

TRUST

a building block of human
socio-economic dynamics

Andrew Popp

andrew.popp@ca.ibm.com





to gain trust
in a model
we need to
understand
why it made
a prediction

just because
you are
using an
algorithm
doesn't
mean it's
going to
be fair
and just



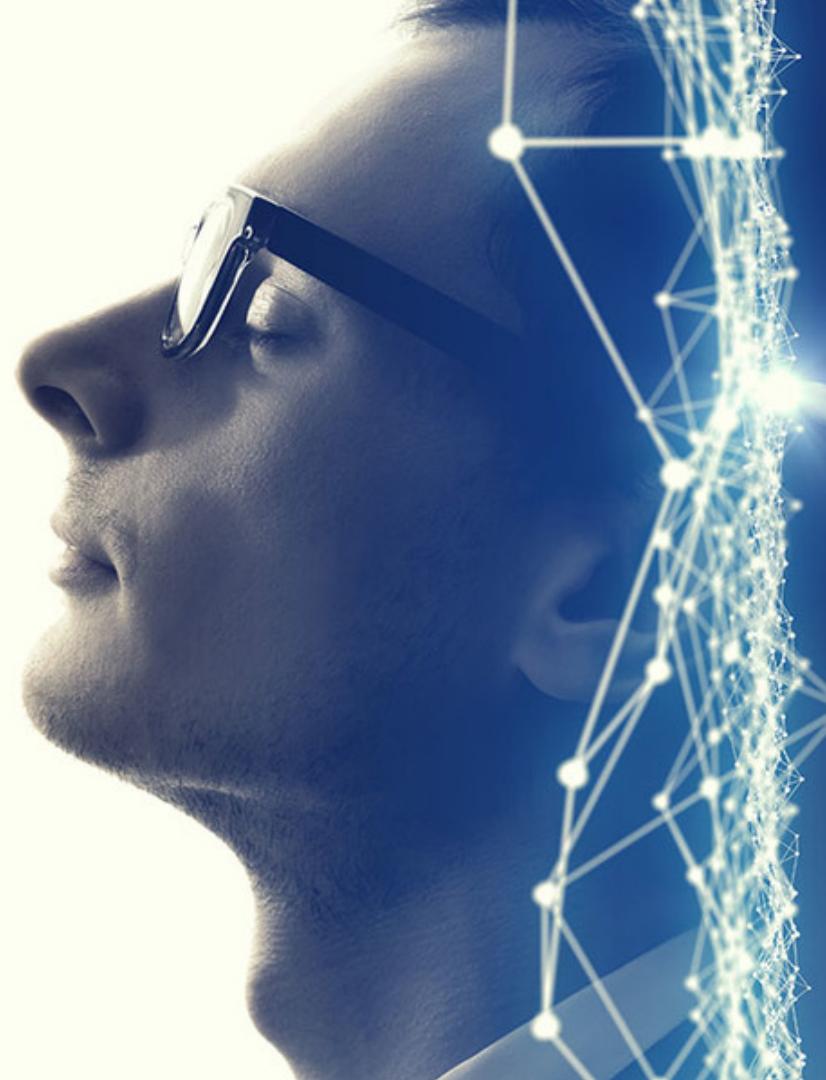


FAIRNESS

AI systems should ...

use training data and models that
are free of bias, to avoid unfair
treatment of certain groups.

The screenshot shows a user interface for a translation application. At the top, there are language selection buttons for English, Spanish, French, and Detect language, followed by a dropdown menu and a 'Translate' button. Below the buttons is a text input area with a character count of 0/5000. The input field contains a single space character. To the right of the input field is a large, empty rectangular box representing the output or translated text area.





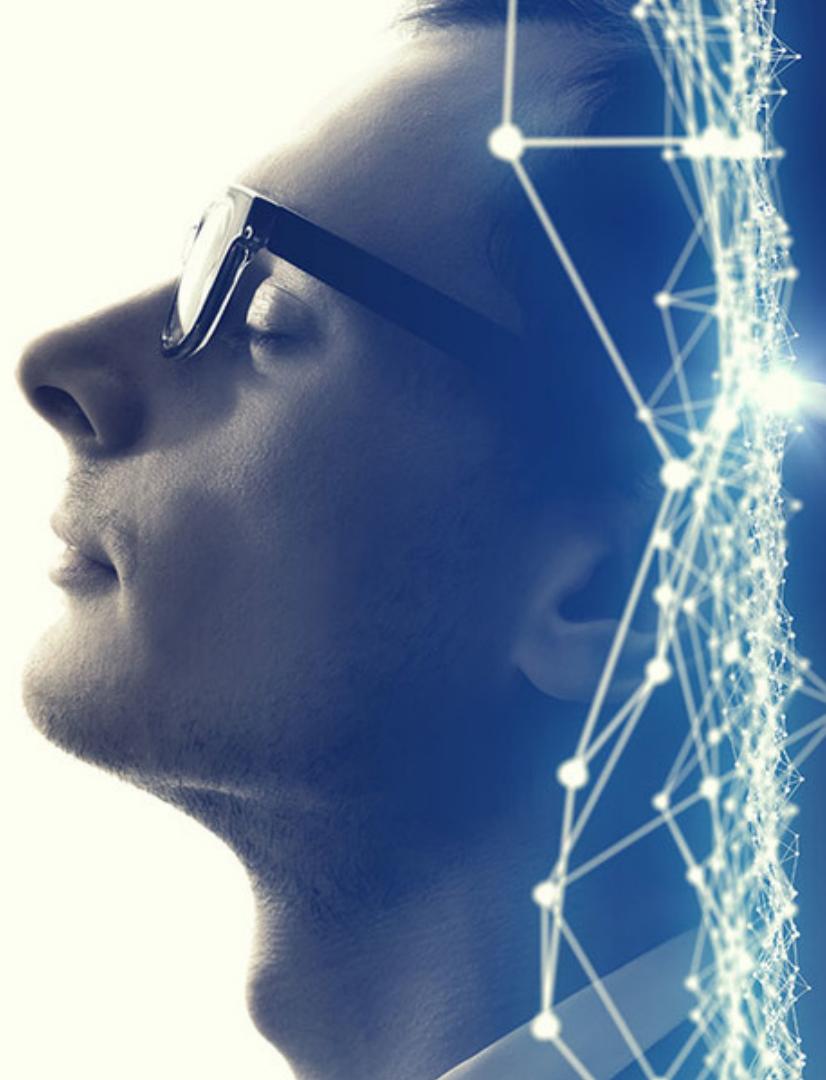
Supplier's Declaration
of Conformity (SDoC)

FAIRNESS

EXPLAINABILITY

AI systems should ...

provide decisions or suggestions
that can be understood by their
users and developers.





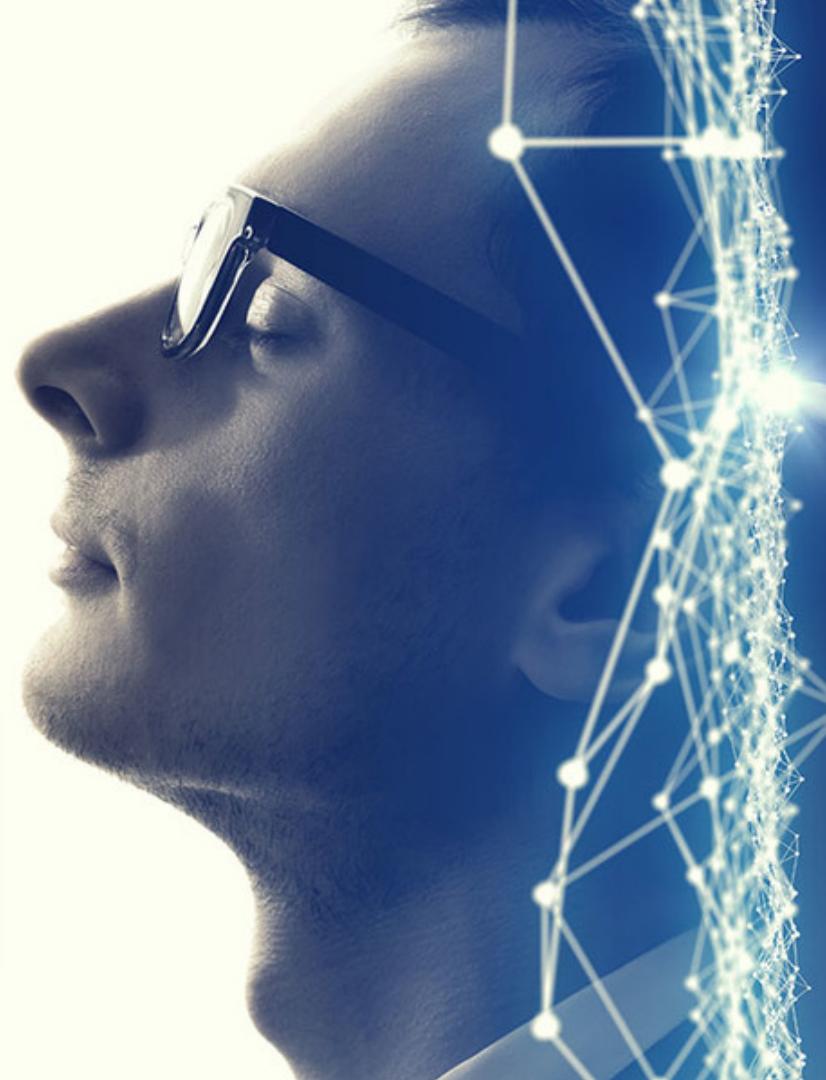
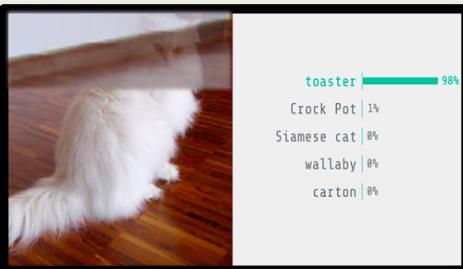
Supplier's Declaration
of Conformity (SDoC)

FAIRNESS EXPLAINABILITY

ROBUSTNESS

AI systems should ...

be safe and secure, not vulnerable
to tampering or compromising the
data they are trained on.







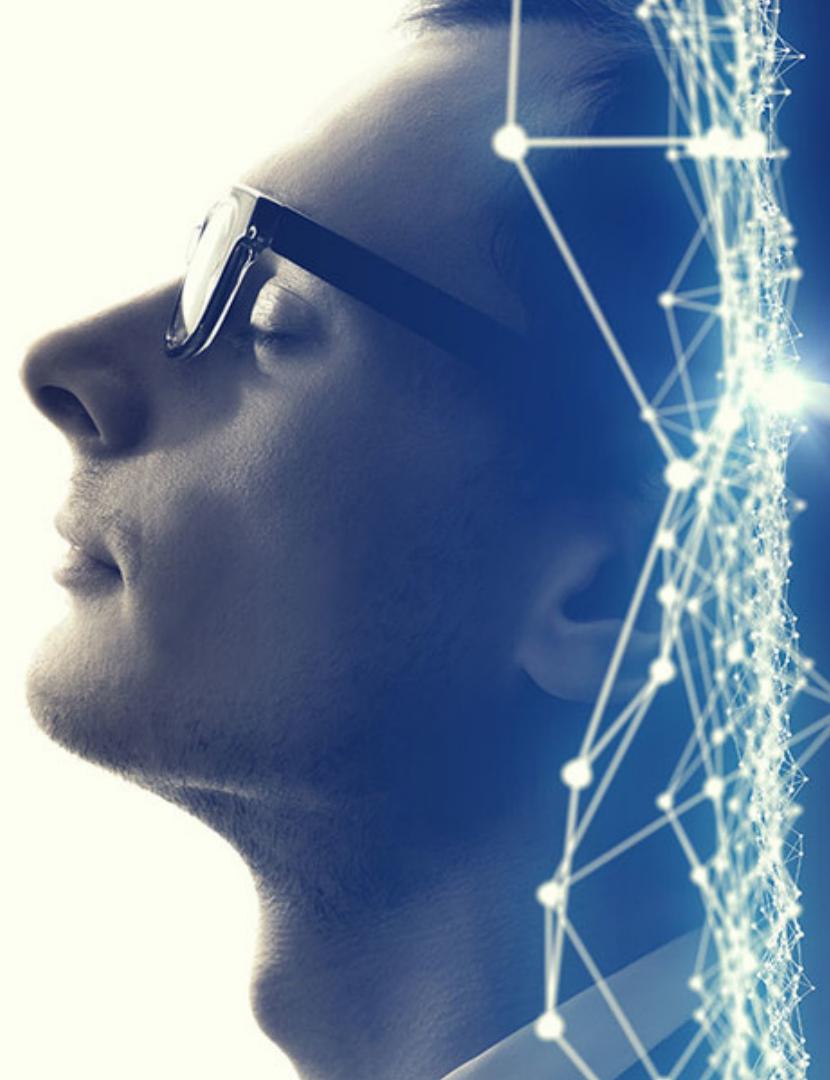
Supplier's Declaration
of Conformity (SDoC)

FAIRNESS EXPLAINABILITY ROBUSTNESS

LINEAGE

AI systems should ...

include details of their development,
deployment, and maintenance so they
can be audited throughout their lifecycle.





Supplier's Declaration of Conformity (SDoC)

FAIRNESS EXPLAINABILITY ROBUSTNESS LINEAGE

Nutrition Facts

Serving Size 2/3 cup (55g)
Servings Per Container About 8

Amount Per Serving

Calories 230 Calories from Fat 40

% Daily Value*

Total Fat 8g **12%**

Saturated Fat 1g **5%**

Trans Fat 0g

Cholesterol 0mg **0%**

Sodium 160mg **7%**

Total Carbohydrate 37g **12%**

Dietary Fiber 4g **16%**

Sugars 1g

Protein 3g

Vitamin A 10%

Vitamin C 8%

Calcium 20%

Iron 45%

* Percent Daily Values are based on a 2,000 calorie diet.
Your daily value may be higher or lower depending on
your calorie needs.

	Calories:	2,000	2,500
Total Fat	Less than	65g	80g
Sat Fat	Less than	20g	25g
Cholesterol	Less than	300mg	300mg
Sodium	Less than	2,400mg	2,400mg
Total Carbohydrate		300g	375g
Dietary Fiber		25g	30g

AI Model Facts

Build Time 1 hour 45 (1000 epochs)

Number of features 56

Determines credit worthiness for loans under \$5000

Scores Approved (A) | Denied (D) | Further Review (FV)

Top feature weightings

Late Payments (engineered) **56%**

R3 – (90 days) **42%**

R2 – (60 days) **32%**

R1 – (30 days) **26%**

Time at residence (engineered) **17%***

* Scored by "Residential Determination" algorithm which creates blended overall score that penalizes multiple moves within 3 years. Algorithm ID# E423. Details bitly.housing-algo.com.

Time at employer (engineered) **13%***

* Scored by "Employer Determination" algorithm which creates blended overall score that penalizes multiple jobs within 2 years. Algorithm ID# E923. Details bitly.job-algo.com.

Current savings **8%**

Current wealth **6%**

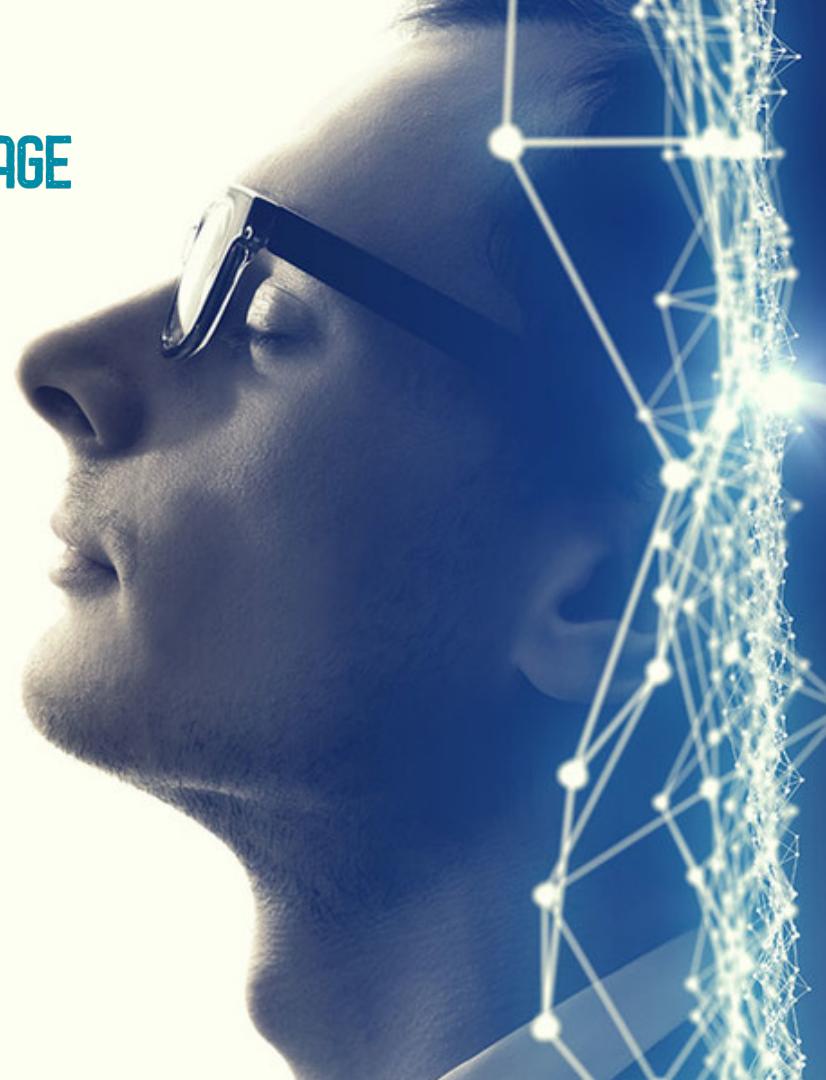
Chief Scientist Rodney Reid

GDPR Compliance Officer Shannon Pickard

Data is source under guidelines of Doc #12233. FICO credit risk score not used per E4233 exemption for low income loan application pool under Rebuild Credit promotion.

Hyperparameters

```
first_neuron:[4, 8, 16, 32, 64], 'hidden_layers':[0, 1, 2], 'batch_size': (1, 5, 5),  
'dropout': (0, 0.2, 0.3), 'weight_regularizer':[None], 'losses':  
[binary_crossentropy], 'optimizer': [Adam, Nadam, RMSprop],  
'activation':[relu, elu], 'last_activation': [sigmoid]
```



DTE: Types of Assets

<http://ibm.biz/dte-live>

The screenshot shows the IBM Planning Analytics landing page. It features a large blue header with the IBM logo and the product name. Below the header, there's a main heading 'IBM Planning Analytics' with the tagline 'Speed, agility and foresight - all-in-one performance management solution'. The page is divided into several sections:

- 5 minute demo:** A video thumbnail showing a dashboard.
- Guided Walkthrough:** A video thumbnail showing a user interface.
- Deep Dive Into IBM Planning Analytics:** A video thumbnail showing a user interface.
- Create a Template with IBM Planning Analytics for Excel and publish it to IBM Planning Analytics:** A section with a form for entering details like Name, Email, Country, Industry, and Number of hours. It also includes a link to 'Learn more about this template'.

At the bottom, there's a copyright notice: 'IBM Cloud / © 2018 IBM Corporation'.

Video Demos

Gives a brief glimpse into the product

- Leverage in expansion conversations
- Address questions on functionality
- Show customers new functionality

Product Tour

Experience the functionality interactively

- Sales training
- Initial client demo (webex)
- Progress the expansion conversation

Hands-on Lab

Full product experience

- Workshops
- Tech sales training
- Proof of technology
- Drive adoption and in turn, consumption