

Poster Session I - Room Silves | 3 May 13:15 - 14:15

- **107 - Francisco Couto** - [Data and Text Processing for Development and Stem Cells](#)

Authors

Francisco M. Couto,

Affiliation

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

Abstract Body

Cell Ontology and Ontology for Stem Cell Investigations are two examples of the current interest of the scientific community in providing comprehensive and accurate knowledge bases that help us more effectively to navigate and retrieve information from literature, find evidence, integrate data, generate hypothesis, or even discern relevant from irrelevant data. However, most of these benefits requires an in-depth understanding of complex and sophisticated technologies, that are constantly evolving. This manuscript proposes an alternative solution that is available to every Health and Life specialist without becoming dependent on continuous financial support, third-party applications, or advanced computer skills. This alternative may not offer the full state-of-the-art potential, but offers a feasible and efficient starting solution to explore the main potential of automatically processing data and text using knowledge bases. The solution is to understand the basics of shell scripting in order to perform named-entity recognition and linking by following the book, intituled Data and text processing for health and life sciences, which shows how we can process data and text the same way we conduct a laboratory protocol i.e. testing and understanding its multiple procedural steps, variables, and intermediate results. Additionally, a fully functional shell script, MER - Minimal Entity Recognizer, (available also as a web tool) can be explored to recognize concepts related to Development and Stem Cells (text from PMID:35833364) The open source software is available at: <https://github.com/lasigeBioTM>; the web tools and all the open access book material at <http://labs.rd.ciencias.ulisboa.pt/>

CLOSE