



cy. center for
algorithmic
transparency

Document Title	White Paper Reports
Project Title and acronym	Cyprus Center for Algorithmic Transparency (CyCAT)
H2020-WIDESPREAD-05-2017-Twinning	Grant Agreement number: 810105 — CyCAT
Deliverable No.	D2.11
Work package No.	WP2
Work package title	Dissemination and Outreach
Authors (Name and Partner Institution)	Styliani Kleanthous (OUC) Jahna Otterbacher (OUC)
Contributors (Name and Partner Institution)	Fausto Giunchiglia (UNITN) Frank Hopfgartner (USFD) Tsvi Kuflik (UH) Michael Rovatsos (UEDIN)
Reviewers	Maria Kasinidou (OUC)
Status (D: draft; RD: revised draft; F: final)	F
File Name	D2.11_White_paper_reports
Date	22 November 2021



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 810105.

Draft Versions - History of Document				
Version	Date	Authors / contributors	e-mail address	Notes / changes
v1.0	02/11/2021	Styliani Kleanthous (OUC)	styliani.kleanthous@ouc.ac.cy	Initial version
v2.0	10/11/2021	Maria Kasinidou (OUC)	maria.kasinidou@ouc.ac.cy	Reviewed Draft
v3.0	22/11/2021	Styliani Kleanthous (OUC)	styliani.kleanthous@ouc.ac.cy	Final Version

Abstract

This document describes the White Papers written by members of the CyCAT consortium. White Paper reports were part of WP2 and aimed to disseminate the knowledge developed within CyCAT to key stakeholders (e.g., teachers, developers) and to the wider public. There are 11 white papers in total targeting different audiences. In this deliverable, a brief description of each white paper is provided along with a link to the CyCAT website where all White Papers are made publicly available.

Keyword(s):

Developers training, Evaluation, Educators, Watchdog, Tools for Transparency, Policy Makers

Contents

1. Executive Summary
2. List of White Papers

1. Executive Summary

As described in the CyCAT DoA, the consortium has collaboratively written 11 White Papers addressing the needs of different audiences including developers, educators at different levels and policy makers. Each manuscript was the outcome of the activities that took place in CyCAT at different WPs. A list of the White Papers follows. For each one, we provide the title, authors, the audience that the paper refers to, and a brief description. All White Papers are publicly available through the CyCAT Website at: <http://www.cycat.io/white-papers/>

2. List of White Papers

- **Practitioners' Education About Algorithmic Fairness, Accountability and Transparency** by Veronika Bogina, Alan Hartman, Tsvi Kuflik, Avital Shulner-Tal - University of Haifa

Audience: Developers

This white paper provides a basis for educating software practitioners about Algorithmic Fairness, Accountability and Transparency (FAT) concepts. The document is based on the work carried out in the context of the CyCAT project on educating software and AI stakeholders about FAT. There are several primary stakeholders and in this white paper we refer only to software practitioners. We present five educational modules and elaborate on their application with respect to software practitioners.

- **Αλγοριθμική Παιδεία στην Εκπαίδευση: Οδηγός για Εκπαιδευτικούς - in Greek** by Styliani Kleanthous, Maria Kasinidou, Michalinos Zembylas, Miranda Christou, Jahna Otterbacher - Open University of Cyprus

Audience: Educators (Cyprus)

Ο οδηγός αυτός απευθύνεται στους εκπαιδευτικούς όλων των βαθμίδων και αποσκοπεί στην προώθηση της αλγοριθμικής παιδείας στα σχολεία. Ο οδηγός είναι βασισμένος σε έρευνες που διεξήχθησαν στα πλαίσια του ερευνητικού έργου CyCAT, και στηρίζεται σε ερευνητικά αποτελέσματα που λήφθηκαν από την διεξαγωγή σεμιναρίων με Κύπριους εκπαιδευτικούς. Ο οδηγός παρέχει τις βασικές έννοιες που χρειάζονται για την κατανόηση της λειτουργίας των αλγοριθμικών διαδικασιών. Στην συνέχεια, παρουσιάζεται ο ρόλος που έχουν οι εκπαιδευτικοί στην προώθηση της αλγοριθμικής παιδείας. Τέλος, ο οδηγός παρέχει το πρόγραμμα δραστηριοτήτων που αναπτύχθηκε από την ερευνητική ομάδα CyCAT για το σκοπό αυτό.

- **FATE in Algorithmic Systems: What Higher Education Practitioners Should Know?** by Styliani Kleanthous, Maria Kasinidou, Kalia Orphanou, Pinar Barlas, Jahna Otterbacher - Open University of Cyprus

Audience: Higher Education Practitioners

This White Paper provides an overview on the issue of Fairness, Accountability, Transparency and Ethics (FATE) in Algorithmic systems and their importance in degrees related to software development and computer science. The document is based on the work carried out in the context of the CyCAT project to understand how future developers perceive elements of FATE in algorithmic decision-making and to understand whether educating young developers about FATE might affect the way they think about those concepts. Thus, it is important for educators to understand how students in computer science and related degrees perceive concepts such Fairness, Accountability, Transparency and Ethics and their possible consequences in algorithmic systems.

- **FATE in Algorithmic Systems: What Developers Should Know?** by Styliani Kleanthous, Maria Kasinidou, Kalia Orphanou, Pinar Barlas, Jahna Otterbacher - Open University of Cyprus

Audience: Developers

This white paper provides a basis for i) becoming aware of Fairness, Accountability, Transparency (FAT) issues in the development of algorithmic systems; ii) learning the core FAT concepts related to software development; iii) developing appreciation for the role that developers play in mitigating algorithmic bias and in promoting ethical practices. The document is based on the work carried out in the context of the CyCAT project to understand future developers' perception of FAT in algorithmic decision-making and to promote their awareness of FAT concepts. Given that algorithmic systems and their decisions are affecting every aspect of our lives and that they do not always behave as they should, reproducing and/or amplifying social stereotypes and inequalities. Thus, it is important for developers to understand concepts such Fairness, Accountability, Transparency and their possible consequences in algorithmic systems.

- **Tips/Best Practices When Generating Datasets** by Nandu Chandran Nair, Fausto Guinchiglia, Jahna Otterbacher, Kalia Orphanou - University of Trento and Open University of Cyprus

Audience: Researchers / developers / educators

This white paper provides tips and best practices for generating datasets in the context of Machine Learning (ML). The document is based on the work carried out in the context of the CyCAT project to support and inform stakeholders involved in generating datasets (e.g., researchers, developers, or even educators creating datasets for their data science or AI courses). This white paper is written to be accessible to a broad audience, including those new to machine learning. After providing a brief overview of the ML process and the possible sources of biases, we detail the nature of data biases. The final section

provides some concrete guidelines and practical tips for those who create ML datasets to mitigate social biases in the data, and eventually, the resulting algorithms and outputs/decisions.

- **Tips/Best Practices When Using Datasets** by Nandu Chandran Nair, Fausto Guinchiglia, Jahna Otterbacher, Kalia Orphanou - University of Trento and Open University of Cyprus

Audience: Students (of data / computer / information science)

This white paper provides tips and best practices for using datasets in a Machine Learning (ML) context. The document is based on the work carried out in the CyCAT project and is written in a manner as to be as accessible as possible to a wide range of stakeholders, who are not experts in ML, and who wish to exploit existing datasets (i.e., datasets created by others) in their work. To this end, we provide an overview of the crucial role of data in the ML pipeline, as well as some guidelines and tips for effectively using the datasets created by others. The ultimate goal of the white paper is to raise the awareness of readers in terms of the common issues of bias in ML datasets, and the need to be vigilant as to the biases in the resulting ML models and outputs, which may occur.

- **Integration of FATE and Critical Data Studies into Data Science Curricula** by Lena Podoletz, Monica Paramita, Jo Bates, Frank Hopfgartner, Michael Rovatsos - The University of Sheffield and The University of Edinburgh

Audience: Educators

While it is widely acknowledged that future data scientists, and other technologists, need to be better equipped to address the ethical and societal challenges posed by emerging technologies, there is considerable debate about how this can be best achieved and put into concrete educational practice. One approach is to embed research-led teaching into the curriculum, drawing on ideas and practices from both computational and social science fields such as Fairness, Accountability, Transparency and Ethics (FATE), and Critical Data Studies (CDS). This paper reports on one teaching team's reflections on the introduction of critical insights from these two fields into their data science curriculum. Through these reflections, we identify a series of challenges for deeper interdisciplinary integration and present recommendations for progressing it to enable effective cross-disciplinary critical FATE learning environments.

- **Review of tools for assessing data and algorithmic bias** by Kalia Orphanou, Maria Kasinidou, Frank Hopfgartner, Paul D Clough - Open University of Cyprus and The University of Sheffield

Audience: Developers

This White Paper provides a review of popular datasets and tools used for assessing and mitigating data and algorithmic bias. We describe the methodologies developed to identify relevant literature, datasets and tools, the results of our investigation, and how such tools could be integrated into the proposed CyCAT framework for mitigating bias.

- **Algorithm Watchdog** by Lena Podoletz, Michael Rovatsos, Veronika Bogina - The University of Edinburgh, University of Haifa

Audience: Policy, NGO, Regulators

This white paper provides our vision of an algorithmic watchdog - an independent, non-profit entity that provides a transparent service for investigations of citizens' complaints regarding harms caused by algorithmic systems, operates a public portal to track such disputes and their resolution, offers expert policy advice, publishes recommendations and guidelines based on their investigations and case studies, and builds a library of past cases to document and advance public debate regarding the societal implications of algorithms.

- **Tools for informal end-user education on algorithmic bias and transparency** by Lena Podoletz, Michael Rovatsos - The University of Edinburgh

Audience: Practitioners / Industry/Non-technical audience

People encounter algorithmic decision-making systems on a daily basis in their work, education and private lives. They do not only use them as users but also are subjected to algorithmic decisions in processes such as recruitment or credit scoring. Despite this, the majority of end users know very little about how these decisions are made. In this White Paper we evaluate how end-users can be educated on algorithmic decision-making processes in informal settings and propose steps towards creating such educational content in a way that is effective and results in increased awareness.

- **Fairness and Transparency Educational Requirements for All Stakeholders in Algorithmic Systems** by Veronika Bogina, Alan Hartman, Tsvi Kuflik, Avital Shulner-Tal - University of Haifa

Audience: All Stakeholders

This white paper discusses stakeholders of algorithmic systems – who they are and what they need to know. We begin by providing some background on algorithmic systems and the potential risks associated with their use. Such risks arise from the application of opaque reasoning mechanisms for decision making in real life situations. We continue by identifying the stakeholders of such systems – their developers, users, owners and regulators. Finally, we discuss what these stakeholders need to know in order to be aware of and avoid the potential risks of bias and discrimination that may be embedded in such systems.