

# 10th Symposium on Languages, Applications and Technologies

SLATE 2021, July 1–2, 2021, Vila do Conde/Póvoa de Varzim,  
Portugal

Edited by

Ricardo Queirós

Mário Pinto

Alberto Simões

Filipe Portela

Maria João Pereira



#### *Editors*

##### **Ricardo Queirós**

Escola Superior de Media Artes e Design, Politécnico do Porto, Portugal  
ricardoqueiros@esmad.ipp.pt

##### **Mário Pinto**

Escola Superior de Media Artes e Design, Politécnico do Porto, Portugal  
mariopinto@esmad.ipp.pt

##### **Alberto Simões**

Instituto Politécnico do Cávado e do Ave, Portugal  
asimoes@ipca.pt

##### **Filipe Portela**

Universidade do Minho, Portugal  
cfp@dsi.uminho.pt

##### **Maria João Pereira**

Instituto Politécnico de Bragança, Portugal  
mjoao@ipb.pt

#### *ACM Classification 2012*

Theory of computation → Formal languages and automata theory; Information systems → Data management systems; Information systems → World Wide Web; Computing methodologies → Natural language processing

## **ISBN 978-3-95977-202-0**

#### *Published online and open access by*

Schloss Dagstuhl – Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany. Online available at <https://www.dagstuhl.de/dagpub/978-3-95977-202-0>.

#### *Publication date*

August, 2021

#### *Bibliographic information published by the Deutsche Nationalbibliothek*

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <https://portal.dnb.de>.

#### *License*

This work is licensed under a Creative Commons Attribution 4.0 International license (CC-BY 4.0):  
<https://creativecommons.org/licenses/by/4.0/legalcode>.



In brief, this license authorizes each and everybody to share (to copy, distribute and transmit) the work under the following conditions, without impairing or restricting the authors' moral rights:

- Attribution: The work must be attributed to its authors.

The copyright is retained by the corresponding authors.

Digital Object Identifier: 10.4230/OASlcs.SLATE.2021.0

ISBN 978-3-95977-202-0

ISSN 1868-8969

<https://www.dagstuhl.de/oasics>

## OASlcs – OpenAccess Series in Informatics

OASlcs is a series of high-quality conference proceedings across all fields in informatics. OASlcs volumes are published according to the principle of Open Access, i.e., they are available online and free of charge.

### *Editorial Board*

- Daniel Cremers (TU München, Germany)
- Barbara Hammer (Universität Bielefeld, Germany)
- Marc Langheinrich (Università della Svizzera Italiana – Lugano, Switzerland)
- Dorothea Wagner (*Editor-in-Chief*, Karlsruher Institut für Technologie, Germany)

**ISSN 1868-8969**

**<https://www.dagstuhl.de/oasics>**



*To all whose effort keeps SLATE alive.*



## ■ Contents

|   |             |
|---|-------------|
| Preface   |             |
| <i>Ricardo Queirós, Mário Pinto, Alberto Simões, Filipe Portela, and Maria João Pereira</i> ..... | 0:ix        |
| Committees  |             |
| .....   | 0:xi–0:xii  |
| Authors   |             |
| .....   | 0:xiii–0:xv |

### Invited Talk

|  |          |
|--|----------|
| Natural and Artificial Intelligence; Natural and Artificial Language |          |
| <i>Diana Santos</i> .....  | 1:1–1:11 |

### Regular Papers

|   |            |
|---|------------|
| Derzis: A Path Aware Linked Data Crawler  |            |
| <i>André Fernandes dos Santos and José Paulo Leal</i> .....   | 2:1–2:12   |
| Major Minors – Ontological Representation of Minorities by Newspapers   |            |
| <i>Paulo Jorge Pereira Martins, Leandro José Abreu Dias Costa, and José Carlos Ramalho</i> .....  | 3:1–3:13   |
| Lyntax – A grammar-Based Tool for Linguistics   |            |
| <i>Manuel Gouveia Carneiro de Sousa, Maria João Varanda Pereira, and Pedro Rangel Henriques</i> .....   | 4:1–4:13   |
| Programming Exercises Interoperability: The Case of a Non-Picky Consumer  |            |
| <i>Ricardo Queirós, José Carlos Paiva, and José Paulo Leal</i> .....  | 5:1–5:9    |
| DataGen: JSON/XML Dataset Generator   |            |
| <i>Filipa Alves dos Santos, Hugo André Coelho Cardoso, João da Cunha e Costa, Válder Ferreira Picas Carvalho, and José Carlos Ramalho</i> ..... | 6:1–6:14   |
| MUAHAH: Taking the Most out of Simple Conversational Agents   |            |
| <i>Leonor Llansol, João Santos, Luís Duarte, José Santos, Mariana Gaspar, Ana Alves, Hugo Gonçalo Oliveira, and Luísa Coheur</i> .....          | 7:1–7:12   |
| NER in Archival Finding Aids  |            |
| <i>Luís Filipe Costa Cunha and José Carlos Ramalho</i> .....  | 8:1–8:16   |
| Mooshak’s Diet Update: Introducing YAPEXIL Format to Mooshak  |            |
| <i>José Carlos Paiva, Ricardo Queirós, and José Paulo Leal</i> .....  | 9:1–9:7    |
| LeMe-PT: A Medical Package Leaflet Corpus for Portuguese  |            |
| <i>Alberto Simões and Pablo Gamallo</i> .....   | 10:1–10:10 |
| Towards Automatic Creation of Annotations to Foster Development of Named Entity Recognizers   |            |
| <i>Emanuel Matos, Mário Rodrigues, Pedro Miguel, and António Teixeira</i> .....   | 11:1–11:14 |

10th Symposium on Languages, Applications and Technologies (SLATE 2021).

Editors: Ricardo Queirós, Mário Pinto, Alberto Simões, Filipe Portela, and Maria João Pereira

OpenAccess Series in Informatics



OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

|   |            |
|---|------------|
| Semantic Search of Mobile Applications Using Word Embeddings<br><i>João Coelho, António Neto, Miguel Tavares, Carlos Coutinho,<br/>Ricardo Ribeiro, and Fernando Batista</i> .....                      | 12:1–12:12 |
| Command Similarity Measurement Using NLP<br><i>Zafar Hussain, Jukka K. Nurminen, Tommi Mikkonen, and Marcin Kowiel</i> .....  | 13:1–13:14 |
| Using Machine Learning for Vulnerability Detection and Classification<br><i>Tiago Baptista, Nuno Oliveira, and Pedro Rangel Henriques</i> .....   | 14:1–14:14 |
| NetLangEd, A Web Editor to Support Online Comment Annotation<br><i>Rui Rodrigues, Cristiana Araújo, and Pedro Rangel Henriques</i> .....  | 15:1–15:16 |
| Intelligent Query Answering with Contextual Knowledge for Relational Databases<br><i>Dietmar Seipel, Daniel Weidner, and Salvador Abreu</i> .....   | 16:1–16:15 |
| Sentiment Analysis of Portuguese Economic News<br><i>Cátia Tavares, Ricardo Ribeiro, and Fernando Batista</i> .....   | 17:1–17:13 |
| Bootstrapping a Data-Set and Model for Question-Answering in Portuguese<br><i>Nuno Ramos Carvalho, Alberto Simões, and José João Almeida</i> .....  | 18:1–18:5  |
| Development of Self-Diagnosis Tests System Using a DSL for Creating New Test Suites for Integration in a Cyber-Physical System<br><i>Ricardo B. Pereira, José C. Ramalho, and Miguel A. Brito</i> ..... | 19:1–19:16 |



## ■ Preface

This book of proceedings compiles the accepted papers for the Tenth edition of the Symposium on Languages, Applications and Technologies, SLATE'2021. It is an honour to organize and host this edition, being a special milestone for the conference.

In the last decade it has shown that there is, in fact, similarities in the techniques used for processing the different types of languages we use daily, ranging from the natural languages used for the communication between humans, the programming languages used to instruct the behavior of computers, and finally the interchange languages, used in the communication between machines.

The Tenth edition bring us some nostalgic, by this way we want to share the conference with all. The conference is held in a hybrid mode, as COVID-19 pandemic looks controlled. We are organizing the conference online and opened the possibility of attending in person at the School of Media Arts and Design, in Vila do Conde/Póvoa de Varzim, which kindly hosts this tenth edition.

In this tenth edition of SLATE, we received several interesting contributions which demonstrate the relevance and interest in the languages processing research and which cover almost equally the three facets of the languages materialized in the three conference tracks.

The tracks are addressing several applications of the languages like data crawlers, sentimental analysis, machine learning applied to languages processing, code generators and parsers.

We want to thank the many people without whom this event would never have been possible: all the Members of the Scientific Program Committee for their valuable effort reviewing the submissions, contributing with corrections and new ideas, and helping on deciding the final list of accepted paper; the members of the Organizing Committee, for the help on dealing with the bureaucratic issues; the invited Speakers (*Ricardo Correia* and *Diana Santos*) for accepting our invitation to share their knowledge; and finally, thank you all the Authors for their contributions to SLATE with their current projects and research problems.





## ■ List of Committees

### Program Chairs

Ricardo Queirós  
ESMAD/P.PORTO, uniMAD &  
CRACS/INESC TEC

Mário Pinto  
ESMAD/P.PORTO & uniMAD

Alberto Simões  
2Ai-School of Technology, IPCA

Filipe Portela  
Universidade do Minho

Maria Joao Pereira  
CeDRI, Instituto Politécnico de Bragança

Antonio Leitao  
Universidade de Lisboa

António Teixeira  
University of Aveiro

António Miguel Cruz  
Instituto Politécnico de Viana do Castelo

Arkaitz Zubiaga  
Queen Mary University of London

Barrett Bryant  
University of North Texas

Bostjan Slivnik  
University of Ljubljana

### Organization Committee

Ricardo Queirós  
ESMAD/P.PORTO, uniMAD &  
CRACS/INESC TEC

Mário Pinto  
ESMAD/P.PORTO & uniMAD

Alberto Simões  
2Ai-School of Technology, IPCA

Teresa Terroso  
ESMAD/P.PORTO & uniMAD

Jorge Lima  
ESMAD/P.PORTO

Brett Drury  
LIAAD-INESC-TEC

Daniela Da Cruz  
Checkmarx

Diana Santos  
University of Oslo

Dietmar Seipel  
University of Wuerzburg

Dušan Kolář  
Brno University of Technology

Fernando Batista  
Instituto Universitário de Lisboa

### Scientific Committee

Alberto Simões  
Instituto Politécnico do Cávado e Ave

Alda Gancarski  
Institut Mines-Telecom

Alexander Paar  
Duale Hochschule Schleswig-Holstein

Alexandre Rademaker  
IBM Research Brazil

Ana Alves  
University of Coimbra

Filipe Portela  
University of Minho

Geylani Kardas  
Ege Üniversitesi

Hugo Gonçalo Oliveira  
University of Coimbra

Irene Rodrigues  
Universidade de Évora

Irene Castellón  
Universitat de Barcelona

Ivan Luković  
University of Belgrade

10th Symposium on Languages, Applications and Technologies (SLATE 2021).  
Editors: Ricardo Queirós, Mário Pinto, Alberto Simões, Filipe Portela, and Maria João Pereira  
OpenAccess Series in Informatics



OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

## 0:xii Participants

Jakub Swacha  
University of Szczecin

Jan Janousek  
Czech Technical University Prague

Jan Kollar  
Technická univerzita v Košiciach

Jaroslav Porubän  
Technical University of Košice

Jean-Christophe Filliatre  
Université Paris-Saclay

João Saraiva  
University of Minho

João M. Lourenço  
Universidade NOVA de Lisboa

José Carlos Ramalho  
University of Minho

José Carlos Paiva  
University of Porto

José João Almeida  
Universidade do Minho

Jose Luis Sierra  
Universidad Complutense de Madrid

José Paulo Leal  
University of Porto

Josep Silva  
Universitat Politècnica de València

Luis Morgado Da Costa  
Nanyang Technological University

Luís Ferreira  
Instituto Politécnico do Cávado e Ave

Luísa Coheur  
Instituto Superior Técnico

Maria João Pereira  
Instituto Politécnico de Bragança

Mario Berón  
National University of San Luis

Mário Pinto  
ESMAD, Politécnico do Porto

Marjan Mernik  
University of Maribor

Mikel Forcada  
Universitat d'Alacant

Pablo Gamallo  
University of Santiago de Compostela

Pedro Rangel Henriques  
University of Minho

Ricardo Queirós  
ESMAD, Politécnico do Porto

Ricardo Rodrigues  
University of Coimbra

Ricardo Rocha  
University of Porto

Salvador Abreu  
University of Évora

Simão Melo de Sousa  
Universidade da Beira Interior

Tomaz Kosar  
University of Maribor

Xavier Gómez Guinovart  
Universidade de Vigo

## ■ List of Authors

Alberto Simões  
2Ai Lab, School of Technology, IPCA,  
Barcelos, Portugal  
asimoes@ipca.pt

Ana Alves  
CISUC, Instituto Politécnico de Coimbra,  
Portugal  
ana@dei.uc.pt

André Santos  
CRACS & INESC Tec LA  
University of Porto, Portugal  
andrefs@andrefs.com

António Neto  
Caixa Mágica Software  
Lisboa, Portugal  
antonio.neto@caixamagica.pt

António Teixeira  
IEETA, DETI  
University of Aveiro, Portugal  
ajst@ua.pt

Carlos Coutinho  
Caixa Mágica Software  
ISTAR-IUL, ISCTE  
Instituto Universitário de Lisboa  
carlos.coutinho@caixamagica.pt

Cátia Tavares  
ISCTE, Instituto Universitário de Lisboa  
Av. das Forças Armadas, Portugal  
catiattavares@gmail.com

Cristiana Esteves Araújo  
Centro ALGORITMI  
Universidade do Minho, Braga, Portugal  
decristianaaraujo@hotmail.com

Daniel Weidner  
University of Würzburg  
Würzburg, Germany  
daniel.weidner@stud-mail.  
uni-wuerzburg.de

Dietmar Seipel  
University of Würzburg  
Würzburg, Germany  
dietmar.seipel@uni-wuerzburg.de

Emanuel Matos  
IEETA, DETI  
Universidade de Aveiro, Portugal  
easm@ua.pt

Fernando Batista  
INESC-ID & ISCTE-IUL  
Instituto Universitário de Lisboa, Portugal  
fernando.batista@iscte-iul.pt

Filipa Santos  
University of Minho  
Braga, Portugal  
a83631@alunos.uminho.pt

Hugo Cardoso  
University of Minho  
Braga, Portugal  
a85006@alunos.uminho.pt

Hugo Gonçalo Oliveira  
CISUC, DEI  
Universidade de Coimbra, Portugal  
hroliv@dei.uc.pt

João Coelho  
Caixa Mágica Software  
Instituto Superior Técnico  
Lisboa, Portugal  
joao.coelho@caixamagica.pt

João Costa  
University of Minho  
Braga, Portugal  
a84775@alunos.uminho.pt

João Santos  
INESC-ID & Instituto Superior Técnico  
Lisboa, Portugal  
joao.l.santos@tecnico.ulisboa.pt

José Carlos Paiva  
CRACS: INESC-Porto LA  
DCC - FCUP, Porto, Portugal  
jose.c.paiva@inesctec.pt

10th Symposium on Languages, Applications and Technologies (SLATE 2021).

Editors: Ricardo Queirós, Mário Pinto, Alberto Simões, Filipe Portela, and Maria João Pereira

OpenAccess Series in Informatics



Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

José Carlos Ramalho  
Department of Informatics  
University of Minho  
Braga, Portugal  
jcr@di.uminho.pt

José João Almeida  
Centro Algoritmi  
Universidade do Minho  
Braga, Portugal  
jj@di.uminho.pt

José Paulo Leal  
CRACS & INESC Tec LA  
University of Porto, Portugal  
zp@dcc.fc.up.pt

José Santos  
CISUC, DEI  
Universidade de Coimbra, Portugal  
santos@student.dei.uc.pt

Jukka Nurminen  
Department of Computer Science  
University of Helsinki, Finland  
jukka.k.nurminen@helsinki.fi

Leandro Costa  
University of Minho  
Braga, Portugal  
leandro.costa16@hotmail.com

Leonor Llansol  
INESC-ID & Instituto Superior Técnico  
Lisboa, Portugal  
leonor.llansol@tecnico.ulisboa.pt

Luís Duarte  
CISUC, DEI  
Universidade de Coimbra, Portugal  
lduarte@student.dei.uc.pt

Luís Cunha  
University of Minho  
Braga, Portugal  
a83099@alunos.uminho.pt

Luísa Coheur  
INESC-ID & Instituto Superior Técnico  
Lisboa, Portugal  
lcoheur@edu.ulisboa.pt

Manuel Sousa  
University of Minho  
Braga, Portugal  
manuelgcsousa@gmail.com

Marcin Kowiel  
F-Secure Corporation, Poland  
marcin.kowiel@f-secure.com

Maria João Pereira  
Research Centre in Digitalization and  
Intelligent Robotics  
Instituto Politécnico de Bragança, Portugal  
mjoao@ipb.pt

Mariana Gaspar  
INESC-ID & Instituto Superior Técnico  
Lisboa, Portugal  
mariarpx@gmail.com

Mário Rodrigues  
IEETA, ESTGA  
Universidade de Aveiro, Portugal  
mjfr@ua.pt

Miguel Brito  
Centro Algoritmi  
University of Minho  
Guimarães, Portugal  
mab@dsi.uminho.pt

Miguel Tavares  
Caixa Mágica Software  
ULHT - Universidade Lusófona de  
Humanidades e Tecnologias  
miguel.tavares@caixamagica.pt

Nuno Oliveira  
Checkmarx, Braga, Portugal  
nuno.oliveira@checkmarx.com

Nuno Ramos Carvalho  
Guimarães, Portugal  
narcarvalho@gmail.com

Pablo Gamallo  
CiTIUS  
Universidade de Santiago de Compostela  
Galiza, Spain  
pablo.gamallo@usc.es

Paulo Martins  
University of Minho  
Braga, Portugal  
paulo.jorge.pm@gmail.com

Válter Carvalho  
University of Minho  
Braga, Portugal  
a84464@alunos.uminho.pt

Pedro Miguel  
IEETA, DETI  
Universidade de Aveiro, Portugal  
pedro.miguel.rafael@gmail.com

Zafar Hussain  
Department of Computer Science  
University of Helsinki, Finland  
zafar.hussain@helsinki.fi

Pedro Rangel Henriques  
University of Minho  
Braga, Portugal  
pedrorangelhenriques@gmail.com

Ricardo Pereira  
University of Minho  
Braga, Portugal  
ricardo-97-pereira@hotmail.com

Ricardo Queirós  
CRACS: INESC-Porto LA, Porto  
uniMAD: ESMAD, Polytechnic of Porto  
Porto, Portugal  
ricardo.queiros@gmail.com

Ricardo Ribeiro  
ISCTE - Instituto Universitário de Lisboa  
INESC-ID Lisboa, Portugal  
ricardo.ribeiro@iscte-iul.pt

Rui Rodrigues  
University of Minho  
Braga, Portugal  
a74572@alunos.uminho.pt

Salvador Abreu  
Nova-Lincs, University of Évora  
Portugal  
spa@uevora.pt

Tiago Baptista  
University of Minho  
Braga, Portugal  
a75328@alunos.uminho.pt

Tommi Mikkonen  
Department of Computer Science  
University of Helsinki, Finland  
tommi.mikkonen@helsinki.fi

