



## Solution for Integrating Educational Platforms into Brazilian CRIS Model

---

Harrysson Gilgamesh de Medeiros Nobrega, Jean Guerethes,  
Talita Moreira de Oliveira and Thiago Adelino de Melo

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

March 24, 2020

## SOLUTION FOR INTEGRATING EDUCATIONAL PLATFORMS INTO BRAZILIAN CRIS MODEL

Harrysson Nobrega, Jean Guerethes, Talita Oliveira, Thiago Melo

### Objective

Considering a CRIS implementation in Brazil, we aim to present a technological solution to interoperate and interact with a diverse scope of systems used by many educational institutions. We designed a customizable product, being able to adapt and be used in any ecosystem regardless the different range of data and technological maturity. Preliminary tests showed a robust and satisfactory solution.

### Highlights

- Different academic systems used by each institution, with different levels of maturity.
- Necessity to have an integrating platform to interoperate data in order to reduce operational effort, to improve data quality and to have information for educational evaluation and funding.
- A customizable solution to collect and expose data.
- Large scale data interchange

### Background

Conecti Brasil project (Brazilian National Consortium for Education, Science, Technologic and Innovation) is a multi-institutional effort to create an ecosystem of information, integrating federal and state systems. The main goal is to provide significant gains in efficiency and operational struggles, and to improve data quality, management and evaluation. Expected benefits of CONECTIBR are:

- **Researches:** Less administrative effort to inform data into several systems
- **Universities:** Interoperability and management
- **Funders:** More efficient budget allocation
- **Scientific Production:** Assurance in publishers and institutional repositories
- **Research Projects:** Partnership and networking
- **Knowledge management:** Complete research reports for a national management.

### Methodology

The first technological solution was a data display, called Heimdall, which can be implemented in any institution interested in providing data and the second solution developed was an API's consumption platform, called Morbius, which will consume the data shared via Heimdall or via

another proprietary solution or even a standardized file.

In Morbius, it is possible to translate the data received into semantic standard adopted by Conecti, following patterns of a CRIS system. This translation is configured only once, by the IT teams of both involved. Both solutions were developed dynamically enough to be customized according to the needs of each institution, self-documented, scalable, and adherent to any security protocol (Figure 1). They are independent and may or may not be implemented together.

They require few computational resources, and can also be encapsulated in Docker, Kubernetes or Openshift. They were built in Java 11 - LTS, using the framework Spring Boot 2.2.1.RELEASE, Spring cloud 2.2.1.RELEASE, Microservices with validation FK, AOP (aspect-oriented programming), Service Mesh, Hateoas, GraphQL, Query Search, Rancher v2.3.2, MongoDB, PostgreSQL, Oracle, Kong, Jfrog, Gitlab Pipeline, Harbor, Graylog 2, Jwt and Jwe, Grafana and Prometheus.

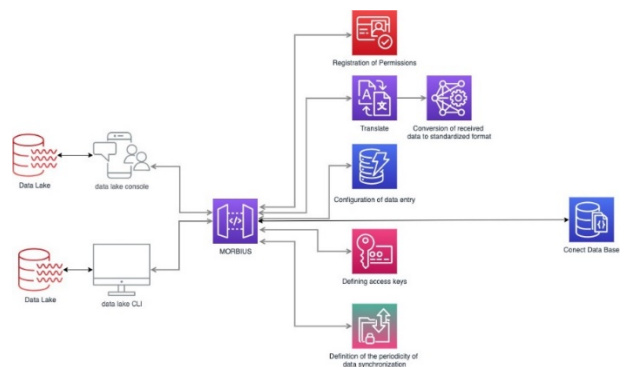


Figure 1. Morbius functionalities

### Results and discussion

Figure 2 shows Morbius and Heimdall operating scheme inside Conecti Brasil context.

Conecti Brasil participants can be universities, government agencies, funders and their respective operating systems, such as research curricula database. Heimdal can be installed to expose data from partners that do not have technological mechanisms for doing this (It is represented in green). If the partner has some other way to expose data it is not necessary to use Heimdall. Then, Morbius will consume data from all partners assuring the correct translation from operational systems and Conecti patterns (represented in red squares). Each partner data will be held in separate staging areas. An artificial neural network will

disambiguate data and maintain in Conecti database the most reliable sources of information.

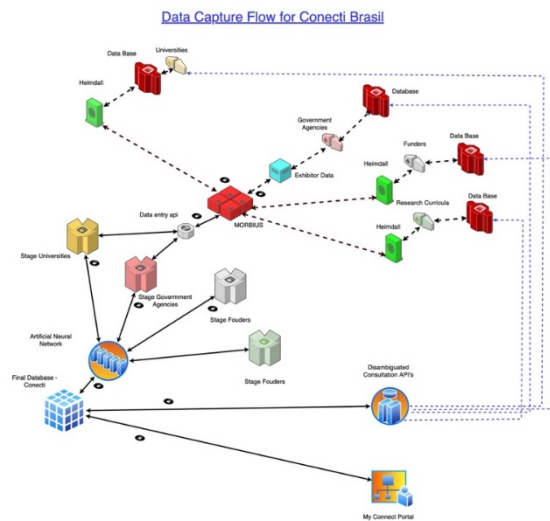


Figure 2. Morbius and Heimdall operating scheme

We initially tested the solutions in the databases of Capes (Brazilian funder and graduate programs evaluator), National Library, Google Scholar, Open Library, ISBN DB, Scopus, ORCID, Google Books, Lattes (national CV), and a partner University to obtain data from the academic system.

The solution supported 60,000 simultaneous requests per second, with the possibility of doubling the capacity by scaling more computers in the cluster.

Both solutions were able to receive any existing data source provided by participants, either via API or file upload, among others whether structured or not. All configuration for receiving and displaying data was simple and made by themselves.

## Conclusion and future work

Preliminary tests showed dynamic and robust solutions to solve any data interchange between institutions regardless of the degree of technological maturity of each one. They will be the technological base to collect, disambiguate and synchronize data related to education, science, technologic and innovation.

The solution will replace web services used by Brazilian government agencies in order to exchange data. We will enlarge our tests with more 12 higher education institutions aiming at promote data enrichment.

## References

1. Abu-Elkheir, Mervat ; Hayajneh, Mohammad ; Ali, Najah. Data Management for the Internet of Things: Design Primitives and Solution. 2013 Sensors, Vol.13(11), pp.15582-15612.
2. Kapur, Deepak ; Narendran, Paliath. Matching, unification and complexity. ACM SIGSAM Bulletin, 01 November 1987, Vol.21(4), pp.6-9.
3. Wasserman, Anthony. The extension of data abstraction to database management. ACM SIGPLAN Notices, Jan. 1981, Vol.16(1), pp.198-200.