

INLG 2017

**Proceedings of the 1st Workshop on  
Explainable Computational Intelligence**

**XCI 2017**

September 4, 2017  
Santiago de Compostela, Spain

©2017 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
209 N. Eighth Street  
Stroudsburg, PA 18360  
USA  
Tel: +1-570-476-8006  
Fax: +1-570-476-0860  
[acl@aclweb.org](mailto:acl@aclweb.org)

ISBN 978-1-945626-79-1

## Preface

In recent years, a torrent of Computational Intelligence (CI) applications with an outstanding autonomous degree in their behaviour has been successfully developed (e.g., fuzzy controllers, neural networks, genetic algorithms, etc). These systems can operate without human intervention and achieve unbeatable results; however, at the same time and as a result of this success, their complexity has been dramatically increased. Thus, for the typical user, these systems become *black boxes* and he or she has to blindly trust in them.

Given this context, the problem of **explainability** arises. Its main goal can be described as transforming these *black-box* systems into *glass-box* ones, where the end user can understand the reasons that support the system's decisions. For instance, when an expert system in medicine advises a patient to take a particular drug for treating his disease, he or she needs to know why these are the right drugs for his or her disease.

During the last two years several workshops has been organised around the topic of explainability in computational systems, mainly in Artificial Intelligence (known as Explainable Artificial Intelligence - XAI). Mostly of them have been focused on machine learning, one of the current hot topics in AI, and how these techniques can produce more *explainable models*. In the case of CI, this issue is already in its agenda from some time ago and has been addressed by the studies in *interpretability*, which main goal is to keep the inner structure of the CI systems as clear as possible both for engineers and users. Thus, for that reason, in the Workshop on Explainable Computational Intelligence (XCI), we focus on other two challenges typical in explainability studies: *explanation interface* and *psychology of explanation*.

Explanation interface is directly related with the techniques to generate effective explanations for human users. Therefore, the use of **natural language**, as the main tool for human communication, perfectly suits to this aim. Holding the XCI during the International Natural Language Generation (INLG) brings us the chance of tackling this hurdle from a multidisciplinary perspective and put the grounds for a new collaboration space between both communities.

Psychology of explanation points out, precisely, to the necessity of a computational theory of explanation. Currently, there is a gap between the **machine logic**, which underlies this type of systems, and **human logic** and, consequently, building a bridge between them is a necessary step to make them more explainable and understandable. In CI, a computational theory of perceptions has been developed and successfully applied, and, for that reason, the experience gathered by the researchers in this field will provide some relevant clues for the development of the aforementioned computational theory of explanation.

In this first edition of the XCI, we have received seven submissions of short papers and four of them appear in this volume. In addition, we have an invited talk by Dr. Jose M. Alonso (University of Santiago de Compostela), who has a broad experience in the topic of interpretability in CI systems.

We would like to thank the Program Committee members who reviewed the papers and helped to improve the overall quality of the workshop. We also thank the General Chairs and Workshop Chairs of INLG Conference the given us the chance to organise this workshop. Last, a word of thanks goes to our invited speaker, Dr. Jose M. Alonso.

August 31, 2017  
Dundee (Scotland)

Martín Pereira-Fariña  
Chris Reed  
Co-Organizers of XCI 2017



**Organizers:**

Martin Pereira-Fariña	University of Dundee & University of Santiago de Compostela
Chris Reed	University of Dundee & Polish Academy of Sciences

**Program Committee:**

João Paulo Carvalho	Instituto Superior Tecnico / INESC-ID
Pablo Gamallo	University of Santiago de Compostela
Albert Gatt	University of Malta
Lluís Godó	Artificial Intelligence Research Institute, IIIA - CSIC
Floriana Grasso	University of Liverpool
Helmut Horacek	Saarland University
Martín Pereira-Fariña	University of Santiago de Compostela
Chris Reed	University of Dundee
Patrick Saint-Dizier	IRIT-CNRS
Alejandro Sobrino	University of Santiago de Compostela
Adam Wyner	University of Aberdeen

**Invited Speaker:**

Jose M. Alonso	University of Santiago de Compostela
----------------	--------------------------------------



## Table of Contents

<i>Two Challenges for CI Trustworthiness and How to Address Them</i> Kevin Baum, Maximilian A. Köhl and Eva Schmidt .....	1
<i>A Simple Method for Clarifying Sentences with Coordination Ambiguities</i> Michael White, Manjuan Duan and David L. King .....	6
<i>Requirements for Conceptual Representations of Explanations and How Reasoning Systems Can Serve Them</i> Helmut Horacek .....	11
<i>An Essay on Self-explanatory Computational Intelligence: A Linguistic Model of Data Processing Systems</i> Jose M. Alonso and Gracian Trivino .....	16





## Workshop Program

**08:30 - 09:30**    **Registration**

**09:00 - 09:10**    **Presentation**

**09:10 - 10:10**    **Invited talk by Jose M. Alonso (University of Santiago de Compostela)**

**10:10 - 11:00**    **Logic and explanation in XCI and NLG**

- **10:10 - 10:35**    *Two Challenges for CI Trustworthiness and How to Address Them*  
Kevin Baum, Maximilian A. Köhl and Eva Schmidt
- **10:35 - 11:00**    *A Simple Method for Clarifying Sentences with Coordination Ambiguities*  
Michael White, Manjuan Duan and David L. King

**11:00 - 11:30**    **Coffee Break**

**11:30 - 12:20**    **Session 2: Challenges in XCI and NLG**

- **11:30 - 11:55**    *Requirements for Conceptual Representations of Explanations and How Reasoning Systems Can Serve Them*  
Helmut Horacek
- **11:55 - 12:20**    *An Essay on Self-explanatory Computational Intelligence: A Linguistic Model of Data Processing Systems*  
Jose M. Alonso and Gracian Trivino

**12:20 - 13:20**    **Open discussion and future plans**

**13:20 - 13:30**    **Wrap-up**

## Invited talk

# eXplainable Computational Intelligence: paving the way from Smart to Cognitive Cities

Jose M. Alonso  
josemaria.alonso.moral@usc.es

*Centro de Investigación en Tecnologías da Información (CiTIUS)  
Universidade de Santiago de Compostela*

## Abstract

In the era of the Internet of Things, there has been a huge effort connecting all kind of devices to Internet.

Accordingly, in modern cities, everything is connected to Internet. Thus, human beings face two main challenges:

- to extract valuable knowledge from the given Big Data (Data Scientists are more and more demanded by companies);
- to become part of the equation, i.e., to become active actors in the Internet of Things (in our daily life).

Researchers and developers have already created more and more intelligent devices which populate the so-called smart cities.

Moreover, nowadays the focus is set on knowledge representation and how to enhance human-machine interaction, i.e., it is time to address the effective interaction between intelligent systems and citizens with the aim of passing from Smart to Cognitive Cities. Non-expert users, i.e., users without a strong background on Artificial Intelligence (AI), require a new generation of eXplainable AI (XAI) systems. They are expected to naturally interact with humans, thus providing comprehensible explanations of decisions automatically made. In this talk, we sketch how certain Computational Intelligence (CI) techniques, namely interpretable fuzzy systems, are ready to play a key role in the development of XCI systems, i.e., CI-based XAI systems.