Document Title	**User-focused solutions**
Project Title and acronym	Cyprus Center for Algorithmic Transparency (CyCAT)
H2020-WIDESPREAD-05-2017-Twinning	Grant Agreement number: 810105 — CyCAT
Deliverable No.	D4.5
Work package No.	WP4
Work package title	Promoting algorithmic transparency
Authors (Name and Partner Institution)	Lena Podoletz (UEDIN) Michael Rovatsos (UEDIN)
Contributors (Name and Partner Institution)	
Reviewers	Jo Bates (USFD) Maria Kasinidou (OUC)
Status (D: draft; RD: revised draft; F: final)	F
File Name	D4.5_User_focused_solutions_M18
Date	31 March 2020

Draft Versions - History of Document				
Version	**Date**	**Authors / contributors**	**e-mail address**	**Notes / changes**
v1.0	17/12/2019	Lena Podoletz	lena.podoletz@ed.ac.uk	Initial draft
v2.0	06/01/2020	Michael Rovatsos	michael.rovatsos@ed.ac.uk	Revised draft
v3.0	18/01/2020	Lena Podoletz	lena.podoletz@ed.ac.uk	Second revised draft
v4.0	01/03/2020	Michael Rovatsos, Lena Podoletz	michael.rovatsos@ed.ac.uk, lena.podoletz@ed.ac.uk	Final version

Abstract

This section examines user-focused solutions on algorithmic transparency, bias and fairness. We will explore the goals and desired outcomes of educating end users in these topics, concentrating on the core concepts defined in D4.1. In this deliverable, we will concentrate on informal forms of user education as more formal mechanisms are covered in D4.2. We will also define the requirements of tools and materials that aim to educate end users in the aforementioned topics. The document examines existing tools in this field and concludes with recommendations for improving informal user education regarding algorithmic transparency.

Keyword(s):	User-focused solutions, educating users, informal education, available solutions, challenges of user education, requirements of informal user education

Contents

1. Executive summary 5

2. Introduction 5

3. Who do we want to teach? 6

3.1. The challenges of educating users in algorithmic transparency, fairness and biases 7

3.2. Proposed solutions 14

4. What do we want to teach and how? 15

4.1. Existing forms of informal user education 15

4.2. Characteristics of existing user-focused solutions 30

4.3. User interface development: a way for developers to support end users in learning about algorithmic systems and building systems that explain how they work to the user and offer reasonings for decisions via user interfaces 31

4.4. What concepts are necessary for the end user to know? 32

4.5. Desirable learning outcomes 38

4.6. Requirements of user-focused solutions 42

5. Further tasks and conclusions 42

6. References 44

1. Executive summary

This section examines user-focused solutions on algorithmic transparency, bias and fairness. We will explore the goals and desired outcomes of educating end users in these topics, concentrating on the core concepts defined in D4.1. In this deliverable, we will concentrate on informal forms of user education as more formal mechanisms are covered in D4.2. We will also define the requirements of tools and materials that aim to educate end users in the aforementioned topics. The document examines existing tools in this field and concludes with recommendations for improving informal user education regarding algorithmic transparency.

2. Introduction

The topics of algorithmic bias and the potential individual and societal harms of algorithmic decision-making have slowly started to capture the interest of expert and non-expert audiences alike in the past few years.[1] It has also been reported that people feel concerned when it comes to the use of algorithmic systems in certain areas. For example, the Pew Research Centre has found that, in a survey conducted in the USA in 2018, that in the cases of criminal risk assessment for people up for parole, automated resume screening of job applicants, and automated video analysis of job interviews, 56%, 57%, and 67% of the respondents found the practice in question 'unacceptable', respectively.[2] The same study found that only 40% of the respondents believed that it is possible "for computer programmes to make decisions without human bias".[3]

Interestingly, it has also been reported that in specific situations (for example when calculating the base of bonus pay) sometimes people are more likely to rely on advice if they believe it came from and algorithm instead of another human, even when they only received a 'minimal description' of said algorithm (Logg *et.al.* 2019). Requirements such as being non-biased, non-discriminatory, and fair are now written into several guidelines and recommendations on trustworthy and ethical AI.[4]

[1] See for example: https://www.bbc.co.uk/news/technology-49717378, https://www.mtu.edu/magazine/2019-1/stories/algorithm-bias/, https://www.forbes.com/sites/bernardmarr/2019/01/29/3-steps-to-tackle-the-problem-of-bias-in-artificial-intelligence/#7719fb2a7a12, https://medium.com/mit-media-lab/the-algorithms-arent-biased-we-are-a691f5f6f6f2, https://unidir.org/sites/default/files/publication/pdfs/algorithmic-bias-and-the-weaponization-of-increasingly-autonomous-technologies-en-720.pdf, https://www.finextra.com/blogposting/17864/fair-ai-how-to-detect-and-remove-bias-from-financial-services-ai-models, https://www.anaconda.com/machine-learning-bias-fairness/

[2] https://www.pewresearch.org/internet/2018/11/16/public-attitudes-toward-computer-algorithms/

[3] https://www.pewresearch.org/internet/2018/11/16/attitudes-toward-algorithmic-decision-making/

[4] See for example: European Commission: Trustworthy AI – Joining for strategic leadership and societal prosperity (https://ec.europa.eu/digital-single-market/en/news/trustworthy-ai-brochure), AI Now Report 2018 (https://ainowinstitute.org/AI_Now_2018_Report.pdf), Amnesty International and Access Now: The Toronto Declaration – Protecting the right to equality and non-discrimination in machine learning systems (https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf), FAT/ML: Principles for Accountable Algorithms and a Social Impact Statement for Algorithms

Regardless of this, there have not been many papers or studies that concentrated on educating users on these topics in a more comprehensive way.

In this deliverable, we will explore user-focused educational solutions regarding algorithmic transparency. We will concentrate on the education of end users on algorithmic transparency, bias, fairness and the real-life implications these phenomena have on the users of the systems.

When talking about education, the most important questions to ask are:

1. Who do we want to teach and why?
2. What do we want to teach and how?

These questions are important when we are determining what form of teaching to use, how to support people in learning, and what the content of the teaching materials should be.

First and foremost, we need to determine our audience, i.e. the people we wish to educate. This is important because a lot depends on this factor, starting with the style of teaching, the complexity of concepts we want to get across and the level of language used. This will also aid us in establishing the particular needs of the group in question and the difficulties that come with their education.

It is also essential to determine what the aims we wish to achieve by educating that particular group, the motivations for the educational activity, and the desired learning outcomes are. Therefore, the next step is to establish the teaching topics and the academic content of our educational materials.

We will do this by reviewing existing informal educational content regarding algorithmic bias, identifying their common characteristics and using the core concepts from deliverables D4.1 and D4.2. We will also identify the key requirements for informal educational materials and offer some general methods for measuring learning outcomes.

3. Who do we want to teach?

The group we are concentrating on in this work is end users. First, we will explore the challenges of end user education, then we will determine the desirable learning outcomes of it. It needs to be noted here that we are talking about end users, namely people who use the algorithmic systems themselves. In the case of most systems, however, there are people who cannot be qualified as ‘users’ per se as they do not use the system even if they can still suffer from certain types of system biases.

(https://www.fatml.org/resources/principles-for-accountable-algorithms), EU High-level Expert Group on AI: Ethics Guidelines for Trustworthy AI (https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai), IBM: Everyday Ethics for Artificial Intelligence (https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf), Microsoft: AI Principles (https://www.microsoft.com/en-us/AI/our-approach-to-ai)

3.1. The challenges of educating users in algorithmic transparency, fairness and biases

a) Heterogeneity of end users as a group

One of the main challenges in the education of users is that they do not constitute a homogenous group. They have different existing technical knowledge, different past experiences and have very different needs depending on what type of systems they use and what reason they are using it for (e.g. private use for entertainment, commercial use or the use of algorithmic systems in the delivery of public services such as crime control). We will use the specific group of senior people as an example to demonstrate the existing digital skills gap in society and to illustrate the potentially different needs of groups in society.

Senior people qualify as a vulnerable group when it comes to the threats that can result from the use of technology. It has been reported in the literature that they, on average, usually have a lower level of understanding and working knowledge when it comes to technology, especially in relation to new technologies and applications (Peacock *et.al.* 2007:191, Roupa *et.al.* 2010, Garattini et.al. 2015:9, Vacek *et.al.* 2017:758, Vaportzis *et.al.* 2017). Thus, their user experiences are different from younger users'. While attempting to use certain new devices (e.g. tablets), amongst other negative feelings they can experience lack of confidence, inability to learn how to operate the device due to lack of clear instructions and guidance, feeling of inadequacy, and – possibly as a result of the aforementioned factors – scepticism about using technology in general (Vaportzis *et.al.* 2017). We can theorize that the same is true when it comes to using algorithmic decision-making systems as they share many features with new devices when it comes to the difficulties first-time users face, such as the lack of guidance or the complexity of the system. This can result in senior people rejecting the use of such technology. However more and more senior people are now adapting to using technological devices and applications due to needs regarding communication, health, wellbeing and general lifestyle. This development is slowly closing the digital skills gap between younger and senior people but at the same time, it has the potential to generate a vulnerable user group that only knows the basics of applying the technology but does not understand the risks and dangers that come with it. The general scepticism many of them feel when it comes to new technologies can affect their education about algorithmic systems because the perceived complexity and rapidly changing nature of the topic can deter them from attempting to learn more about it, even when they use the systems themselves.

The example above shows how the needs for one group in society can differ from that of another. It is also true, however, that it would be a mistake to sort all people sharing one particular characteristic – in this case, age – into the same group and claim they all have similar needs. For this reason, we suggest a three-tier approach towards informal education regarding algorithmic transparency:

1. General level education. It seems reasonable that education on certain key topics – such as the concept of bias or algorithms - should happen on a general level from which all users may benefit.

2. Special content created for specific groups. Even though we cannot state that every person in a specific group has the same level of knowledge, abilities or past experiences regarding technology and the use of algorithmic systems, it can be beneficial to also create content that targets a particular group in society (for example young adults or senior people). Here we also have to consider that these groups and the individuals in any particular group may have very different experiences regarding issues such as bias and discrimination in society.

3. Educational content created for users of particular types of algorithmic systems or one specific system. Almost everybody who has access to technology or online content uses some type of algorithmic system, but some people do so much less frequently than others. People who use a particular system may not have similar levels of background knowledge or IT skills but when it comes to the internal mechanisms of an algorithmic system, it can be beneficial for all to receive explanation during the course of human-computer interaction. For this reason, it seems logical to put significant efforts into the development of systems that offer this type of guidance.

b) Most end users have a very minimal understanding of IT

The 2015 OECD Programme for International Student Assessment (PISA) showed that "one in five pupils in the EU has insufficient proficiency in reading, mathematics or science".[5] Also, in the EU-28 countries the share of households with internet access reached 89% in 2018, and "the proportion of individuals aged 16-74 (…) who ordered or bought goods or services over the internet for private use" was 60%.[6] These statistics suggest that most likely a very high percentage of the population in question have used or are actively using information access systems. Despite this high level of online activity, according to the European Commission, more than 44% of Europeans between the age of 16 and 74 did not have basic digital skills in 2017.[7] Moreover, a study published in early 2019 reported that in the EU "79% of lower secondary school students and 76% of upper secondary school students never or almost never engage in coding or programming at school" and "more than 4 out of 5 female European students attending secondary schools never or almost never engage in coding at school".[8] A study conducted by the Pew Research Centre instructed participants to view their Facebook 'ad preferences' page and

[5] https://ec.europa.eu/education/policies/school/key-competences-and-basic-skills_en

[6] Digital economy and society statistics – households and individuals (https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals)

[7] The Digital Skills Gap in Europe (https://ec.europa.eu/digital-single-market/en/news/digital-skills-gap-europe)

[8] 2nd Survey of Schools: ICT in Education (https://ec.europa.eu/digital-single-market/en/news/2nd-survey-schools-ict-education)

showed that 74% of the participants 'did not know Facebook maintained this list of their interests and traits, and 50% of the participants were 'not comfortable' with it.[9] However, another survey on US adult social media users undertaken by the same institute suggested that the majority of participants believed it to be very easy or at least somewhat easy for social media sites to figure out characteristics such as race (84%), hobbies and interests (79%), political affiliation (71%) and religious beliefs (65%).[10] The latter result suggests at least some degree of awareness when it comes to data and how it can be used.

The 'Essential Digital Skills Framework' in the United Kingdom published in September 2018 and updated in April 2019 distinguishes between five groups of essential skills: 1. Communication, 2. Handling Information and Content, 3. Transacting, 4. Problem Solving, 5. Being Safe and Legal Online.[11] The individual, concrete skills listed in each group cover a great deal of online and offline digital activities, but the document does not mention skills such as awareness of potential bias and discrimination. The framework does refer to certain skills that are connected to these issues in a broader sense (for example the ability to evaluate the reliability of online content, the ability to make use of search terms to generate better search results or the understanding of the more-or less permanent nature of any online activity) but overall the problem of algorithmic bias and the question of transparency are missing from this government framework.

According to the 'Digital Economy and Society Index (DESI)', in 2019 Cyprus ranked 22nd of 28 EU Member States.[12] In 2018 Cyprus' ranking was 21st.[13] The index takes five dimensions into account: 1. Connectivity, 2. Human Capital, 3. Use of Internet Services, 4. Integration of Digital Technology, 5. Digital Public Services. Between 2018 and 2019 the country's performance improved in all areas, except for the dimension of 'Human capital' where they performed less well in 2019.[14] Despite the increased performance index Cyprus still remains under the EU average in all five areas. In the 'Human capital' dimension the country was ranked 24th amongst all EU countries.[15] According to the findings described in the report "almost a sixth of Cypriots have never used the internet and half lack basic digital skills".[16] It is also important to note that ICT specialists are 2.3% of the workforce and this is below the

[9] https://www.pewresearch.org/internet/2019/01/16/facebook-algorithms-and-personal-data/
[10] https://www.pewresearch.org/internet/2019/01/16/facebook-algorithms-and-personal-data/
[11] UK Government: Essential Digital Skills Framework (https://www.gov.uk/government/publications/essential-digital-skills-framework)
[12] https://ec.europa.eu/digital-single-market/en/desi
[13] https://ec.europa.eu/information_society/newsroom/image/document/2018-20/cy-desi_2018-country-profile_eng_B43F6E93-DC41-A4D3-6FEDC85F4EC8246B_52217.pdf
[14] Digital Economy and Society Index (DESI), 2019 Country Report, Cyprus, p3 (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)
[15] Digital Economy and Society Index (DESI), 2019 Country Report, Cyprus, p7 (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)
[16] Digital Economy and Society Index (DESI), 2019 Country Report, Cyprus, p3 (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)

EU average (3.7%).[17] The proportion of female ICT specialists ("expressed as a share of total female employment") is only 0.7% which is also under the EU average which is 1.4% in 2019.[18] When it comes to using the internet, according to the findings Cypriots are far below the EU average in the areas such as online banking (39% in Cyprus, 64% in EU), online shopping (38% in Cyprus, 69% in EU) or selling items online (3% in Cyprus, 23% in EU). However, they perform much better than the EU average in areas like making video calls (74% in Cyprus, 49% in EU) and the use of social networks (82% in Cyprus, 65% in EU). Regarding the use of social network the findings show that even though the use of these platforms is more common than in the EU average, when it comes to professional social networks Cyprus has a lower performance than the EU average (11% in Cyprus, 15% in EU).[19] Within the dimension of 'Integration of digital technology' Cyprus is performing better than the EU average in the social media use of enterprises (37% in Cyprus, 21% in EU) and the electronic information sharing of enterprises (35% in Cyprus, 34% in EU) but they underperform in e-commerce (6% in Cyprus, 10% in EU).[20] The 'Digital public services domain is one where Cyprus performs better than in other dimensions as their ranking in this area is the 19th among all EU countries. However, this is still below EU average and the data shows that the performance is on a slight decline since 2017.[21]

The 'Women in Digital Scoreboard 2019', where Cyprus is ranked 22nd amongst EU Member States, shows that while the country is above the EU average in regular internet use of women (84% in Cyprus, 82% in EU), when it comes to female users, the country underperforms in key areas such as online baking (36% in Cyprus, 63% in EU), using professional networks online (9% in Cyprus, 13% in EU), doing an online course (5.5% in Cyprus, 8.1% in EU) or having above basic digital skills (19% in Cyprus, 28% in EU).[22]

The '2018 Global Digital IQ Survey'[23] conducted by PricewaterhouseCoopers found that in the digital era the majority of CEO-s from Cyprus, who participated in the survey "introduced a digital strategy in their business plans"(80%) and also "encourage their workforce to take initiatives" in that area (67%). At the same time, the survey also found that they identified "deficiencies in their employees'" basic digital skill, with only 47% of them stating that "their labour force is familiar with technology" and

[17] Digital Economy and Society Index (DESI), 2019 Country Report, Cyprus, p7 (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)
[18] Digital Economy and Society Index (DESI), 2019 Country Report, Cyprus, p7 (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)
[19] Digital Economy and Society Index (DESI), 2019 Country Report, Cyprus, p9 (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)
[20] Digital Economy and Society Index (DESI), 2019 Country Report, Cyprus, p10 (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)
[21] Digital Economy and Society Index (DESI), 2019 Country Report, Cyprus, p12 (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)
[22] Women in Digital 2019 Scoreboard (https://ec.europa.eu/digital-single-market/en/scoreboard/cyprus)
[23] https://www.pwc.com.cy/en/press-releases/press-releases-2019/deficiencies-in-the-basic-digital-skills-of-the-cypriot-workforce.html

with 37% of them reporting that "they have changed hiring processes in order to achieve a greater level of familiarity with technology in their workforce". As a comparison, the former percentage was 60% and the latter was 66% on average in all participating countries.[24]

Even though for all intents and purposes here, even developers can qualify as end users when it comes to the systems they use in their daily lives, it seems safe to state that the average end user of any algorithmic system possesses only the minimum level of working knowledge when it come to the system they use (for example what type of input data is required, what the required method of data input in, where and in what format the output data will be accessible for the user). In order to acquire a sufficient level of understanding of algorithmic decision-making systems and how embedded bias and discrimination work, users first need to possess basic digital literacy skills. Basic digital literacy skills are the knowledge one needs to function in today's digital society. These skills include core hardware and software skills and the ability to perform simple, everyday online tasks (for example handling a touch-screen device, sending an email or setting the privacy settings on their devices).[25] These abilities however are not sufficient for empowerment regarding the harms and threats that can arise from algorithmic bias. End users need to have certain intermediate-level digital skills to be able to counter some of the effects of these biases. The reason for this is because intermediate level digital skills are the ones that can give people the ability to "critically evaluate technology or create content".[26] These are the skills that can give a user deeper understanding regarding how a particular system works and how system bias is generated. While it is true that more and more people are acquiring this type of knowledge, it still can be stated that a high percentage of users do not have adequate understanding of algorithmic systems.

The findings described above suggest that the education of users should start with introducing the very basic concepts of algorithmic decision-making, such as algorithm, input data, output data and how these relate to each other in that particular system.

c) Opacity of algorithmic system

As has been reported by many researchers (see for example Kroll *et al* 2016, Paudyal *et al* 2018, Goldenfein 2019),[27] most algorithmic systems have a very high level of opacity, whether these be risk

[24] https://www.pwc.com.cy/en/press-releases/press-releases-2019/deficiencies-in-the-basic-digital-skills-of-the-cypriot-workforce.html
[25] ITU: Digital Skills Toolkit, p6 (https://www.itu.int/en/ITU-D/Digital-Inclusion/Documents/ITU%20Digital%20Skills%20Toolkit.pdf)
[26] ITU: Digital Skills Toolkit, p7 (https://www.itu.int/en/ITU-D/Digital-Inclusion/Documents/ITU%20Digital%20Skills%20Toolkit.pdf)
[27] See also European Parliament: A governance framework for algorithmic accountability and transparency (https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf), Algorithmic Fairness and Opacity Working Group (https://afog.berkeley.edu/), epic.org: Algorithmic Transparency: End Secret Profiling (https://epic.org/algorithmic-transparency/)

scoring systems (Tan *et al* 2018) or algorithms employed in news media in the dissemination of stories (Diakopoulos *et al* 2016).

Opacity makes user-focused education significantly more difficult. It results in user-focused solutions not being able to explain certain aspects of a particular system and only distributing information on biased output and how the use of algorithmic systems in general can lead to discriminatory results and other individual and societal harms. Therefore, when it comes to user education, an important piece of information to pass on to end users is that some of the decisions made by these systems are unpredictable even for experts or for the developers themselves. So essentially the opacity makes it very difficult to provide anything else but a black-box explanation for users without the cooperation of the developers, copyright owners, etc. Naturally, when talking about non-expert end users it is worth discussing whether there is even a need to present them with white-box explanations. Chang *et al* (2019:7-8) found of a non-representative sample of users that while white-box explanations do not provide a greater level of self-reported understanding of the algorithm compared to black-box explanations, they did help to reach a higher level of objective knowledge about them. The authors noted that one explanation for the similar levels of self-reported understanding could have been due to the more complex nature of white-box explanations where the users only perceived they understood them less since their objective knowledge was indeed higher (Chang *et al* 2019:9). This means that there should be greater emphasis on working together with these stakeholders and on developer education regarding how they can make their algorithmic systems more transparent and intelligible.

Explainability has indeed been an important part of work on ethical on trustworthy AI,[28] and is considered a crucial requirement for the future development of these systems. Another aspect of how opacity makes user education more difficult is that the opaque nature of these systems also has a risk of making users dubious toward whether it is possible at all for them to understand algorithmic decision-making.

Opacity does not only make the understanding and therefore education about algorithmic decision-making significantly more difficult; it also contributes to other practical problems such as the appealability of the decisions made by these systems. As one motivation for user education is to enable them to identify where and when there might have been a violation of their rights (for examples in cases

[28] See for example: Access Now: Human Rights on the Age of Artificial Intelligence (https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf), AI4People: AI4People's Ethical Framework for a Good AI Society: Opportunities, Risks, Principles and Recommendations (https://www.eismd.eu/ai4people-ethical-framework/), IBM: IBM's Principles for Trust and Transparency (https://www.ibm.com/blogs/policy/trust-principles/), Amnesty International and Access Now: The Toronto Declaration – Protecting the right to equality and non-discrimination in machine learning systems (https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf)

of discrimination), it seems crucial to be able to provide more detailed, white-box explanations of these systems.

d) Fairness is relative

The concept of fairness and fair decisions is relative not only to the specific context the decision was made in, but also to each person or group affected by it. The particular definition of fairness applied in a system can and most likely will depend on the particular system in question. This makes education about fairness very difficult as it seems impossible to create educational content regarding fairness in general when it comes to algorithmic systems. As stated before, fairness can be relative to the perspective of the individual or group affected by the decision in question, so at this point in time it is also very difficult to educate people regarding what is fair in the context of a particular algorithmic system.

For this reason, we suggest that fairness education takes this into account and raises awareness regarding this precise issue, namely that fairness can be a very relative concept that is dependent on context, perspective, space and time and that whether a system is fair can change with the changing circumstances within which it is deployed and used. At the same time, fairness education should also cover examples that can qualify as discriminatory practice to make users realise how their choices – even with the best intentions – can lead to discrimination of others (for example in the case of a recruitment algorithm) and how they can recognise when their rights have been violated.

e) Conflict of interest between transparency and intellectual property rights

The potential conflict of interest between transparency and intellectual property rights poses one of the biggest challenges for algorithmic transparency as trade secrecy and copyrights, for example, are deeply embedded and highly regarded values in the industrial, commercial and creative sectors. Technology companies may be reluctant to disclose their proprietary algorithmic techniques, especially if these have been costly to develop and/or provide them with competitive advantage over their rivals. Naturally, it is possible to explain the basic operations of algorithms and algorithmic systems without being able to see inside the black box but raising awareness of the specific biases and risks that arise from the use of a particular system can become increasingly difficult.

It is safe to predict that any regulation or court ruling that ends in the mandatory declaration of the exact mechanisms of algorithmic decision-making systems can lead to negative consequences. For one, any such ruling or regulation will unavoidably meet backlash from the concerned industries as it could threaten their most valuable assets, unique innovations and applications. Secondly, it can lead to very serious concerns regarding the security and safety of systems and the privacy of their users. It also worth pointing out that when it comes to artificial intelligence techniques such as deep learning providing a

full reasoning and explanation for decisions is not always possible. Transparent algorithms can also lead to the unintended consequence of users 'gaming the system'.

In the case of information access systems, it is not only the case that the value of a particular system lies in the unique decision-making system but, in many cases, it comes from the information the system learns about the specific user in order to present them with more relevant content.

To resolve this problem, we suggest a form of a partial 'transparency-by-design' approach where the user receives at least a black-box type explanation during the use of the programme. By this we mean that in the cases when it is impossible to provide a white-box explanation to the users where they can see and understand exactly how the given algorithmic system works, the developers should aim for an explanation that makes it clear to users what the exact relationship between input and output is and how differences in input may influence the output.

3.2. Proposed solutions

Based on the above considerations, we propose the following general solutions to the challenges:

CHALLENGE	PROPOSED SOLUTION
End users are a heterogenous group	1. General education is provided on a basic level which is intelligible for all end users 2. Specific education concentrates on a specific system and its mechanisms
Lack of IT knowledge	1. User education is provided in a way that includes information regarding the basics of computing (for example what is an algorithm) 2. Education should be delivered in a way which is more practical than theoretical, and includes real-life examples and visualisations

Opacity of algorithmic systems	Promoting the use of 'transparency-by-design' elements which can help users to: 1. Understand the consequences of their actions while interacting with the system 2. Tailor the system more to their actual needs
Fairness is relative	1. User education should cover the fact that fairness is relative 2. It should also illustrate with examples what can qualify as discrimination and what steps are available in case of potential violation of rights
Conflict of interest between transparency and intellectual property rights	1. Regulation or judicial order mandating full transparency of algorithmic systems and AI may not lead to the desired results therefore solutions need to be invented for cases where full transparency is not possible 2. A partial 'transparency-by-design' approach that gives at least a black-box explanation to the users of the system

4. What do we want to teach and how?

4.1. Existing forms of informal user education

In the following subsection we are going to describe a few, selected means and examples of informal end user education. The aim is to identify what methods and material are already available for this group and what tools and ways are employed in order to raise awareness and share relevant information with users. Some of the examples mentioned below do not have a sole focus on algorithmic bias but also

explore related issues like privacy and data protection. Despite this we thought it necessary to include these as well so that we could get a more comprehensive picture of available tools and all possible methods of user-focused solutions.

4.1.1. Exhibitions

Exhibitions, especially of an interactive kind, can serve as a very innovative and useful way to educate users on many different topics. They can be a very good form of presenting real-life scientific problems and making them more relatable to the visitors, and can also induce discussions and debates. Art installations and hands-on experiments can make it easier and more relatable for users to visualise the scientific problems and otherwise very complex questions that the exhibition pieces represent. In case of problematic issues in the field of data science and artificial intelligence it is particularly important to transform the problems embedded in abstract concepts such as transparency, fairness and bias to a language non-expert user find comprehensible. It has been reported that non-interactive exhibitions may only assist in 'superficial learning' whereas interactive, 'hands-on' exhibitions are more helpful in gathering deeper understanding and meaningful knowledge, especially if they are accompanied by a debate-session or some other discussion-based event (Groundwater-Smith *et.al.* 2003, Allen 2002).

Example 1. The Glass Room

One such existing exhibition is 'The Glass Room' exhibition sponsored by Firefox and 1Password amongst others which operates as a pop-up exhibition and has been on tour in major cities including London and San Francisco in 2019.[29] The exhibition focuses on educating end users on the impacts of some of the most common technological tools in day-to-day life (e.g. face recognition and search histories) and examines the possible solutions users can apply to mitigate these risks and negative effects (e.g. password protection). The exhibition consists of a selection of installations that help users to visualise and understand the underlying implications and hidden consequences of the sharing of personal data, the usage of social media, dating apps and search engines. The exhibition not only makes it possible for non-expert audiences to get a clearer picture of the – mostly unintended – results of their online and in cases of many applications offline actions but also offers easily comprehensible and accomplishable actions to take in order to mitigate the risks and harms they are being subjected to.[30]

One example of the installations was the 'Quick Fix' vending machine (see Figure 1.) which demonstrated the ease of gaining 'fake' likes and followers on social media platforms through the use of auto-generated accounts or hired commenters via paid suppliers.[31] Another such exhibited apparatus was the 'White Collar Crime Risk Zones' display (see Figure 2.) which was meant to show that while

[29] https://theglassroom.org/
[30] Data Detox Kit (https://datadetoxkit.org/en/home)
[31] https://theglassroom.org/object/dries_depoorter-quick_fix

the police and predictive policing tools often concentrate on violent crimes, they tend to ignore certain types of offences such as white collar crimes.[32] Both of these pieces are good examples of ways of raising awareness of the biased nature of data.

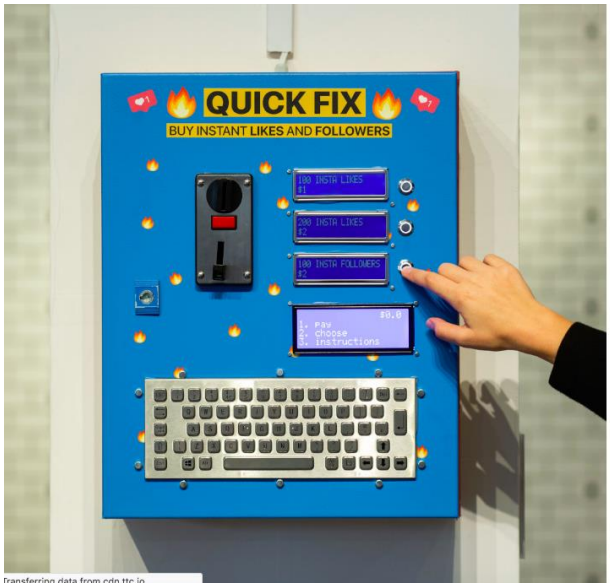


Image credits: Boris Zharkov

Dries Depoorter @driesdepoorter

Quick Fix

If you wish you had more followers on Instagram, Twitter, Facebook, or YouTube, Quick Fix lets you buy followers or likes in just a few seconds.

Choose your product, insert a coin, and fill in your social media user-name. You'll receive the likes or followers just a few seconds later. (The accounts that like or follow you are auto-generated accounts.)

Your order is saved in a database with the location of the exhibition, date, city, and country where it was purchased.

In an age where the number of followers you have can translate to how much influence you have, what price are you willing to pay for that status? And if those followers might be 'fake', what is it really worth?

Figure 1. Quick Fix[33]

Image credits: Boris Zharkov

Sam Lavigne Brian Clifton Francis Tseng @sam_lavigne @BrianClifton_ @frnsys

White Collar Crime Risk Zones

The news is full of reports about where violent crimes have taken place, but we rarely hear about which neighborhoods are experiencing the most tax fraud or embezzlement. White Collar Crime Risk Zones aims to make those areas visible by predicting where financial crimes are most likely to occur across the U.S.
Predictive policing systems used by law enforcement rely on data about past crimes to help predict where future crimes may occur. These predictive systems are based on police data and may serve to reinforce the existing biases of police departments. White Collar Crime Risk Zones uses machine learning to invert those systems, to find out which neighborhoods are considered risky when you focus on financial crimes. The video you see here is a tour of zones in the U.S. The faces are composite images based on LinkedIn profiles of top executives in the area. What are the consequences of policing based on one data-set of crimes, while excluding another?
See more information or download the app at:
https://whitecollar.thenewinquiry.com

Figure 2. White Collar Crime Risk Zones[34]

In the first case non-expert users were able to familiarise themselves with the fact that the amount of likes or followers on social media does not necessarily mirror the actual popularity of something. Users

[32] https://theglassroom.org/object/lavigne-clifton-tseng-white_collar_crime_risk_zones
[33] https://theglassroom.org/object/dries_depoorter-quick_fix
[34] https://theglassroom.org/object/lavigne-clifton-tseng-white_collar_crime_risk_zones

can thus learn that when they see certain seemingly popular opinions/posts/individuals appearing on their feed which are based on system-bias that preferentiates content with high levels of engagement, the popularity data may not be genuine, with the consequence that the algorithm may expose them to a highly biased composition of the online world.

In the case of the crime mapping algorithm, people and even expert users of such algorithms like police officers can be made aware of the fact that these systems are biased to a very high degree as to which types of crimes they show and are also based on statistical data that has long been proven to be biased in various ways (see for example Skogan 1977, Quinney 1970, McClintock 1970, Young 2004, Tombs 2014).

Example 2. Big Bang Data

The 'Big Bang Data' exhibition was on display in Barcelona in 2014.[35] The installations that were featured were built around the theme of introducing the concept of big data and the ways people – sometimes unintentionally – share data in their everyday lives through the use of devices such as mobile phones or sensors and through activities like online purchases or social networking. The intention was to raise awareness to these ways data can be shared that are often unknown for the non-expert users and also to bring to their attention the implications of having these multiple different sorts of data available for companies and other people to use. The science-based art projects were aiming for the visualisation of the potential effects and dangers that are embedded in data sharing and also to show how our data can be collected and used for different purposes. Apart from installations, the exhibition space also hosted events such as thematic workshops, hackathons, educational programmes and meetups.[36]

4.1.2. Interactive content online and offline (e.g. games, seminars)

Another approach to support users learning about how algorithms work is to create interactive content for them to use and discover either by themselves or in a group. Such interactive content can exist in both digital and analogue forms. Examples of interactive content can include games, quizzes, or guidelines for roleplaying of relevant situations, etc.

*Example 1. UnBias Awareness Cards (*http://unbias.wp.horizon.ac.uk/fairness-toolkit)

The 'UnBias Awareness' playing cards (see Figure 3a and 3b), which were created in 2018, focus on introducing the most important concepts relating to biases that are present in algorithmic systems and how unfairness can rise from these biases. The cards can help stimulate different discussions on these topics and through those 'help develop key skills, such as critical thinking about how decisions come

[35] https://www.cccb.org/en/exhibitions/file/big-bang-data/45167, https://www.inexhibit.com/case-studies/big-bang-data-exhibition-cccb/
[36] https://www.cccb.org/en/exhibitions/file/big-bang-data/45167

to be made'.[37] As well as explaining basic concepts like human rights and data, the cards also use examples and questions to ask questions regarding hypothetical situations (e.g. in relation to algorithms that are used in recruitment). They also offer exercises to act out. The cards seem to be a very practical tool to use for example in a classroom setting, educational events or in a scientific/educational exhibition. They are created in a way that is comprehensible for an average user, probably above the age of early/mid-teens.

Figure 3a. UnBias Awareness Cards, examples

[37] UnBias Awareness Cards: http://unbias.wp.horizon.ac.uk/fairness-toolkit

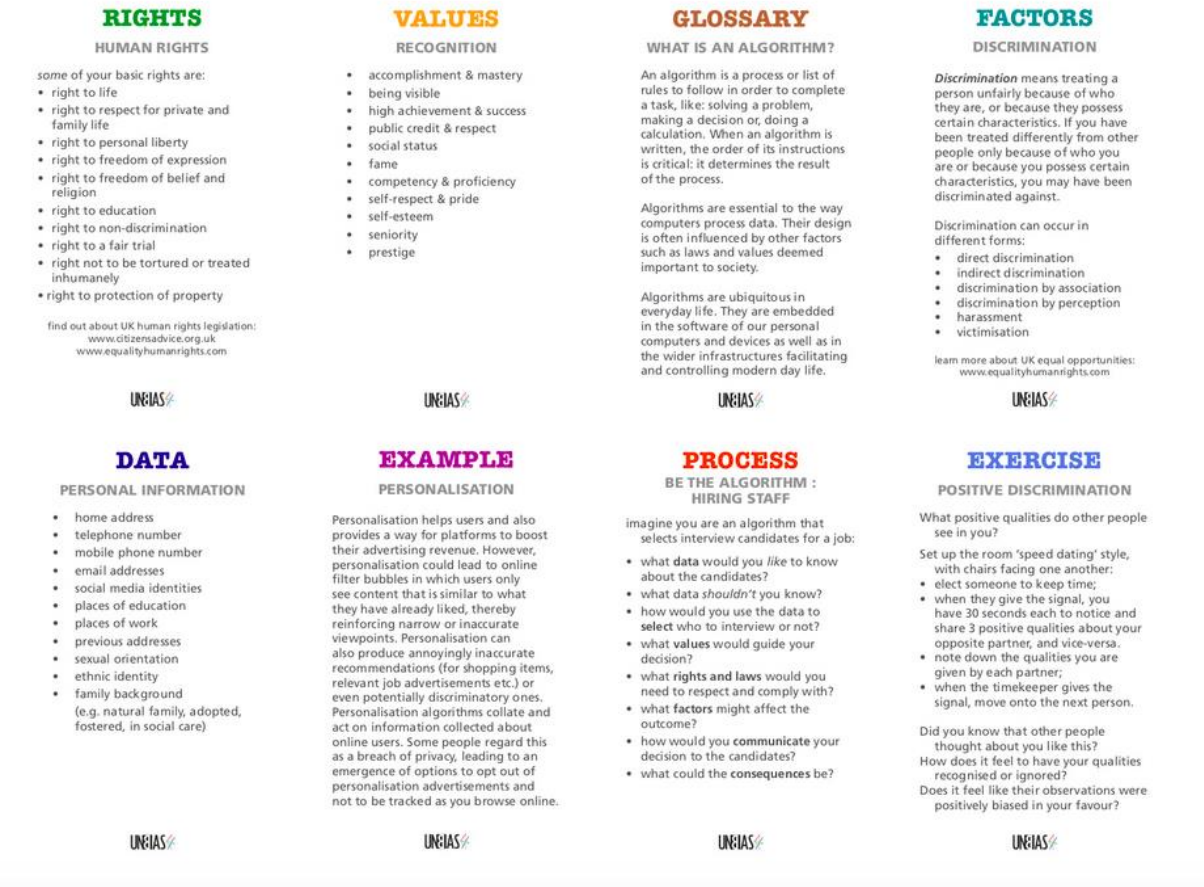


Figure 3b. UnBias Awareness Cards, examples

Example 2. 'Survival of the best fit' game (https://www.survivalofthebestfit.com/)

Another example of online interactive content is the game called 'Survival of the best fit' (see Figure 4a and 4b). It is designed to show people how biases find their way into the automation of recruitment. The game was developed by student alumni of NYU Abu Dhabi, a group of software engineers, designers and technologists.[38]

In the game the player is asked to hire a certain number of employees about whom they know certain characteristics, such as their level of skill, work experience, ambition and the prestige of their place of education. Each fictional applicant is also assigned a colour, orange or blue. First, the player has to make the hiring decisions themselves which is made more difficult by the very short time limit given to complete each phase of hiring. Then the task then gets automated based on the previous decisions of the player (e.g. if their hiring-rejection pattern shows that they hold skill in higher regard than ambition, then the recruiting algorithm will do the same and select similar candidates). Then the game quickly

[38] https://www.survivalofthebestfit.com/about

demonstrated how such hiring methods, even with the best intentions can turn into a discriminatory practice. One of the examples used by the game is if the blue candidates are underrepresented in the beginning of the process where the training data comes from, this can lead to creating and reproducing a pattern where proportionately fewer blue people will be hired regardless of their potentially high levels of skill or work experience.

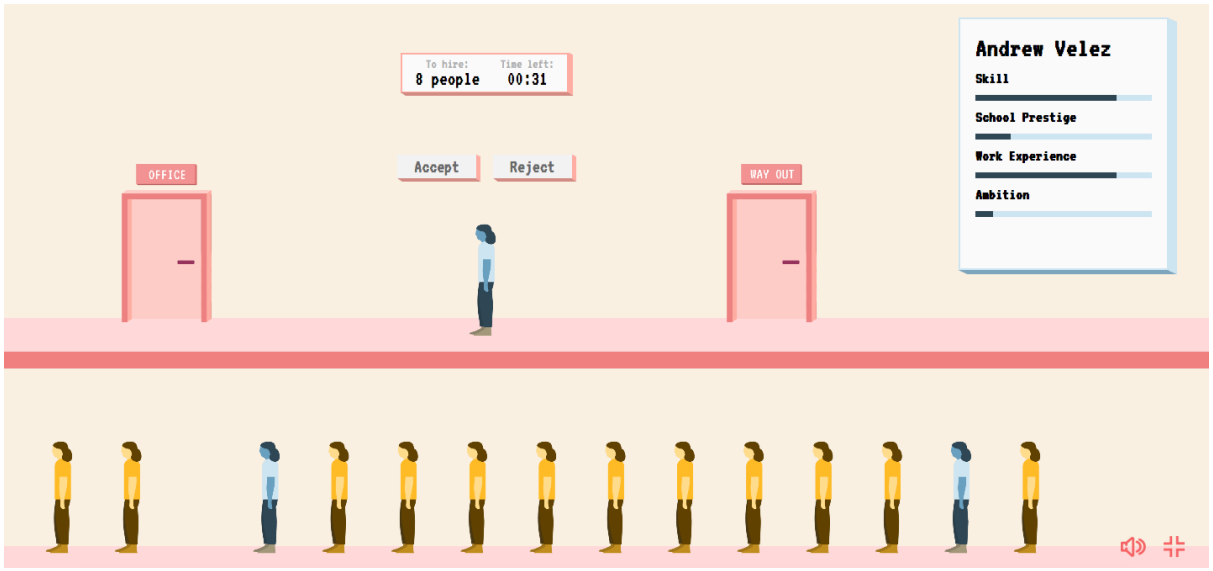


Figure 4a. 'Survival of the best fit game': The player chooses applicants based on their level of skill, etc

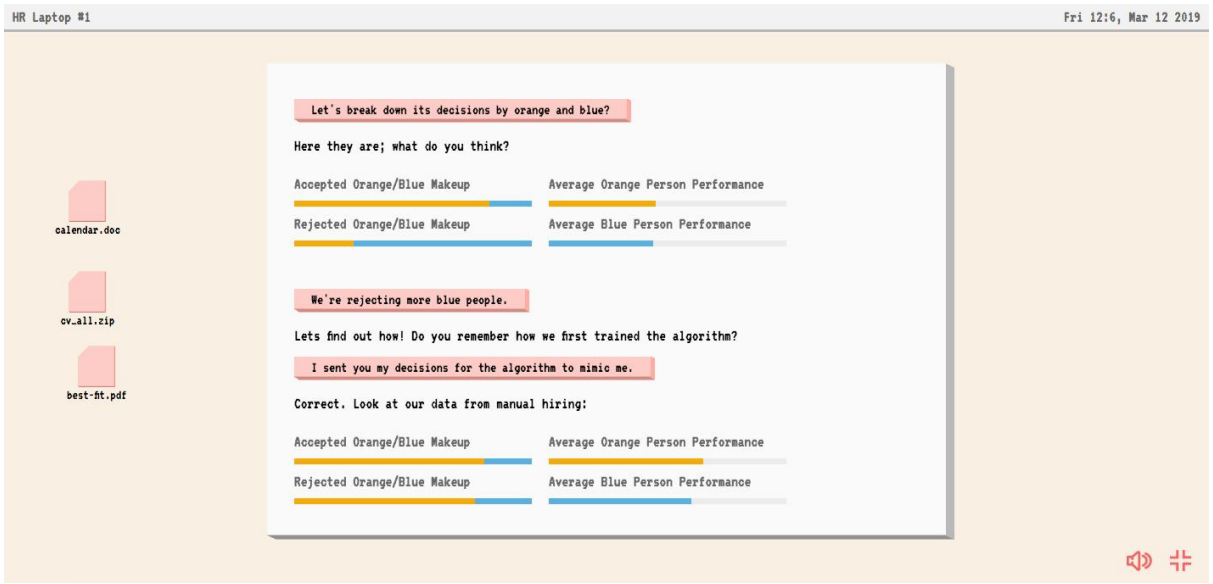


Figure 4b. 'Survival of the best fit game': The game shows the player the 'colour breakdown' of their chosen team

The game then offers a more detailed explanation on 'what went wrong' in the hiring process, what 'training' means in this context, and how previous employment practices can lead to biased data and discriminate, even if the CVs of the applicants do not contain sensitive data such as race or gender. The game's website also links to many scientific reports and studies on the subject for those who are interested in further information (https://www.survivalofthebestfit.com/resources).

Example 3. 'AI fairer than a judge' (https://www.technologyreview.com/s/613508/ai-fairer-than-judge-criminal-risk-assessment-algorithm/)

This 'courtroom algorithm game' (see Figure 5a and 5b) examines the risk assessment tool called COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) which is an algorithmic tool used in multiple jurisdictions in the USA to assess the likelihood of recidivism in an individual offender. The assessment is based on characteristics such as current charges, criminal history, family criminality, and indicators of non-compliance. The use of COMPAS has been critically assessed by experts in crime and penology and its reliability is debated regarding the ability to predict the chances of reoffending (see for example Brennan *et.al.* 2009, Blomberg *et.al.* 2010, Flores *et.al.* 2017, Kehl *et.al.* 2017, Dressel *et.al.* 2018, ProPublica: Machine Bias[39]).

The article that presents the game to the player briefly explains the function of COMPAS and points out crucial problems, for example that even though race is not included as a factor in the risk-score calculation, the algorithm still seems to be biased against black people. In the game, the task of the player is to 'make COMPAS better'.

[39] ProPublica: Machine Bias: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

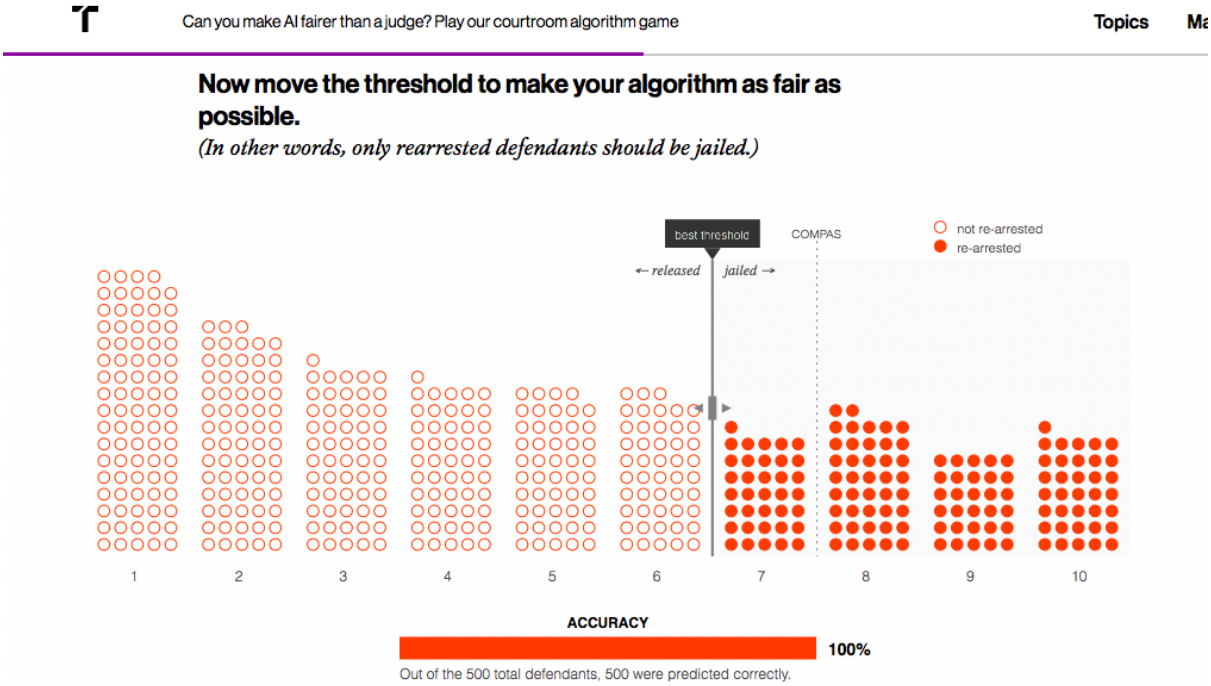


Figure 5a. *'AI fairer than a judge'*

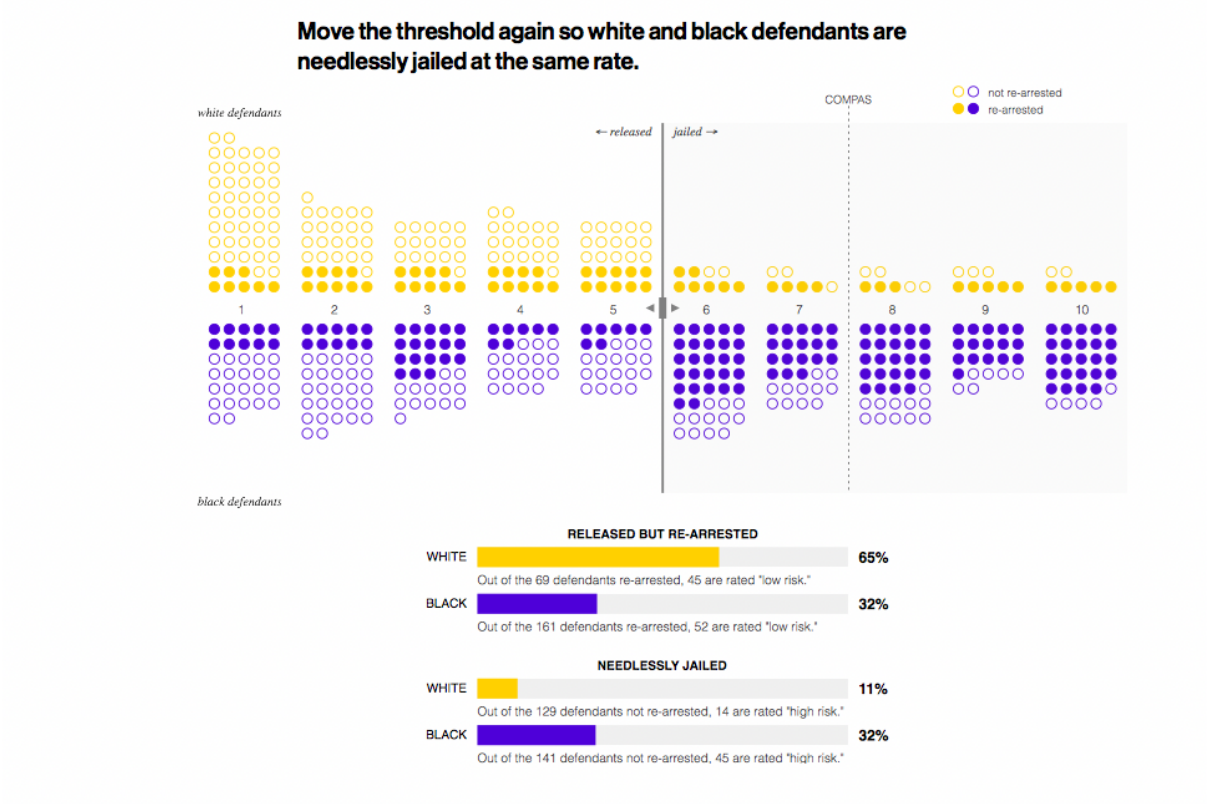


Figure 5b. *'AI fairer than a judge'*

The game and article present the topic in a mostly comprehensible way, but it fails to disclose certain pieces of information that would be necessary for the player to understand more about the algorithm and the issue in question. For example, it never explains how COMPAS works exactly and how it makes its predictions therefore the players only know it is an algorithm that somehow predicts the chances of re-offending in case of individual offenders. It also does not give any information on what 'high-risk' or 'low-risk' mean in this context and what characteristics or circumstances make an offender qualify as 'high-risk' in terms of re-offending. This results in playing the game according to the instructions given by the creators but not getting a full picture of why the player has to perform an in-game task a certain way.

Example 4. 'Monster Match' game (https://monstermatch.hiddenswitch.com/)

The online mini-game 'Monster Match' (see Figures 6a, 6b, 6c and 6d) simulates a typical dating app to illustrate the player how they work. The website also offers some background information on online dating and dating apps based on studies and links to several reports and papers on the topic.[40]

In this game, the player first has to create a dating profile which asks for information that is common to dating apps or websites, such as a picture (in this case it is of a 'monster' put together by the player) and an introduction. Then the player needs to start training the app by accepting or rejecting each potential match. When a match is accepted, a chat is simulated between the player's character and the matched monster in which the player can choose between lines offered by the game.

During the gameplay, the game educates users on the things that they experience and also what is happening in the background that they might not know about, for example the fact that the app is 'reading' their chat conversations to create a more accurate profile of them and to be able to give then more suitable matches. As a result, the game helps users learn more about some of the non-transparent ways in which real version of such apps may be functioning, and hence possibly reconsider whether or how best to use such apps in the future.

[40] https://monstermatch.hiddenswitch.com/online-dating, https://monstermatch.hiddenswitch.com/algorithms

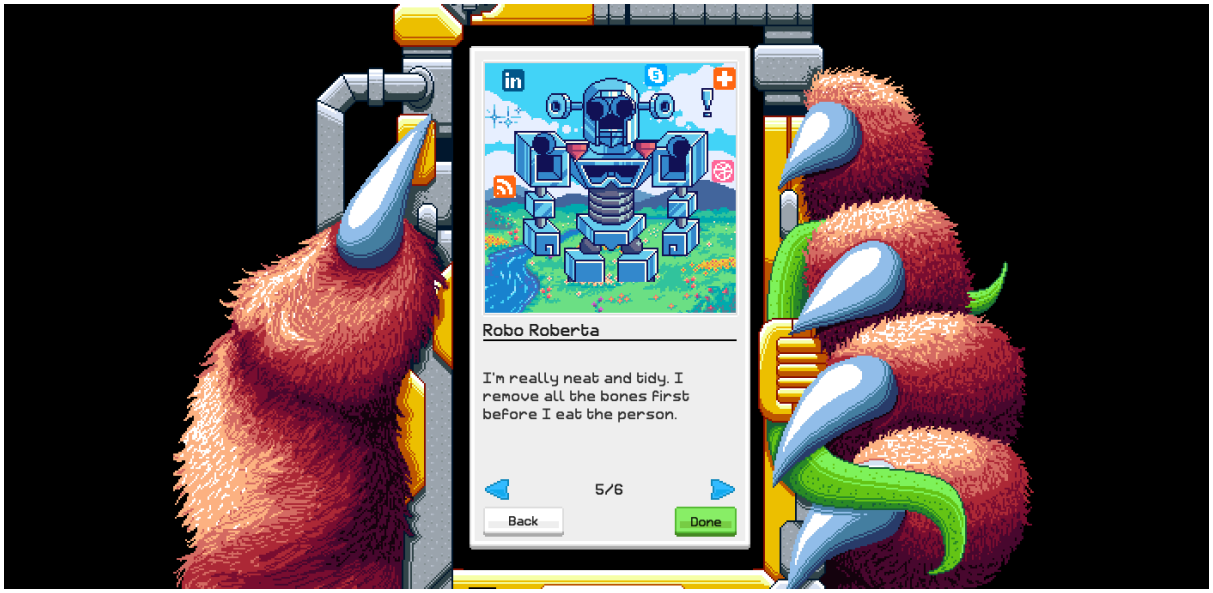


Figure 6a. 'Monster Match': The player creates their profile

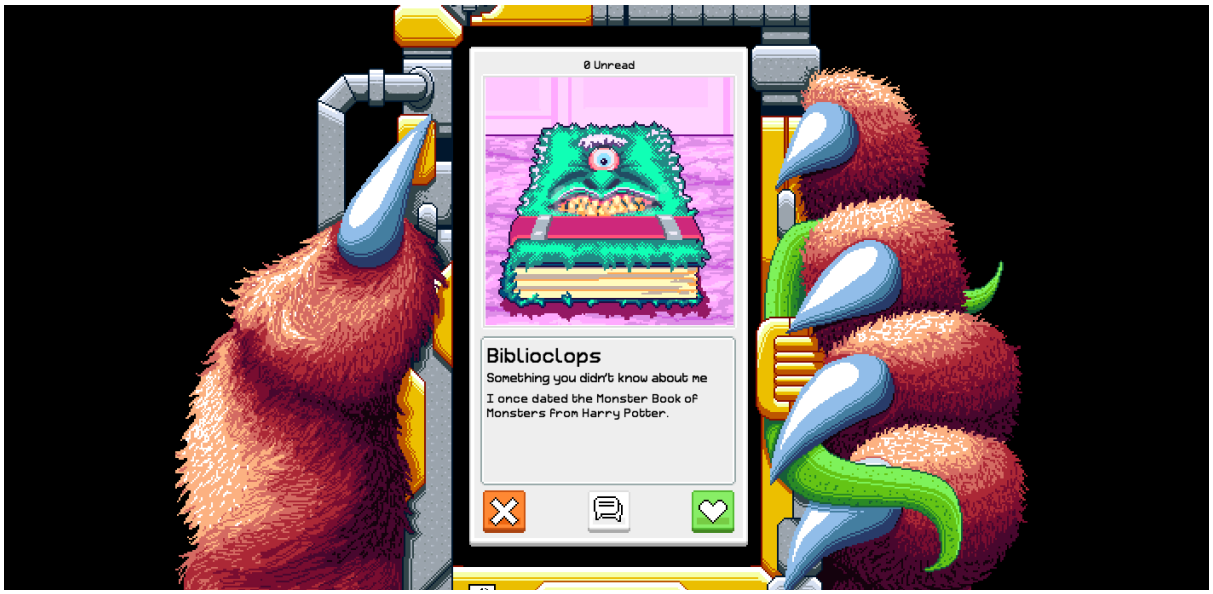


Figure 6b. 'Monster Match': The player chooses to accept or to reject a possible match

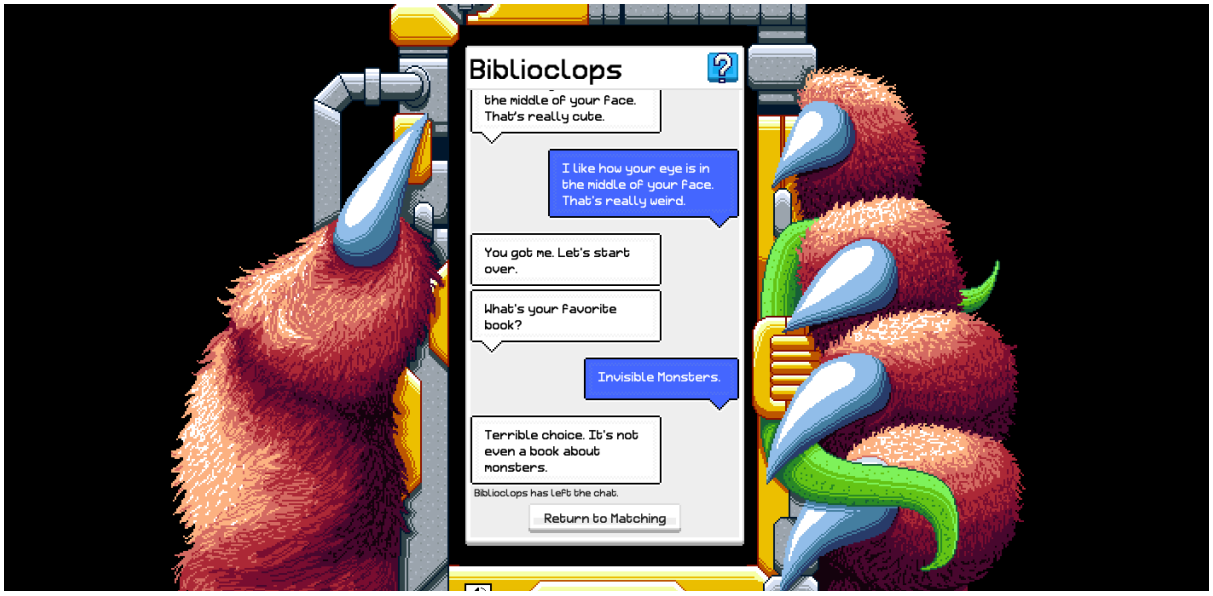


Figure 6c. 'Monster Match': The player interacts with a chosen match

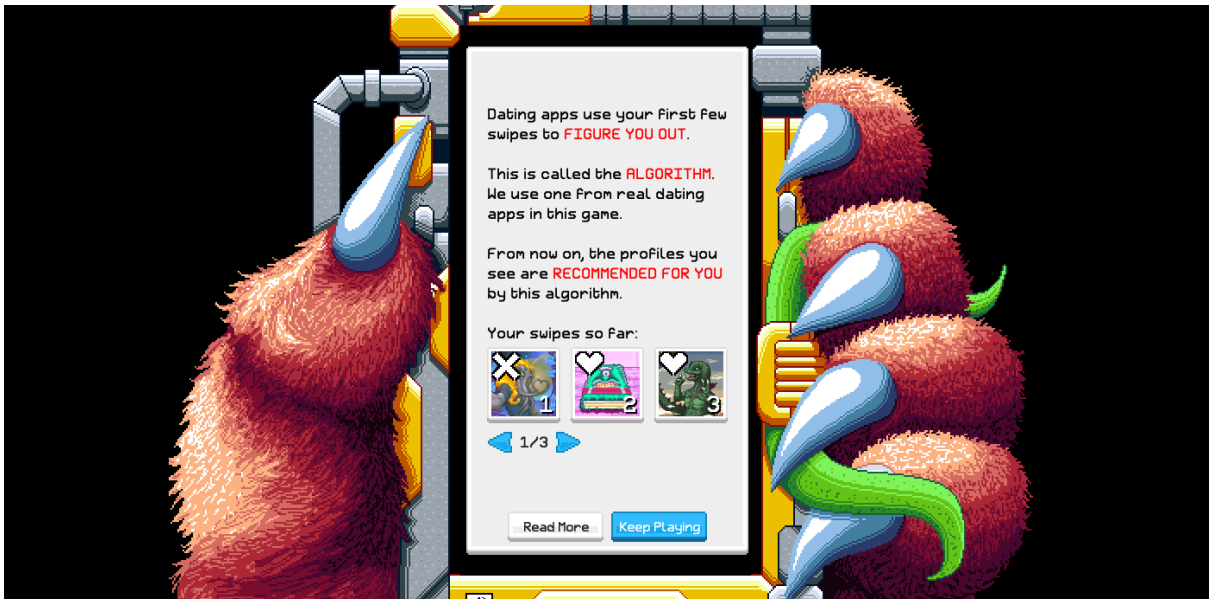


Figure 6d. 'Monster Match': The game illustrates how in-game choices (similar to ones made when using dating-apps) influence outputs

4.1.3. Non-interactive or static content online and offline (e.g. information packages, videos)

Booklets, leaflets, videos, some forms of exhibitions are examples of non-interactive content.

Example 1. 'Data Detox Kit' (https://datadetoxkit.org/en/home)

The 'Data Detox Kit' created by Tactical Tech presented in 'The Glass Room' exhibition (see above) is an excellent example of educating users on the dangers that come from using certain type of

technological devices and of interacting with the online world and utilising its resources.[41] The Kit uses the metaphor of 'Detox' which many users may already be familiar with in relation to healthy eating or lifestyle. The structure of the kit follows a very simple template:

1. Task, 2. Description of threat/harm/risk, 3. Detailed description of basic solution, 4. Further, more detailed solution with more tasks

Example (see Figure 7a and 7b):

1. Clear your location footprints, 2. '…they could reveal important details about you and your habits…', 3. '…go through each app's permissions and turn off the location services…'. Then follows a detailed description of how to do this on different mobile devices., 4. 'Degooglise your life'.[42]

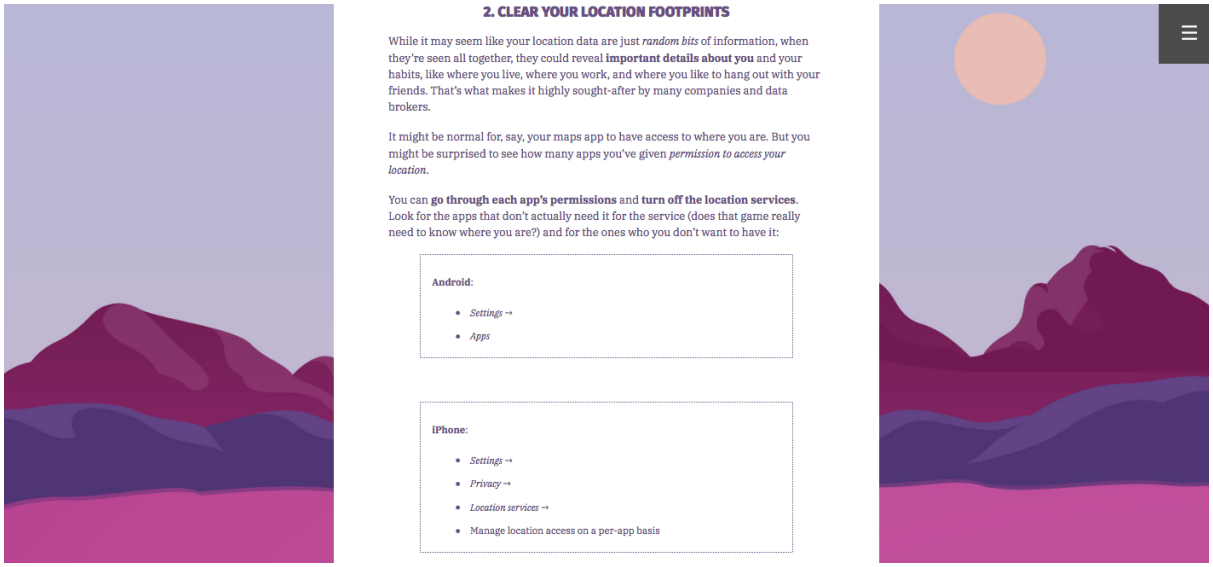


Figure 7a. 'Data Detox Kit': Clear your location footprints 1.

[41] Data Detox Kit: https://datadetoxkit.org/en/home
[42] https://datadetoxkit.org/en/privacy/essentials#step-1

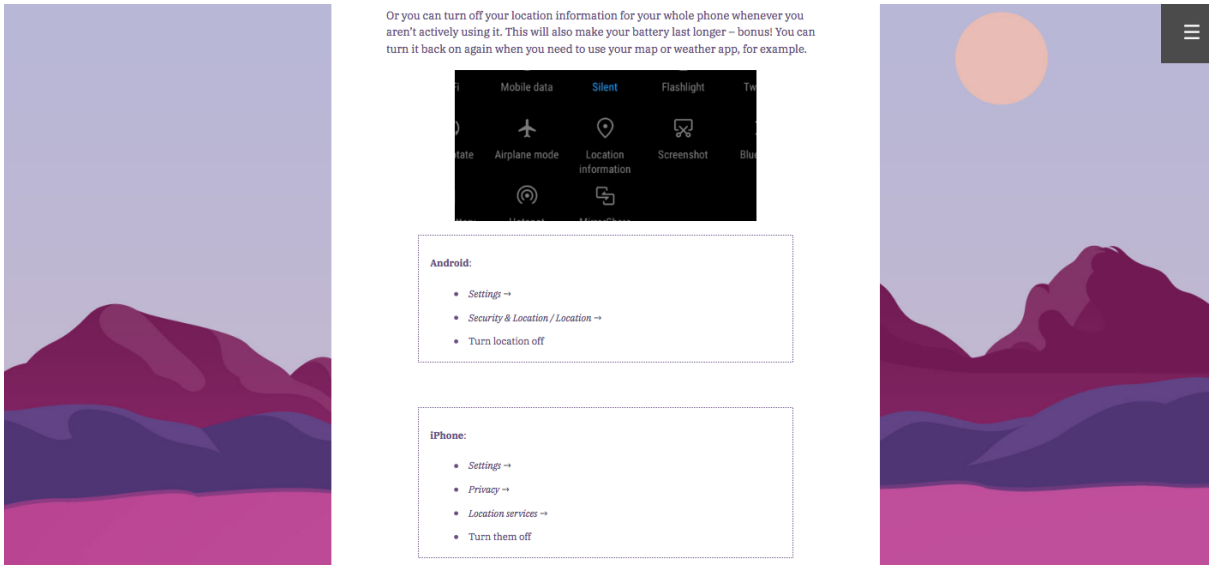


Figure 7b. 'Data Detox Kit': Clear your location footprints 2

This structure makes it easy for non-experts to quickly grasp the essence of what risk they are subjected to, what the potential consequences of their actions/disinterest to change their user habits would be, and what they need to do in order to achieve a higher level of online security or privacy. The template described above also breaks down the potential actions into two categories: a) actions that concerned users should do at a minimum level and could accomplish very quickly, and b) actions that might take more time and consideration but could further increase online safety.

In the Kit the advisory part titled 'Untag your life' educates users on some of the risks of having a social media profile that contains a lot of personal details and most of all, other people's personal details. Under the subheading 'Renovate your social media profile' (see Figure 7c), along with the aforementioned description of potential threats users can find some examples of how the contents of their social media accounts or posts by other people in which they were included in can influence their lives (e.g. insurance apps that attempt to predict a person's driving style based on the style of their online communication).[43]

[43] https://datadetoxkit.org/en/privacy/profile

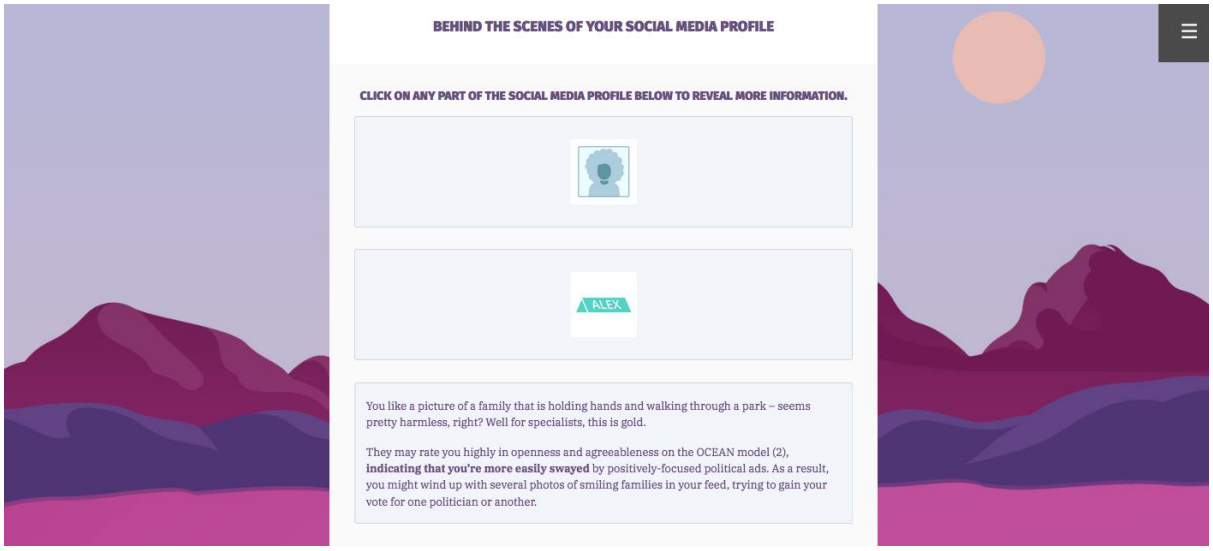


Figure 7c. 'Data Detox Kit': Renovate your social media profile

There are two key characteristics of this kit that should be noted and applied in any such user-focused solution. One is the fact that it does not only describe the problem and raise awareness of the most problematic and risky issues in the online world, but also offers detailed advice suggesting easily manageable, quick solutions. This is crucial because if this step is missing, the average user is not very likely to try to find a solution themselves. The second important attribute is its simplicity, namely that both the summaries of core issues and the proposed solutions are written in a common language that is easily understandable for all users, even for those who are the most inexperienced. These two features make it more likely that users will employ the suggested defensive methods and will do so in full understanding of the reasons behind them.

Example 2. Videos

There are several videos available on the internet that explain algorithmic bias, including many TED and TEDx Talks given by experts.[44] A noteworthy fact regarding these videos is that most of them do not have what we can describe as 'a lot of views'. If we search YouTube for the words 'algorithm bias', the most watched video has roughly 813k views and all other hits have less than 350k.[45] If we search for 'artificial intelligence bias', we get the same video as most viewed and roughly the same view counts

[44] See for example: How I'm fighting bias in algorithms (https://www.youtube.com/watch?v=UG_X_7g63rY), Can we protect AI from our biases? (https://www.youtube.com/watch?v=eV_tx4ngVT0), Machine learning and human bias (https://www.youtube.com/watch?v=59bMh59JQDo), How to keep human bias out of AI (https://www.youtube.com/watch?v=BRRNeBKwvNM), Algorithmic biases in AI and machine learning (https://www.youtube.com/watch?v=q1u6Q_jch9A), Race, technology and algorithmic bias (https://www.youtube.com/watch?v=Y6fUc5_whX8), The moral bias behind your search results (https://www.youtube.com/watch?v=_vBggxCNNno), Beware online 'filter bubbles' (https://www.youtube.com/watch?v=B8ofWFx525s)

[45] On 9th December 2019 the video titled Machine Learning and Human Bias (https://www.youtube.com/watch?v=59bMh59JQDo) has around 813k views.

for other videos. In comparison, the most popular TED talks on 11 December 2019 were ’This is what happens when you reply to a spam email’ with 45M views and ’Inside the mind of a master procrascinator’ and ’How to speak so that people want to listen’, both with 25M views. On the TEDx channel these were ’Beatbox brilliance’ with 71M views and ’My philosophy for a happy life’ with 36M views.

This shows that – at least compared to other topics on which we can find TED and TEDx Talks – there is relatively little interest in society around matters of algorithmic bias when it comes to watching educational videos. Of course, multiple factors play a part in this, for example the most popular videos having been uploaded years earlier than the ones about algorithmic bias, and the fact that it is still not a widely known issue in society. Nevertheless, the relatively low number of views these important videos receive seem imbalanced compared to the level of the effect the issues these videos address have on our everyday lives. It is also worth noting that despite the relatively marginal position that such videos occupy online in terms of their popularity, the number of views is nonetheless large in comparison with the audience potentially reached by traditional outreach methods such as exhibitions. Another issue related to videos is that it is really hard to evaluate whether they had an impact on the audience (e.g. have they changed their behaviour or views as a result of watching the video).

4.2. Characteristics of existing user-focused solutions

The following common characteristics were identified after the examination of existing user-focused solutions regarding algorithmic bias (see Figure 8):

1. Intelligibility: The solutions were using everyday, real-life examples to explain the problem to the non-expert users. While doing this they applied language that was free of high-level technical terms but at the same time the solutions provided a basic understanding of the most important concepts such as ‘algorithm’ or ‘bias’.

2. Visualisation: The reviewed solutions did not only provide explanations but also illustrated the problem with the use of interactive games, figures.

3. Explanation: The above solutions all provided a description of the tool they were illustrating, a basic explanation of the core problems regarding bias (i.e. where the bias comes from and how it appears in the outputs) and short commentary on how this affects individuals, groups and the whole society.

4. Reference to further sources: Most of the referred solutions also have cited a list of useful resources for those who wish to educate themselves further on the respective topics.

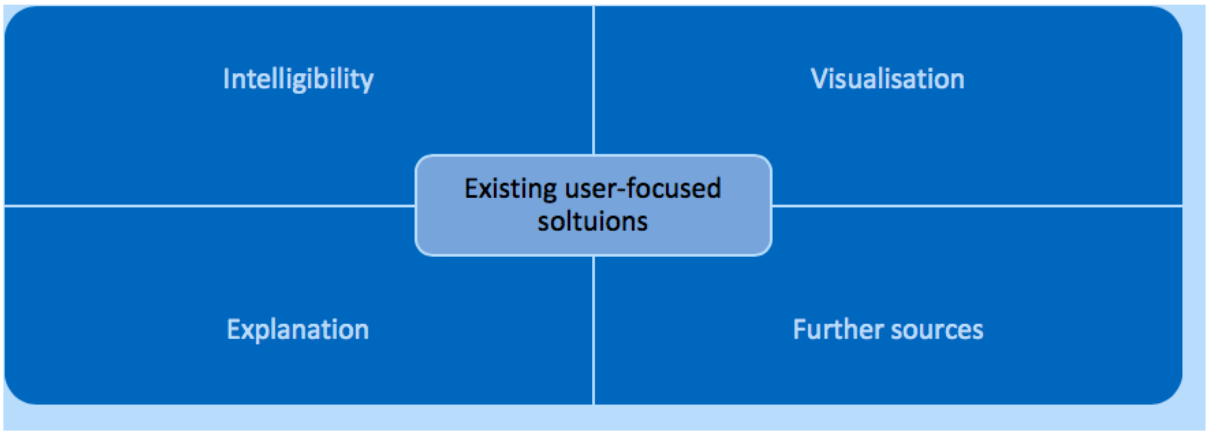


Figure 8. Characteristics of existing user-focused solutions

4.3. User interface development: a way for developers to support end users in learning about algorithmic systems and building systems that explain how they work to the user and offer reasonings for decisions via user interfaces

The previously discussed user-focused solutions mostly aim at raising awareness of the issues of algorithmic bias and their goal is to explain how algorithms affecting our daily lives work in general, how system-biases are created, and how they impact our lives and society. As opposed to this, the last group of solutions aims to allow users a better understanding of the particular system they are using, the connection between input and output, and thus empower them to shape the output more according to their needs or desires. The explanations provided by the user interfaces can either be interactive or non-interactive. Even though in this chapter we focus on informal tools of education, since we concluded that certain challenges that stand in the way of effective user education can only be solved in cooperation with the developers of algorithmic systems, it is worth mentioning user interface development as a method for user education.

An experiment conducted on university admission algorithms by Cheng *et.al. (2019)*, found that "both interactive explanations and 'white-box' explanations" contributed to the users' objective and self-reported understanding of the algorithm. Black-box and static explanations also contributed to these factors, yet to a lesser extent (Chang et.al. 2019). This study seems to support the claim that explanations could increase understanding which could lead to a greater level of trust in a given system. Another possible way of increasing trust and also performance at the same time would be to offer the user tools using which they would be able to have a greater influence on outputs, for example in case of output-ranking in an information access system. These types of solutions pose several open questions which should be studied in order to be able to make more grounded claims when it comes to user interface development.

4.4. What concepts are necessary for the end user to know?

End users today can access more information regarding the different ways in which algorithmic decision making can affect their lives than ever, but the availability of this information does not equal comprehensibility, and it also does not ensure that users will actively look for, find and actually read and digest all relevant pieces of information necessary for them in order to have an understanding of these systems and the potential consequences of their uses.

Some of the widely publicised, high profile cases[46] related to Google's search engine and hate speech detector or algorithms used by policing and insurance companies showed that there is a great need for not only informing end users about these potential issues and threats but also educating them about the workings of these systems in a fashion that could empower them to be pro-active and informed system users. Because of the nature and complexity of algorithmic systems it is also crucial to distribute this knowledge in a shape and form that is easily comprehensible for the everyday user and, if there is need, target specific user-groups with tailored educational solutions, content and methods. Without this type of education end users have a very good chance of being stuck in an informational 'echo chamber' (see for example Bozdag 2013:209, Mittelstadt 2016:4992) or being misled by the results of a search in a search engine. The education should not only cover the end users' day-to-day activity that is influenced by these systems, but should also distribute information on existing biases and discriminatory practices that are being incorporated into particular algorithmic systems' operations through biased training data, developer bias, input data bias, etc.

What is identified by 5Rights Foundation as 'the right to know'[47] for children and young people can be generalised and applied to all members of society. People should not only have the right to have information on who processes their personal data and why but also should be given opportunities to understand how algorithmic decisions concerning their lives are being made. Expert groups and organisations such as UnBias have made it their mission to educate the public or certain user groups with special attributes (e.g. young people) on algorithmic decision-making and the online world.

---

[46] Does Google Manipulate Your Search Results? Sundar Pinchai's Rival Says Yes, Explains How (https://observer.com/2018/12/google-search-algorithm-bias-duckduckgo-ceo/), Google's Artificial Intelligence Hate Speech Detector Is Racially Biased (https://www.forbes.com/sites/nicolemartin1/2019/08/13/googles-artificial-intelligence-hate-speech-detector-is-racially-biased/), Google's Algorithm Is Not Biased, It's Just Not Human (https://www.wired.com/story/google-algorithm-conservatives-biased-its-just-not-human/), Predictive policing poses discrimination risk, thinktank warns (https://www.theguardian.com/uk-news/2019/sep/16/predictive-policing-poses-discrimination-risk-thinktank-warns), Police officers raise concern about 'biased' AI data (https://www.theguardian.com/uk-news/2019/sep/16/predictive-policing-poses-discrimination-risk-thinktank-warns), Supposedly 'Fair' Algorithms Can Perpetuate Discrimination (https://www.wired.com/story/ideas-joi-ito-insurance-algorithms/), Racial Bias Found in a Major Health Care Risk Algorithm (https://www.scientificamerican.com/article/racial-bias-found-in-a-major-health-care-risk-algorithm/)

[47] 5Rights Foundation: The 5Rights, https://5rightsfoundation.com/the-5-rights/

Ethically driven computer science does not only mean trying to avoid causing harm to individuals, groups and society as a whole. We can argue that it also comes with the duty to inform people on how the systems they use work, what are the potential consequences of their actions while using these systems and how can they contribute to the prevention of existing risks.

When it comes to education, one of the most important questions is what we want to teach. As stated in D4.1/ Glossary of terms and need to know, the end user needs to be educated on the following key topics:

Concepts			**Owner**	**End user**	**Developer / Auditor**	**Observer /Regulator**	**Educator**
General		Algorithmic System Model (Components)	x	x	x	x	x
		Algorithmic Transparency	x	x	x	x	x
		Computational Regulations (GDFR)	x	x	x	x	x
		Diversity	x	x	x	x	x
		Explainability	x	x	x	x	x
		F&T Awareness	x	x	x	x	x
Problem Space	**Biases**	Algorithmic Bias		x	x	x	x
		Algorithmic processing Bias			x	x	
		Developer Bias			x	x	
		Input Bias	x	x	x	x	
		Perceived bias	x	x	x	x	x
		Third Party Bias	x		x	x	
		Training Data Bias			x	x	
Solution Space	**Discrimination Discovery**	Discrimination Correction			x		
		Discrimination Detection	x	x	x	x	x

		Explicit Discrimination			x		
		Implicit Discrimination			x		
		Protected (sensitive) Attributes			x		
		Protected Group			x		
		Proxy Attributes			x		
	Fairness Promotion	F&T Requirements	x		x	x	x
		Fairness Certification			x	x	
		Fairness Formalization			x	x	
		Fairness learning			x	x	
		Fairness Sampling			x	x	
		Group Fairness/Parity			x		
		Individual Fairness			x		
	Auditing	Auditing methods			x	x	
	Explainability Management	Black Box Explanation	x	x	x	x	x
		White box explanation	x	x	x	x	x

In order for end users to have the necessary level of comprehension to use algorithmic systems responsibly and to be able to give what we consider as informed consent to the use of their data they first need to have a basic understanding of core concepts mentioned in D4.1 such as ‘algorithm’, ‘input’, ‘output’ and the connection between these elements of the system. Building on that knowledge it will be possible for them to be successfully learn about the above-mentioned key topics. In short, end users should be educated on the meaning, concept and certain matters of algorithmic transparency, explainability, fairness, bias and accountability.

The explanation of the core concepts, as we will see later in this section, can be undertaken in many different ways depending on the target audience and the context. Here, in this section we are exclusively focusing on informal means of end user education. More specifically, we are covering possible ways in which end users can be taught about the topics described above outside the traditional classroom settings or a formal curriculum.

4.4.1. Example of explanation of core concepts

Algorithmic System Model

Here end users need to familiarise themselves with the concept of an algorithmic system. For this, they need to comprehend the components of the system described in D4.1:

- Input
- Output
- Algorithm
- Training Data
- Third Party Constraints

In D4.2 an '**Algorithm'** is explained as:

"The **algorithmic** model of the system is its core. This is the system component that, after having been trained from the data, maps a given input to a given output. In the case of a modern, proprietary search engine, the Algorithmic Model is actually not a single algorithm, but rather, an entire set of algorithmic processes. However, perhaps the most frequently discussed algorithm used in a search engine is the *ranking algorithm*, which assigns a score to each Web page retrieved in response to the user's keywords, such that the pages can then be presented to the user in ranked order."

'**Input'** and '**Output'** are explained as:

'**Input'**: "When using the system, the user inputs some particular value(s) in order to run a given instance of the system. In a Web search, the user provides a set of keywords, which express her need for information on a given topic."

'**Output'**: "The output is the information produced by the system, in response to the user's input. In the Web search case, this is the ranked set of Web pages that the algorithms have deemed to be the most likely to be useful for the user."

For this information to be comprehensible for most end users, we need to use some examples, metaphors or visualisation. For example, Khan Academy explains algorithms in a video the following way (see Figure 9.): "You might have an algorithm for getting from home to school, for making a grilled cheese sandwich. (…) In computer science, an algorithm is a set of steps for a computer program to accomplish a task. (…) So, what makes a good algorithm? The two most important criteria are that it solves a

problem and that it does so efficiently. Most of the time, we want an algorithm to give us an answer that we know is always correct. Sometimes we can live with an algorithm that doesn't give us the correct answer or the best answer because the only perfect algorithms that we know for those problems take a really, really long time."[48]

BBC Bitesize, that targets children, approaches the explanation in a cartoon as follows (see Figure 9.): "If we want a computer to understand how to do something, we need to give it an algorithm. (…) An algorithm is a list of steps you could give to computers to solve a problem or get something done. Imagine that you need to show someone how you brush your teeth so they can learn how to do it themselves. You would need to explain all the little steps you do in the right order. (…) Instructions would go like this: 1. Open the toothpaste. (…) It is important to explain the right steps in the right order."[49]

Both sources also use illustrations or animation to help viewers visualise what is being explained.

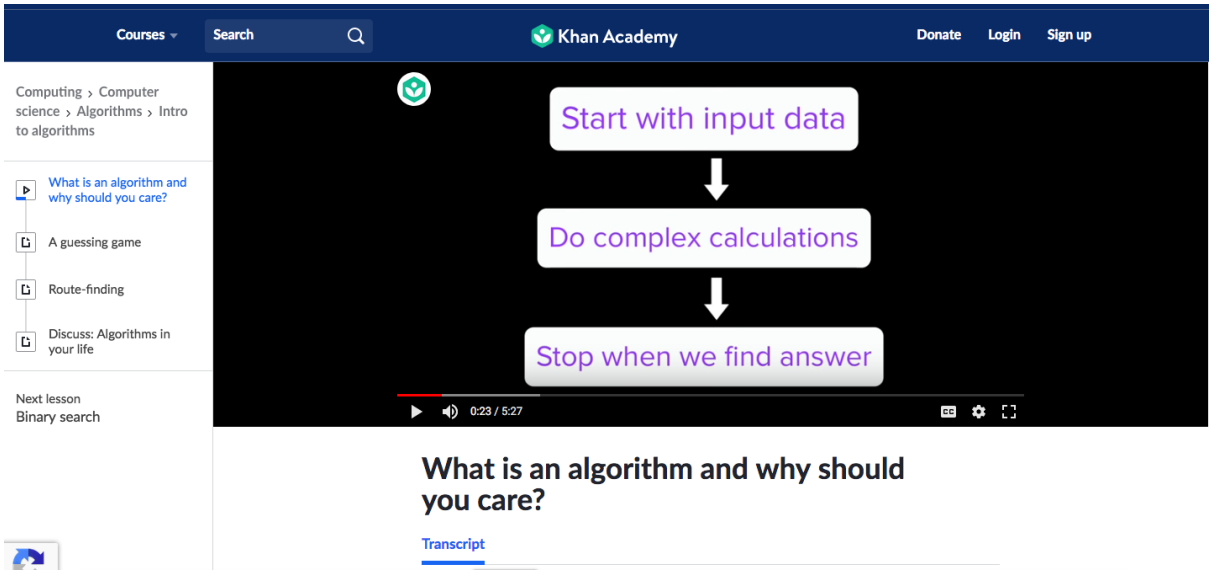
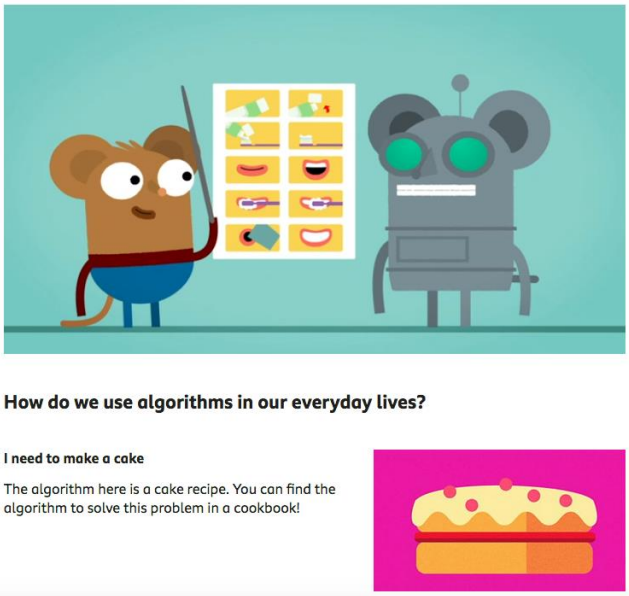


Figure 9. Khan Academy: What is an algorithm and why should you care?[50]

[48] https://www.khanacademy.org/computing/computer-science/algorithms/intro-to-algorithms/v/what-are-algorithms
[49] https://www.bbc.co.uk/bitesize/topics/z3tbwmn/articles/z3whpv4
[50] https://www.khanacademy.org/computing/computer-science/algorithms/intro-to-algorithms/v/what-are-algorithms



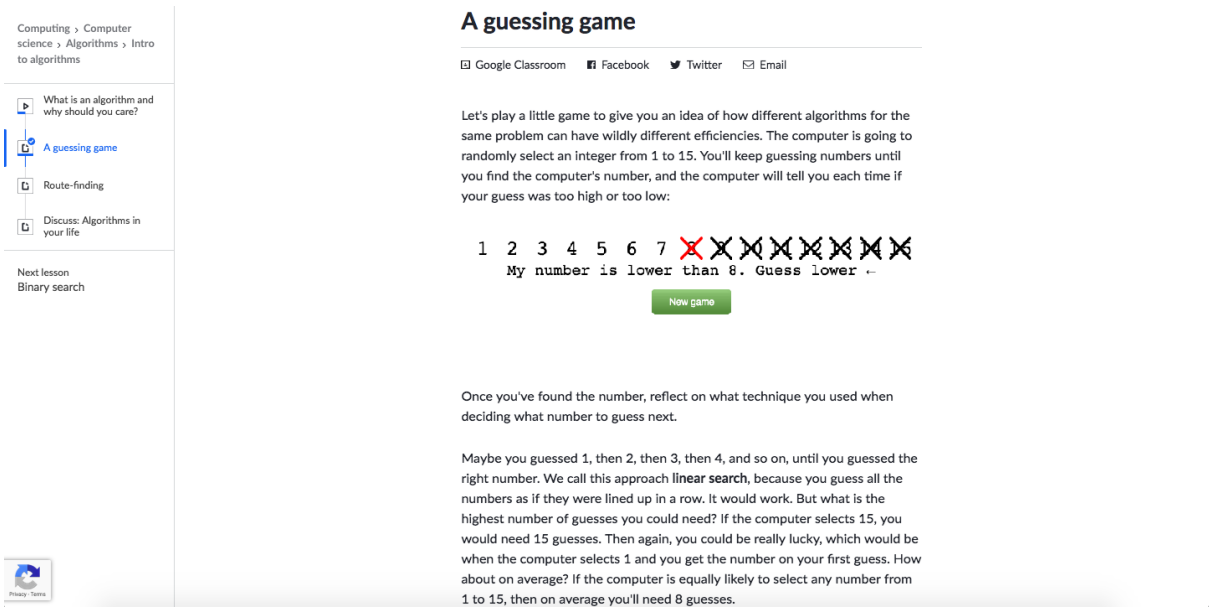
How do we use algorithms in our everyday lives?

I need to make a cake

The algorithm here is a cake recipe. You can find the algorithm to solve this problem in a cookbook!

Figure 10. BBC Bitesize: What is an algorithm?[51]

Khan Academy also has mini-games to illustrate what can an algorithm be used for and how users can visualise its workings (see Figure 11).



Computing › Computer science › Algorithms › Intro to algorithms

What is an algorithm and why should you care?

A guessing game

Route-finding

Discuss: Algorithms in your life

Next lesson
Binary search

A guessing game

Google Classroom Facebook Twitter Email

Let's play a little game to give you an idea of how different algorithms for the same problem can have wildly different efficiencies. The computer is going to randomly select an integer from 1 to 15. You'll keep guessing numbers until you find the computer's number, and the computer will tell you each time if your guess was too high or too low:

1 2 3 4 5 6 7 ~~8~~ ~~9~~ ~~10~~ ~~11~~ ~~12~~ ~~13~~ ~~14~~ ~~15~~

My number is lower than 8. Guess lower ←

New game

Once you've found the number, reflect on what technique you used when deciding what number to guess next.

Maybe you guessed 1, then 2, then 3, then 4, and so on, until you guessed the right number. We call this approach **linear search**, because you guess all the numbers as if they were lined up in a row. It would work. But what is the highest number of guesses you could need? If the computer selects 15, you would need 15 guesses. Then again, you could be really lucky, which would be when the computer selects 1 and you get the number on your first guess. How about on average? If the computer is equally likely to select any number from 1 to 15, then on average you'll need 8 guesses.

Figure 11. Khan Academy: A guessing game[52]

By quoting from the two sources above we wish to illustrate that it seems that intelligibility and examples are key in this context. In case of tools to educate about algorithmic systems in general, any such examples could be very helpful to users to understand how algorithms work. When it comes to the

[51] https://www.bbc.co.uk/bitesize/topics/z3tbwmn/articles/z3whpv4

[52] https://www.khanacademy.org/computing/computer-science/algorithms/intro-to-algorithms/v/what-are-algorithms

explanation of a particular system, of course it needs to feature what that specific system does and what input and output mean in that specific context, just like we saw with the examples of 'Monster Match' and 'Survival of the best fit' games. Both of these games together with the mini-game from Khan Academy illustrate how interactive content can enhance intelligibility and support learning.

When explaining concepts such as input or output – the comprehension of which later will be necessary to be able to learn about bias –, it seems necessary to explain it via specific examples. For instance, to let users know exactly what qualifies as input in case of the algorithmic system used as the example. Just like shown in D4.2, where the explanation of the term input uses web searches and the keywords used by the user to illustrate what a real-life, everyday example for an input is.

In case of interactive educational tools, it is important to provide the explanation while the user is interacting with the content because that can support them in learning as they go through the content (for example as they play through the scenario portrayed by the game).

4.5. Desirable learning outcomes

a) A basic understanding of the core concepts and key topics

The first expected learning outcome is for the end users to have a basic understanding of the core concepts to be described below. This means they should not only be able to understand the terms and their definitions but also the connections between concepts such as input and output and be able to apply the knowledge to particular systems.

b) A general awareness regarding the potentially biased nature of algorithmic systems

The second learning outcome to aim for is users to have a general awareness regarding the potentially biased and discriminatory nature of algorithmic systems whether they are encountering Google searches, using a recruitment algorithm for their company, or a dating app. This may empower certain users to be able to make responsible and informed decisions regarding their online activities and their use of algorithmic systems. It can also help them 'to lower the expectations' they have towards the outputs of these systems. As all user-groups and individual users are different from each other, naturally it is also possible that general awareness would have little to no effect on some of them therefore this needs further investigation and studies of the empirical nature.

c) The ability to critically review the outputs of information access systems and algorithmic decision-making systems in general

The ability to critically assess data (e.g. to critically analyse and "evaluate the credibility and reliability of sources of data, information and digital content"[53]) has been featured in documents describing the content of digital literacy, such as UNESCO's 'A Global Framework of Reference on Digital Literacy Skills for Indicator 4.4.2.' from 2018 and 'DigComp 2.0: The Digital Competence Framework for Citizens' which was published by the Joint Research Centre in 2016.[54] This competence should include awareness of potential bias and discriminatory practices when using or being subjected to algorithmic decision-making as the bias can heavily influence the reliability of information the users are presented with in form of outputs.

Educating users on transparency and fairness can make it more likely that they will either abandon biased systems or will campaign for more fair applications. This, in turn, can encourage on developers and vendors to place greater emphasis on transparency and fairness, and also may encourage regulators to put limits on where such systems can be applied.

Being educated on fairness and accountability will empower users to not only view the outputs critically but also take action in cases on perceived unfairness.

An important question to ask when discussing the ability to critically assess a system is whether user-focused solutions could increase trust in the system and whether they should do that at all. According to Chang *et al*'s study (2019), neither black-box nor white-box explanations in user interfaces were able to increase trust in the system, users remained wary and relatively untrusting of them. This, of course, doesn't mean that other forms of user-focused solutions or even user interface-based solutions would not be able to increase trust in the system if developed to meet certain criteria. The question is whether these solutions should in fact aim to increase trust. The term 'being able to critically assess systems' does not answer this question directly as a critical perspective could simply mean that users should be able to determine which system to trust (for example which system is capable of producing unbiased output) and which systems to avoid or where to raise questions regarding the trustworthiness of the received outputs.

4.5.1. The limits of informal end user education regarding algorithmic bias and transparency

Some users are overly trusting in the systems they are using, some do not have a choice but to use or be subjected to them in a work or any other context, and many simply do not care about the potential bias because the gain deriving from using the systems is greater than the potential risk of bias or

[53] UNESCO: A Global Framework of Reference on Digital Literacy Skills for Indicator 4.4.2., p23 (http://uis.unesco.org/sites/default/files/documents/ip51-global-framework-reference-digital-literacy-skills-2018-en.pdf)

[54] DigComp 2.0: The Digital Competence Framework for Citizens (https://publications.jrc.ec.europa.eu/repository/bitstream/JRC101254/jrc101254_digcomp%202.0%20the%20digital%20competence%20framework%20for%20citizens.%20update%20phase%201.pdf)

discrimination. This means that despite any available information or education there will be people who will continue to use biased systems.

According to a 2016 Eurostat poll, at the time 37% of internet users claimed that they "read privacy policy statements before providing personal information" and 31% of them stated they restrict access to their geographical location.[55] The same survey also found that "71% of people aged 16 to 74 in the EU-28 countries who had used the internet in the previous 12 months knew that cookies can be used to trace people's online activities", the ratio of awareness being slightly higher than average amongst younger users (74% amongst 16-24 year olds). Despite the alleged knowledge only a little more than one third (35%) of the aforementioned users stated that they limited or prevented the use of cookies by changing their settings.[56] These results suggest that even users who are aware of certain risks emanating from the use of online systems can remain passive. There can be multiple possible explanations for this. One is that they have indeed heard about the threat cookies can pose to privacy, but they do not comprehend the connection between the use of cookies and the potential loss of privacy. It is also possible that they are not aware of the possibility to change their settings or they do not possess the required knowledge, skills, time or inclination to manage it. A third possibility is that they are not concerned by the potential invasion of privacy as 'they have nothing to hide' which is a commonly heard argument. Similarly to this, simple awareness of the potential bias, discrimination and knowing the advantages of transparent systems will not necessarily result in all users actively looking for less opaque systems, anti-bias certificates or will campaign for explainability. This is why it is important not only to educate users on the potential issues but also to present them with easily comprehensible and applicable tools which can assist them in recognising and potentially eliminating biases and taking steps when they were subjected to harm as individuals or as a group as a result of the operation of an algorithmic decision. Even if this is achieved, the possibilities of user education in changing biases are limited.

As stated above in the desirable learning outcomes, informal end user education is capable of distributing knowledge regarding the most important concepts of algorithmic systems and biases, it can raise awareness of the problematic issues and it can empower end users to view these systems and their outputs critically and to potentially recognise biases and discrimination. In addition, it can offer certain types of solutions, for example regarding the available legal steps in case of discrimination. However, neither of these are capable of solving most of the individual, group-level or societal problems that arise

[55] Privacy and protection of personal identity (2016 survey) (https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals#Internet_usage)

[56] Privacy and protection of personal identity (2016 survey) (https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Digital_economy_and_society_statistics_-_households_and_individuals#Internet_usage)

from algorithmic bias. Furthermore, they can only offer knowledge and guidance for users, but nothing will serve as a guarantee that they will follow the advice or will internalise the knowledge or use it in their everyday lives.

4.5.2. Measuring learning outcomes in case of informal methods of user-education

When it comes to the development of any system, tool or method, the most important question is whether it actually works and delivers on the promised results. In case of education, determining this is a very complex issue. In case of informal user-focused solutions we propose that effectiveness should be decided based on whether the proposed learning outcomes have been met. In order to decide this, there are multiple methods we can apply either by themselves or in combination with each other. When measuring acquired knowledge, we have to distinguish between objective knowledge gained as a result of the solution and subjective feelings of understanding the topic.

To test objective knowledge, we can ask users to complete quizzes, solve practical problems or perform a task. To measure subjective feelings of understanding the topic, we can conduct interviews or questionnaires that cover this subject.

It is important to take the passing of time into account so another method to measure effectiveness we suggest is to conduct another survey some time after the testing to determine whether the acquired knowledge is lasting and whether the education caused any changes in the users' online behaviour (for example higher level of awareness of bias in the long term, abandonment of biased systems, seeking out further information, becoming wary or scared of algorithmic systems).

It could also be useful to introduce a control group, both when measuring immediate and long-term effects. It would help to determine whether for example the ability to solve a particular task came from the use of the educational tool, or long-term effects such as the heightened levels of awareness or unintended consequences like becoming scared of algorithmic systems happened due to the education tool or external factor (e.g. a high-profile case of an unfair algorithm that happened since the education).

The details of performance evaluation naturally depend on the exact nature and tasks of a particular user-focused solution, but we believe that the following tests are necessary to determine whether the tool in question is effective or not:

- Testing the tool on a group of users to measure changes both in their subjective feelings of their understanding and in their objective knowledge
- Re-evaluating the same group of users after some time has passed
- Using a control group

4.6. Requirements of user-focused solutions

a) Description of the problem

It is very important to describe the problem in a way that is relatable for end users. It seems therefore crucial to adopt everyday examples regarding algorithmic tools that people use in their lives. When talking about user-focused solutions, the description of the problem ideally should contain the following elements:

- Description of the core problem
- Showing the course of events that lead to the occurrence of the problem
- The effects of the problem on individuals, groups and society.

b) A solution to said problem (where possible)

It seems that at this point it is almost impossible to suggest to end users a course of action that can offer a comprehensive solution to the societal and personal problems that can arise from algorithmic bias. The user-focused solutions examined by us have not done so either, although some of them mentioned some possible ways forward.

Due to the lack of straightforward and easy ways, offering at least a partial-solution for end users can mean several things in this context:

- Suggesting ways for the user to make for example their search results less biased
- Suggesting an alternative, less biased way to achieve the same result that the original algorithm offers
- Suggesting ways to make the user's training data less biased
- Promoting long-term, societal solutions, for example diverse hiring practices or unconscious bias trainings
- Promoting/developing systems that have more frequent human-computer interaction and allow a higher level of user-impact on the final results (decreasing the level of automation)
- Offering advice on what steps are available if someone's or a group's rights have been (possibly) violated.

5. Further tasks and conclusions

Naturally, there is only so much users can achieve in terms of solving the problems that emerge from algorithmic bias as other stakeholders, such as developers and policy makers have significantly more tools in their hands to work towards a solution. For now, this is mirrored in the available examples of user-focused education we examined, as they aim mostly to raise awareness and to educate end users on existence of these biases and the importance of algorithmic transparency.

Despite this, it seems crucial to support end users in learning about algorithmic systems, biases and the impact that biased systems can have on their lives and on society as a whole. For instance, it can raise awareness regarding the potentially problematic issues that can arise from neglecting the requirement of unbiased and discrimination-free systems during the development or testing process. Users' ability to recognise bias and discrimination that can put pressure on vendors and developers to try to create less biased systems and even to deliver more transparent software.

As seen in this section, end user education can be provided in various forms. We believe that all forms can be useful, as all can target different audiences and be helpful in different contexts. Having diverse ways available for users to learn about these topics would also mean that every user could find the tool and method that is suitable for their needs.

While providing general educational content regarding algorithmic bias that show the meaning of core concepts and even the weaknesses of particular types of systems seems possible for external content-creators, explaining the exact mechanisms of a system can only be done by the developers of that particular system. As stated above, this calls for cooperation with system developers and other stakeholders, such as vendors.

6. References

1. Allen, S. (2002) Looking for learning in visitor talk: A methodological exploration, Learning Conversations In Museums, G. Leinhardt, K. Crowley and K. Knutson, Mahwah, Lawrence Erlbaum Associates, p259-303.
2. Blomberg. T., Bales, W., Mann, K., Meldrum, R., Nedelec, J. (2010) A Validation of the COMPAS Risk Assessment Classification Instrument, Center for Criminology and Public Policy Research, College of Criminology and Criminal Justice, Florida State University http://criminology.fsu.edu/wp-content/uploads/Validation-of-the-COMPAS-Risk-Assessment-Classification-Instrument.pdf)
3. Bozdag, E. (2013) Bias in algorithmic filtering and personalization, in Ethics and Information Technology 2013:15, p209-227
4. Brennan, T. Dieterich, W., Ehret, B. (2009) Evaluating the predictive validity of the COMPAS risk and needs assessment system, in *Criminal Justice and Behavior* 36:1, p21-40
5. Chang, H.F., Wang, R., Zhang, Z., O'Connell, F., Gray, T., Harper, F.M., Zhu, H. (2019) Explaining Decision-Making Algorithms through UI: Strategies to Help Non-Expert Stakeholders, CHI 2019, May 4-9, 2019, Glasgow, Scotland, UK, https://www.cs.rochester.edu/u/zzhang95/doc/pub/algorithm_explanation_nonstakeholder.pdf
6. Diakopoulos, N., Koliska, M. (2016) Algorithmic Transparency in the News Media, in: News Media, Digital Journalism, http://www.nickdiakopoulos.com/wp-content/uploads/2016/07/Algorithmic-Transparency-in-the-News-Media-Final.pdf
7. Dressel, J., Farid, H. (2018) The accuracy, fairness and limits of predicting recidivism, in *Science Advances* 4:1
8. Flores, A.W., Lowenkamp, C.T., Bechtel, K. (2017) False Positives, False Negatives, and False Analyses:A Rejionder to "Machine Bias: There's a Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks." (http://www.crj.org/assets/2017/07/9_Machine_bias_rejoinder.pdf)
9. Garattini, C., Prendergast, D. (2015) Introduction. Ageing and Digital Life Course (eds. Prendergast, D., Garattini, C.). New York: Berghahn Books. p1.
10. Goldenfein, J. (2019) Algorithmic Transparency and Decision-Making Accountability: Thoughts for buying machine learning algorithms' in Office of the Victorian Information Commissioner (ed), Closer to the Machine: Technical, Social, and Legal aspects of AI
11. Kehl, D., Guo, P., Kessler, S. (2017) Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing, Responsive Communities Initiative, Berkman Klein Center for Internet and Society, Harvard Law School
12. Kroll, J.A., Huey, J., Barocas, S., Felten, E.W., Reidenberg, J.R., Robinson, D.G., Yu, H. (2016) Accountable Algorithms, in: University of Pennsylvania Law Review, p633.

13. Logg, J.M., Minson, J.A., Moore, D.A. (2019) Algorithm appreciation: People prefer algorithmic to human judgement, in *Organisational Behavior and Human Decision Processes* 151, p90-103
14. Mittelstadt, B. (2016) Auditing for Transparency in Content Personalization Systems, in: International journal of Communication 2016:10, p4991-5002
15. Paudyal, P., Wong, B.L. W. (2018) Algorithmic Opacity: Making Algorithmic Processes Transparent through Abstraction Hierarchy, Proceedings of the Human Factors and Ergonomics Society 2018 Annual Meeting, p192-196, https://journals.sagepub.com/doi/pdf/10.1177/1541931218621046
16. Groundwater-Smith, S., Kelly, L. (2003) As we see it: improving learning in the museum, Paper presented at the British Educational Research Association Annual Conference, Heriot-Watt University, Edinburgh, 11-13 September 2003, http://www.leeds.ac.uk/educol/documents/00003271.htm
17. McClintock, F.H. (1970) "The Dark Figure", in Collected Studies in Criminological Research 4, Strasbourg, France: Council of Europe, p7-34
18. Peacock, S.E., Künemund, H. (2007) Senior Citizens and Internet Technology. European Journal of Ageing. 4:4. p191.
19. Quinney, R. (1970) The Social Reality of Crime, Boston: Little Brown
20. Roupa, Z., Marios, N., Gerasimou, E., Zafeiri, V., Giasyrani, L., Kazitori, E., Sotiropoulou, P. (2010) The use of technology by the elderly. Health Science Journal. 4:2. p118.
21. Skogan, W. (1977) Dimensions of the Dark Figure of Unreported Crime, in Crime and Delinquency 1977/Jan, p41-50
22. Tan, S., Caruana, R., Hooker, G., Lou, Y. (2018) Distill-and-Compare: Auditing Black Box Models Using Transparent Model Distillation With Side Information, https://arxiv.org/pdf/1710.06169.pdf
23. Tombs, S. (2014) Health and Safety 'Crimes' in Britain: The Great Disappearing Act, in: Invisible Crimes and Social Harms. Critical Criminology Perspectives (eds: Davies, P., Francis, P., Wyatt, T.), London: Palgrave Macmillan, p199-220
24. Vacek, P., Rybenská, K. (2017) Digital Technology in the Contemporary Lives of Senior Citizens. International Journal of Information and Education Technology. 7:10. p758.
25. Vaportzis, E., Clausen, M.G., Gow, A.J. (2017) Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Study Group. *Frontiers in Psychology*. 8:1687