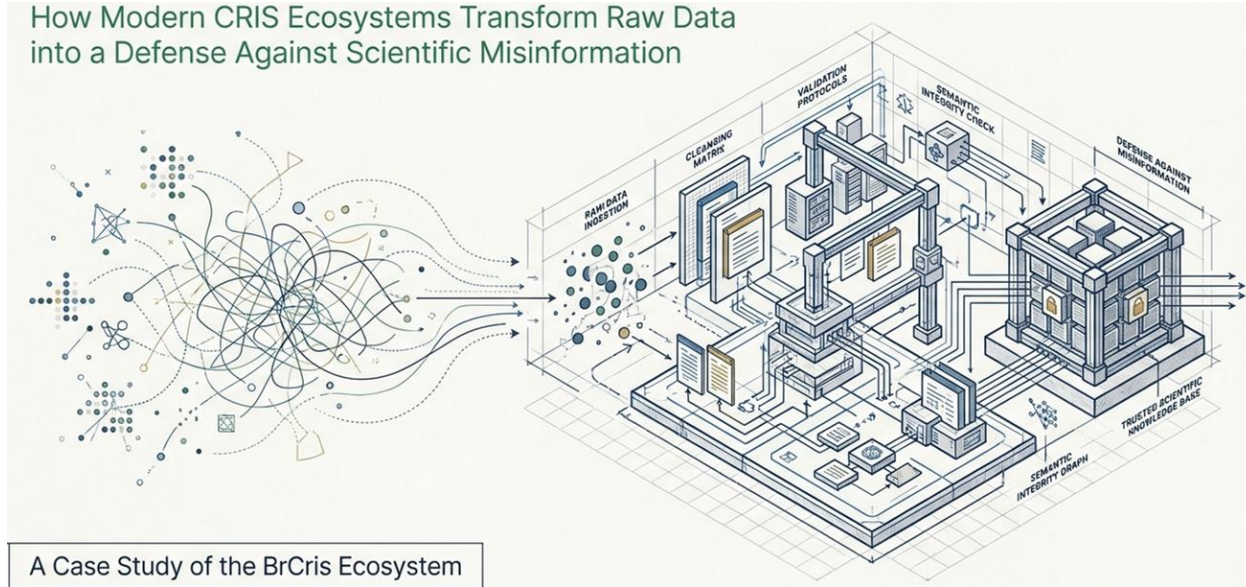


Advancing Research Information Interoperability

How Modern CRIS Ecosystems Transform Raw Data into a Defense Against Scientific Misinformation



A Case Study of the BrCris Ecosystem

CRIS 2026

17th CRIS Conference on Current Research Information Systems

May 20-22, 2026





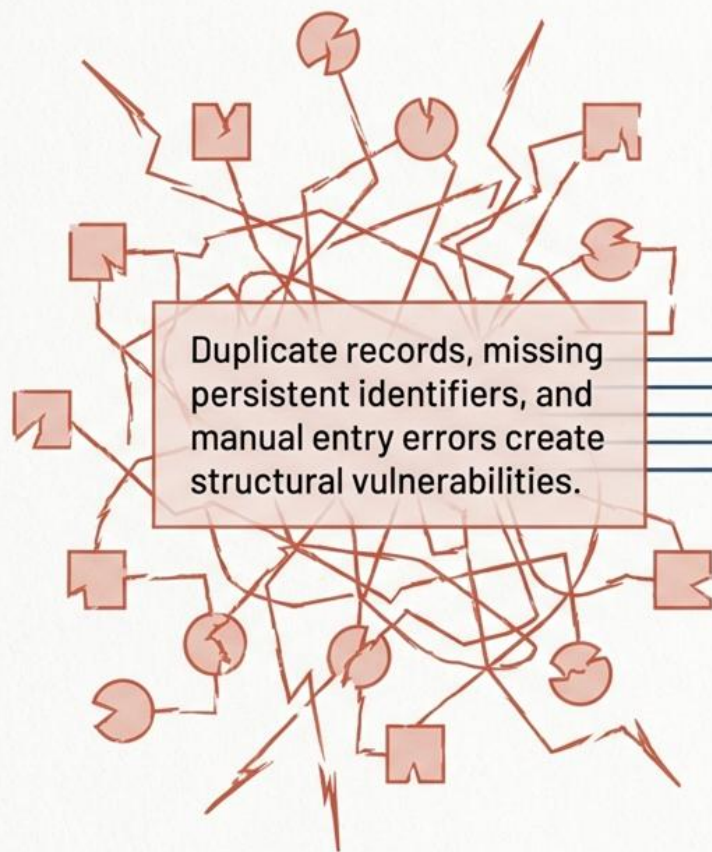
Thiago Dias



Marcel Souza



Washington Segundo



Distorted Indicators

Inconsistencies warp evaluations of national science, technology, and innovation capabilities.

Exploitation

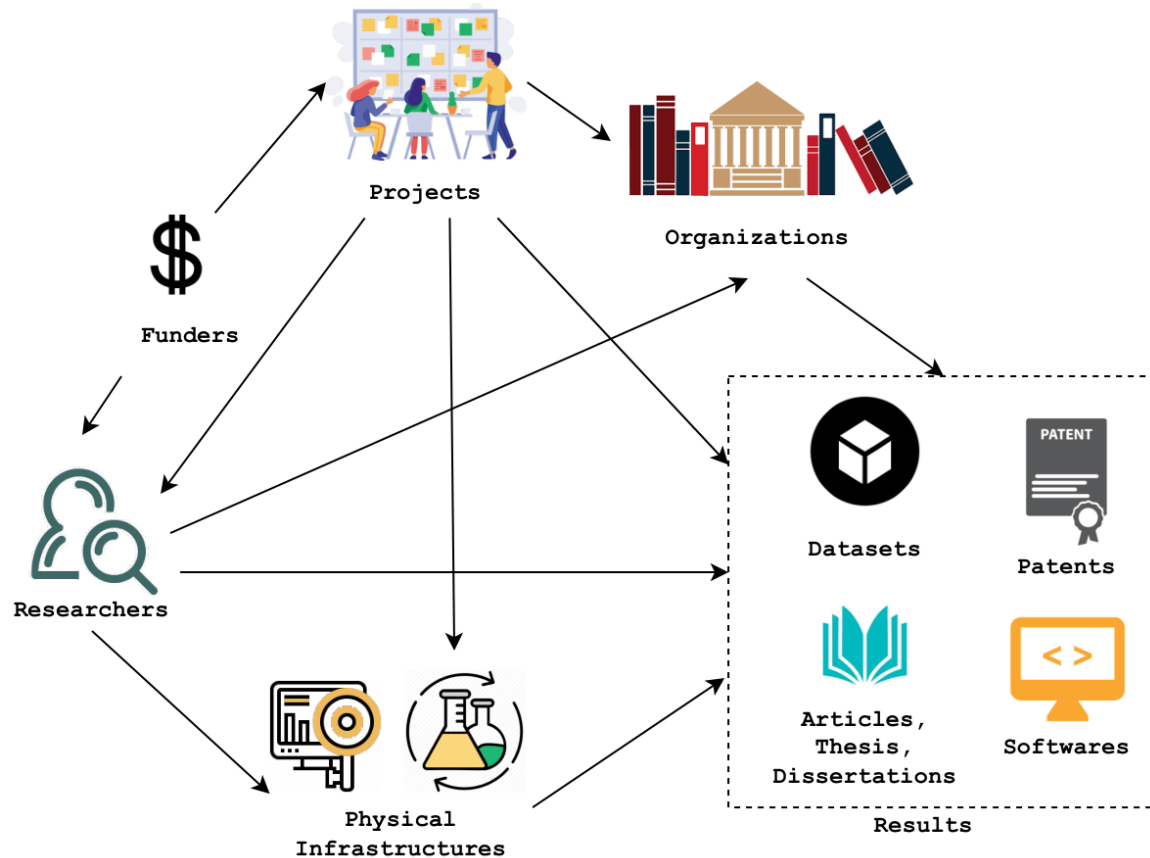
Informational gaps are leveraged—inadvertently or intentionally—for disinformation practices.

Policy Failure

Bad data fundamentally compromises the construction of evidence-based public policies.

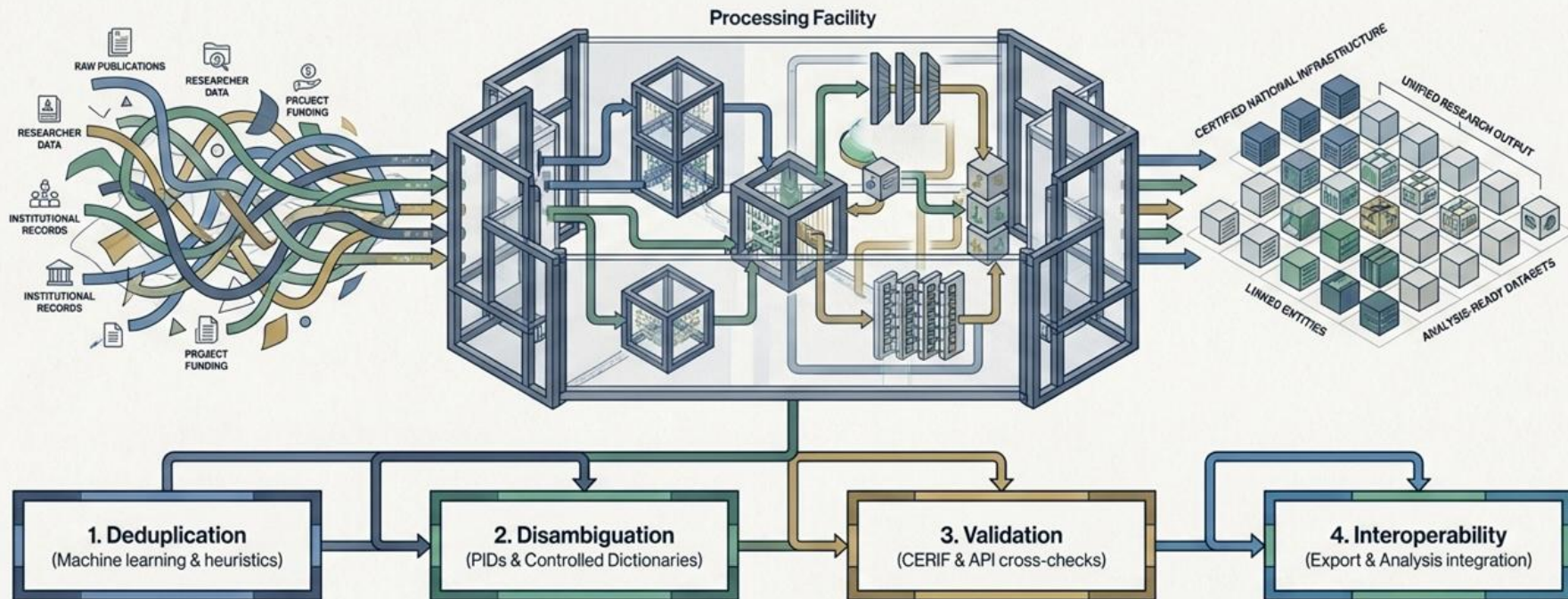
CRIS systems must transcend administrative storage to become critical defensive infrastructures.

BrCris



The BrCris Architecture: Systematizing national scientific production

Brazilian Scientific Research Information Ecosystem (BrCris)



A technological and methodological engine designed to unify publications, researchers, institutions, and projects into a certified national infrastructure.

Introducing BrCris

Brazilian Scientific Research Information Ecosystem — developed to unify and systematize national scientific production.

Scope

Researchers, institutions, publications, projects, and research outputs

Integrations

Lattes Platform, OpenAlex, Oasisbr, and more

Purpose

Bibliometric & scientometric analyses at broad and detailed scope



Key Data Quality



Duplicate Records

Same entity represented multiple times



Missing Identifiers

Lack of persistent DOI, ORCID references



Manual Entry Errors

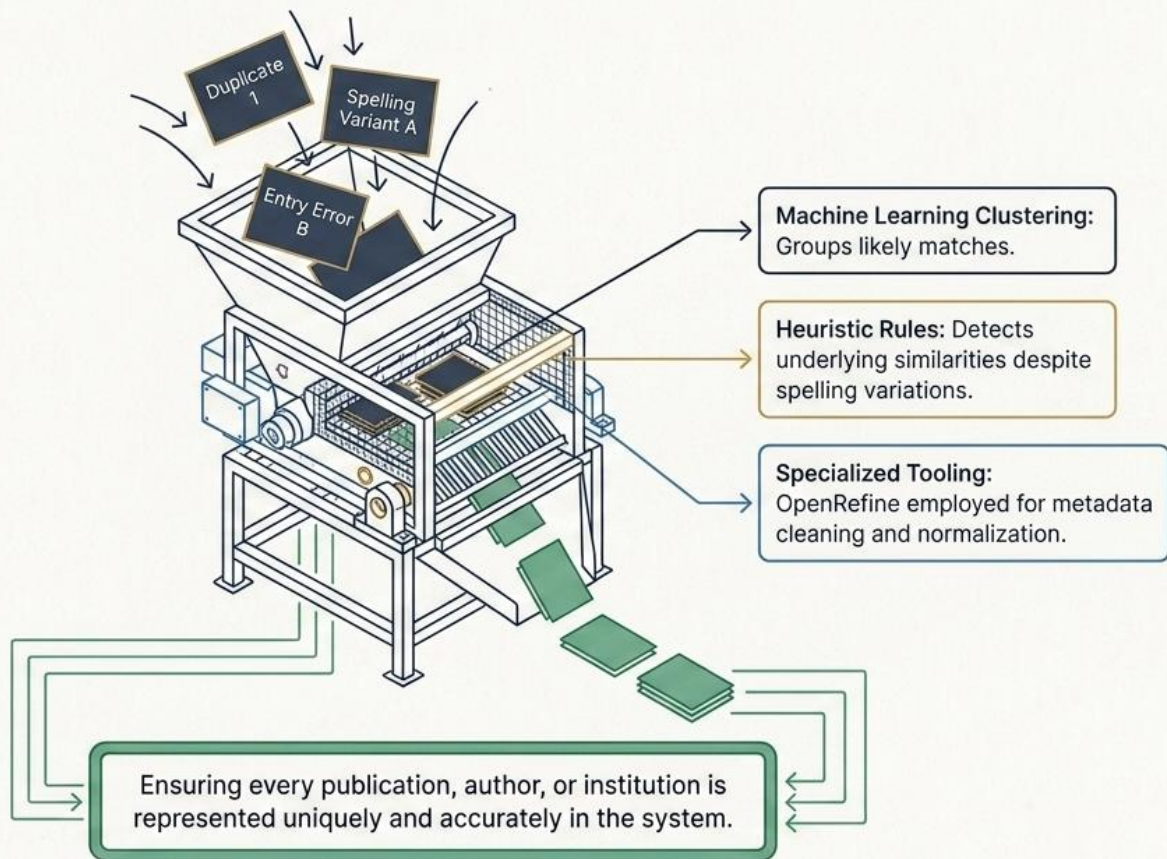
Spelling variations, inconsistent abbreviations



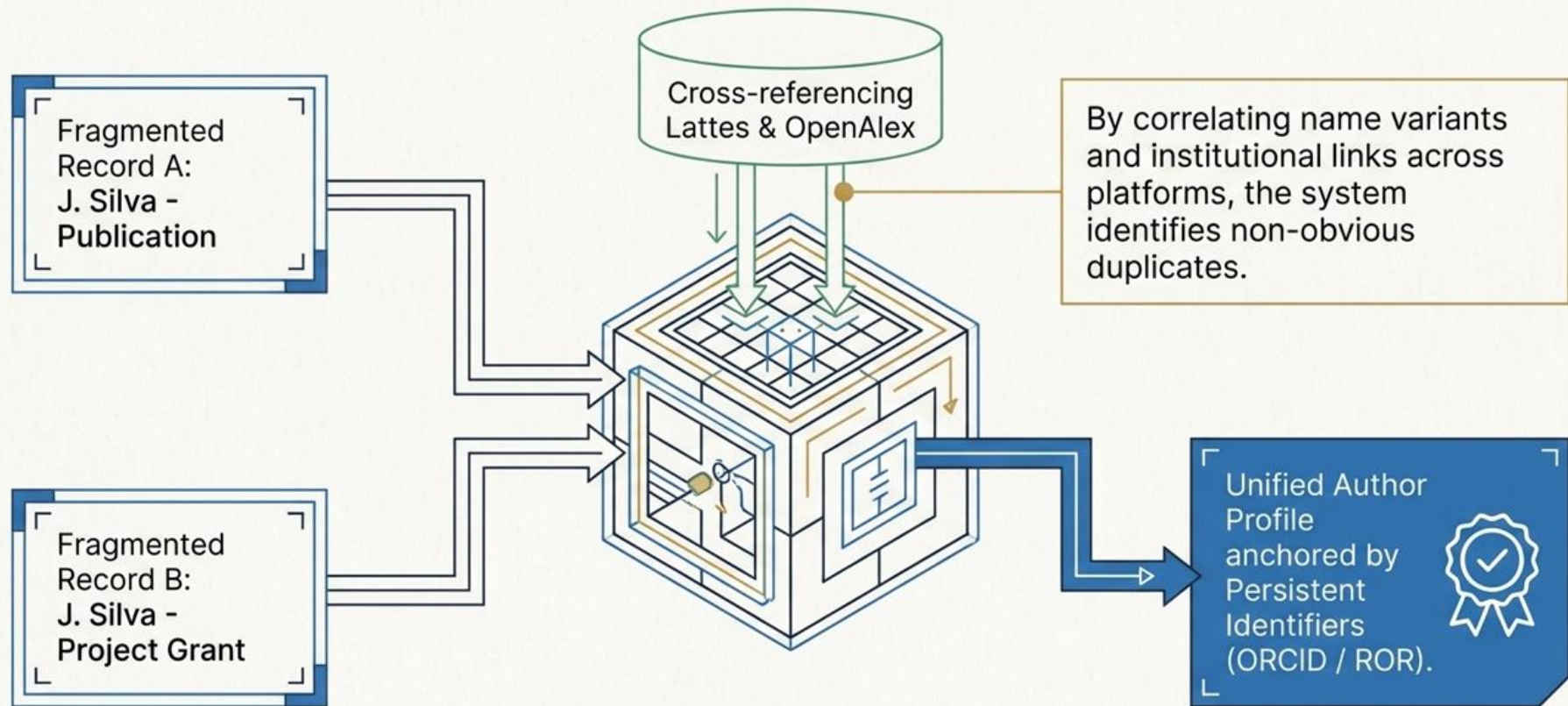
Informational Vulnerabilities

Exploitable gaps in disinformation practices

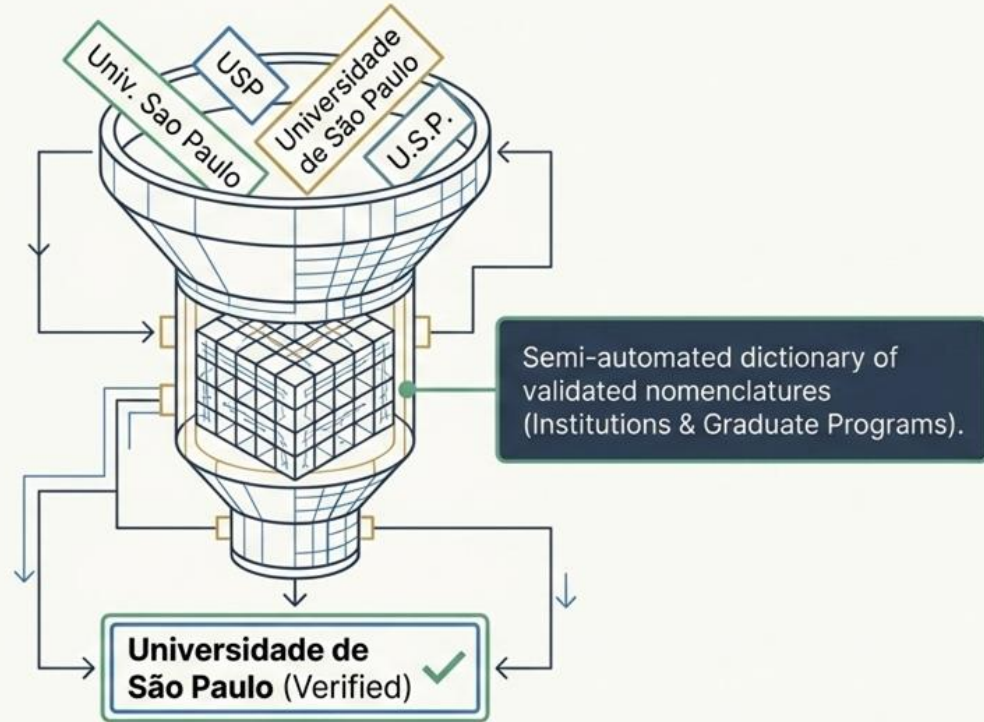
Stage 1: Algorithmic Deduplication



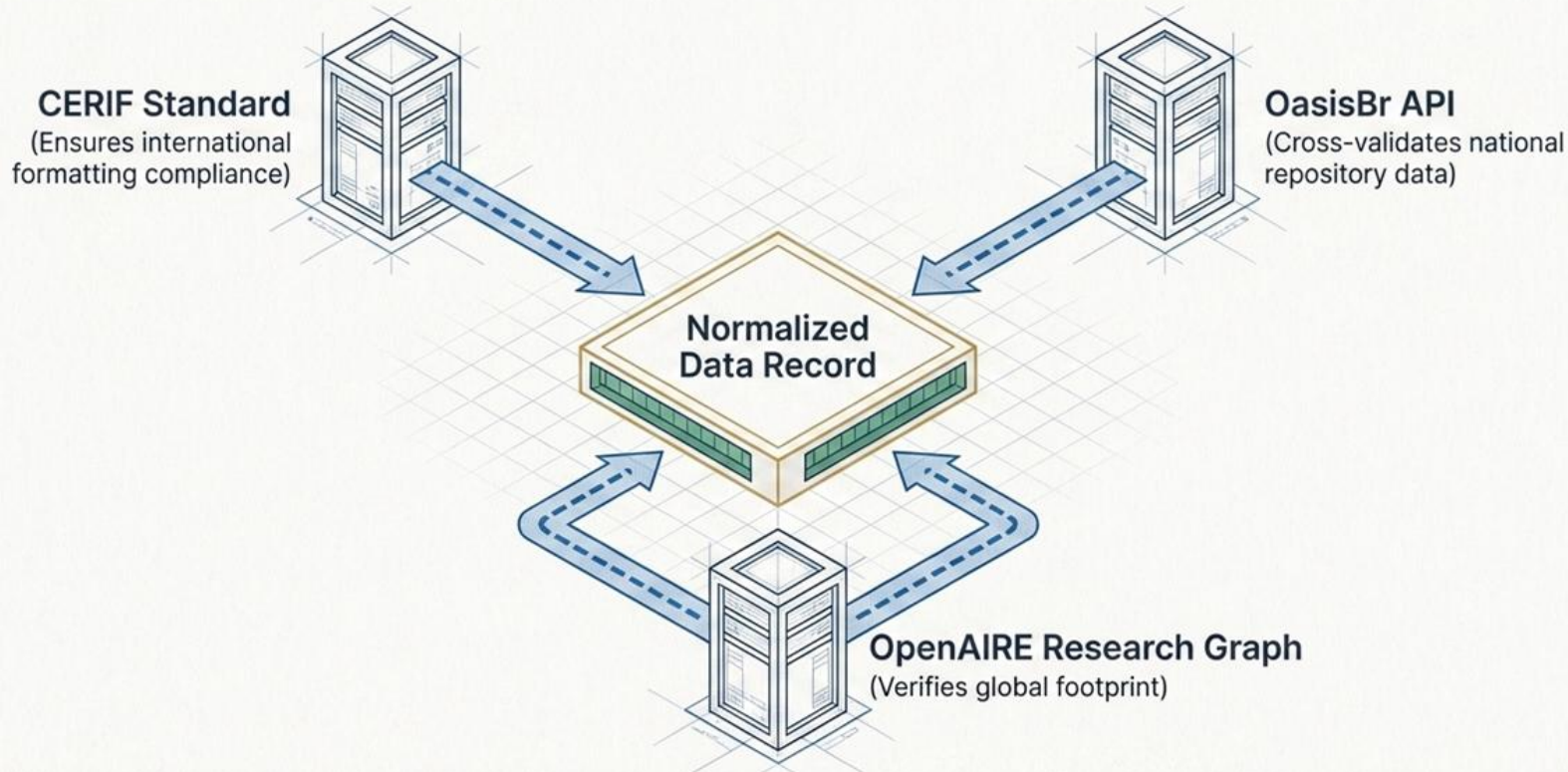
Stage 2: The Disambiguation Engine



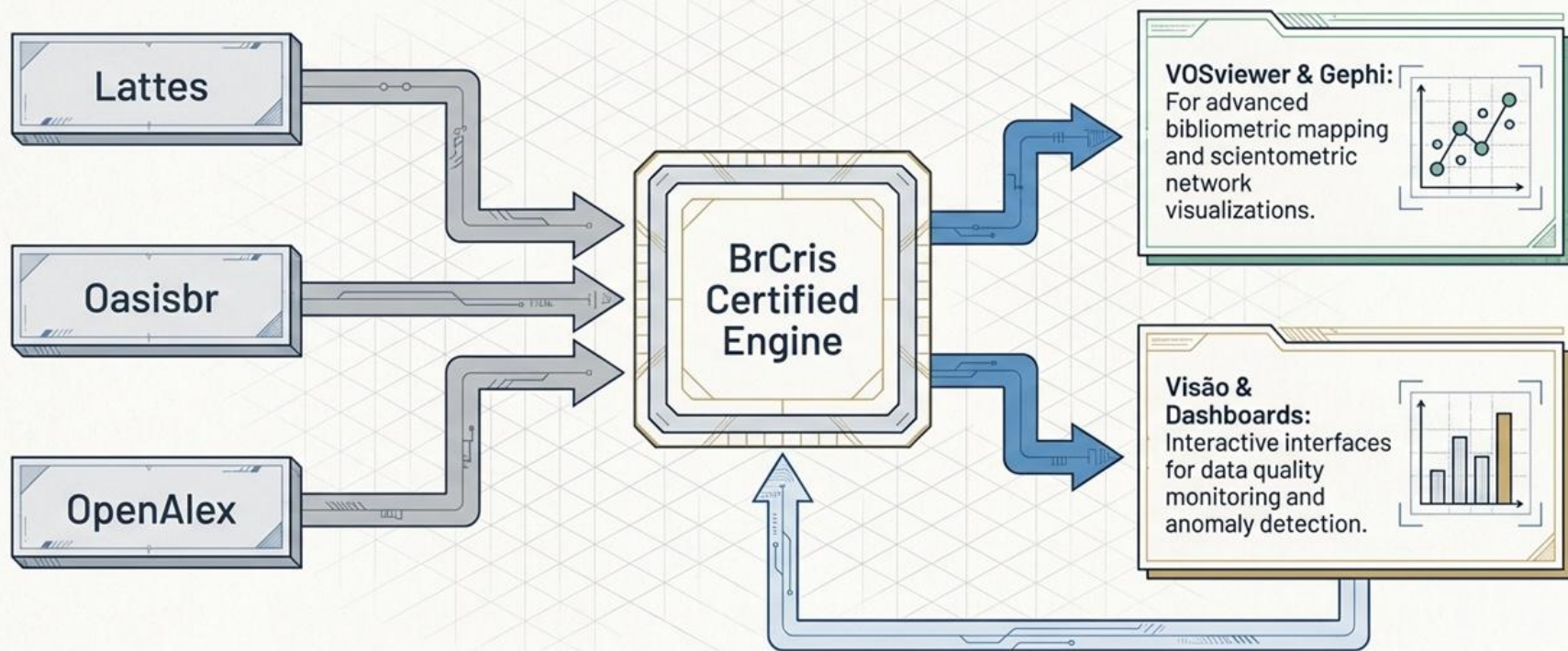
Pre-Entry Standardization: The Controlled Dictionary Filter



Stage 3: Automated Validation & Metadata Certification

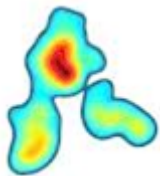


Stage 4: The Interoperability Hub



STAGE 5

Integration with Bibliometric Tools



VOSviewer

Network mapping of research outputs



Gephi

Graph-based anomaly detection



Dashboards

Real-time data quality monitoring

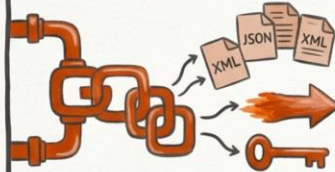
FAIR Principles in BrCris

Findable



**persistent identifiers
and metadata**

Accessible



**open APIs
and export**

Interoperable



**CERIF standards and
cross-platform links**

Reusable



**controlled
vocabularies
and
controlled vocabularies
and certified records**

Achieving FAIR Data

BrCris aligns with FAIR principles through systematic adoption of:

- Persistent identifiers (DOI, ORCID, ROR)
- International metadata standards
- Open infrastructure integrations

Results: What BrCris Achieves



Reduced Ambiguity

Cleaner author and institutional records enable coherent aggregation



Enhanced Interoperability

Alignment with international standards improves global visibility



Better Policy Indicators

More reliable longitudinal and comparative analyses

Structural Limitations Identified

Data Completeness

Variations across
sources and
institutions

Update Delays

Dynamic system
requires continuous
effort

Uneven Adoption

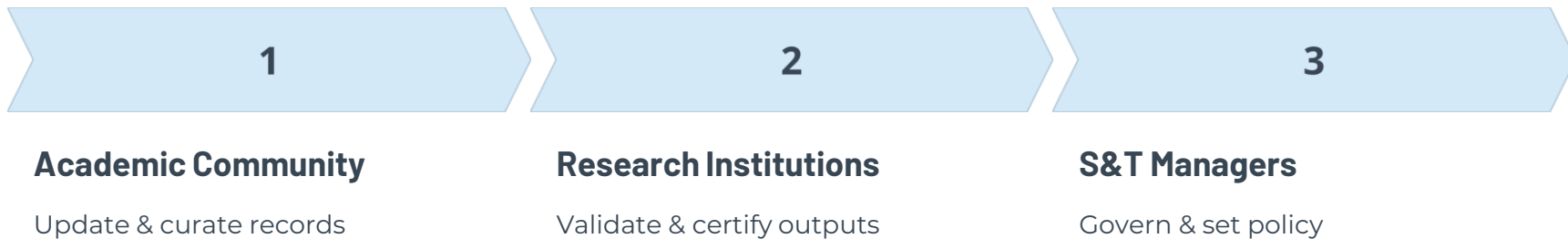
Persistent identifiers not uniformly used

A Valuable Outcome

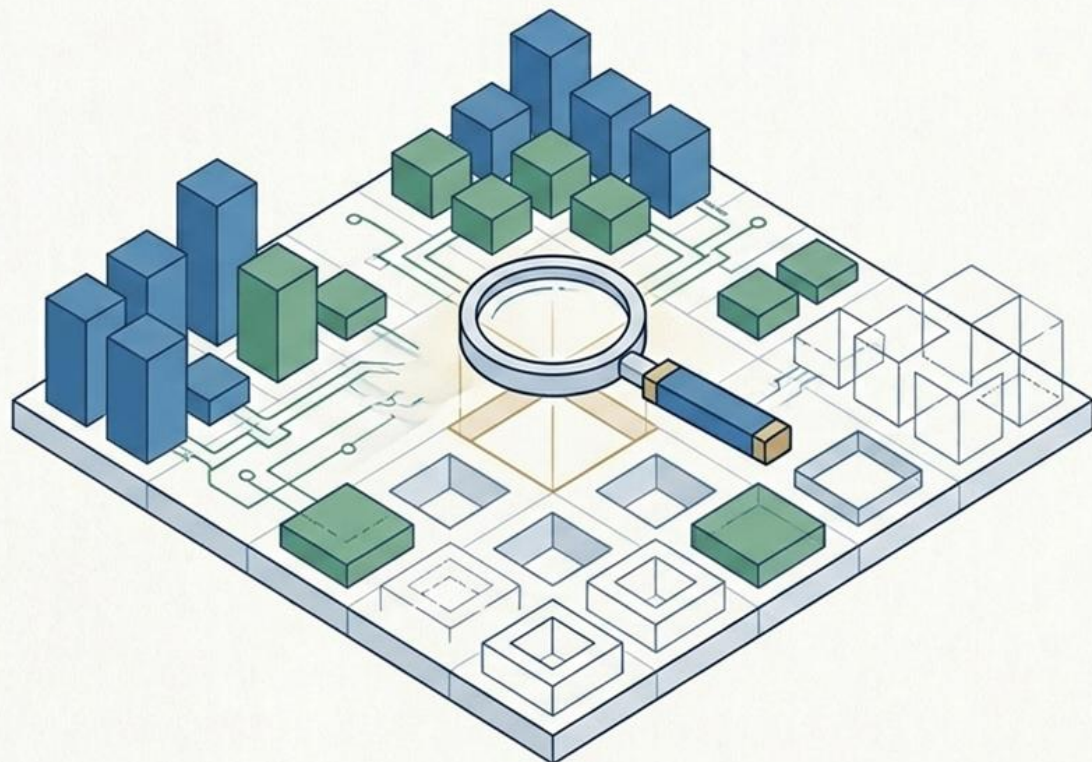
Identifying limitations **is itself a result**
it informs continuous improvement and supports
evidence-based governance.

Shared Responsibility

Data quality governance cannot rest on technology alone.



Illuminating the Gaps: Structural limitations as governance signals



The Reality of Multi-Source Systems

Variations in data completeness, delayed updates, and uneven PID adoption persist across institutions.

Exposing Vulnerabilities

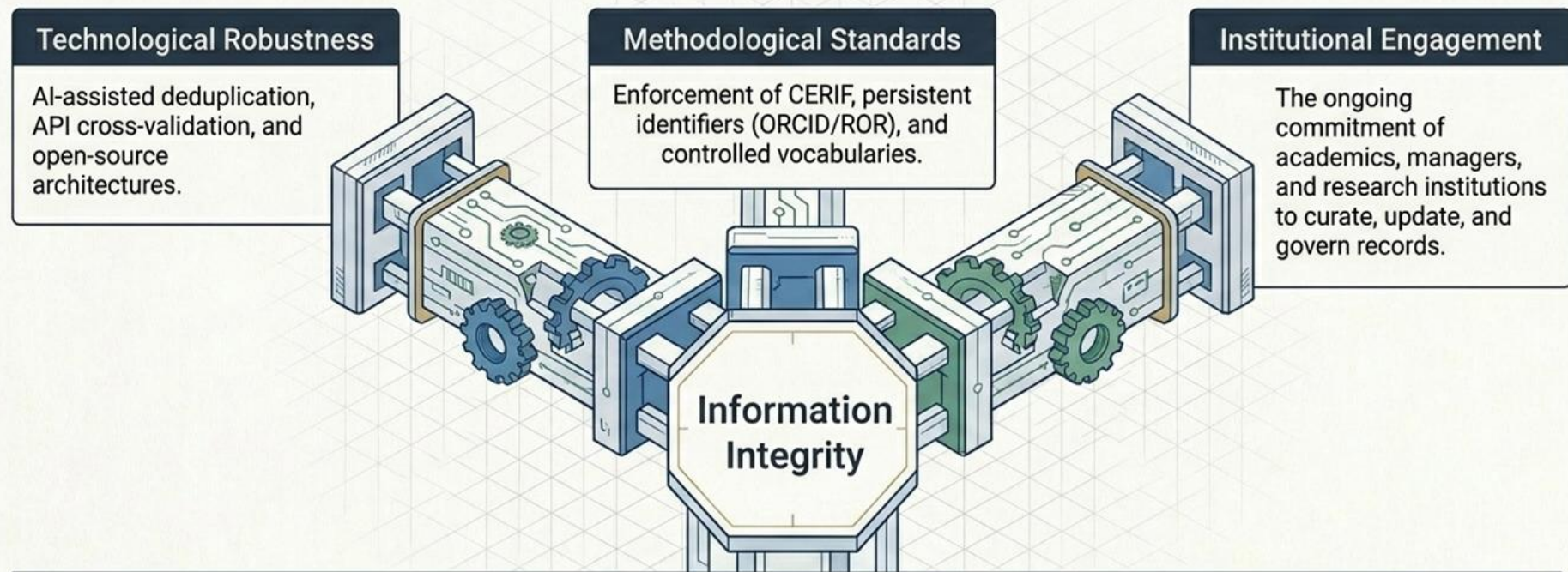
The system's ability to aggregate data effectively highlights these very informational gaps.

Actionable Intelligence

Identifying systemic limitations isn't a failure—it informs continuous improvement strategies and targets institutional interventions.

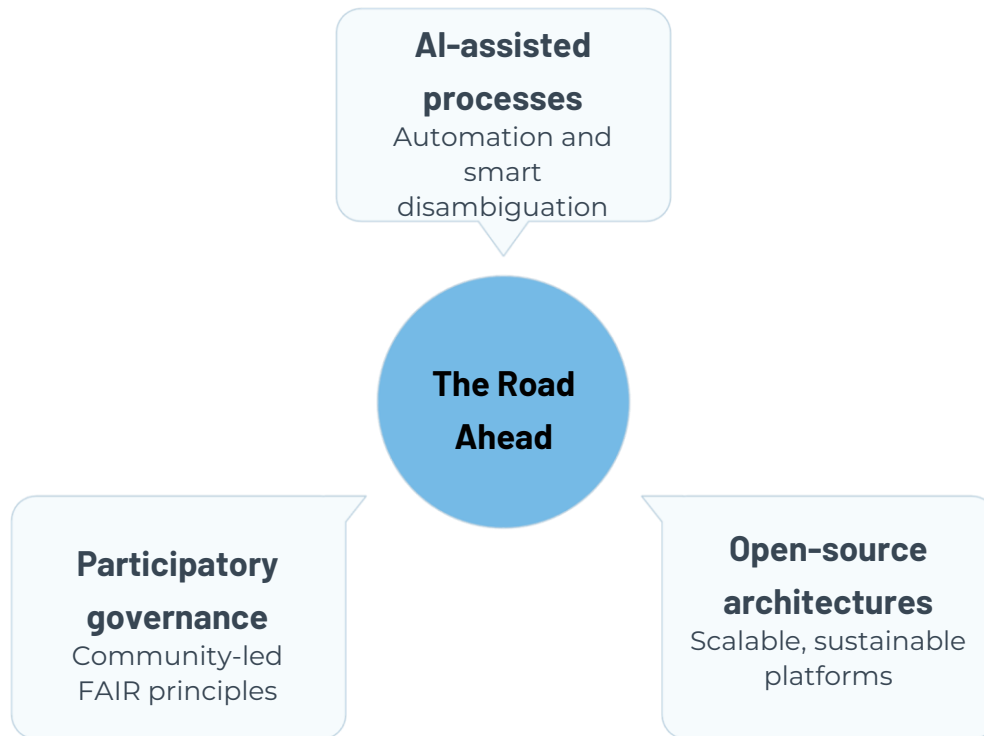
Identifying systemic limitations isn't a failure—it informs continuous improvement strategies and targets institutional interventions.

The Synthesis: CRIS as a Socio-Technical Infrastructure



BrCris illustrates that data quality is deeply intertwined with human governance. Without active community engagement, technological infrastructure eventually decays.

The Road Ahead: Future Directions



Strengthening collaboration between national and international research information infrastructures is essential to support FAIR principles and trustworthy metrics.



BrCris & Global Positioning

BrCris contributes not only to internal management and policy formulation, but also to positioning Brazilian research within **transnational analytical contexts**.

CRIS systems act as mediators between **local data production** and **global research assessment systems**.

Key Takeaways

01

Data Quality is Strategic

Deduplication, disambiguation, and certification are foundational

03

Governance is Shared

Technology + institutions + community = trustworthy ecosystem

02

Interoperability Matters

CERIF, ORCID, ROR, and open APIs connect national to global

04

AI & Open Source Are Next

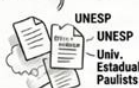
Future CRIS must embrace AI-driven, participatory models

BrCris: Architecting a Reliable Research Information Ecosystem

Building a trustworthy, transparent infrastructure for Brazilian science.



The Challenge: Informational Vulnerabilities



✓ The Risk of Data Fragmentation

Reliance on manual entry and diverse sources leads to duplicate records, spelling variations, and inconsistent abbreviations that compromise data integrity.



✓ Lack of Persistent Identifiers (PIDs)

The absence of universal IDs like DOIs and ORCIDs makes it difficult to track academic output accurately across different platforms.



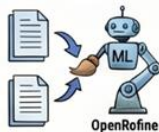
✓ Vulnerability to Misinformation

Data distortions and entry errors can be exploited in misinformation practices, threatening the credibility of institutions and distorting performance indicators.



Strategic Pillars of BrCris Reliability

Smart Record Deduplication



BrCris utilizes machine learning algorithms, clustering techniques, and tools like OpenRefine to identify and merge duplicate records for authors and publications.

Researcher & Institution Disambiguation



By integrating PIDs such as ORCID (for authors) and BOR (for institutions), the system eliminates ambiguity in institutional affiliations and research identities.

Controlled Dictionaries & Validation



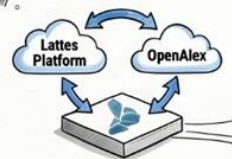
Semi-automated dictionaries for names and graduate programs ensure records are compared against validated nomenclatures before being finalized.

International Standards (CERIF)

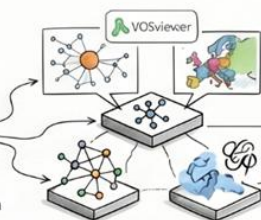


Adopting the Common European Research Information Format (CERIF) guides data modeling and ensures BrCris remains interoperable with global infrastructures.

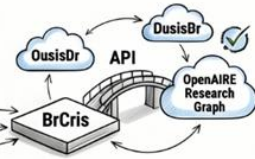
Integration & Visualization



Cross-Platform Synergy
Integration between the Lattes Platform and OpenAlex significantly increases accuracy in identifying institutional links.



Advanced Bibliometric Tools
BrCris exports data to specialized tools like VOSviewer, Visko, and Gephi to identify inconsistent patterns through network visualization.



Automated Cross-Validation
The system uses APIs to cross-reference metadata with reliable external sources like OuisDr and the OpenAIRE Research Graph to automatically correct errors.

The Result: Impact on Open Science



Enhanced Global Visibility

Alignment with international standards positions Brazilian research more effectively within transnational analytical contexts and evaluation frameworks.



Evidence-Based Governance

High-quality, longitudinal data allows for more coherent aggregation of research outputs, supporting fair and reliable policy formulation.



A Shared Responsibility

Effectiveness depends on the ongoing engagement of the academic community in curating and validating records to maintain the ecosystem's integrity.

Washingtonsegundo@ibict.br

marcelsouza@ibict.br

Thank you! Obrigado!

CRIS 2026

17th CRIS Conference on Current Research Information Systems

May 20-22, 2026

