

Turing Commons

Scaling our Skills and Training Platform



Presentation

Overview

1

What is the Turing Commons?

2

Centres for Doctoral Training

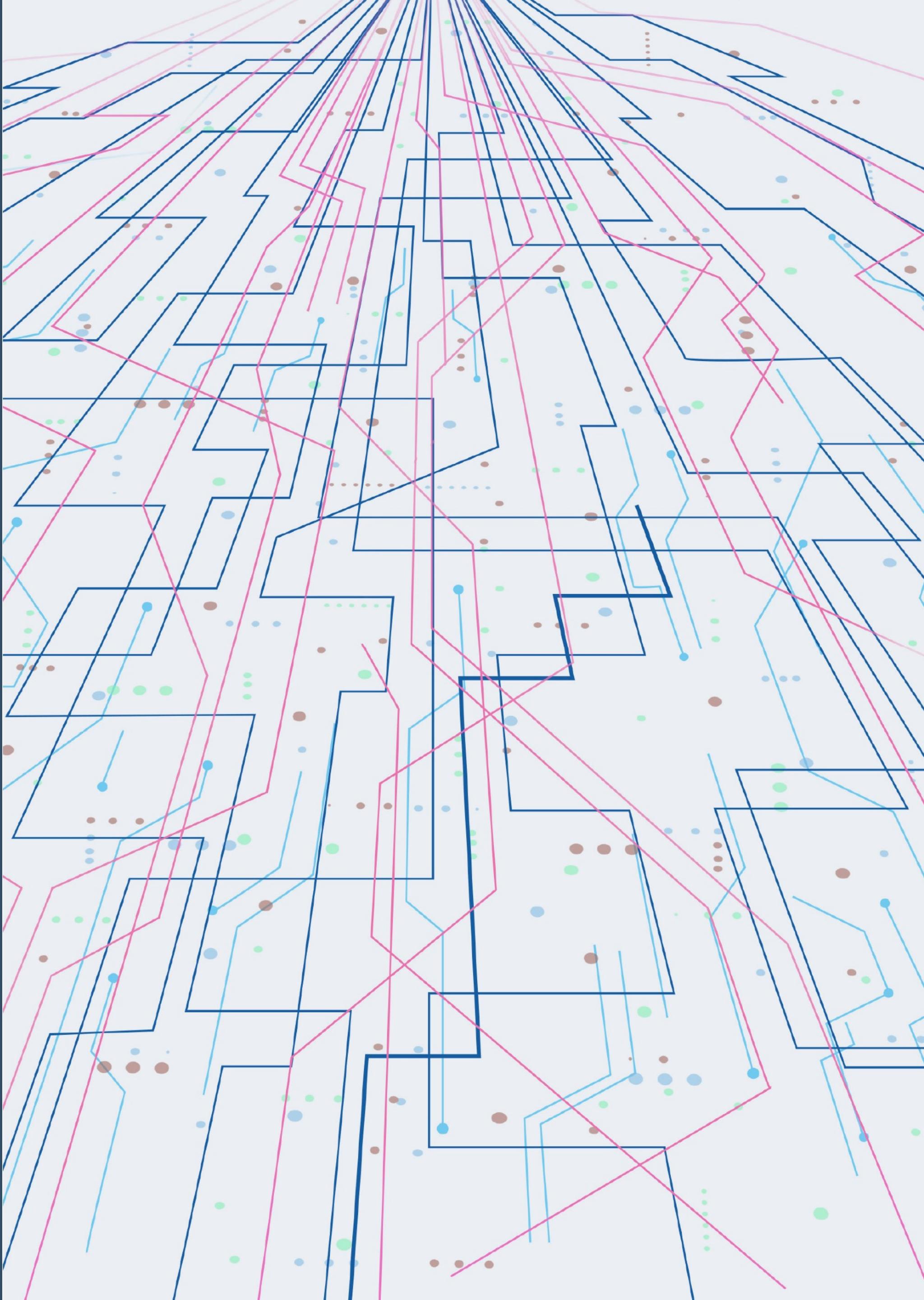
3

Platform Development



Section 1

What is the Turing Commons?



The screenshot shows the homepage of the Turing Commons website. The header features a dark teal bar with the "Turing Commons" logo, a search bar, and a GitHub link. Below the header, a navigation menu includes "Home", "Welcome", "Responsible Research and Innovation", "Public Engagement of Data Science and AI", and "Blog". The main content area has a white background with a large, central, circular illustration depicting a complex network of people, buildings, vehicles, and data storage units. To the left of the illustration, the text "Welcome to the Turing Commons" is displayed, followed by a description: "An online platform to support open dialogue and reflection about the responsible design, development, and deployment of data-driven technologies." A green "Get Started" button is located below this text. At the bottom of the page, a black footer bar contains the copyright notice "Copyright © 2022 Alan Turing Institute" and "Made with Material for MkDocs Insiders".

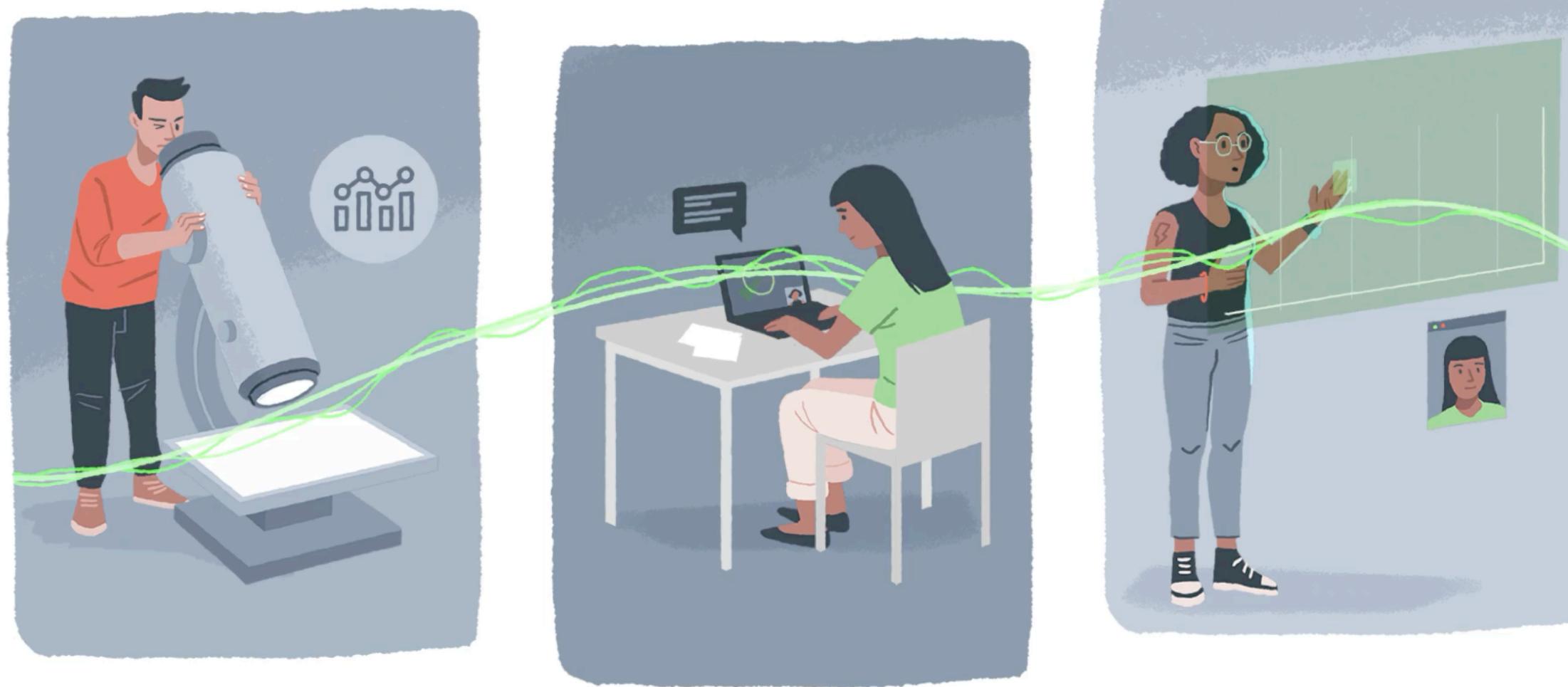
Online Platform

Turing Commons

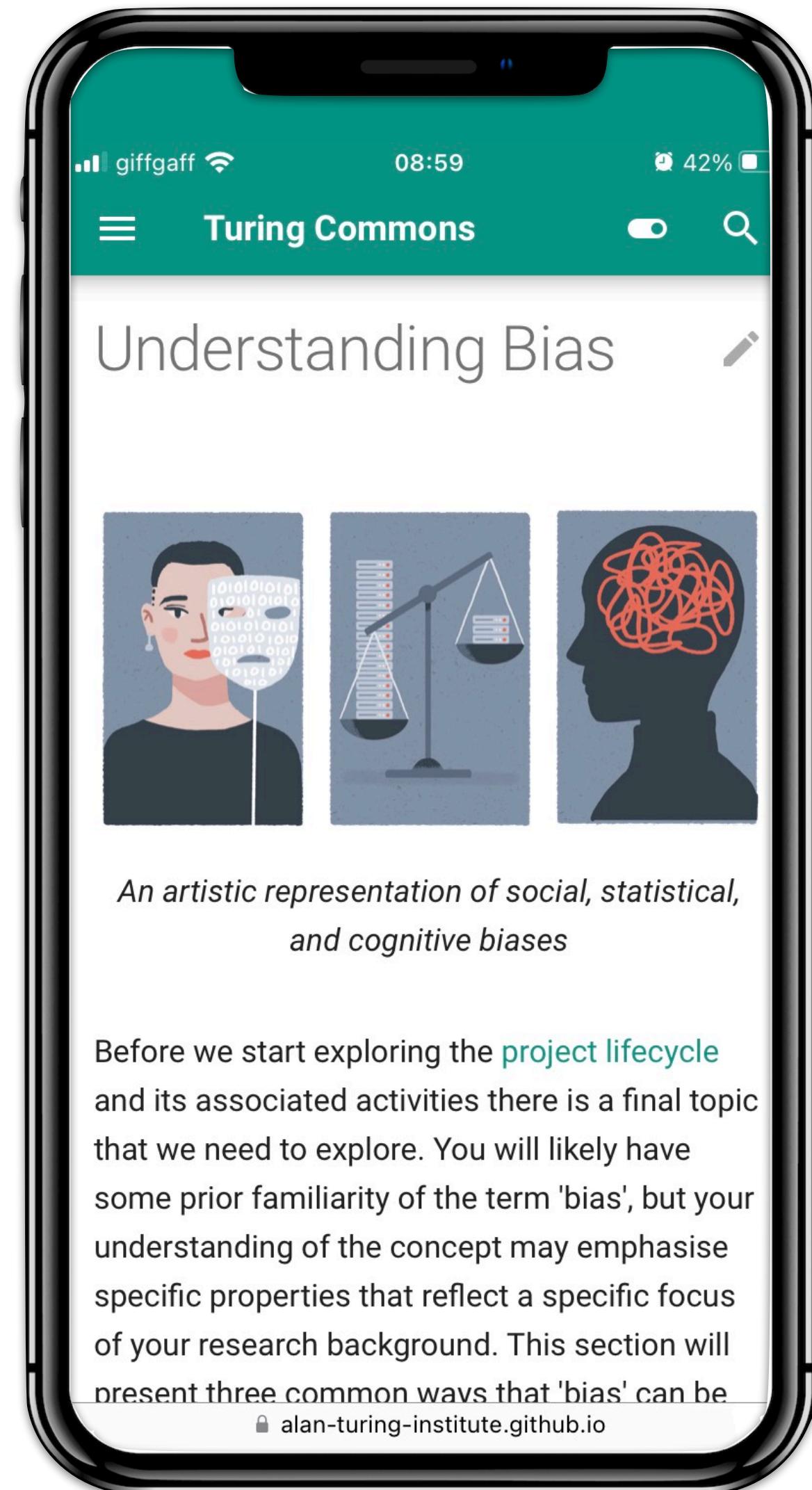
An online platform to support open dialogue and reflection about the responsible design, development, and deployment of data-driven technologies.

Courses

Researchers and Developers



- ① Responsible Research and Innovation
- ② Public Engagement of Data Science and AI
- ③ AI Ethics and Governance



The browser window has a teal header with the 'Turing Commons' logo and a search bar. The main content area is titled 'Data Analysis'. On the left is a sidebar with a tree view of topics: Responsible Research and Innovation, Responsible Data Science and AI, and The Project Lifecycle. The 'Data Analysis' section contains text about EDA, a note about Python/Jupyter notebooks, and a 'launch binder' button. To the right is a 'Table of contents' sidebar listing various sub-topics from 'What is Exploratory Data Analysis?' to 'Forms of Selection Bias (and more missing data)'.

Data Analysis

This module will introduce and examine the project lifecycle stage, 'data analysis' (or, exploratory data analysis). However, rather than focusing solely on the statistical or technical techniques employed in modern data analysis, we will approach this stage with a contextual, bias-aware perspective. That said, we will make use of [Jupyter notebooks](#)—a popular tool in data science—to aid our exploratory data analysis by visualising some data.

You do not need to be familiar with either Python or Jupyter Notebooks if you just want to gain an understanding of how social, cognitive, and statistical biases interact and affect downstream stages in the research and innovation lifecycle. But the code is presented for those who wish to get more "hands-on". You can also explore this code in an interactive Jupyter notebook: [launch binder](#)

Table of contents

- What is Exploratory Data Analysis?
- COVID-19 Hospital Data
- Importing Data
- Describing the Data
- Cleaning the Data
- Removing Unnecessary Variables
- Removing Outliers
- Handling Missing Data
- Analysing the Data
- Visualising the Data
- Dependency on other factors
- Forms of Selection Bias (and more missing data)

What is Exploratory Data Analysis?

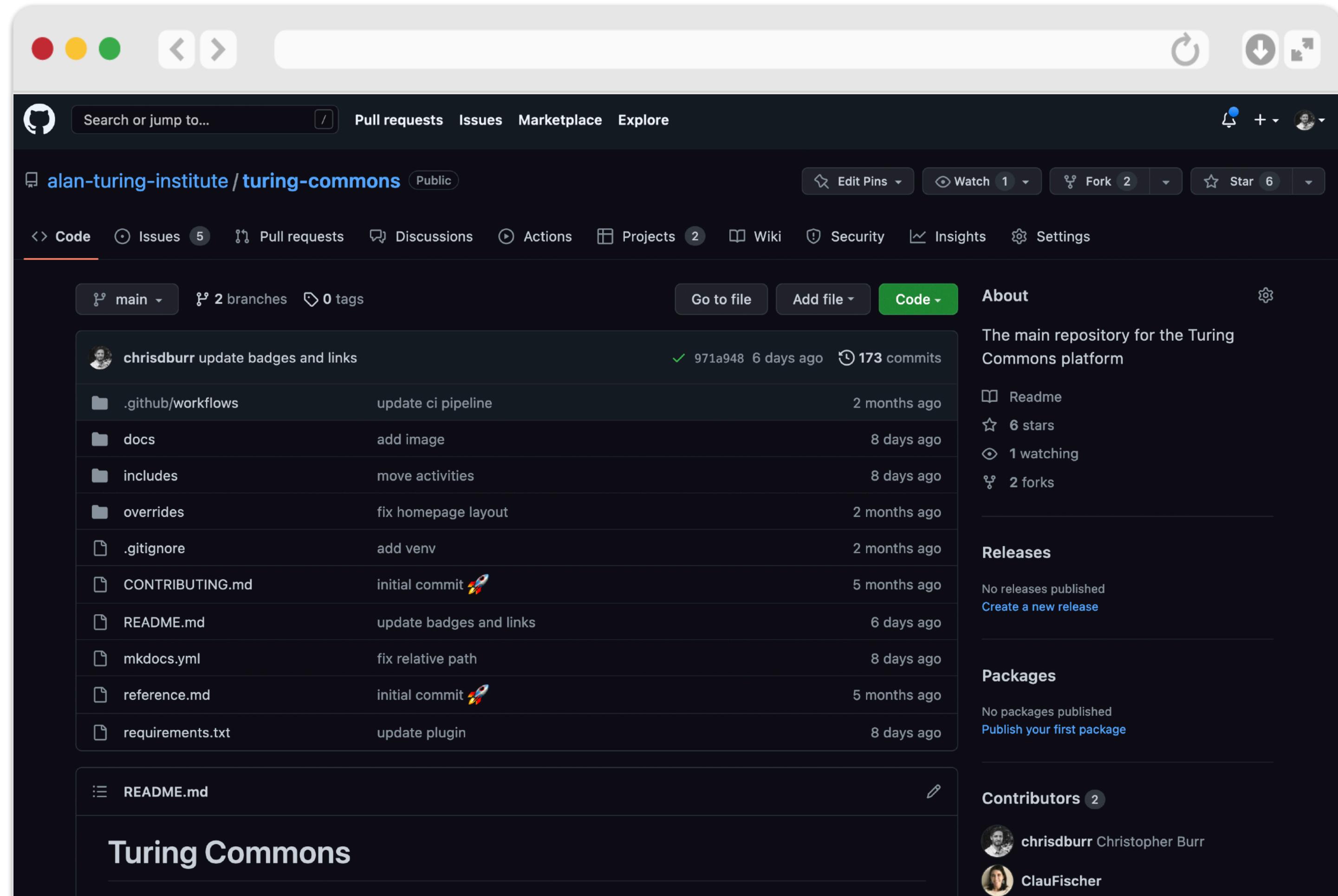
Exploratory data analysis (EDA) is a crucial stage in the project lifecycle. It is where a number of techniques are employed for the purpose of gaining a better understanding of the dataset and any relationships that exist between the relevant variables. Among other things, this could mean,

- Describing the dataset and important variables
- Cleaning the dataset
- Identifying missing data and outliers, and deciding how to handle them
- Provisional analysis of any relationships between variables
- Uncovering possible limitations of the dataset (e.g. class imbalances) that could affect the project

We will cover each of these sub-stages of EDA briefly, but to reiterate, our primary focus in this

Feedback

“I feel the case study aspect added a lot of value—it allowed us to apply the principles we had learnt within the course and strengthen our understanding of these principles.”



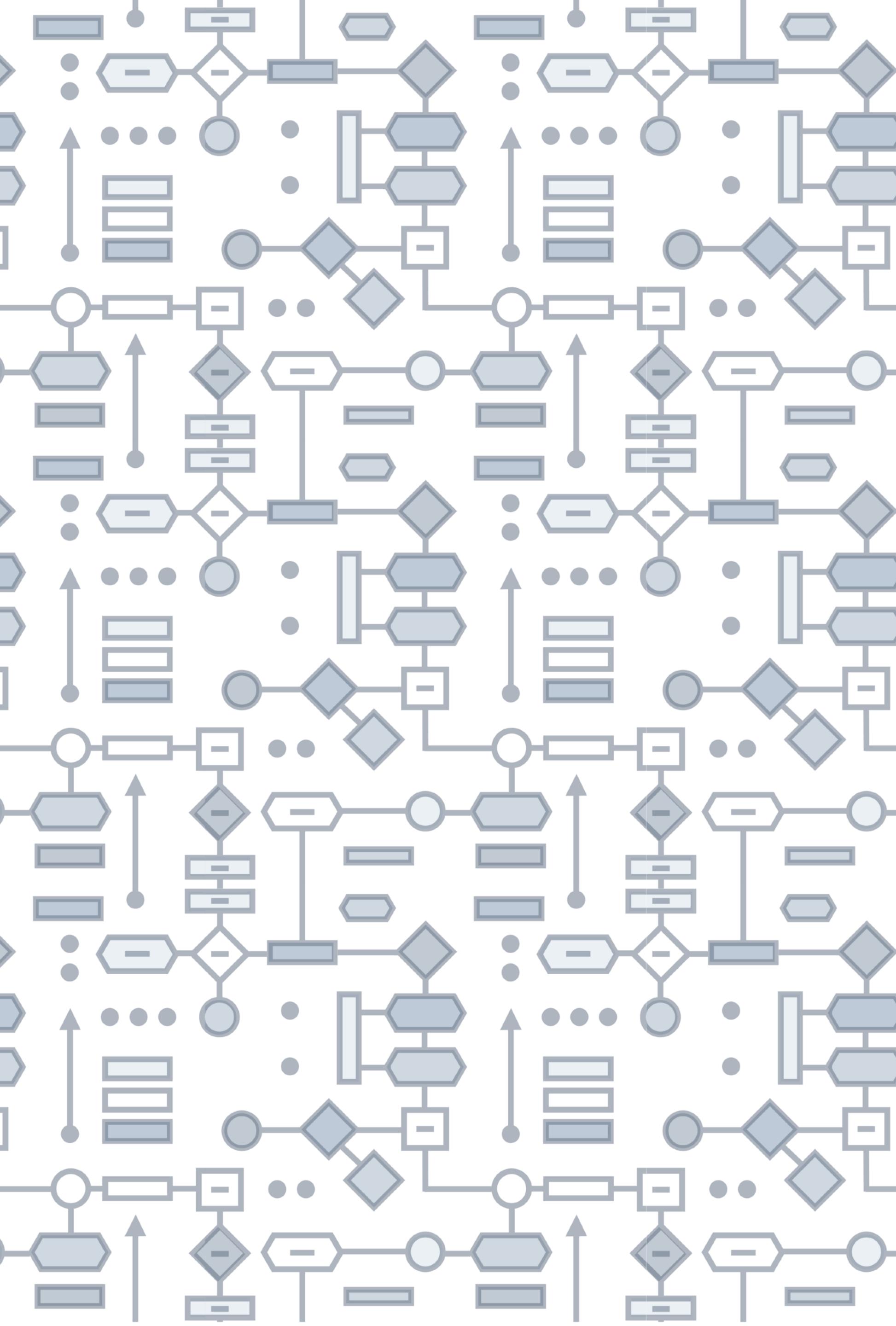
Open Infrastructure

GitHub and Open Access

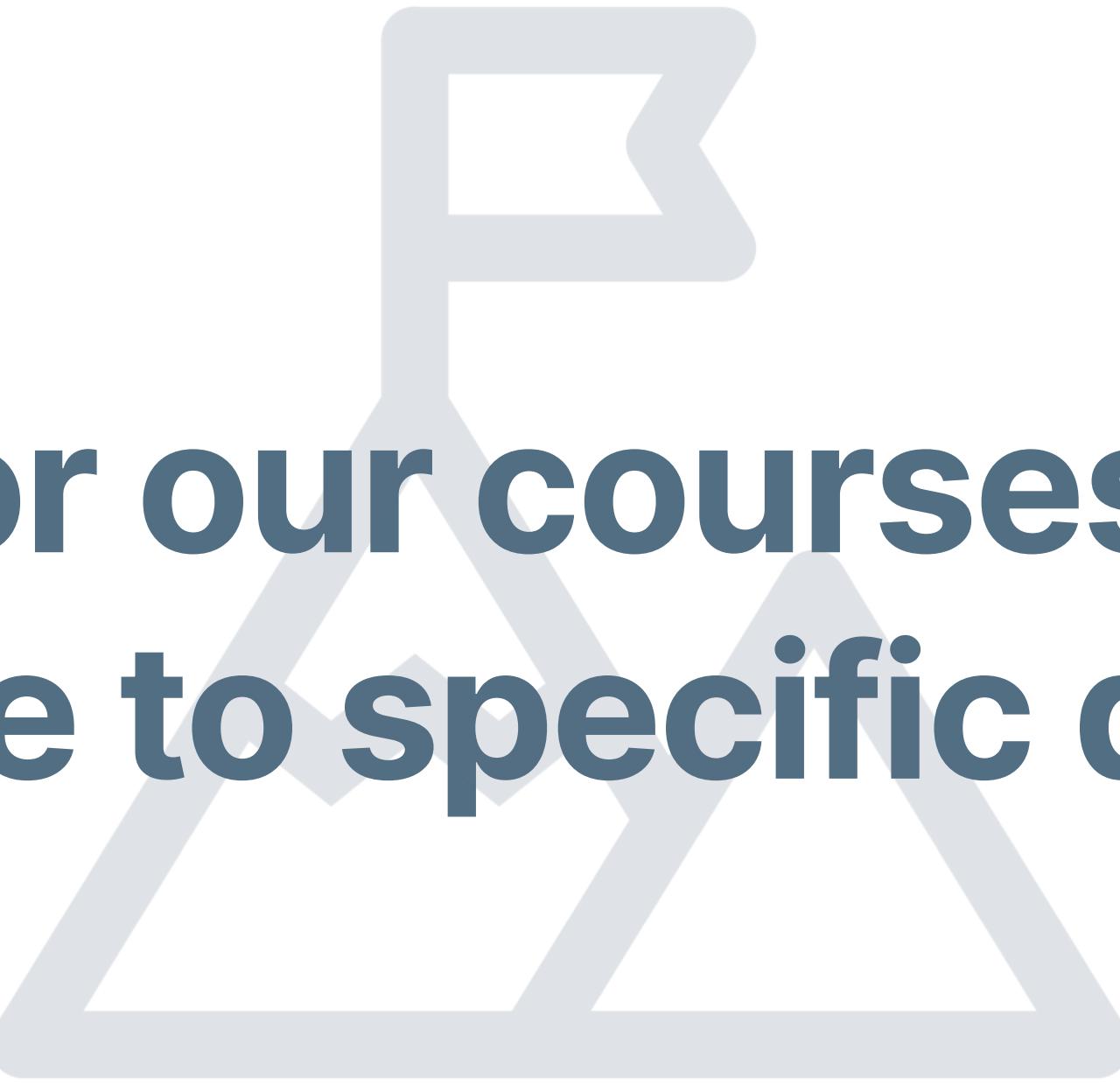
- All resources are freely available through repository (CC BY-4.0 license)
- Community building and engagement through public repository
- Also allows specific contribution to platform through pull requests
- Blog articles to improve public impact and engagement

Section 2

Centres for Doctoral Training



Challenge



**How do we tailor our courses to make them
more accessible to specific disciplines?**



Core Modules

Responsible Research and Innovation

1. What is Responsible Research and Innovation?
2. The Project Lifecycle Model
3. Fair Machine Learning and AI
4. Explaining Automated Decision-Making
5. Responsible Communication and Open Science

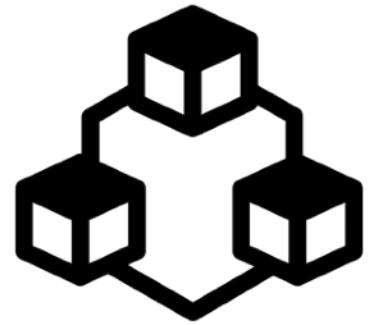


Structure Skills Tracks

- Adaptable and scalable versions of our courses to meet domain-specific needs:
 - Core modules
 - Activities
 - Case Studies

Skills Track

Core Modules



Activities

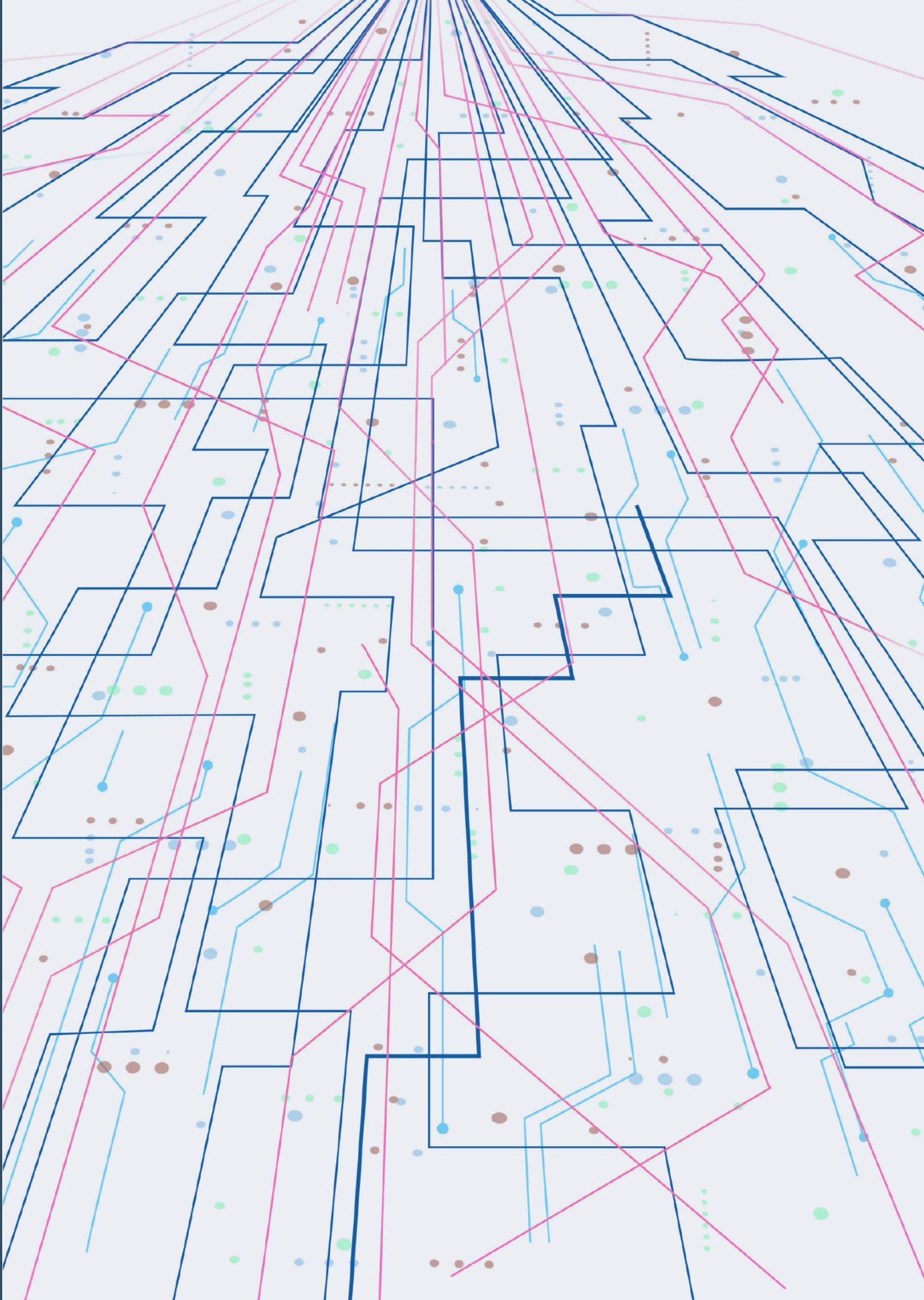


Case Studies



Section 3

Platform Development

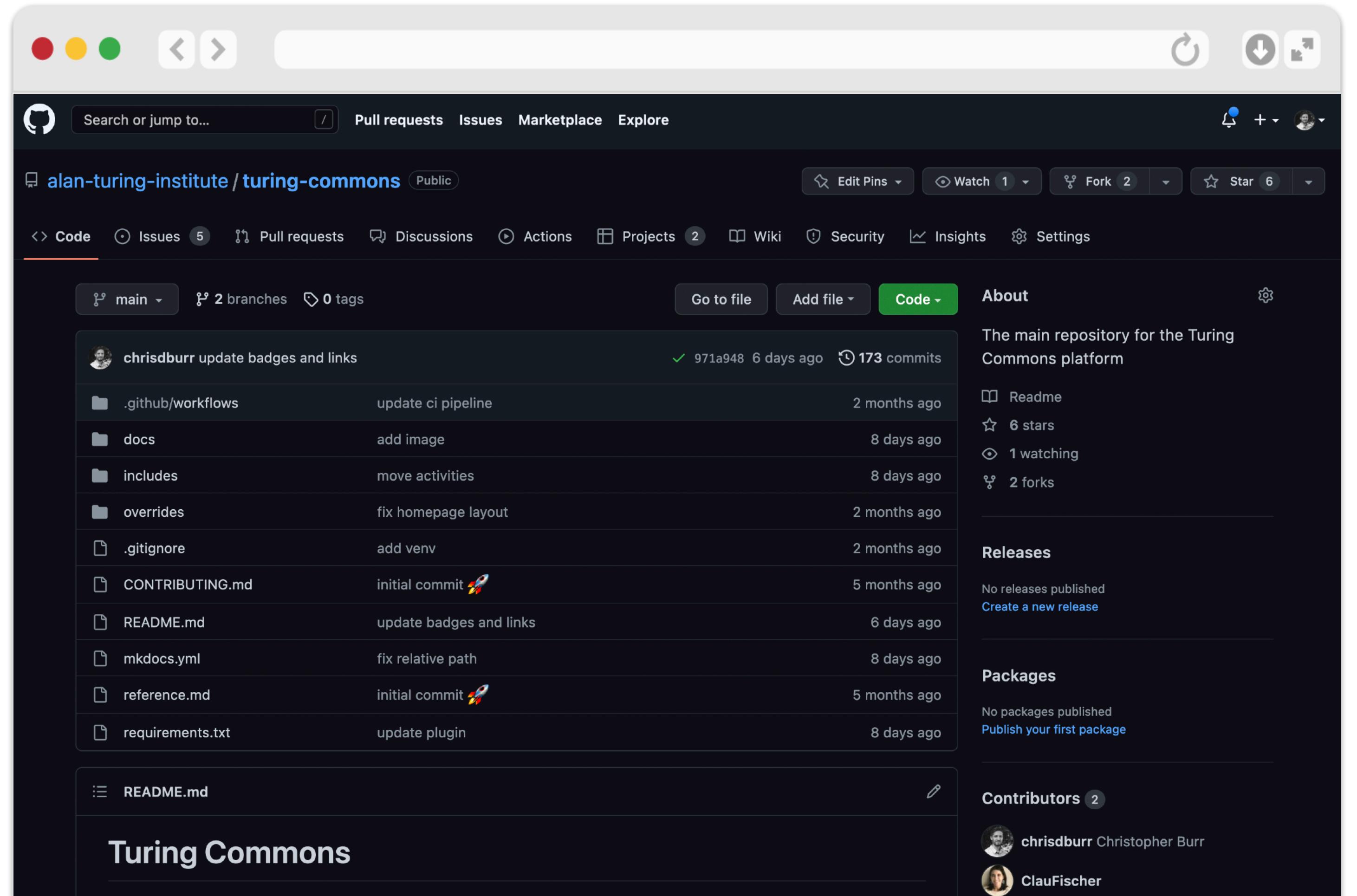




Next Steps

Work Streams

- 1 Increased Engagement
- 2 Improved Open Infrastructure
- 3 Platform Evaluation



Open Infrastructure

GitHub and Open API

- Active research project to support open infrastructure and open science
- Development of API to enable community contributions (e.g. design of new case studies)
- Support for ongoing best practices in open data science
- Support for interactivity and self-directed learning (e.g. online leaning environment)

Decision Support System

The Alan Turing Institute

Using natural language processing to help psychiatrists assess patients

The system uses a modern form of natural language processing that is reliant on neural networks to identify salient features of a patient's speech. This includes the words and phrases the patient states as well as extra-linguistic properties, such as intonation or pace.

The system operates in real-time to make recommendations to the psychiatrist. The recommendations can include suggestions about specific topics that were raised by the patient, which may require follow-up (e.g. reference to problematic relationships), as well as generic features that may be informative. Many recommendations require additional interpretation from the psychiatrist, and no automated decisions are made by the system.

The system's outputs can be shared with the patient where explanations are required. As such, the system has been designed to support patient-psychiatrist communication and participatory decision-making. For instance, phrases that were flagged as relevant are emphasised and specific recommendations are also accompanied by confidence ratings that can be explained by a trained psychiatrist.

© The Alan Turing Institute

```
{
  "title": "Decision Support System",
  "domain": "healthcare",
  "userGroup": 2,
  "personas": [
    "Psychiatrist",
    "Patient",
    "Developer"
  ],
  "isLive": true
}
```

Overview

Psychiatrists working for a national healthcare system have been provided access to an AI system that can support decisions made during initial patient assessment and diagnosis.

The system uses a modern form of natural language processing that is reliant on neural networks to identify salient features of a patient's speech. This includes the words and phrases the patient states as well as extra-linguistic properties, such as intonation or pace.

The system operates in real-time to make recommendations to the psychiatrist. The recommendations can include suggestions about specific topics that were raised by the patient, which may require follow-up (e.g. reference to problematic relationships), as well as generic features that may be informative. Many recommendations require additional interpretation from the psychiatrist, and no automated decisions are made by the system.

The system's outputs can be shared with the patient where explanations are required. As such, the system has been designed to support patient-psychiatrist communication and participatory decision-making. For instance, phrases that were flagged as relevant are emphasised and specific recommendations are also accompanied by confidence ratings that can be explained by a trained psychiatrist.

© The Alan Turing Institute

Key Consideration

The system has been designed for use as a *decision support system*.

Are there properties of the system that could negatively impact the ability for the psychiatrist to perform their clinical duties effectively?

Deliberative prompts

- If the system was designed from scratch, how should psychiatrists and patients be involved in its design, development, and deployment?
- Why does it matter that the system functions as a decision support tool and not an automated decision-making system?
- Should patients have full access to the outputs of the decision support system?

© The Alan Turing Institute

Datasheet

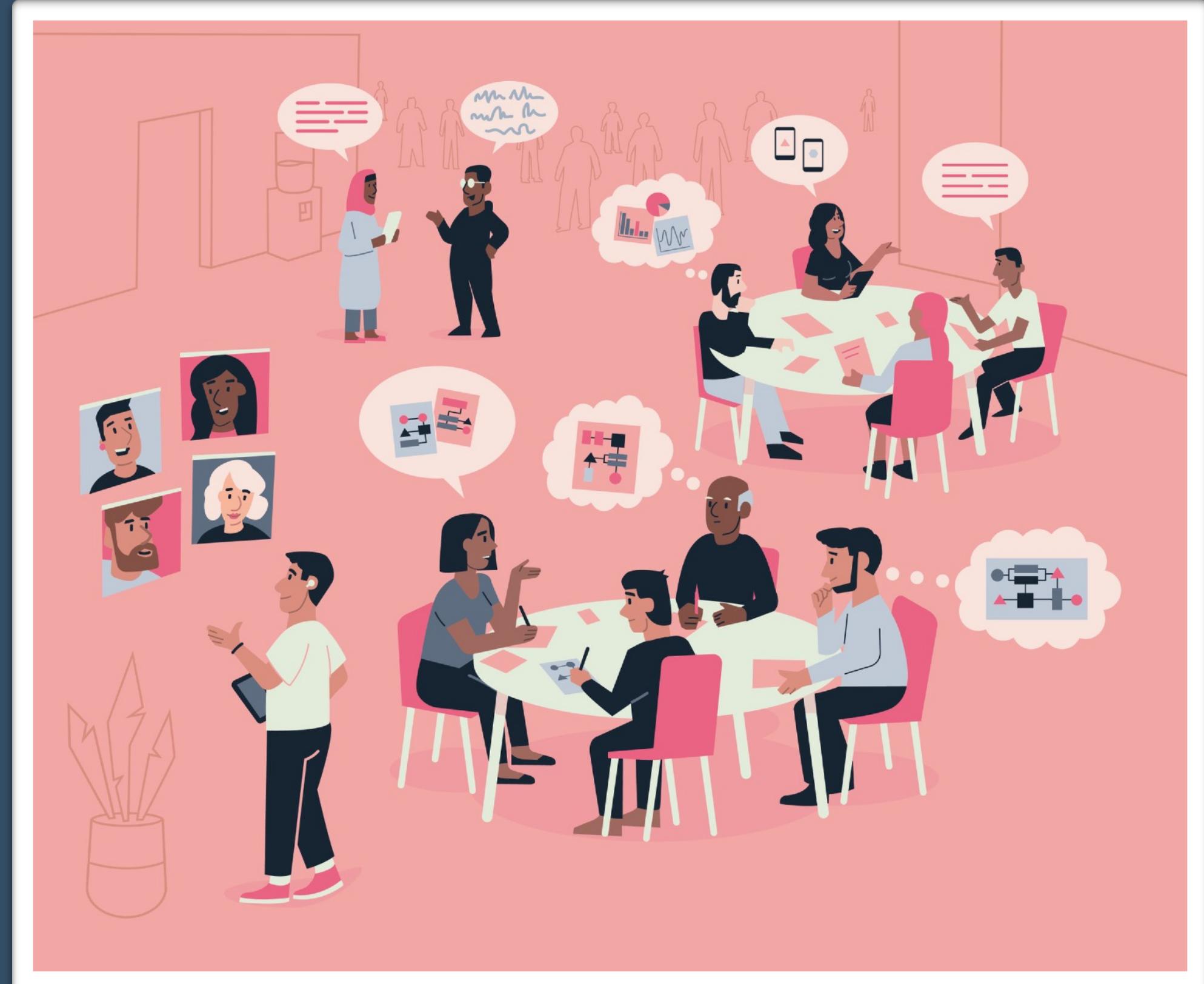
Category	Details
Available Data	<ul style="list-style-type: none"> Automated transcription of the conversations between patient and psychiatrist Extracted extra-linguistic features from audio recording of the conversation Relevance feedback from the psychiatrist about salience of specific recommendations Electronic health record of patient, including any initial assessment data (e.g. PHQ-9) or prescriptions.
Analysis Techniques	<ul style="list-style-type: none"> Natural language processing (NLP): <ul style="list-style-type: none"> System uses convolutional neural networks for speech recognition and to identify features from audio recording, which are then classified according to whether they are informative (i.e. exceeding some relevance threshold) Visualisation techniques support explainability by highlighting salient features

© The Alan Turing Institute

Groups, Organisations and Affected Individuals

- Patients
- Psychiatrists

© The Alan Turing Institute



Engagement Wider Reach and Impact

- Researchers (e.g. additional disciplines, additional stages of education)
- Members of the public
- Developers and Industry
- Public Sector
- Wider groups and professions (e.g. journalism)

Evaluation



Is our platform working for our partners and audience? How can we measure its impact?

Thank you!

The
Alan Turing
Institute