



Simon Aubury- Principal Data Engineer

# Alice has a Blue Car:

Beginning the  
Conversation Around  
Ethically Aware  
Decision Making

BY **ThoughtWorks®**

Product - we're  
an insurance co



Alice wants  
some insurance



# What should Alice's premiums be?

A machine can do this



$$\text{Premium (\$)} = \text{Risk (\%)} \times \text{Value (\$)}$$

- Features
- Algorithms

# UX

## What are our assumptions?

- Consent - meaningful consent
  - Gathering
  - Storage
  - Aligned to 3rd parties
- TOS - did you read it?

Email address

Password

☐ I have read and agree to the [privacy policy](#), [terms of service](#) and [community guidelines](#).

Sign up

# UX - simple example

What are our assumptions?

**Sex**

☐ male

☐ female

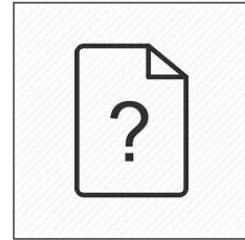
Your responses to these questions will be kept private.

Why do we ask for this info ?

Birth Month and Year

Gender ☐ Male ☐ Female

Country



# UX - simple example

What are our assumptions?



**Application Form**

Please fill in ALL the boxes below in BLOCK CAPITALS using black ink.  
If you miss something out it might delay your application.  
Remember to read the Agreement conditions booklets enclosed and the

**Step 1 – Your personal details**

Title  Surname  Middle name

First name

Are you? Male ☐ Female ☐ Date of birth

# UX - protect the individual

## Randomized Response Technique

**Table 1.** Example of randomized response technique (RRT) application by using Warner's method.

Probability	Assignment by Rolling a Dice <sup>1</sup>	Question	Answer
$\frac{2}{3}$	1-4 → Sensitive Question:	I have an intention to dispose of waste illegally.	<input type="checkbox"/> Yes <input type="checkbox"/> No
$\frac{1}{3}$	5-6 → Inverse Question:	I do not have any intention to dispose of waste illegally.	

<sup>1</sup> Respondents would answer either a sensitive question or the inverse question, according to the outcome of rolling a die.



# Do you drive drunk?

Collect useful insights across a population

## Short-list group

- Have you ever been abroad?
- Have you ever used a taxi?
- Have you been using a plane this week?
- Did you wash your car this week?

## Long-list group

- Have you ever been abroad?
- Have you ever used a taxi?
- Have you been using a plane this week?
- Did you wash your car this week?
- Have you ever been driving a car although you had drunk too much alcohol?

# What data to calculate risk?

What are our assumptions?

Who	Gender	Car	Model	Colour	Email	Postcode	DOB	Risk
Alice	a	Holden	Barina	Blue	<a href="mailto:alice@hotmail.com">alice@hotmail.com</a>	2000	12/09/1990	?
Bob	b	BMW	5s	White	<a href="mailto:bob@gmail.com">bob@gmail.com</a>	2010	8/02/1972	?
Carole	c	Toyota	Carola	Red	<a href="mailto:carole@yahoo.com">carole@yahoo.com</a>	2115	9/05/1998	?
Derek	d	Honda	CRv	Brown	<a href="mailto:derek@aol.com">derek@aol.com</a>	2069	5/03/1985	?

# Dev

## What are our assumptions?

DOB	Star Sign
April	Aries
April	Aries
May	Taurus
June	Gemini
May	Taurus
May	Taurus
April	Aries
May	Taurus
July	Cancer
August	Leo

DOB	Star Sign
April	?
May	?
June	?
July	?
August	?

DOB	Star Sign
April	Aries
May	Taurus
June	Gemini
July	Cancer
August	Leo

# Dev

## What are our assumptions?

The screenshot shows the Google Translate web interface. At the top, the Google logo is on the left, and a grid icon with a red circle containing the number '2' is on the right. Below the logo, the word 'Translate' is displayed in a reddish-brown font, and a link 'Turn off instant translation' is on the right. The main interface has two language selection bars. The left bar shows 'English', 'German', 'Spanish', and 'Malay - detected' with a dropdown arrow. The right bar shows 'German', 'English', and 'Russian' with a dropdown arrow, followed by a blue 'Translate' button. Below these bars, there are two text input areas. The left area contains a list of Malay words: 'dia adalah doktor', 'dia adalah parajurit', 'dia adalah profesor', 'dia adalah pelacur', 'dia adalah jururawat', and 'dia adalah pembantu rumah'. The right area contains a list of English translations: 'he is a doctor', 'he is a soldier', 'he is a professor', 'she is a prostitute', 'she is a nurse', and 'she is a maid'. This list of English translations is enclosed in a red rectangular box. At the bottom left of the input area, there is a small pencil icon. At the bottom right, there is a character count '124/5000'.

Google

Translate

Turn off instant translation

English German Spanish Malay - detected

German English Russian Translate

dia adalah doktor  
dia adalah parajurit  
dia adalah profesor  
dia adalah pelacur  
dia adalah jururawat  
dia adalah pembantu rumah

he is a doctor  
he is a soldier  
he is a professor  
she is a prostitute  
she is a nurse  
she is a maid

124/5000

# Infra & security & storage

What are our assumptions?

Architecting for HIPAA  
Security and Compliance

# Legal - Aust

## What are our assumptions?

### Notifiable Data Breaches scheme

On 22 February 2018, new privacy laws come into effect to regulate the reporting and notification of eligible data breaches to the Office of the Australian Information Commissioner (OAIC) and impacted individuals.

# Legal - UK

What are our assumptions?

## EU Gender Legislation

### Information centre

News

Press releases

About us

Contact us

Our partners

### What is it and what does it mean?

The gender directive contains a requirement across the EU for equal treatment between men and women. Following a European Court of Justice ruling in March 2012, an exemption which previously allowed gender specific pricing for insurance contracts because of statistical differences in risk was removed. As a result, from **21 December 2012**, insurers can no longer consider gender when calculating an insurance quote or any benefits.

# Expectations

Is this fair?

[example@hotmail.com](mailto:example@hotmail.com)

[example@gmail.com](mailto:example@gmail.com)



The screenshot shows the Express website. At the top is the Express logo with the tagline "Home of the Daily and Sunday Express". Below the logo is a navigation bar with links: HOME, NEWS, SHOWBIZ & TV, SPORT, COMMENT, FINANCE. A secondary navigation bar includes: LIFE, STYLE, HEALTH, GARDEN, CARS (highlighted in red), PROPERTY, FOOD, TECH, DIETS. Below the navigation bar is a breadcrumb trail: Home > Life & Style > Cars. The main headline reads: "How your email address could increase your car insurance premium". Below the headline is a sub-headline: "DRIVERS are being warned that their car insurance premium could cost more depending on what type of email they use."

**EXPRESS** Home of the Daily and Sunday Express

HOME NEWS SHOWBIZ & TV SPORT COMMENT FINANCE

LIFE STYLE HEALTH GARDEN **CARS** PROPERTY FOOD TECH DIETS

Home > Life & Style > Cars

## How your email address could increase your car insurance premium

DRIVERS are being warned that their car insurance premium could cost more depending on what type of email they use.



# Expectations

Is this fair?



## MONASH UNIVERSITY ACCIDENT RESEARCH CENTRE REPORT DOCUMENTATION PAGE

Report No.	Date	ISBN	Pages
262	May 2007	0 7326 3232 2	20
<b>Title</b> An Investigation into the Relationship between Vehicle Colour and Crash Risk			

**Author(s):** Stuart Newstead & Angelo D'Elia

**Sponsoring Organisation(s):** This project was funded as contract research by the following organisations: New South Wales Roads and Traffic Authority, NRMA Motoring and Services, Royal Automobile Club of Victoria, Transport Accident Commission, VicRoads.

### **Abstract:**

This study has assessed the relationship between vehicle colour and crash risk through the analysis of real crash outcomes described in mass crash data reported to police in two Australian states. A stratified induced exposure study design was employed identifying vehicle to vehicle crashes and crashes involving unprotected road users as those having a risk dependent on vehicle colour whilst exposure was induced from single vehicle crash involvement. Analysis was stratified by vehicle type, light conditions and jurisdiction of crash.

# Back to our risk algorithm

How's this looking?

$$\text{Premium (\$)} = \text{Risk (\%)} \times \text{Value (\$)}$$

- Features

- ~~Gender~~ if ! EU
- Car Model
- Colour (maybe)
- Postcode
- Customer Name
- ~~Email~~

# What data to calculate risk?

What are our assumptions?

Who	Gender	Car	Model	Colour	Email	Postcode	DOB	Risk
Alice	a <b>X</b>	Holden	Barina	Blue	<a href="mailto:alice@hotmail.com">alice@hotmail.com</a> <b>X</b>	2000	12/09/1990	?
Bob	b	BMW	5s	White	<a href="mailto:bob@gmail.com">bob@gmail.com</a>	2010	8/02/1972	?
Carole	c	Toyota	Carola	Red	<a href="mailto:carole@yahoo.com">carole@yahoo.com</a>	2115	9/05/1998	?
Derek	d	Honda	CRv	Brown	<a href="mailto:derek@aol.com">derek@aol.com</a>	2069	5/03/1985	?

# What data to calculate risk?

What are our assumptions?

Who	Gender	Car	Model	Colour	Email	Postcode	DOB	Risk
Alice	a	Holden	Barina	Blue	<a href="mailto:alice@hotmail.com">alice@hotmail.com</a>	2000	12/09/1990	?
Bob	b	BMW	5s	White	<a href="mailto:bob@gmail.com">bob@gmail.com</a>	2010	8/02/1972	?
Carole	c	Toyota	Carola	Red	<a href="mailto:carole@yahoo.com">carole@yahoo.com</a>	2115	9/05/1998	?
Derek	d	Honda	CRv	Brown	<a href="mailto:derek@aol.com">derek@aol.com</a>	2069	5/03/1985	?

# Alice has a blue car

Even if we don't explicitly set out to capture the data; there's implicit relationships in our feature set

Name	Is it Female?
Carol	Yes
Betsy	Yes
Hazel	Yes
Ike	No
Sandy	Yes
Agnes	Yes
Connie	Yes

