Building a better working world

What if the start of your career could be the end of a global problem?

Don't just predict the future.
Build a better one.

The better the question.
The better the answer.
The better the world works.

Welcome to the Better Working World Data Challenge!

This year, thousands of people around the world will take part in this challenge using data science and remote sensing to help bushfire authorities streamline fire mapping.

You're in good company
Thank you for joining a passionate group of people on this purpose driven mission. In addition to helping bushfire authorities build capacity, you are helping to build digital skills to take on the world's toughest challenges.

Team up
We hope you enjoy this opportunity to make friends and learn new skills. You are welcome to form a team of up to four people with other participants and collaborate on the challenge.

Benefits for you
Aside from up to $10,000 USD cash prize money, you have an opportunity to access up to $850 Azure credits and a voucher for achieving a Microsoft Azure certification.

On your marks, get ready!
On 24 March 2021, submissions are now open on the EY Data Science Platform. When you submit results, you will earn a place on the leaderboard!

This Welcome Booklet provides an overview of the 2021 Data Challenge and the steps you can take to get started.

Good luck!

The Better Working World Data Challenge team

This document contains a description of the 2021 Challenge and links to the resources you'll need to take part.
There are two related challenges available for you to complete. Both challenges are designed to help streamline the process of bushfire mapping for bushfire authorities.

- Challenge 1: Fire mapping from airborne imagery
- Challenge 2: Fire behavior modelling

Each challenge has a separate leaderboard. You can either compete in both challenges or focus on Challenge 1.

Winners of the 2019 NextWave Data Science Challenge in New York City with Beatriz Sanz Saiz (Global Data & Analytics Leader, EY) and Carmine Di Sibio (CEO, EY).

Challenge 1. Fire mapping from airborne imagery

During bushfire season, bushfire managers require timely, accurate information about the location and rate of spread of active fires. To collect this information, aircraft carrying infrared cameras fly over and record the intensity and location of fires. The camera scans the fire in lines to construct an image. This image is known as an 'infrared linescan' and are currently considered one of the best sources of information about fire intensity and location.

Upon receipt of these images, the Country Fire Authority (CFA) label the fire boundaries in each linescan image manually by hand-drawing polygons around the edges of the fire using geospatial software. During times of intense firefighting activity, this process can create a bottleneck in delivering timely information to operational firefighting teams.

Your first challenge is to create a system that performs this mapping exercise automatically, allowing for more effective allocation of human CFA resources. In computer vision, this task is known as 'image segmentation'.

To help you, the mapping officers at Victoria's CFA have shared some fire-edge polygons from the 2018/19 fire season. An area enclosed by any one of these polygons has been deemed on fire by the CFA. You can use the polygons provided by the CFA as calibration and validation data to build your model.

We will test your system by giving you a set of coordinates and asking you to tell us which coordinates are on fire, and which are not on fire.

Step	Topic	What we'll provide	Finding the resources	What you'll do	What you'll achieve
1.1	Accessing infrared linescan images	▸ A guide to using Jupyter and our custom python libraries to query infrared linescan images for selected geographic regions.	▸ 01_Beginners_guide/01_Jupyter_notebooks.ipynb ▸ 03_EY_challenge1/Challenge1_Getting_started.ipynb	Work through example Jupyter notebooks that showcase the basic features of Jupyter before exploring how our custom python libraries can load and display selected linescan images.	Capability to collect and visualize infrared linescan images
1.2	Detecting fires in infrared linescan images	▸ Suggestions for getting started with analysing linescan images to detect fire.	▸ 03_EY_challenge1/Challenge1_Getting_started.ipynb	Explore the fire mapping polygon dataset and develop a basic model to detect fire areas in infrared linescan images.	An understanding of the fire mapping polygon dataset and its limitations. A basic approach to detecting fire in infrared linescan images
1.3	Submitting results	▸ A guide to making a submission on the EY Data Science platform.	▸ 03_EY_challenge1/Challenge1_Getting_started.ipynb	Submit results from a basic fire detection model.	A place on the leaderboard. Refine your model over time to achieve higher scores.

Challenge 2.
Fire behavior modeling

While infrared linescan imagery is still one of the best sources of information about fire intensity and location, there are times when it is not possible to acquire. This may be because there are not enough aircraft available to map all actively burning fires, or because the weather conditions are too poor, making it unsafe to deploy an aircraft. An alternative source of images for fire mapping is via satellite. The availability and resolution of satellite imagery has increased substantially in recent years. This makes it possible to monitor bushfires from space. Satellite imagery is not always available, and won't replace airborne imagery soon, but is a valuable complement.

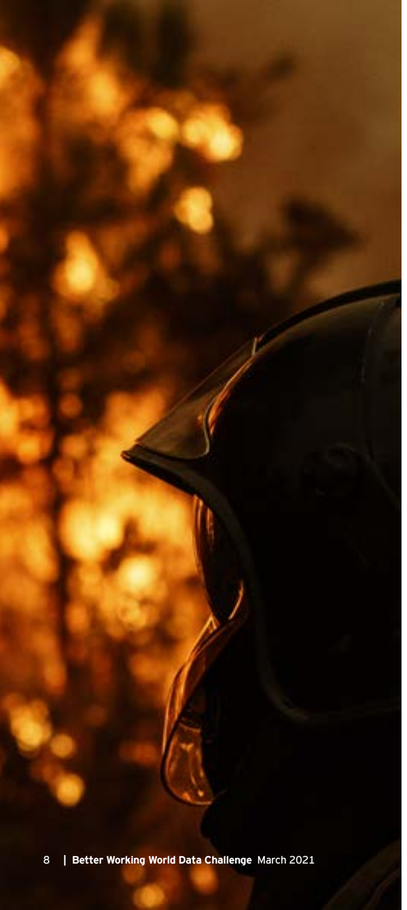
While the number of satellites in orbit is continuing to increase, especially with commercial operators expanding their service offerings, there are still long periods when there is no coverage of bushfire affected areas. For times when neither linescan nor satellite data are available, it is possible to extrapolate from previous observations to forecast the current location of the fire. It can also be useful for firefighting teams to forecast future locations of fires based on current observations. In computer vision, this task is known as 'next frame prediction'.

Your challenge is to create a model capable of predicting the evolution of a fire region from a given time sequence of linescan images. You are encouraged to complement this sequence with satellite imagery as well as any external data of your choosing.

There are many sources of satellite data available, so we have curated a set of useful satellite imagery for you. However, you may use other sources of satellite imagery if you like. You can use your model developed in challenge 1 in addition to the CFA polygons as calibration and validation data to build your model.

We will test your system by giving you a set of coordinates and asking you to tell us which coordinates are on fire, and which are not on fire.

Step	Topic	What we'll provide	Finding the resources	What you'll do	What you'll achieve
2.1	Accessing satellite images	▸ A guide to using our custom libraries to query the Open Data Cube (ODC) and return satellite images for selected geographic regions	▸ All notebooks in the '01_ Beginners_guide/' directory	Load selected satellite images following an example Jupyter notebook.	Capability to query and visualize satellite images
2.2	Exploring real-world examples and starting challenge 2	▸ Suggestions for getting started with analyzing satellite images to detect fire.	▸ 04_EY_challenge2/ Challenge2_Getting_started.ipynb ▸ All in '02_Real_world_examples/' especially 'Burnt_area_mapping. ipynb'	Explore examples of satellite data being used to solve real world problems. Load satellite data to begin challenge 2.	A starting point to begin solving challenge 2.
2.3	Submitting results	▸ A guide to making a submission on the EY Data Science platform.	▸ 04_EY_challenge2/ Challenge2_Getting_started.ipynb	Submit results from your fire movement prediction model.	A place on the leaderboard. Refine your model over time to achieve higher scores.



Getting Started

Before you begin digging into the data, there are a few steps you need to take.

0.1 Learn about bushfire mapping and how you can help (see pages 4-7 of this booklet)

0.2 Deploy your computational environment in Azure
<via the official GitHub repository>

0.3 Sign up for the EY Data Science Platform. This is where you'll submit your results
<EY Data Science Platform>

Preparing for the challenge

Step	Topic	What we'll provide	What you'll do	What you'll achieve
0.1	Bushfire mapping and how you can help	▸ An overview of the bushfire mapping challenge	Learn about the current approach to fire mapping and the opportunity to help streamline the system.	An understanding of your opportunity to help improve bushfire management using Data Science
0.2	Deploying your Data Science environment	▸ A magic link and instructions to deploy your computational environment in Azure	Set up an environment in Azure and open Jupyter notebooks.	Access to a suitable compute environment
0.3	Competing and winning	▸ Link to the EY Data Science Platform used to submit and score your results	Learn how your results will be assessed on a leaderboard. Log in to the EY Data Science platform and form a team if desired.	An opportunity to have fun, learn new skills and maybe even win a place in the finals!

2021 Better Working World Data Challenge

Participant Journey Map

Participant finds out about challnge

Register

- Participant arrives at the **EY Challenge website** and is provided with an overview of the challenge purpose

EY Data Science Platform

- Participant is directed to the EY Data Science Platform to register for the challenge.

Get Started

- Participant is provided with a link to a **Github page** that contains instructions for set up.

Github "getting started" page

- Azure subscription
- Virtual Machine
- Information on Open Data Cube

- Participant **forms / joins a team** on the Data Science Costramer platform

Experiment

- Participant attempts challenge
- **FAQ** page for frequently asked questions; **additional support** available for other questions (eg - infrastructure FAQs)

Submit

- **Submissions** are made on the challenge platofrm (one submission per day)
- The top 200 scores and awarded **additional credit**.

Submission page

Finish

- Technical challenge is closed
- **Finalists are announced** and final steps for video presentation are communicated.

2021 Better Working World Data Challenge

Support Ecosystem

0 Participant finds out about challnge

1 Register

- Welcome booklet
  - this booklet
- Videos
  - Inspirational content (eg. Why is the challenge important?)

2 Get Started

- Getting started docs
  - **GitHub** (eg. Setting up Azure student subscription, deploying data science environment)
- Videos
  - **Educational content** (eg. image analysis in Jupyter Notebook)

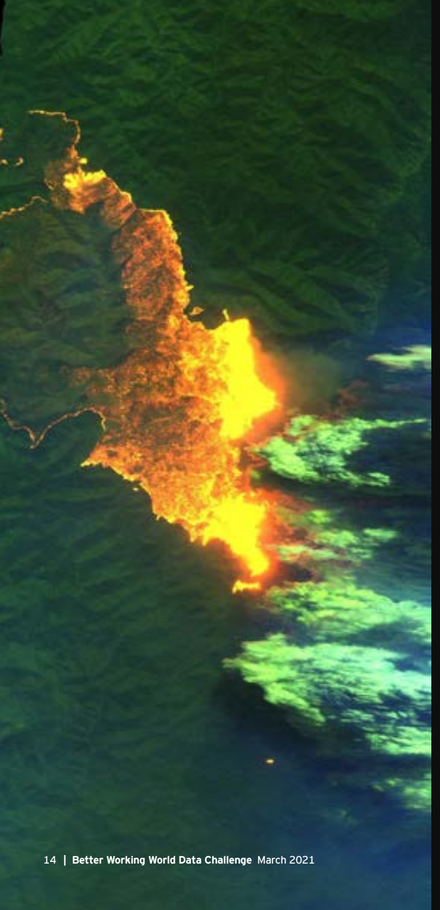
3 Experiment

- Challenge FAQs
  - **EY Data Science Platform** (eg. Where do I find the datasets?)
- Infrastructure FAQs
  - GitHub (eg. How do I turn off the VM?)
- Example Jupyter notebooks
  - **Your data science environment** (eg. How do I load images for analysis?)

4 Submit

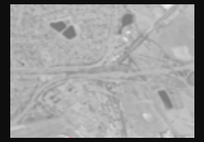
- Example Jupyter notebooks
  - **Jupyter Notebooks deployed via** GitHub (eg. How do I submit my answers to the platform?)

5 Finish

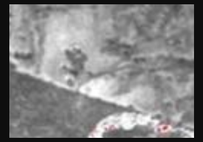


Linescan Interpretation Tips

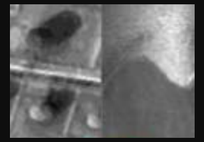
This guide, provided by Australia's Country Fire Authority, will help you interpret the infrared linescan images you will work with during the challenge.



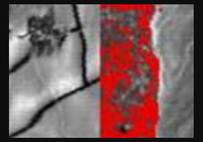
Roads, Streams, Trees Pasture
Streams and roads can be identified in Firescan images as linear features of a different temperature to the surrounding terrain. In general during the day roads will appear hotter than their surrounds. This contrast may last for much of the bight as the roads gives off its heat gained during the day. In general flowing steams will be cooler than their surrounds during the day and warmer at night, Standing water may be warm or cool.



Solar Heating
Areas of open ground and road surfaces will rise in temperature during the day, sometimes to temperatures well in excess of the daily maximum temperature. These solar heated areas can become so warm they may lead to false events and false fire indication. In the image shown fire is visible in the right corner. A solar heated field can be seen just above fire towards centre of image.

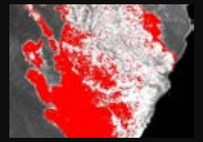


Sun Shadows and Reflected Radiation
The more sun light incident on a object the hotter it will appear. Objects which are in shadow will there appear cooler. During the day tall objects on the target surface will cast thermal shadows. Heavy smoke will also cast shadows in strong sunlight. Water and metal structures may reflect sunlight. This can be the cause of false events, and so false fire indications. Reflected solar illuminations is worst around midday.

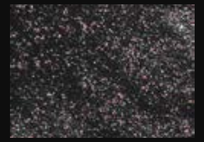


Wind Steaks and Smear
Wind Streaks occur down wind from obstructions on flat terrain and typically occur as warm patterns on an image. Wind velocity is lower down winds from obstructions which reduced the cooling effect of the wind. Surface wind may produce parallel curved lines of alternating lighter and darker intensity that may extend over large portions of an image. This is called wind smear.

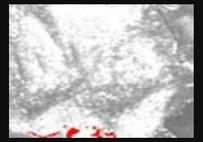
Cloud
Infra-red radiation is absorbed by water molecules. Any cloud or fog between the scanning aircraft and the ground will attenuate the infra-red radiation passing up from the ground. Heavy cloud may block out all such radiation. Clouds are usually seen in the imagery as areas of very low apparent temperature.



Active Fire
Any flame in the field of view of the scanner will raise the apparent temperature of the portion of the image in which they fall. If flames covers an entire pixel then the apparent temperature of the pixel will be that of the flame. Small fires can be overlooked as they may not raise the pixels apparent temperature by much. Active fire will tend to 'saturate' the infra-red detector. This will mean that the fire will be seen as a blob of one colour.



Burnt Area
One the fire front has past, the area burnt through will still contain a number of burning areas, stumps, tree boles, fence posts, etc. Heavily timbered dry fuel will continue to smoulder for weeks while a grass or heath land may burn out in minutes. These remaining fires appear as speckles on a background that appears washed out and low contrast. Fire burn patterns are sometimes visible in these areas.



Fire Cooled Area
Once area is burnt considerable ash and charcoal may be left behind, the more severe the fire the less vegetation will be left.

Good luck and have fun!

Thank you to our collaborators for making this challenge possible.

EY | Building a better working world

EY exists to build a better working world, helping to create long-term value for clients, people and society and build trust in the capital markets.

Enabled by data and technology, diverse EY teams in over 150 countries provide trust through assurance and help clients grow, transform and operate.

Working across assurance, consulting, law, strategy, tax and transactions, EY teams ask better questions to find new answers for the complex issues facing our world today.

ey.com