






V. Moitinho de Almeida ^{1,2}, C., Silva ^{3,4}, L., Trigo ^{1,4} (2023). “A integração de Ciência Cidadã no fluxo da produção científica nas áreas das humanidades e ciências sociais”. In *4º Encontro Nacional de Ciência Cidadã (ENCC2023)*, Coimbra. <https://doi.org/10.5281/zenodo.10263605>

¹ Centre for Digital Culture and Innovation (CODA), Faculdade de Letras da Universidade do Porto (FLUP)

² Centro de Investigação Transdisciplinar Cultura, Espaço e Memória (CITCEM), FLUP

³ Departamento de Estudos Portugueses e Estudos Românicos (DEPER), FLUP

⁴ Centro de Linguística da Universidade do Porto (CLUP), FLUP

RESUMO

O CreoPhonPt é um repositório de acesso livre que reúne vários *datasets* com informação fonológica, lexical e sociolinguística dos crioulos do português. Foi criado no âmbito de um projeto de doutoramento (Silva, 2023), para facilitar a análise quantitativa destas línguas e tornar acessíveis dados que anteriormente estavam fechados e/ou dispersos. Dois novos ramos do projeto – CreoPhonPt_classes (2022-23) e CreoPhonPt_field (2023-24) – surgiram na Unidade Curricular (UC) de Fonologia I, Mestrado em Linguística, FLUP. Estes projetos interdisciplinares e colaborativos têm como base a integração dos princípios de Ciência Cidadã, da Ciência Aberta e dos dados FAIR, no fluxo da produção científica nas áreas das humanidades e ciências sociais. Ou seja, com e para a academia e a sociedade. Os seus objectivos assentam sobre o seguinte tripé:

- 1) Comunidade: envolver comunidades luso-africanas do Porto, entre outras, na recolha de campo de dados e no seu arquivo, (re)utilização e divulgação;
- 2) Investigação: enriquecer o repositório existente com mais dados/metadados sócio-históricos, culturais e linguísticos;
- 3) Aprendizagem: facultar aos estudantes capacidades de recolha estruturada de dados científicos, acompanhamento rigoroso de protocolos de investigação, processamento manual/automático dos dados, ligação aberta dos dados/metadados, estimular a co-criação de dados científicos.

Para tal, fazem parte da metodologia de trabalho, a utilização e alimentação simultânea de sistemas eminentemente de cariz cidadã, colaborativos, de acesso livre, abertos e interoperáveis – como sejam, bases de dados, programas, redes, e-infraestruturas de dados, entre outros recursos digitais (p.ex., OpenRefine, Wikidata, GeoNames, Lingua Libre, GitHub, Colab). O sucesso do CreoPhonPt_classes traduziu-se no bom aproveitamento dos estudantes, nas classificações atribuídas à UC nos inquéritos pedagógicos e numa Menção Honrosa do Prémio Prática Pedagógica Inovadora, atribuída pela U.Porto. Quanto ao projeto CreoPhonPt_field, encontra-se já em curso desde este ano lectivo, estando a ser implementado em parceria com investigadores de diferentes áreas das humanidades digitais e ciências sociais.

A integração de Ciência Cidadã e da Ciência Aberta no fluxo da produção científica nas áreas das humanidades e ciências sociais

Vera Moitinho de Almeida, Carlos Silva, Luís Trigo (CODA, CLUP, DEPER, CITCEM - FLUP)

Rationale

Data collection methods

Data processing methods

WHY

WHAT

HOW

Language data
Language metadata

Phonological data
Phonological metadata

Word
Syllables

CreoPhonPt



BIBLIOGRAPHY

- . Data insertion
- . Phonetic transcriptions
- . Syllable/word processing
- . Data cleaning



FIELD

- . Community
- . Data recording
- Linked Open Data



- . Mapping tables for multiple substitutions (REGEX)
- . String replacement
- . Normalisation of phonetic symbols
- . Format conversion - IPA CV



- . Data wrangling
- . Normalisation of data
- . Data reconciliation and matching
- Linked Open Data



- . Repository - **Portophon**
- . Contribution tracking
- . Version control
- **Citizen Science**
- **Open Science**
- **FAIR data**

