

Development of an information system for gathering and analysis of scientific production data

Bruno Boaventura Scholl^{1,2,3}, Sandra C.S. Vessoni⁴, Ivo LeBrun⁵,
Maria Carolina Elias^{1,2}, and Marcelo S. Reis^{1,2}

¹ Center of Toxins, Immune-response and Cell Signaling (CeTICS), Instituto Butantan
² Laboratório de Ciclo Celular, Instituto Butantan
³ Instituto de Matemática e Estatística, Universidade de São Paulo
⁴ Laboratório de Fisiopatologia, Instituto Butantan
⁵ Laboratório de Bioquímica, Instituto Butantan

Background

Scientific productivity evaluation is a crucial part of scientific knowledge production control. The Center of Scientific Development (CDC) of the Butantan Institute produces yearly a detailed productivity report based on the information provided by researchers from various laboratories of the Institute. However, despite recent improvements, the process of producing this report demands too much time and labor from both the researchers and the committee nominated by CDC to compile and analyze the collected data.

Objectives

- **General objective:** To speed up and to automate the scientific report production through the development of an information system.
- **Specific goals:**
 1. Designing and implementation of a database to store and handle years of data
 2. Usage of the database to carry out the 2018 scientific production report
 3. Development of a user interface for data insertion into database

Methods

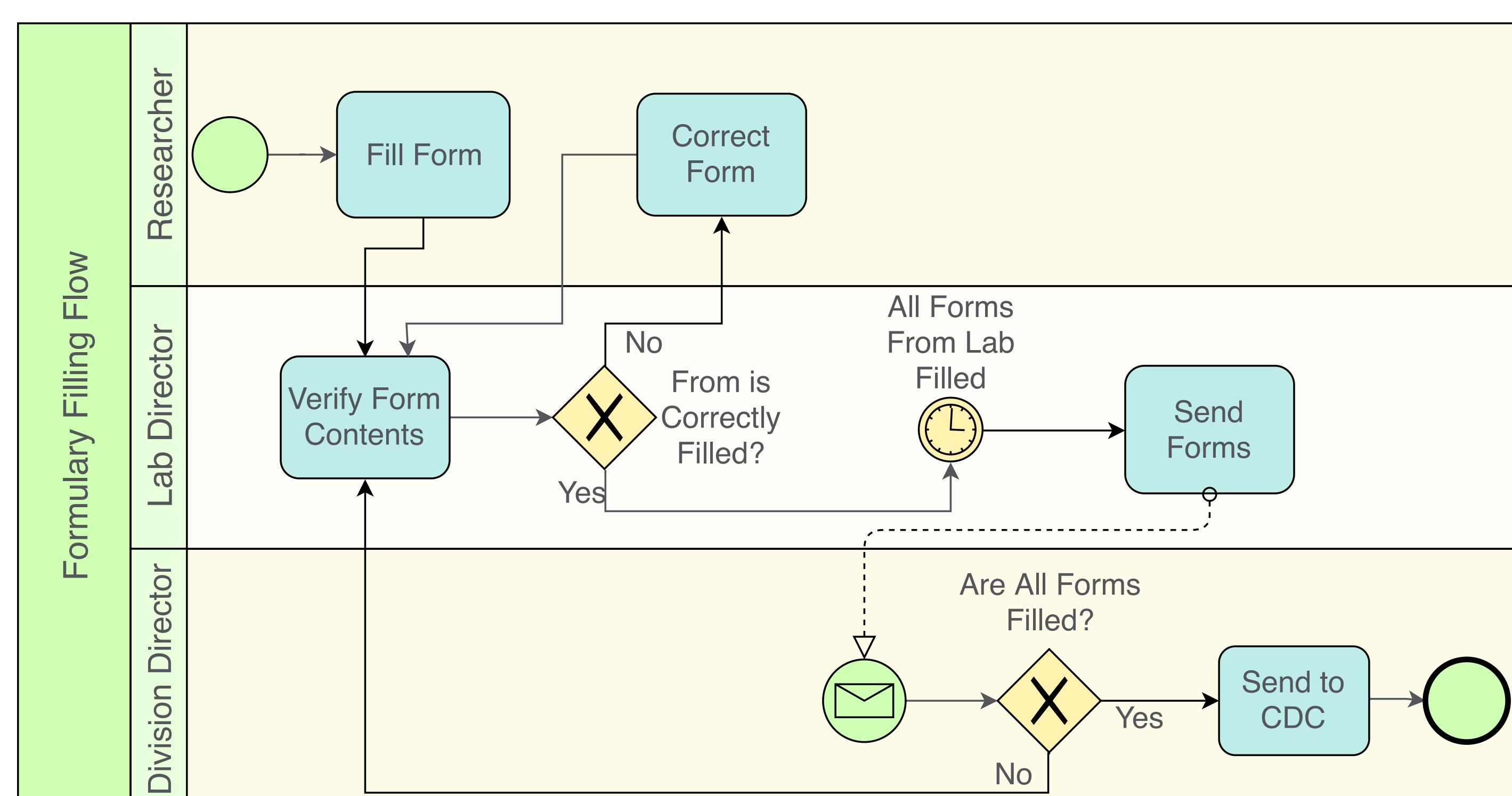


Fig. 1: Data acquisition flow. Researchers fill the forms, then lab directors assess report correctness and approve the submission. After all forms of a lab are filled, they are evaluated by the director of the division. Once all forms from a division are filled, they are ready for analysis.

System interface

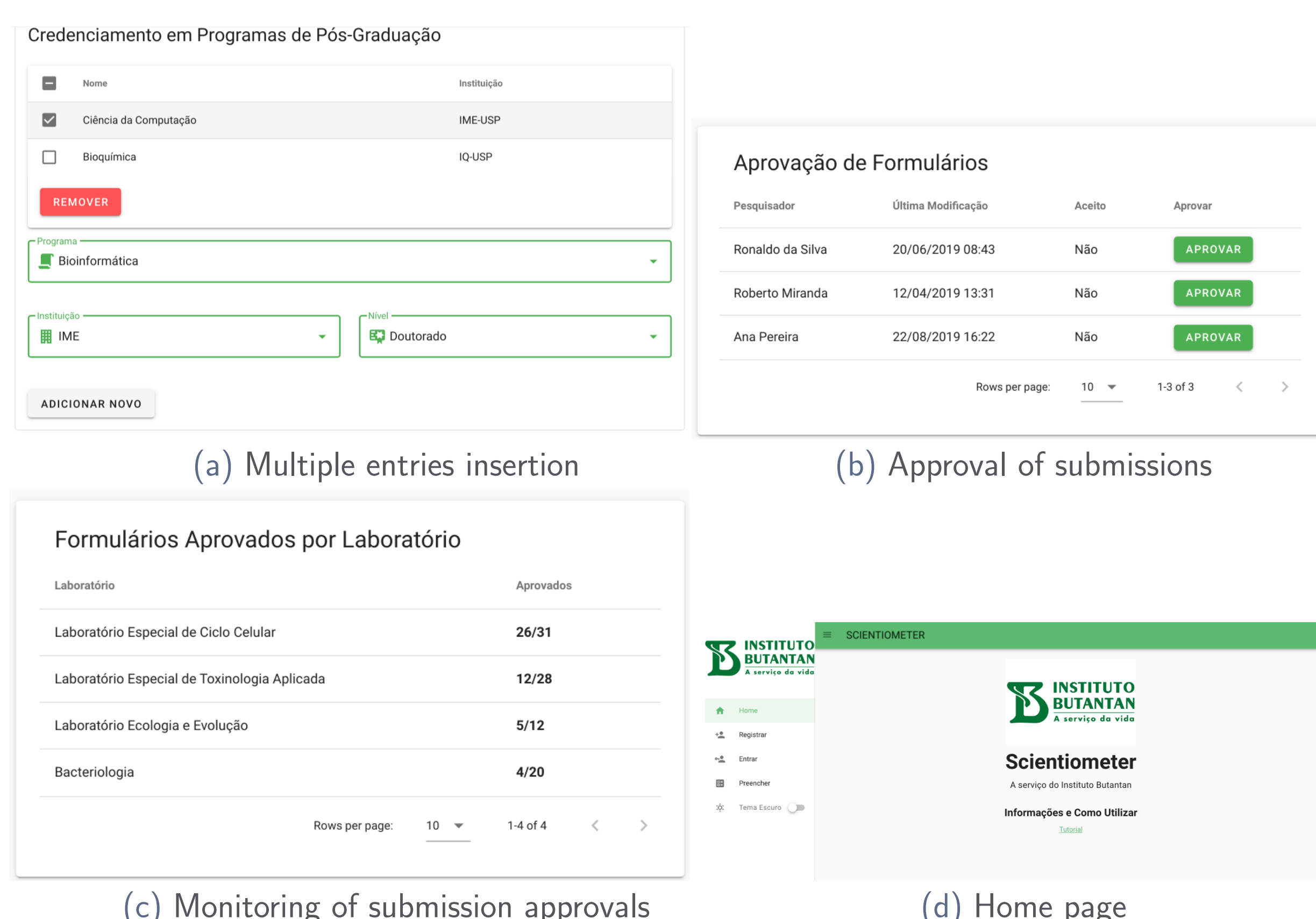


Fig. 2: Interface elements of the interface and homepage. Those elements appear throughout the form filling process and submission approval.

System interface



Fig. 3: First section of the form. In this section, data about the researcher are inserted. The other sections of the form receive data about students, grants, scholarships, publications and activities.

Designed database

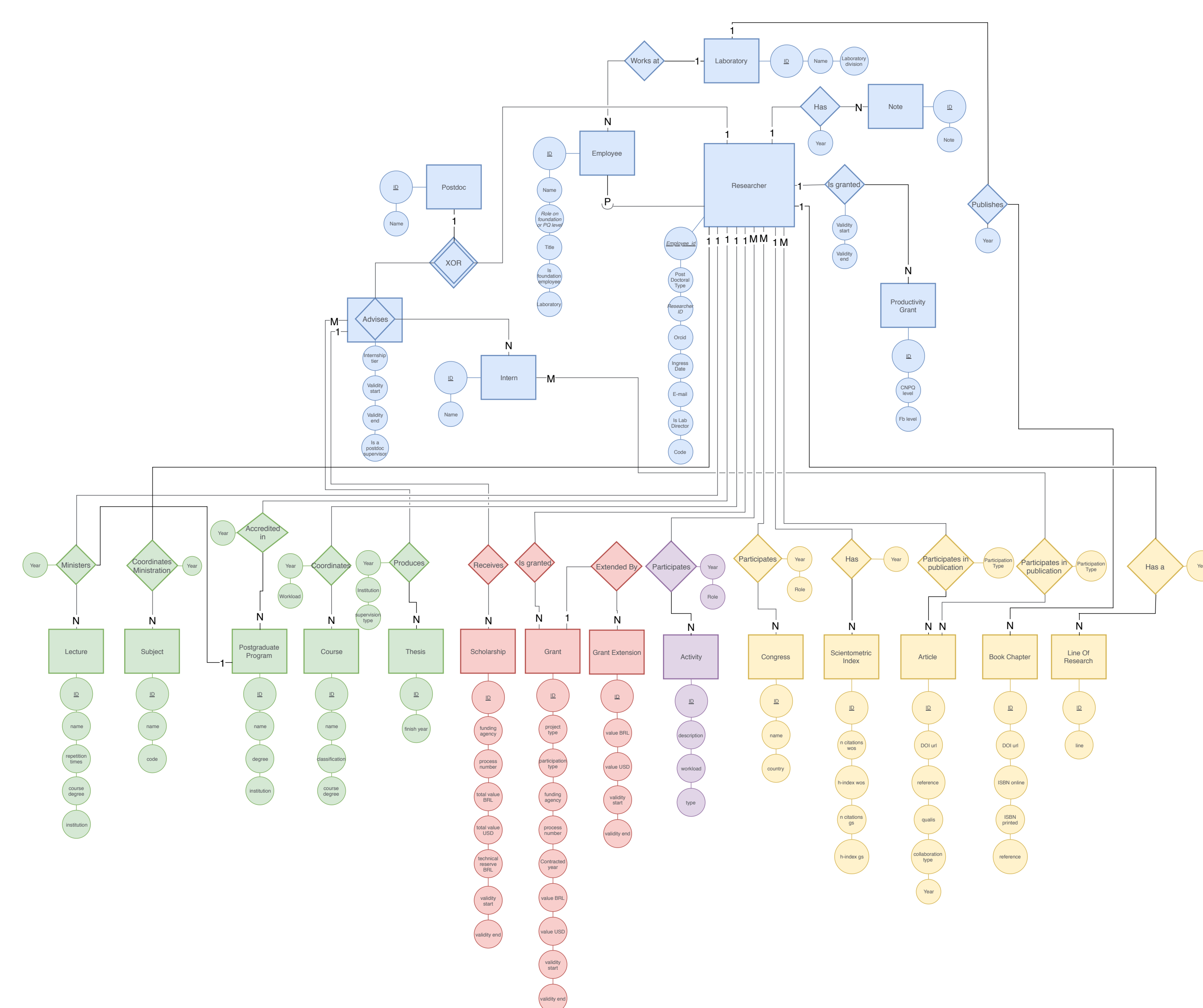


Fig. 4: Entity-Relationship Model of the database. Colors indicate different sections that handle different data kinds (e.g. scholarships, articles). Lines are links between sections and entities; observe that “researcher” is heavily connected with the remaining entities.

Conclusions and future work

- The database was successfully used to carry out the 2018 report
- It was also used to assess individual researcher performance
- Full system will speed-up report production from 2019 onwards
- Next step is the deployment of the system
- Future work can include an interface for data analysis.