key entities that constitute the semantic core of the certificate request process: Public Administration, representing issuing authorities; Certificate Request, the central node linking requesters, certificates, and institutions; Certificate, with specialized subclasses such as Certificate for Non-EU Citizens and Certificate for Minors; Requester and Requesting Group; and domain-specific entities including Company, Property, and Address, each modeled through subtypes that reflect administrative scenarios.



The ontology [10] was serialized in Turtle (TTL) [14] format and published through OnToology [10], ensuring URI persistence [15] [16] [17] and alignment with Linked Open Data standards [6].

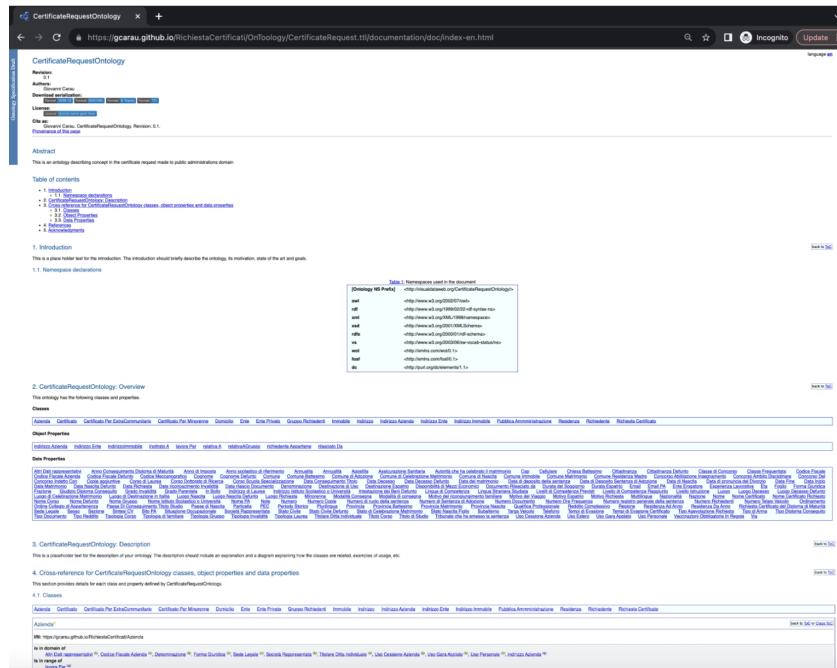


Figure 2: Certificate Request Ontology: publication on a specific URI.

Data preparation activities involved configurable Excel templates completed by Public Administration operators, defining both the catalogue of certificate types and the configuration of entities, attributes, and navigation rules. These templates were then transformed into CSV datasets and processed through a Python-based architecture adopting the Model–View–Controller (MVC) pattern [18]. As shown in Figure 3, RDFLib [5] was used for loading the graph through a rule engine, NetworkX [11] for traversal and partitioning, providing the basis for centrality analysis, and Pyvis [12] for rendering and interactive visualization via VisJS.

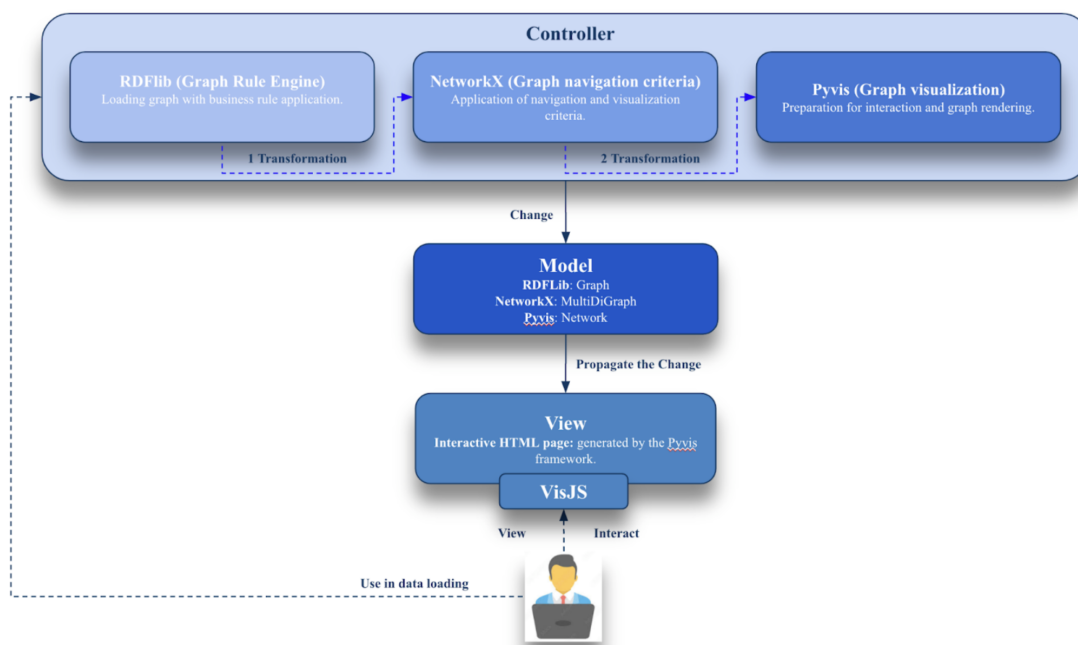


Figure 3: Solution Architecture: Python Frameworks used in the MVC architecture of the solution.

2.3. Transitioning from Knowledge Graphs to Distributed Linked Open Data: Leveraging Graph-Driven Methods for Microservice Prioritization

The centralized knowledge graph is designed to be decomposed into subgraphs corresponding to the domain of competence of each Public Administration, enabling distributed governance while preserving semantic coherence. These subgraphs are published as Linked Open Data, enabling controlled inter-administrative interaction and supporting integration into the national semantic interoperability framework, including National Semantic Catalog SCHEMA.gov.it. The full Linked Data transformation process—graph partitioning, data publication, and SPARQL [19] based access—is summarized in Figure 4.

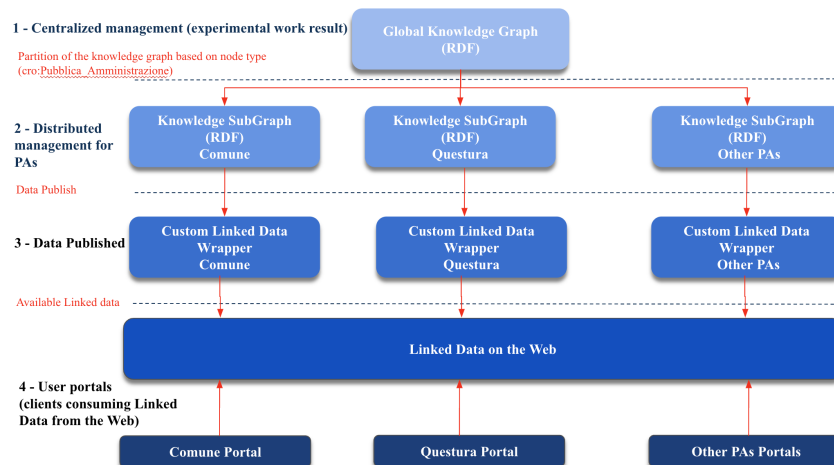


Figure 4: Pathway to translate the knowledge graph, the work result, into a Linked Data perspective.

These custom wrappers perform real-time retrieval of semantic data via SPARQL queries, ensuring the digital library consistently serves up-to-date administrative information directly from the source.

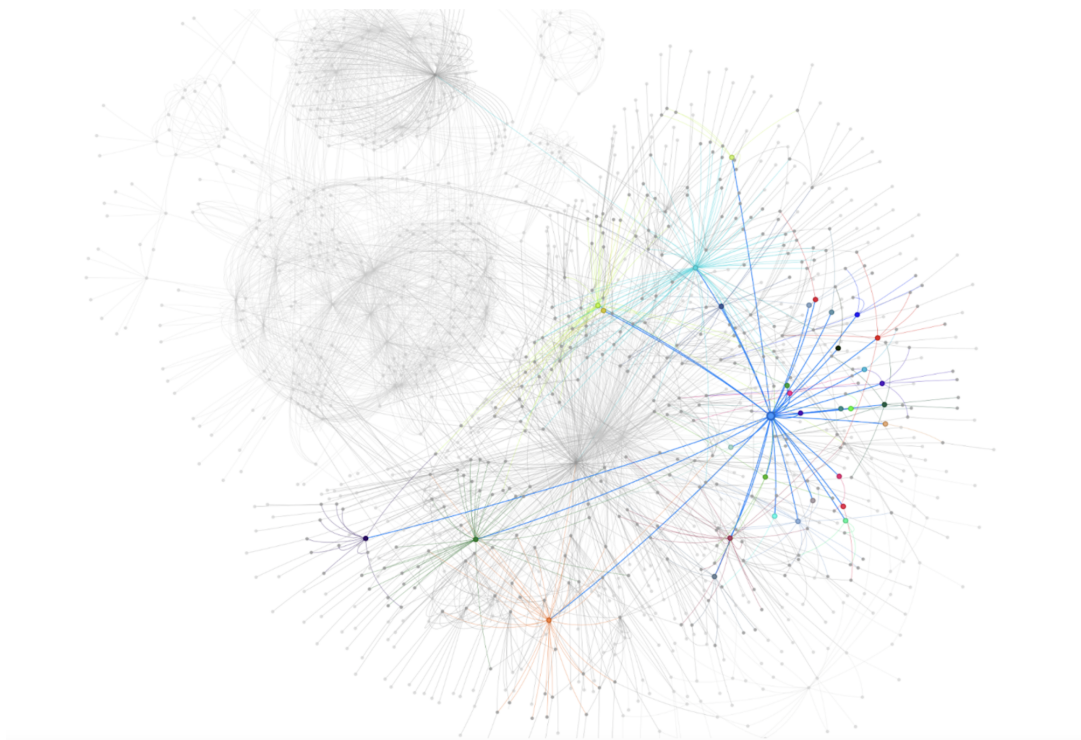


Figure 5: Knowledge Graph: visualization of Public Administrations as colored nodes (1st level classification).

Building on this operational implementation, the present work leverages the realized semantic infrastructure [20] as the basis for developing an empirical method for assessing the structural relevance of certificate types within the administrative domain. The knowledge graph (Figure 5), with its formalized classes, relationships, and navigation paths, provides the coherent semantic environment required for computing and comparing centrality measures.

In this phase, the graph becomes an analytical instrument for determining which centrality metric aligns most closely with real-world request volumes. A representative subset of certificate types, for which empirical usage statistics are available, is employed to validate the approach.

Once identified, the most representative centrality measure is applied to the full set of certificates modeled in the graph. The underlying assumption is that certificate types that are structurally central in the knowledge graph are more likely to correspond to high-demand services; this hypothesis is empirically tested by comparing graph centrality measures with real request volumes published by ANPR.

As a preparatory step toward implementing microservices for the digital delivery of certificates, this study introduces a method to identify the nodal relevance measure that most accurately and representatively reflects the empirical distribution of certificate request volumes. The analysis focuses on a subset of certificates, using data published by the "The National Registry of Resident Population" (ANPR) [8] (Figure 6).

It is important to note that the ANPR dataset used for validation is numerically limited compared to the comprehensive scope of the knowledge graph; however, it represents a highly significant and authoritative sample of national-level administrative activity.

Given the complexity of the certificate domain and the limited resources available for digital transformation, it is not feasible to implement all certificate-related microservices simultaneously; therefore, a prioritization strategy based on nodal relevance measure is required.

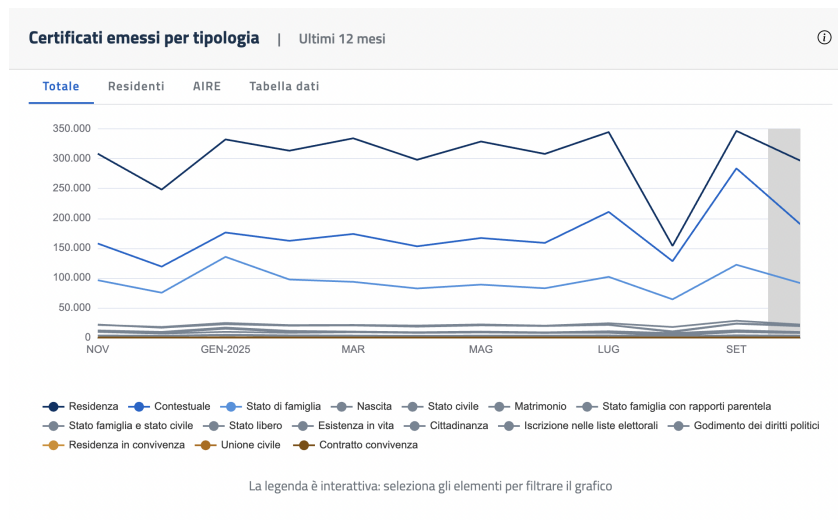


Figure 6: Emission volumes by certificate type.

These data, which indicate the types of certificates most frequently requested at the national level (number of certificates limited compared to that of the knowledge graph), served as an empirical benchmark for validating the proposed model. The analysis [8] (Figure 7) reveals that approximately 30% of certificate requests exhibit anomalies in their processing workflows. These irregularities stem in part from inconsistencies and lack of uniformity in information retrieval and inter-entity relationships challenges. Such issues could be mitigated and resolved through the adoption of an ontology, such as the one developed in this work.

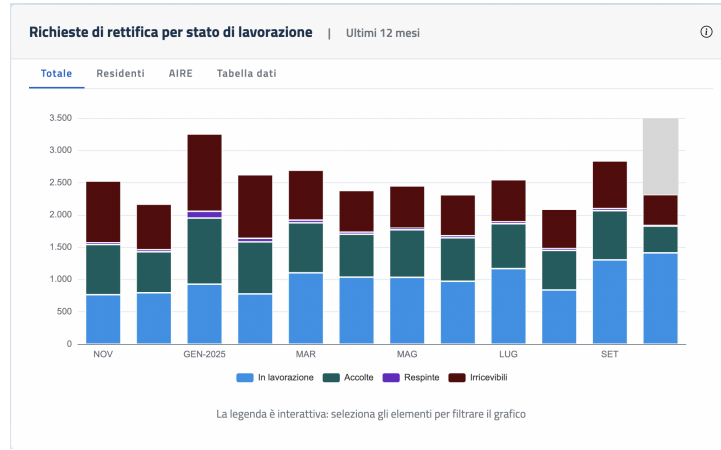


Figure 7: Processing status of certificates with status evidence.

The proposed method enables the selection of the centrality measure that best aligns with observed behaviors, thereby extending the analysis to the entire domain of certificate requests represented within the knowledge graph. This extension supports the definition of prioritization criteria for planning and developing microservices, guiding the design of digital service delivery systems according to a data-driven approach grounded in the relational structure of the graph.

The objectives of this study are threefold: (i) to identify the most relevant certificates and attributes within the administrative domain, (ii) to identify key attributes and semantic elements that should be standardized and shared across microservices, (iii) to provide an empirical foundation for the automation and orchestration of digital services related to certificate issuance and management, following a prioritization based on observed request volumes.

Leveraging datasets published by ANPR [8], this study developed a distributed reference library to model the semantic structure of the public certificate microservices ecosystem and to determine the most suitable measure method for assessing certificate relevance. Structural analysis techniques grounded in centrality and node-ranking algorithms were applied to support the prioritization of digital services. The analysis incorporated the principal centrality measures—Degree, Betweenness, Closeness, and Eigenvector—alongside the PageRank algorithm [21], which assesses topological importance of node through a probabilistic transition model.

As widely documented in the literature [22], these measures provide complementary perspectives on nodes' roles within complex networks: from direct connectivity (Degree), to intermediation capacity (Betweenness) [23], to average proximity (Closeness) [22], to influence derived node from highly connected neighbors (Eigenvector), and to the probability of access via random walks (PageRank) [21]. The following are the centrality measures applied to the certificate request knowledge graph.

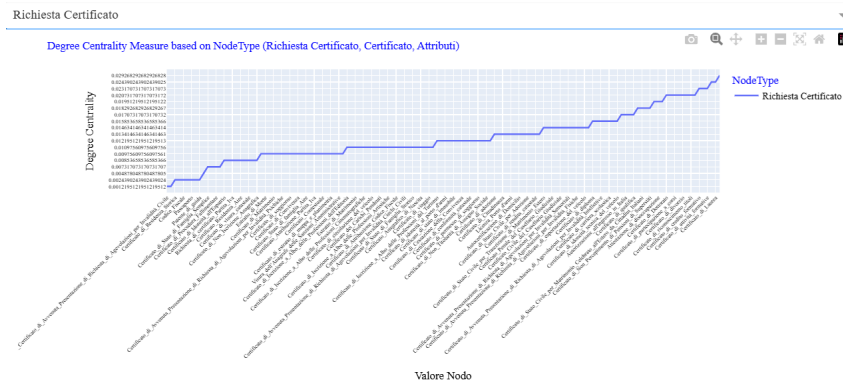


Figure 8: Degree Centrality of Certificate Request.

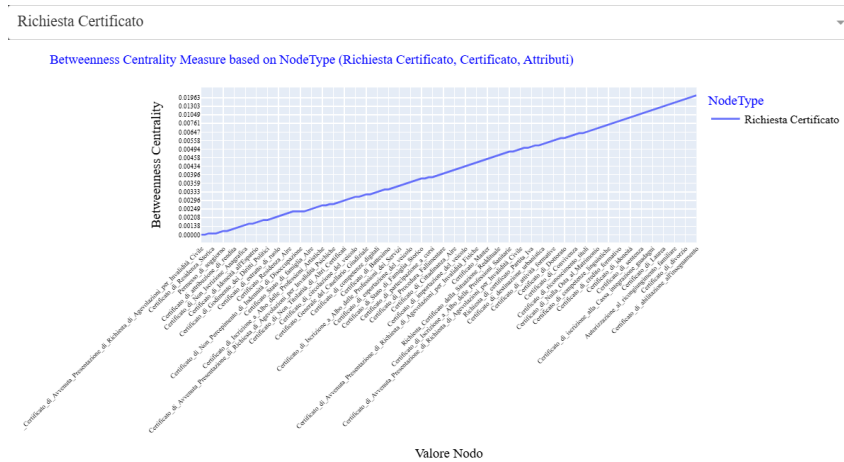


Figure 9: Betweenness Centrality of Certificate Request.

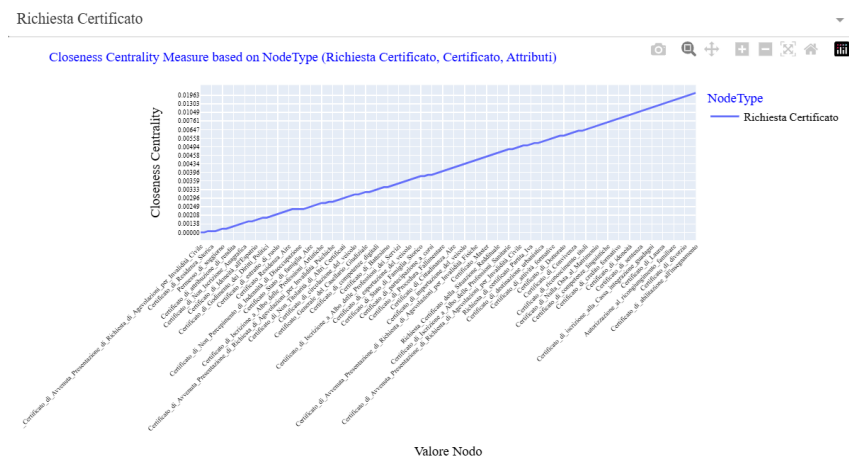


Figure 10: Closeness Centrality of Certificate Request.

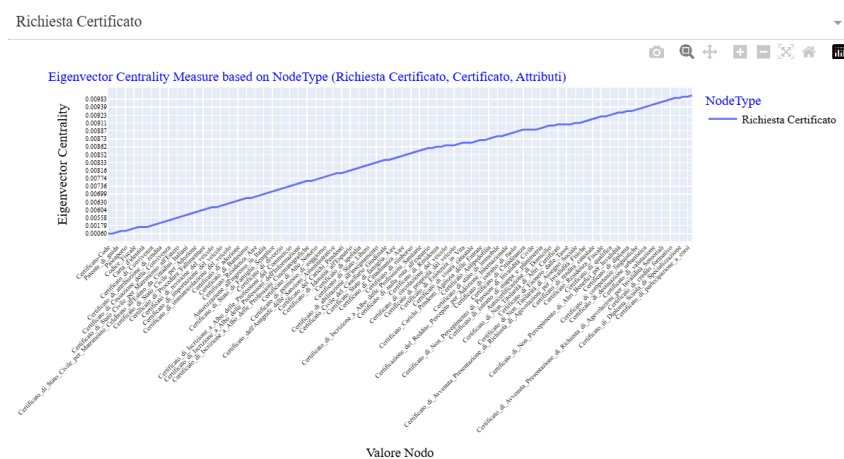


Figure 11: Eigenvector Centrality of Certificate Request.

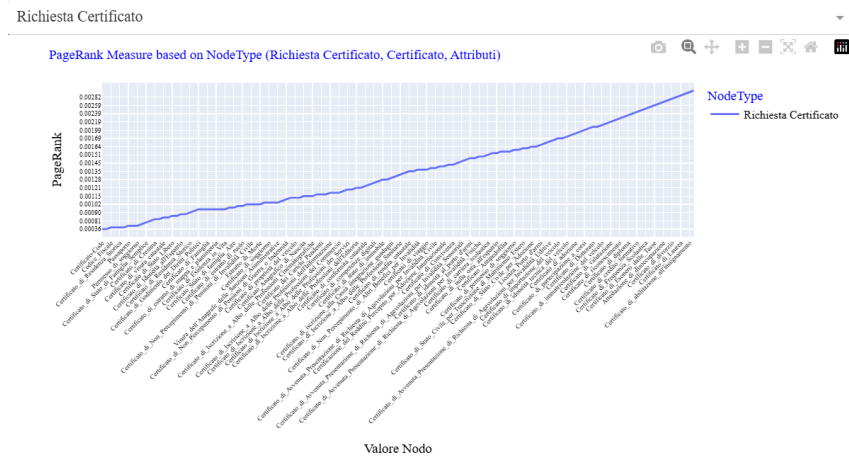


Figure 12: Page Rank of Certificate Request.

To evaluate the consistency of the model with empirical data, a Pearson correlation analysis was performed between each nodal relevance measure and the actual request volumes recorded in the ANPR dataset [8]. This rigorous selection ensures that our prioritization criteria for developing microservices are not just theoretical, but are grounded in the actual structural reality of Public Administration.



Figure 13: Correlation (R) between ANPR's Certificate Volume and Degree-Centrality score: 0.594765333638955.



Figure 14: Correlation (R) between ANPR's Certificate Volume and Betweenness-Centrality score: 0.6386054969399237.



Figure 15: Correlation (R) between ANPR's Certificate Volume and Closeness-Centrality score: 0.6386054969399237.

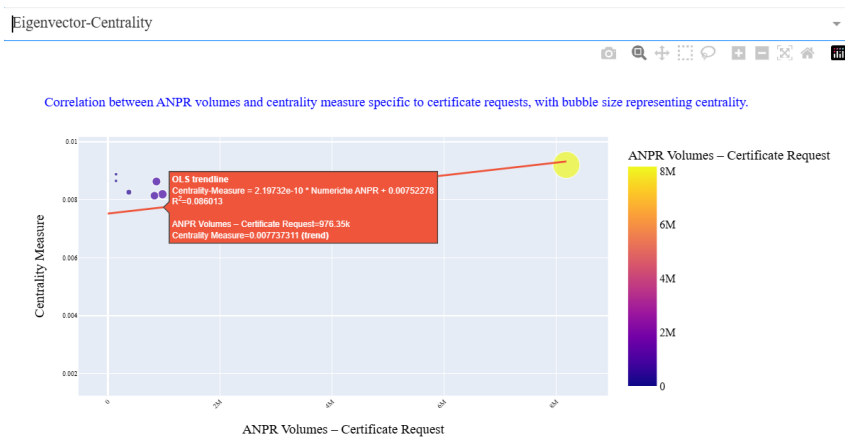


Figure 16: Correlation (R) between ANPR's Certificate Volume and EigenVector-Centrality score: 0.29327994370824584.



Figure 17: Correlation (R) between ANPR's Certificate Volume and PageRank score: 0.6455037022006045.

The results are summarized below:

Measure / Algorithm	Correlation with ANPR Volumes
PageRank (stochastic ranking)	0.6455
Betweenness Centrality	0.6386
Closeness Centrality	0.6386
Degree Centrality	0.5948
Eigenvector Centrality	0.2933

Table 1

Correlation between different centrality measures and ANPR volumes.

The PageRank algorithm [21] emerges as the most effective metric for capturing the empirical distribution of certificate requests, followed closely by Betweenness Centrality [23]. Despite the noise inherent in legacy data, this outcome suggests that certificate relevance depends not only on direct connectivity but also on the strategic and intermediary position of nodes within the network. The superior performance of PageRank suggests that certificate relevance depends not only on direct associations but also on indirect dependencies and navigational paths within the administrative knowledge structure, which are effectively captured by random-walk-based ranking models. While Betweenness Centrality identifies certificates that act as 'bridges' between different administrative sectors, PageRank highlights those that serve as the structural foundation for critical workflows. PageRank's superior performance proves that administrative relevance is defined by being a key dependency rather than just a connector. By pinpointing the 'root' certificates upon which both dependent certificates and administrative processes rely, PageRank provides a clear roadmap for prioritization: building the foundations first ensures the stability of the entire digital ecosystem. From an operational perspective, the selected measure is employed to define development priorities for microservices dedicated to the digital management and delivery of certificates. Specifically, the application of PageRank supports:

1. Planning, design, and implementation of microservices according to the topological relevance of certificates, including identification of central and mandatory certificates, which are consequently prioritized for automation.
2. Recognition and centralization of shared attributes across different Public Administrations [2], in alignment with the once-only principle, thereby preventing redundant data entry once information has been acquired.

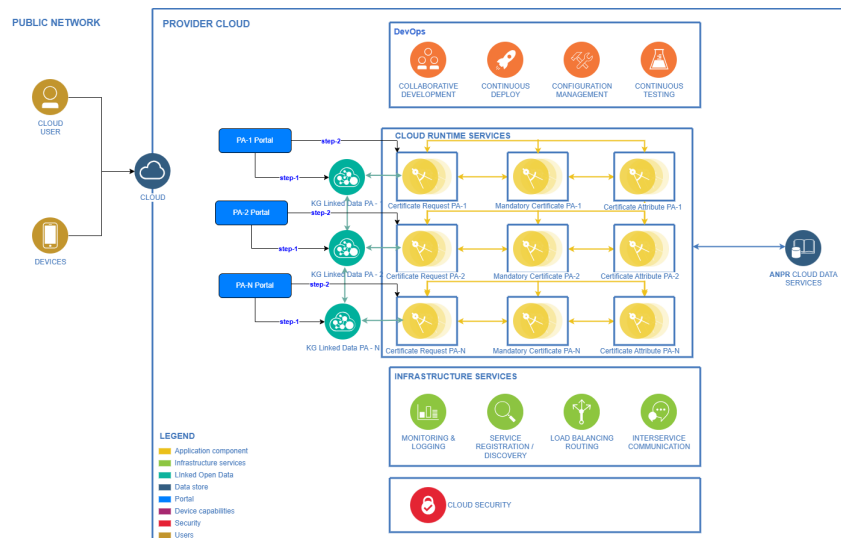


Figure 18: Architecture showing the portal's interaction with the knowledge graph (1) and its subsequent communication with dedicated microservices (2).

The proposed architecture (Figure 18) illustrates the portal's interaction with the knowledge graph (Step 1) to define and structure certificate requests. Subsequently, the portal communicates with dedicated microservices (Step 2), which first validate the forwarded requests by interacting with the semantic library, and then perform data update operations on live datasets.

In this context, semantic standardization enables the 'Once-Only' principle: data becomes a reusable asset in the digital library, ensuring administrations access up-to-date, verified information. This eliminates the need for citizens to provide the same data multiple times and reduces redundant certificate requests.

Looking ahead, integrating the knowledge graph into SCHEMA.gov.it [3] marks a strategic milestone for public data harmonization and ontology standardization. This transition establishes a national interoperability framework [3] that fosters the reuse of schemas and microservices, promoting an integrated, intelligent administrative infrastructure.

3. Conclusion

In conclusion, this study demonstrates how a knowledge graph evolved as a distributed semantic digital library can act as a foundational infrastructure for the design, validation, and prioritization of microservice-based public services. By combining semantic modeling with quantitative assessment, the approach identifies the PageRank algorithm as the centrality measure that best correlates with actual certificate request volumes—even when validated against a numerically limited but authoritative ANPR dataset. Consequently, this measure is applied across the entire certificate domain to guide microservice prioritization and the standardization of shared attributes.

The knowledge graph's semantic coherence and extensibility into the distributed semantic library enable intuitive navigation, consistent interpretation of information, and the scalable integration of new certificate types. Grounding service planning in empirical relevance aligns development with real administrative demand, enhancing efficiency and user experience.

Concluding the undertaken path, both through a centralized approach with the knowledge graph and through a cooperative and decentralized evolution using Linked Open Data, has fostered collaboration between computers and users. This has allowed the developed application of the semantic library to assist users in automating complex tasks within the certificate processing workflow. Achieving this goal was made possible by defining an ontology of information, ensuring a comprehensible syntax for both users and machines, serving as a connecting element between them. In this perspective, the knowledge graph and its evolution into the distributed semantic library are not merely an analytical instrument but a cognitive infrastructure supporting the evolution of the public information system toward more efficient, interoperable, and citizen-centric digital services [24].

Looking forward, this framework lays the groundwork for integration into a broader Linked Open Data ecosystem—specifically within the national semantic framework of schema.gov.it—promoting interoperability, automated information exchange, and the practical application of the 'Once-Only' principle. By uniting semantic and data-driven approaches, the study outlines a practical path toward a more intelligent, interoperable, and responsive digital administration.

Ultimately, this system transcends the role of a simple management tool, evolving into a comprehensive knowledge infrastructure for Public Administration. By adopting the Digital Library paradigm, a scalable framework is provided where administrative knowledge is not just stored, but semantically organized and curated, enabling more intelligent, transparent, and proactive public services. While this study focuses on certificate management, the framework's modular nature allows for seamless adaptation to other critical domains, such as healthcare data orchestration and public procurement monitoring, where semantic consistency is equally vital.

Future developments could extend this framework toward a multi-dimensional prioritization model, integrating graph-based centrality measures with operational variables such as implementation costs and technical complexity to provide an even more comprehensive decision-support system for digital transformation.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

- [1] Valentina Albano, Giovanni Carau, Donatella Firmani, Elio Gullo, Claudia Ilardi, and Luigi Laura. “Semantic Digital Libraries in Public Administration: A Knowledge Graph Approach to Certificate Request Management”. In: *Proceedings of the 21st Conference on Information and Research Science Connecting to Digital and Library Science (IRCDL 2025)*. Ed. by Marcella Cornia, Giorgio Maria Di Nunzio, Donatella Firmani, Stefano Mizzaro, Giuseppe Serra, Sara Tonelli, and Alessandro Tremamunno. Vol. 3937. CEUR Workshop Proceedings. Published under Creative Commons License Attribution 4.0 International (CC BY 4.0). Udine, Italy: CEUR-WS.org, Feb. 2025, –. URL: <https://ceur-ws.org/Vol-3937/paper11.pdf>.
- [2] (n.d.) *Elenco delle Amministrazioni - SIOPE*. 2023. URL: https://www.rgs.mef.gov.it/VERSIONE-I/e_government/amministrazioni_pubbliche/siope/elenco_delle_amministrazioni/ (visited on 11/16/2023).
- [3] Italian National Catalogue for Data and Information Exchange among Public Administrations. *Italian National Catalogue for Data and Information Exchange among Public Administrations*. <https://schema.gov.it/>. Accessed: 2025-10-26. n.d.
- [4] (n.d.) *Resource Description Framework (RDF)*. 2023. URL: <https://www.w3.org/RDF/> (visited on 10/01/2023).
- [5] (n.d.) *Getting started with RDFLib*. 2023. URL: <https://rdflib.readthedocs.io/en/stable/> (visited on 10/15/2023).
- [6] (n.d.) *Linked Open Data*. 2023. URL: https://www.w3.org/egov/wiki/Linked_Open_Data (visited on 10/01/2023).
- [7] (n.d.) *Microservices*. 2023. URL: <https://www.ibm.com/think/topics/microservices> (visited on 10/25/2023).
- [8] (n.d.) ANPR – *Anagrafe Nazionale della Popolazione Residente, Statistiche sui certificati digitali, Ministero dell’Interno, Roma*. 2024. URL: <https://www.anagrafenazionale.interno.it/anpr/numeri/> (visited on 10/28/2025).
- [9] (n.d.) *WebVOWL (Web-based Visualization of Ontologies)*. 2023. URL: <https://service.tib.eu/webvowl/> (visited on 10/20/2023).
- [10] (n.d.) *Using OnToology: Tutorial*. 2023. URL: <https://ontoology.linkeddata.es/tutorial> (visited on 10/20/2023).
- [11] (n.d.) *Tutorial — NetworkX 3.2 documentation*. 2023. URL: <https://networkx.org/documentation/stable/tutorial.html> (visited on 10/15/2023).
- [12] (n.d.) *Documentation — pyvis 0.1.3.1 documentation*. 2023. URL: <https://pyvis.readthedocs.io/en/latest/documentation.html> (visited on 10/20/2023).
- [13] (n.d.) *LinkedData*. 2023. URL: <https://www.w3.org/wiki/LinkedData> (visited on 10/20/2023).
- [14] (n.d.) *RDF 1.2 Turtle*. 2023. URL: <https://www.w3.org/TR/rdf12-turtle/> (visited on 10/20/2023).
- [15] Giovanni Carau. *Certificate Knowledge Graph Ontology (CKGO) Documentation*. <https://gcarau.github.io/RichiestaCertificati/OnToology/CertificateRequest.ttl/documentation/doc/index-en.html>. Custom web ontology developed for the Certificate Knowledge Graph project. 2023. (Visited on 10/30/2023).

- [16] Giovanni Carau. *Certificate Knowledge Graph Ontology (CKGO) TTL*. <https://gcarau.github.io/RichiestaCertificati/OnToolology/CertificateRequest.ttl/documentation/doc/ontology.ttl>. Custom web ontology developed for the Certificate Knowledge Graph project. 2023. (Visited on 10/30/2023).
- [17] (n.d.) *Universal Resource identifiers in WWW*. 2023. URL: <https://www.w3.org/Addressing/URL/uri-spec.html> (visited on 10/20/2023).
- [18] (n.d.) *Model-view-controller*. 2023. URL: <https://en.wikipedia.org/wiki/Model%E2%80%93view%E2%80%93controller> (visited on 10/20/2023).
- [19] (n.d.) *SPARQL Query Language for RDF - Formal Definitions*. 2023. URL: <https://www.w3.org/2001/sw/DataAccess/rq23/sparql-defns.html> (visited on 10/25/2023).
- [20] T. Berners-Lee, J. Hendler, and O. Lassila. “The Semantic Web”. In: *Scientific American* 284.5 (2001), pp. 34–43.
- [21] L. Page, S. Brin, R. Motwani, and T. Winograd. *The PageRank citation ranking: Bringing order to the web*. Technical Report. Stanford InfoLab, 1999.
- [22] P. Bonacich. “Power and centrality: A family of measures”. In: *American Journal of Sociology* 92.5 (1987), pp. 1170–1182.
- [23] M. E. J. Newman. “A measure of betweenness centrality based on random walks”. In: *Social Networks* 27.1 (2005), pp. 39–54.
- [24] European Commission. *New European Interoperability Framework: Promoting seamless services and data flows for European public administrations*. ISBN 978-92-79-63756-8. Luxembourg: Publications Office of the European Union, 2017. DOI: 10.2799/78681. URL: https://ec.europa.eu/isa2/sites/default/files/eif_brochure_final.pdf.