Data Integration as an Innovation Strategy in Public Management: Challenges and Advances in the Context of TED INCRA/UFPR

Estephanie Daiane Batista da SILVA, Lilian de Fatima BENCZ, Silvana Philippi CAMBOIM, Luis Henrique da COSTA, Vitor Silva de ARAUJO, Alex Sebastiao CONSTANCIO, Davi dos Santos Villela JUNIOR, Valmir Antunes PEREIRA, Edson Flávio de SOUZA and Eduardo VEDOR de Paula, Brazil

Key words: Data Management, Public Sector, Data Integration, Land Policy, Innovation

SUMMARY

The TED INCRA/UFPR programme addresses long-standing fragmentation in Brazil's land administration by transforming heterogeneous and siloed datasets into an integrated, standards-based environment. Through participatory and user-centred methods — including design thinking, requirement validation and iterative prototyping — the initiative engaged INCRA professionals and researchers to co-create solutions aligned with operational needs. The resulting infrastructure consolidates tens of thousands of legal, geodetic, environmental and socio-economic records in an institutional repository and a PostgreSQL/PostGIS spatial database, enriched by a standardised data dictionary and conceptual. National (INDE, ET-EDGV) and international (ISO, OGC) standards guided semantic harmonisation and interoperability, while a public monitoring dashboard promotes transparency for managers, oversight bodies and society. The experience reveals not only technical achievements but also cultural and institutional shifts, advancing a data-centric mindset, encouraging collaboration among multiple TED projects, and setting the foundations for the future integration of legacy INCRA systems into coherent spatial data infrastructures.

Data Integration as an Innovation Strategy in Public Management: Challenges and Advances in the Context of TED INCRA/UFPR

Estephanie Daiane Batista da SILVA, Lilian de Fatima BENCZ, Silvana Philippi CAMBOIM, Luis Henrique da COSTA, Vitor Silva de ARAUJO, Alex Sebastiao CONSTANCIO, Davi dos Santos Villela JUNIOR, Valmir Antunes PEREIRA, Edson Flávio de SOUZA, Eduardo VEDOR de Paula, Brazil

1. INTRODUCTION

Land administration in Brazil is significantly impacted by the historical fragmentation of data and the absence of shared standards among public institutions. Within the National Institute for Colonisation and Agrarian Reform (INCRA) — the federal agency responsible for agrarian reform and national land tenure — multiple legacy systems coexist, such as the SIPRA (Sistema de Informações de Projetos de Reforma Agrária – Agrarian Reform Projects Information System), PGT (Plano de Gestão Territorial – Territorial Management Plan), SIGEF (Sistema de Gestão Fundiária – Land Management System) and SNCR (Sistema Nacional de Cadastro Rural – National Rural Cadastre System). These platforms rarely interoperate, resulting in redundant workflows, inconsistent information, and limited transparency. Such fragmentation has long prevented the consolidation of a Cadastro Territorial Multifinalitário (CTM), or Multifunctional Cadastre, capable of integrating fiscal, legal and physical data to support land governance (Paixão et al., 2012; Amorim et al., 2018).
In recent years, the decentralised funding mechanism known as the Termo de Execução Descentralizada (TED – Decentralised Execution Term) has expanded significantly (INCRA, 2025) . TEDs allow federal resources to be transferred to universities and federal institutes to support applied research, capacity building and technological development. This proliferation has fostered valuable partnerships in areas such as education (e.g., PRONERA – National Program for Education in Agrarian Reform), mapping of quilombola territories, geospatial data governance, rural tourism, land market monitoring and conflict management. However, without a common conceptual model or data standards, the multiplication of TEDs risks producing isolated, non-harmonised solutions, reinforcing rather than solving fragmentation.
The TED INCRA/UFPR programme, a long-term partnership between INCRA and the Universidade Federal do Paraná (UFPR) through its Laboratório de Geoprocessamento e Estudos Ambientais (LAGEAMB – Geoprocessing and Environmental Studies Laboratory), directly addresses this challenge. The programme is structured into eight complementary projects, each generating specific and valuable datasets for land administration:

1. Supervisão Ocupacional (Occupational Supervision) – socio-economic and field monitoring data on families and settlement status.
2. Geodésia (Geodesy) – high-precision geodetic data, parcel boundaries and coordinates for settlement demarcation.
3. Ambiental (Environmental) – environmental assessments, including land use and vegetation cover data.

4. Busca Cartorial (Registry Search) – registry and legal property searches to support tenure regularisation.
5. Gestão e Publicação de Dados (Data Management and Publication) – integration and dissemination of all produced data, including infrastructure for databases, repositories and dashboards.
6. Integração de Resultados e Extensão (Results Integration and Outreach) – scientific outputs, knowledge transfer and outreach activities.
7. Gestão do Programa (Programme Management) – strategic and administrative coordination.

Each of the first five projects generates heterogeneous datasets — from scanned legal titles and field reports to georeferenced parcels and environmental layers — which historically remained isolated "data islands" (Linheira & Oliveira, 2018; Nubiato & Delazari, 2021). From a conceptual perspective, the Land Administration Domain Model (LADM, ISO 19152) provides a critical framework for overcoming fragmentation and enabling interoperability. Internationally, LADM has underpinned integrated land information systems and SDIs, helping align cadastral, legal and environmental data (Lemmen et al., 2015; Van Oosterom et al., 2020; Kalogianni et al., 2020). In Brazil, researchers have suggested adopting LADM to harmonise the cadastre and land registry (Paiva et al., 2018; Souza et al., 2022) and to support transparent and interoperable land governance. The TED INCRA/UFPR initiative envisions these principles, observing LADM concepts and applying UML and OMT-G modelling, a standardised data dictionary, and open-source geospatial technologies.

By combining international standards, collaborative design and institutional articulation, TED INCRA/UFPR demonstrates how the proliferation of TEDs can be channelled into coherent, interoperable and replicable solutions. The initiative strengthens information governance, advances Brazil's digital maturity in land administration and creates a foundation for other public programmes to modernise and integrate their geospatial data.

2. METHODOLOGY

The methodological approach of the TED INCRA/UFPR programme was designed to transform a previously fragmented information ecosystem into an integrated, standards-based environment. Figure 1 summarises the overall workflow, illustrating how heterogeneous datasets produced by different TED subprojects converge into a unified infrastructure comprising an institutional repository, a central spatial database, and an indicator and monitoring environment.

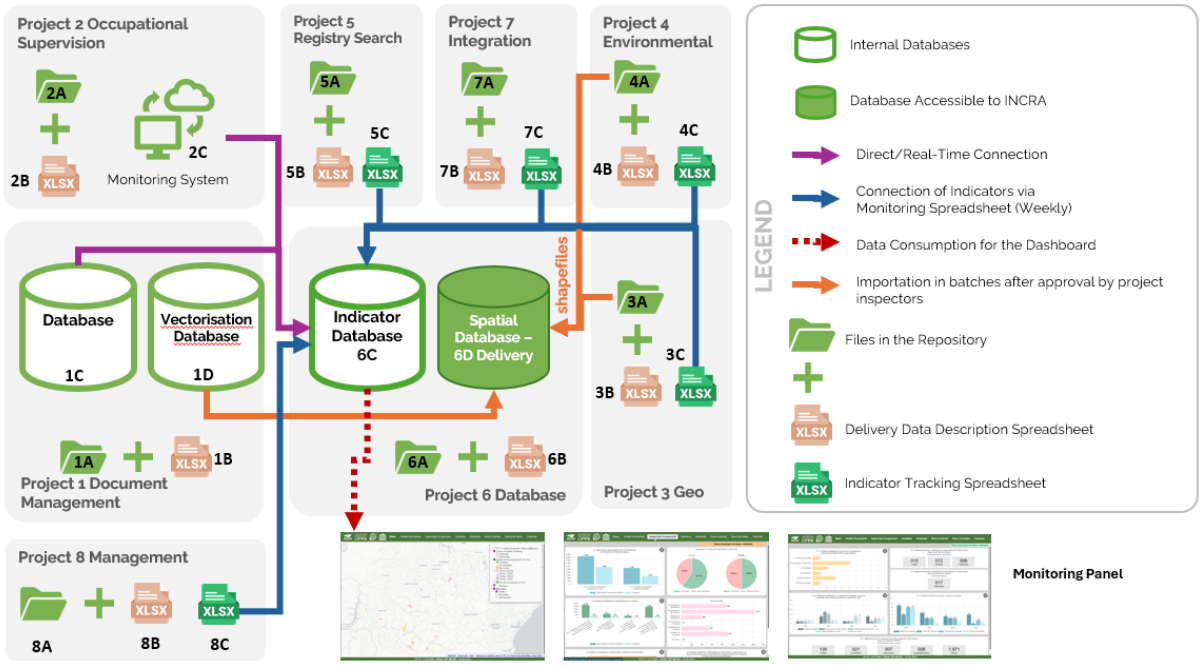

Figure 1. Integrated workflow for data management and monitoring in TED INCRA/UFPR.

2.1 **Institutional and territorial context**

The State of Paraná, in southern Brazil, is a highly diverse agricultural and territorial environment, with over 360 agrarian reform settlements (INCRA, 2025) and an extensive land regularisation agenda. The Superintendência Regional do INCRA no Paraná (INCRA/SR-09) is responsible for implementing agrarian reform, monitoring settlements, registering and titling parcels, supervising occupational conditions and managing environmental and socio-economic information. The variety and volume of records — including legal titles, geodetic measurements, environmental assessments and socio-economic surveys — have historically been handled through separate systems.

The TED INCRA/UFPR partnership was established to bring innovation to respond to these demands, with a multi-year plan including more than 60 deliverables and large volumes of information: tens of thousands of scanned legal documents and maps, more than 3,000 georeferenced parcels, field surveys, legal environmental regularization on 48 settlements and registry checks, totaling large amounts of data to be harmonised.

2.2 Workflow for data integration

Data integration followed a participatory and iterative approach, but rather than isolated tasks, it was structured as a continuous flow connecting all subprojects. Each TED project generated its own outputs, including scanned legal files and spreadsheets from document management systems, socio-economic and occupational monitoring tables, geodetic and cartographic shapefiles, environmental land-use and vegetation layers, and legal registry searches.

Deliverables were first described in standardised data description sheets that captured file origin, attribute structure and geometry types. These sheets unified semantics and supported quality control.
After documentation, the data was incorporated into the Institutional Repository, a structured file system that uses controlled folder hierarchies and documented access policies to secure and preserve records while enabling selective sharing. Spatial and alphanumeric components were ingested into the PostgreSQL/PostGIS database using automated ETL (Extract, Transform, Load) processes. During this step, attribute names and categories were aligned with the ET-EDGV (Brazilian cartographic data standard), and identifiers were created to consistently link parcels, families, and administrative acts. A dedicated Indicators Database was derived from the main spatial database to feed the Monitoring Dashboard, which was built using Python/Django and Vue.js. This dashboard aggregates quantitative and spatial indicators that support not only internal management but also the activities of all stakeholders involved in the programme — including more than 130 professionals and researchers at UFPR — and it plays a key role in providing transparency to society and oversight bodies. The platform is publicly accessible (https://paineltedincra.lageamb.ufpr.br/dashboard/banco_de_dados).
Quality control and validation were central to this workflow. Before each import, data packages were checked by INCRA technical staff for completeness and consistency. Only after this homologation were shapefiles and tables imported in bulk to the unified database, preserving an auditable trail. Weekly updates of indicator spreadsheets kept the monitoring panel synchronised with the evolving repository.
To support the technical processes, the LAGEAMB laboratory maintains an institutional data center environment at UFPR. This infrastructure includes dedicated servers configured to host the TED INCRA/UFPR spatial database and repository securely, ensuring centralised control, robust backup routines, and reduced operational complexity. By concentrating processing and storage in a controlled environment, the team guarantees data consistency, performance, and long-term preservation while facilitating controlled access for INCRA staff and authorised partners.

2.3 Methodological principles

Three interrelated principles guided the entire process. The first was semantic harmonisation and interoperability, achieved by adopting database modelling techniques and the Object Modelling Technique for Geographic applications (OMT-G) — an extension of the Unified Modelling Language (UML) that incorporates geographic primitives to represent spatial entities, relationships and constraints. Figure 2 presents a simplified class diagram illustrating the relationships between key entities, such as legal registrations.

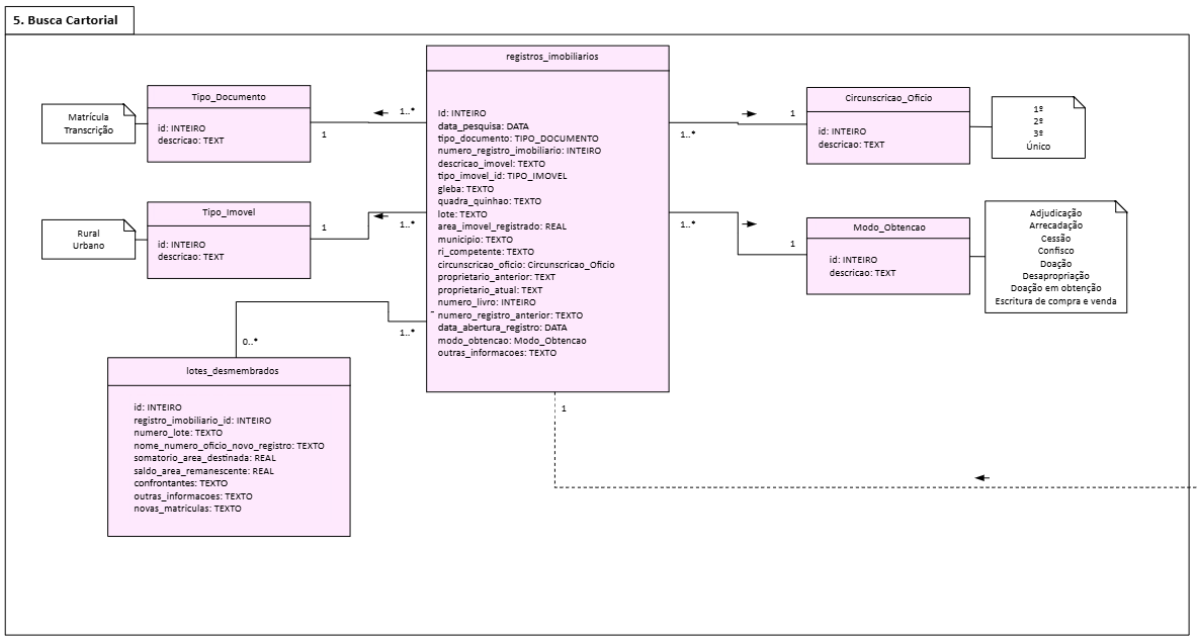

Figure 3. Part of the UML/OMT-G conceptual model

In addition, the project adhered to Brazilian standards for geospatial data and metadata, as defined by the Infraestrutura Nacional de Dados Espaciais (INDE – Brazilian National Spatial Data Infrastructure). INDE establishes guidelines for data structure, metadata documentation and service publication to promote discoverability and reuse. At the same time, the modelling and implementation process observed international standards from the International Organization for Standardization (ISO) — particularly the ISO 19100 series for geographic information — and best practices from the Open Geospatial Consortium (OGC), which define interoperable web services and geospatial data formats. This dual adherence ensures that the environment remains compatible with both national frameworks and widely adopted global interoperability protocols.

The second principle was the use of free and open-source technologies to ensure sustainability, independence from proprietary solutions and replicability by other institutions. The infrastructure is built on PostgreSQL/PostGIS for spatial data management, with Python/Django powering the backend, Vue.js powering the frontend, and QGIS serving as the primary platform for technical staff.

Finally, user-centred design ensured that technical decisions responded to real operational needs. Design thinking sessions with INCRA professionals and UFPR teams supported the refinement of requirements, interface prototyping and usability validation (Pisetta et al., 2024). These sessions fostered a shared understanding of workflows and data flows, thereby reducing resistance to change.

Another key methodological element was the establishment of data governance tools to maintain integrity across the integrated system. Unique identifiers and geocoding schemes were designed to relate multiple representations of the same territorial object — for example, connecting a settlement parcel to its legal registration record and socio-economic profile.

Controlled vocabularies and consistent naming conventions were applied across all tables and repositories.
Additionally, a comprehensive data dictionary was created, serving as a "translation layer" between different technical and administrative domains. This dictionary has been crucial to avoiding semantic ambiguity and enabling interoperability as new datasets are incorporated. Figure 3 illustrates an excerpt of the data dictionary, which documents attribute names, definitions and geometry types to ensure semantic consistency across heterogeneous datasets.

Classe	Descrição			Código	Geometria
instrucao_processual	Planilha de controle interno da equipe de supervisão ocupacional			2.2.1	Não espacial
Atributo	**Tipo**	**Tamanho**	**Descrição**	**Domínio**	**Requisito**
cod_so	Alfanumérico		Indica um código único para identificação no banco de dados, referente aos dados obtidos pela supervisão ocupacional	A ser preenchido	
cod_mun	Inteiro		Código do município	A ser preenchido	
nm_mun	Alfanumérico		Indica o nome completo da instância.	A ser preenchido	
cod_sipra_pa	Alfanumérico		Indica um código único para identificação no banco de dados, baseado no código SIPRA	A ser preenchido	
nome_pa	Alfanumérico		Nome da PA	A ser preenchido	
cod_beneficiario	Inteiro				
titular_ocupante_1	Alfanumérico		Nome completo do Titular 1	A ser preenchido	
cpf_t1	Alfanumérico		Indica o número do CPF do Titular 1	A ser preenchido	
titular_ocupante_2	Alfanumérico		Nome completo do Titular 2	A ser preenchido	
cpf_t2	Alfanumérico		Indica o número do CPF do Titular 2	A ser preenchido	

Figure 3. Data dictionary example

2.4 Data flow analysis

The data flow, illustrated in Figure 1, shows how raw inputs evolve into strategic indicators. Data from each thematic project is first delivered to the repository in its native format, accompanied by metadata and descriptive documentation. The repository serves as a secure, versioned memory of all official deliveries. Once validated, these packages feed the spatial database, where ETL processes harmonise schema, geometry types and attribute semantics. As information accumulates, curated indicators are extracted and stored in the indicators database that powers the monitoring dashboard.
Different types of connections ensure flexibility and reliability. Direct and real-time links enable some continuous updates; weekly synchronisation keeps indicators fresh without compromising validation; and bulk imports guarantee that only checked and approved datasets populate the spatial core. This architecture creates a living, traceable information ecosystem that can scale to new TED projects and future extensions such as the planned Story Map and Spatial Data Infrastructure (SDI).

3. RESULTS

The TED INCRA/UFPR initiative delivered a robust and replicable digital environment for integrated land data governance. From a technical perspective, the unified infrastructure now consolidates approximately 70,000 documents (≈311 GB) in the institutional repository, harmonised using controlled metadata and access policies. These documents include legal instruments, field forms, environmental assessments, and geodetic measurements, which were previously scattered across disconnected systems.

The PostgreSQL/PostGIS database integrates the main thematic layers produced by the different TED projects — parcel boundaries from Geodésia (Geodesy), legal status and land tenure records from Busca Cartorial (Registry Search), socio-economic and field supervision data from Supervisão Ocupacional (Occupational Supervision), and environmental information from Ambiental (Environmental). This consolidation enables cross-domain queries and supports more consistent decision-making.

Building on these data assets, an Indicators Database powers the Monitoring Dashboard (https://paineltedincra.lageamb.ufpr.br/dashboard/banco_de_dados
). The dashboard features interactive pages dedicated to each TED subproject, as well as a central map view that presents the spatial distribution and status of settlements. Indicators track progress in document digitisation, geodetic surveys, environmental analysis and legal registration. The platform supports multiple stakeholder groups, including technical managers at INCRA, more than 130 researchers and staff at UFPR, and external oversight bodies and the broader society, thereby reinforcing accountability and transparency. Figure 4 shows the main interface of the Monitoring Dashboard, where indicators and interactive maps present the progress of data integration and settlement monitoring.

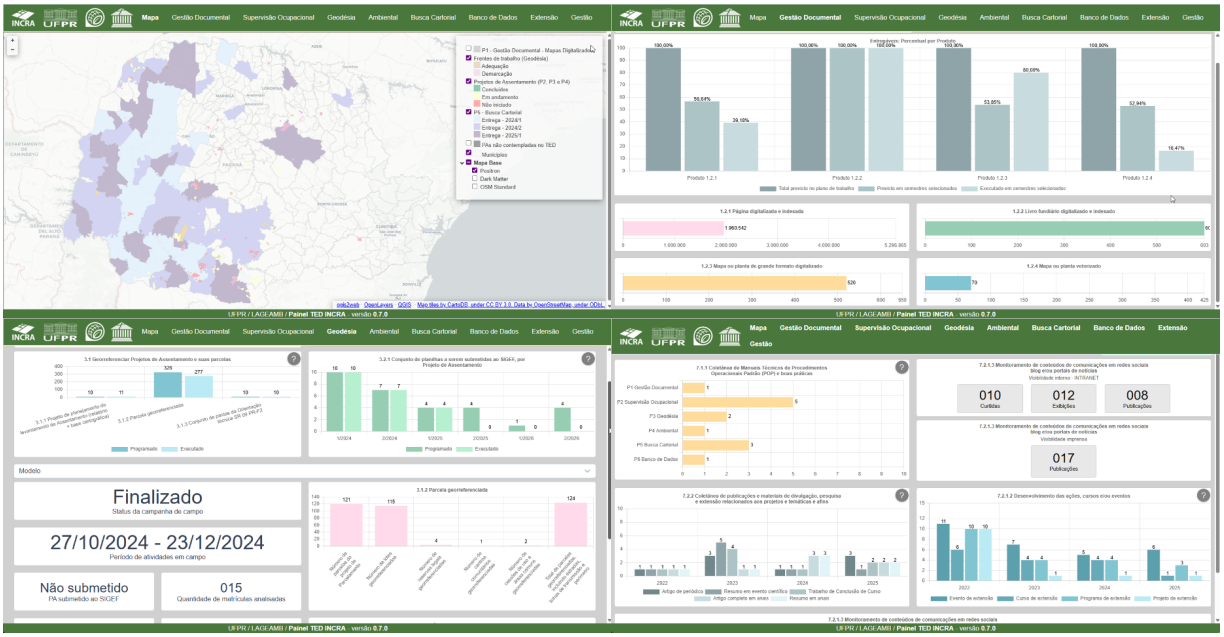

Figure 4. Monitoring Dashboard overview

Beyond data consolidation, the project delivered semantic artefacts — including a comprehensive data dictionary aligned to ET-EDGV and LADM — and a conceptual model (UML/OMT-G) that structures classes and relationships. These artefacts enable continuous system improvement and facilitate the incorporation of new data.

4. DISCUSSION

Developing an open and reproducible infrastructure within TED INCRA/UFPR required more than technical expertise; it demanded sustained collaboration across diverse teams and a clear focus on the people who would use and maintain the system. The user-centred design approach proved essential in this multi-project programme. Workshops and iterative prototypes helped translate operational pain points into solutions, allowing INCRA technicians and university researchers to influence interface design, data models and workflows. This co-creation culture fosters trust and encourages adoption, thereby counteracting the risk that new platforms remain underutilised because they do not align with everyday practices.

Despite significant advances, full integration of INCRA's legacy systems remains a future challenge. The coexistence of platforms such as SIPRA, SIGEF, SNCR and PGT — each with different data structures and institutional routines — complicates the creation of a single, coherent land information infrastructure. The work carried out here shows that applying shared conceptual models (LADM, ET-EDGV), open standards (ISO, OGC), and user-centred methods can guide this integration. Extending these principles to other INCRA systems could progressively reduce redundancy and improve data quality.

Another key reflection is the expansion of TEDs with Brazilian universities and federal institutes. While this decentralised model has fostered innovation and brought technical expertise closer to the field, projects are often developed in parallel, without systematic exchange of solutions or semantic alignment. As a result, INCRA faces the long-term challenge of managing and integrating diverse technological legacies. The TED INCRA/UFPR experience suggests that stronger inter-TED dialogue and reuse of standards, models and tools would reduce fragmentation and accelerate national digital transformation in land administration.

Finally, the initiative reflects a broader cultural shift towards a data-centric mindset in public sector technology. Rather than treating data as a by-product of administrative work, the programme treats it as a strategic asset: carefully modelled, semantically harmonised and designed for interoperability. Embracing this culture is essential for organisations like INCRA to evolve from isolated applications to integrated digital ecosystems, enabling transparency, scalability and innovation.

5. CONCLUSION

The TED INCRA/UFPR programme marks a significant advance in integrated land data governance in Brazil. It demonstrates how a participatory and standards-based approach can

overcome historical fragmentation, enabling a sustainable and reproducible technical environment within a complex institutional setting. While challenges remain — particularly in the full integration of parcel, registry and beneficiary information — the foundation laid by this initiative offers a clear path for scaling, including planned future products such as a Story Map and a broader Spatial Data Infrastructure (SDI). The model is replicable for other public digital transformation initiatives seeking transparency, efficiency and coherent territorial governance.

REFERENCES

AMORIM, Amilton; PELEGRINA, Marcos Aurélio; JULIÃO, Rui Pedro. Cadastro e gestão territorial: uma visão luso-brasileira para a implementação de sistemas de informação cadastral nos municípios. São Paulo: Editora UNESP, 2018. 131 p. E-book. DOI: https://doi.org/10.7476/9788595462823. 24 de setembro de 2025.

COSTA, T. F. et al. Proposta de Perfil de Metadados para o SINTER com base na ISO 19115 e no LADM (ISO 19152). Revista Brasileira de Cartografia, [S. l.], 2020. Disponível em: DOI: 10.14393/rbcv74n4-64008. Acesso em: 25 de setembro de 2025.

INCRA. Termo de Execução Descentralizada. 2025. Incra. Recuperado 30 de setembro de 2025, de https://www.gov.br/incra/pt-br/centrais-de-conteudos/publicacoes/editais-de-licitacoes/termo-de-execucao-descentralizada. Acesso em: 25 de setembro de 2025.
INCRA. Assentamentos—Relação de Projetos. 2025. Incra. https://www.gov.br/incra/pt-br/assuntos/reforma-agraria/assentamentos-relacao-de-projetos. Acesso em: 25 de setembro de 2025.

KALOGIANNI, Eleni; VAN OOSTEROM, Peter; LEMMEN, Christiaan; QUIRING, Jochen. *The LADM: impact and future developments*. Land Use Policy, v. 98, p. 104-146, 2020. doi:10.1016/j.landusepol.2020.104146.

LEMMEN, C. The Land Administration Domain Model (LADM) as an international standard for land administration (ISO 19152). Land Use Policy, v. 49, p. 556–563, 2015.

LINHEIRA, Guilherme; OLIVEIRA, Francisco Henrique de. O LADM aplicado ao Cadastro Nacional de Sítios Arqueológicos. In: CONGRESSO BRASILEIRO DE CADASTRO TÉCNICO MULTIFINALITÁRIO E GESTÃO TERRITORIAL, 13., 2018, Florianópolis.

NUBIATO, Everton Leandro; DELAZARI, Luciene Stamato. O cenário de editais de licitação pública de sistema de informação territorial para administrações municipais: Diagnóstico e proposta de requisitos utilizando técnicas de engenharia de requisitos. Revista Brasileira de Cartografia, [S. l.], v. 73, n. 2, p. 375-389, 2021. Disponível em: [e-journal]. Acesso em: 25 de setembro de 2025.

PAIVA, Caio dos Anjos; ANTUNES, Alzir Felippe Buffara; CAMBOIM, Silvana. A Proposal for Integrating Data of Land Registry and Urban Cadastre. Boletim de Ciências Geodésicas, [S. l.], v. 24, n. 4, e2018004, 2018. DOI: 10.1590/s1982-21702018000400032. 27 de setembro de 2025.

PAIXÃO, S. K. S.; NICHOLS, S.; CARNEIRO, A. F. T. Cadastro Territorial Multifinalitário: Dados e Problemas de Implementação do Convencional ao 3D e 4D. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 14., 2008, Natal. Anais [...]. Natal: INPE, 2008.

PISETTA, Jaqueline Alves; CAMBOIM, Silvana Philippi; COETZEE, Serena. Design Thinking no projeto cartográfico: experiência com usuários do Brasil e da África do Sul. In: COLÓQUIO BRASILEIRO DE CIÊNCIAS GEODÉSICAS, 13., 2024, Curitiba. Anais [...]. Curitiba: UFPR, 2024.

SOUZA, Deborah Valandro de et al. Cadastro Urbano: uma Proposta para Codificação de Parcela Territorial Urbana com Abrangência Nacional. Revista Brasileira de Cartografia, [S. l.], v. 74, n. 4, p. 839–858, 2022. DOI: 10.14393/rbcv74n4-64008. 27 de setembro de 2025.

BIOGRAPHICAL NOTES

Estephanie Daiane Batista da Silva
Mestranda no Programa de Pós-Graduação em Ciências Geodésicas da Universidade Federal do Paraná (UFPR), Brasil, e afiliada Laboratório de Geoprocessamento e Estudos Ambientais - LAGEAMB, que faz parte do Departamento de Geografia, da UFPR e tem como principal linha de pesquisa o planejamento e gestão ambiental, e ao LabGeoLivre, um grupo de pesquisa da UFPR focado em tecnologias geoespaciais abertas.

Lilian de Fatima Bencz
Mestranda no Programa de Pós-Graduação em Ciências Geodésicas da Universidade Federal do Paraná (UFPR), Brasil, graduada em Engenharia Cartográfica e de Agrimensura e afiliada ao Laboratório de Geoprocessamento e Estudos Ambientais - LAGEAMB, que faz parte do Departamento de Geografia, da UFPR e tem como principal linha de pesquisa o planejamento e gestão ambiental.

CONTACTS

Estephanie Daiane Batista da Silva
Universidade Federal do Paraná
Avenida Coronel Francisco Herácito dos Santos, 100
Curitiba, PR
BRASIL

Phone: + 55 41 99626-0731
E-mail: estephaniedaiane@lageamb.ufpr.br
Website: lageamb.ufpr.br

Lilian de Fatima Bencz
Universidade Federal do Paraná
Avenida Coronel Francisco Herácito dos Santos, 100
Curitiba, PR
BRASIL
Phone: + 55 41 99854-6599
E-mail: lilianbencz@lageamb.ufpr.br
Website: lageamb.ufpr.br