Keywords

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

Keywords

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.

What Happens Next

If your TrustScape is shared online, it may be responded to by people working in the technology industry, policymakers or others (see the next section on Stakeholder MetaMaps). The MetaMaps can also provide further opportunities for discussion and debate. Additionally, you might use the Participant Value Perception Workshop (p. 27) to evaluate any MetaMaps and the quality or value of their responses.

We will post it on UnBias’ website and social media (Instagram & Twitter) with the hashtag: #trustscape

Check out UnBias’ website and social media for other TrustScapes and use them to discuss issues in class or group settings.
INTRODUCTION

Increasingly, our everyday interactions in the modern world are mediated and filtered through algorithms: analysing our behaviours, inferring our preferences and channeling the choices we are offered. These algorithms are embedded both within the personal computers and devices we use, as well as within the larger infrastructures that manage and facilitate everyday life.

The UnBias project explores the user experience of algorithm-driven internet sites and the processes of algorithm design. We are particularly interested in circumstances in which algorithmic processes might (intentionally or unintentionally) produce biased or unfair outcomes — for instance in the form of helping fake content to spread on social media, producing search results that reinforce prejudiced attitudes, or the excessive personalisation of content and collection of personal data.

The Fairness Toolkit is one of our project outputs aiming to promote awareness and stimulate a public civic dialogue about how algorithms shape online experiences and to reflect on possible changes that can address issues of online unfairness.

The tools are not just for critical thinking, but for civic thinking and doing: supporting a more collective approach to imagining how algorithms shape online experiences and to imagine our online environments and citizenship in ways that are more inclusive and equitable.

The toolkit consists of three main parts:

- Awareness Cards — to help people to explore and think critically about the issues in an engaging way;
- TrustScapes — to help people express their perceptions and feelings about the issues in a shareable format;
- MetaMaps — to help stakeholders (industry, researchers, policy and decision-makers) respond to the issues raised in the TrustScapes and contribute to an evolving public civic debate.

Sources

UnBias Fairness Toolkit Handbook
Giles Lane
2018-08-31 & CC BY-NC-SA
Published by Proboscis
A project of the Horizon Digital Economy Institute
(University of Nottingham), University of Oxford,
University of Edinburgh & Proboscis
unbias.wp.horizon.ac.uk
Designing the toolkit was a participatory process involving the following people, to whom many thanks are offered:

UnBias Team

Horizon Digital Economy Institute, University of Nottingham
Professor Derek McAuley, Dr Elvira Perez Vallejos, Dr Ansgar Koene, Dr Liz Dowthwaite, Dr Virginia Portillo, Monica Cano Gomez & Dr Helen Creswick.

Department of Computer Science, University of Oxford
Professor Marina Jirotka, Dr Helena Webb, Dr Menisha Patel, Paula Fiddi & Ross Gales.

School of Informatics, University of Edinburgh
Dr Michael Rovatsos, Dr Sofia Ceppi & Dr Alan Davoust.

Co-Design Participants:
Staff & Students at Sir John Lawes School, Harpenden
Staff & Students at Elizabeth Garrett Anderson School, Islington
Mel Jewell, Community EET Worker at Oxfordshire County Council and members of the go_girl:code and create group.

Special thanks to threemaker and improv expert Alex Murdoch for devising the Awareness Cards Exercises.

It also includes additional materials:
- Value Perception Worksheets – for assessing and evaluating the value of using the toolkit;
- Sketch Sheets – illustrations by artist, Alice Angus, which are derived from our co-design workshops with young people visualising many of their concerns and feelings about the internet, data and algorithms;
- Keyword Sheets – keywords also derived from our co-design workshops with young people.

The toolkit has been co-created with young people and stakeholders whose input has greatly informed the design process. We conducted a series of workshops with young people aged twelve to seventeen in two schools and with a group of sixteen to twenty-two year olds in a community centre to understand more about how aware they are of the issues, how relevant to their own lives they perceive them to be and, what they thought should be done.

Our observations revealed a consistent fatalism and lack of a sense of agency linked to the highly personalised nature of young people’s interactions with the technologies. This fostered a sense of isolation and inability to effect change, coupled with a very limited sense of their rights and how the law already protects them in their interactions with service providers and big companies. They often feel that their voice is neither heard nor listened to, even when they are the targets of some of the most aggressive marketing techniques.

Many of the participants were surprised both at the scale of algorithmic systems and how much their data and behaviours are already tracked across the internet by government and industry.
What might the costs be to stakeholders if they ignore or do not act on the perceptions shared in the TrustScapes of values perception?

Stakeholders should care about the perceptions of users shared in TrustScapes because it can stimulate a public civic dialogue.

The toolkit includes two tools designed to stimulate a public civic dialogue: MetaMaps – worksheets for users to share across numerous platforms and services. The Data, Rights, Values and Factors cards can be used as specific discussion topics that encompass wider themes and issues in algorithmic systems with real-world examples. The Exercise cards suggest a range of activities for participants to explore the issues in different ways: from the solitary to the social. Participants can draw upon their own experiences, or on stories and reports they have encountered in the media. The exercise cards provide opportunities for role play, improvisation, collaboration and shared experiences of bias, fairness and trust.

When do algorithms make decisions? Algorithms help analyse and make decisions in automated systems. How do algorithms and data sources play in automated systems? What role do algorithms and data sources have in automated systems? What are algorithms? Algorithms are processes or lists of rules to follow in order to complete a task, such as solving a problem, doing a calculation or making a decision. The cards are designed to help participants to explore the issues in different ways: from the solitary to the social. Participants can draw upon their own experiences, or on stories and reports they have encountered in the media. The exercise cards provide opportunities for role play, improvisation, collaboration and shared experiences of bias, fairness and trust.

**Fairness Toolkit Handbook**

The toolkit includes two tools designed to stimulate a public civic dialogue: MetaMaps – worksheets for users to share across numerous platforms and services. The Data, Rights, Values and Factors cards can be used as specific discussion topics that encompass wider themes and issues in algorithmic systems with real-world examples. The Exercise cards suggest a range of activities for participants to explore the issues in different ways: from the solitary to the social. Participants can draw upon their own experiences, or on stories and reports they have encountered in the media. The exercise cards provide opportunities for role play, improvisation, collaboration and shared experiences of bias, fairness and trust.

**Getting Started**

There are no specific rules to use the cards. You can use the whole deck or pick individual suits or cards and use them in different combinations.

The Example cards are ideal for introducing the issues of bias, trust and fairness in algorithmic systems with real-world examples. The Data, Rights, Values and Factors cards can be used as specific discussion topics that encompass wider themes and issues in algorithmic systems with real-world examples. The Exercise cards suggest a range of activities for participants to explore the issues in different ways: from the solitary to the social. Participants can draw upon their own experiences, or on stories and reports they have encountered in the media. The exercise cards provide opportunities for role play, improvisation, collaboration and shared experiences of bias, fairness and trust.

**Critical Thinking Skills**

Use the cards to help people develop critical thinking skills. The Data, Rights, Values and Factors cards can be used as specific discussion topics that encompass wider themes and issues in algorithmic systems with real-world examples. The Exercise cards suggest a range of activities for participants to explore the issues in different ways: from the solitary to the social. Participants can draw upon their own experiences, or on stories and reports they have encountered in the media. The exercise cards provide opportunities for role play, improvisation, collaboration and shared experiences of bias, fairness and trust.

**Glossary**

Data = 12 cards describing types of personal data we commonly share across numerous platforms and services. Values = 10 cards with over one hundred values that typically motivate people in everyday life. Rights = 6 cards listing legal rights held by linked rights. Factors = 4 cards describing common factors which affect decision-making – whether human or machine-automated – bias, prejudice, discrimination, trust and fairness.
VALUE PERCEPTION

Purpose

The value perception worksheets have two purposes:
1 as a prospective tool both participants and stakeholders to imagine the potential value of participating in the UnBias public civic dialogue.
2 as an evaluation tool for reflecting on what participation has actually resulted in, for instance for participants to evaluate any responses to the TrustScapes in Stakeholders’ MetaMaps.

How to Use

Print out a copy at the size most useful for your purposes, e.g.: for individuals – A3 size for a small group – A2 size for larger groups – A1 poster size. Bigger sizes can be printed cheaply at local copy shops.

The Value Perception sheets are intended as personal or group reflection tools, not as contributions to the public civic dialogue.

However, if you do wish to share a completed – and anonymous – worksheet, please email it to us: unbias@cs.ox.ac.uk We will post it on UnBias’ website and social media (Instagram & Twitter).

Why is this important?

Decisions about our lives are made by systems that are increasingly being automated. The cards can help build awareness of where and how such decisions occur. They also help develop critical thinking capacities about whether or not the decisions may be biased, and what factors might motivate any discriminatory or unfair outcomes.

Risk and trust
Risk and trust are inseparable and intertwined. The more we are at risk of vulnerability, the more important trust becomes. Many of the Exercises rely on and explore trust between participants. They also offer a space of self-reflection about how we trust ourselves, and how we perceive our world-view: as safe or unsafe, hostile or friendly, happy or sad, open or closed etc.

UnBias’ website and social media (Instagram & Twitter)
How to Use

MetaMaps can be completed by an individual or by a group. Print out a copy at the size most useful for your purposes, e.g.:
- for individuals – A3 size
- for a small group – A2 size
- for larger groups – A1 poster size

Bigger sizes can be printed cheaply at local copy shops.

Select a TrustScape from those shared online via the UnBias social media (Instagram, Twitter & Facebook). Print it out and stick it onto the MetaMap in the box provided:
- for A3 MetaMaps – print TrustScape at A5 size
- for A2 MetaMaps – print TrustScape at A4 size
- for A1 MetaMaps – print TrustScape at A3 size

Responding to a TrustScape

Bear in mind that, although anonymous, each TrustScape is the work of a person or group who are sharing their perceptions and feelings about crucial issues affecting our society. Try to respond in both an informative way and an aspirational way.

Please remember to be respectful about the young people’s perceptions and feelings – your response will be a contribution to a public civic dialogue which they will be part of.

Sketches & Keywords

The toolkit also includes sketches and keywords you may wish to use in your own response. Print them out on label sheets and stick them in the boxes or use them as inspiration for your own words and drawings.

Keep Your MetaMap Anonymous

MetaMaps are designed to be shared online so please do not add any personal details or information that could identify you. But don’t worry if you do – we’ll anonymise any we spot that need it.

Games & Play

Adapt games you already know as a way of familiarising yourself and others with the issues. For instance:

HAPPY FAMILIES: Shuffle the deck and deal them out amongst the players. Players then take turns to pick a card from the player next to them. When a player has 4 cards from any of the same suit they can lay them down. The aim is to have as many suits of 4 as possible.

TRUMPS: Pick an Example Card place it face up. Select the Data, Rights, Factors and Values Cards, shuffle them and deal them out among the players. Players then take turns to add their cards on top explaining why they are relevant to the Example. The winner will have put down the most relevant cards per example.

GUESS WHO?: Use the Data, Rights, Factors, Example and Values cards. Shuffle them and deal each player a card. Players then take turns in asking each other one question to try to work out which card each player has. The correct guess wins the card, the aim being to win as many cards as possible.

• Acknowledging your own biases.
• Investigating the consequences of your own bias.
• How do you know if you have acted fairly?
• What makes us trust social media?
• How are algorithms created and used in computer science?
• Addressing bias when writing computer programmes.
• Are search engines trustworthy?
• How can we discern facts from opinions?
• How can you collect, analyse and represent data to prevent bias?
• Can you prevent bias in statistics?
• How do you identify information sources that are reliable and trustworthy?
• How does bias affect the use of data and graphs in different subjects?
How to Use

TrustScape can be completed by an individual or by a group.

1. Select an experience of online bias, unfairness or untrustworthiness you have had or are concerned about.

2. Describe this experience.

3. What would you like to see done?

4. How do you think these issues are being addressed by companies and authorities?

5. Illustrate what is important to you about this experience.

When complete, take a photo of this sheet & email it to: unbias@cs.ox.ac.uk

To help stimulate a public civic dialogue it will be shared on UnBias’ website and social media (Twitter & Instagram) using the hashtags: #unbias & #trustscape.

This work is licensed under a Creative Commons: Attribution Non-Commercial Share-Alike 4.0 International Licence.

Concept & Design

proboscis.org.uk

How to Share

Photograph your completed TrustScape and email it to:

unbias@cs.ox.ac.uk