

Analysing social media

A mixed-methods framework combining
computational and qualitative text
analysis

Matthew Andreotta^{1,2}, Robertus Nugroho^{2,3}, Mark Hurlstone³,
Fabio Boschetti⁴, Simon Farrell¹, Iain Walker⁵, & Cecile Paris²

[1] School of Psychological Science,
University of Western Australia

[2] Data61, CSIRO

[3] Soegijapranata Catholic University

[4] Ocean & Atmosphere, CSIRO

[5] School of Psychology and Counselling,
University of Canberra



<https://psyarxiv.com/bynz4/>



matthew.andreotta@research.uwa.edu.au



@MattAndreotta

Public perceptions & social media

- Members of the public differ in their opinions of important policy issues
- Public commentary can influence policy agenda

Public perceptions & social media

- Members of the public differ in their opinions of important policy issues
- Public commentary can influence policy agenda



Public perceptions & social media

- Members of the public differ in their opinions of important policy issues
- Public commentary can influence policy agenda



Large volume

+

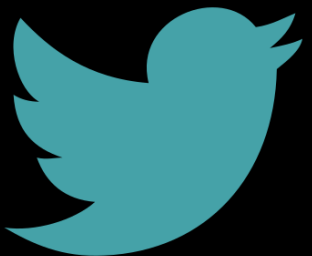
Dynamic content

=

Large and diverse data

Public perceptions & social media

- Members of the public differ in their opinions of important policy issues
- Public commentary can influence policy agenda



Large volume

+

Dynamic content

=

Large and diverse data

AIM

Identify the common commentary of Australian's tweets about climate change

AIM Identify the common commentary of Australian's tweets about climate change

Thematic analysis to identify themes

**DESIRED
ANALYSIS**

AIM Identify the common commentary of Australian's tweets about climate change

Thematic analysis to identify themes

**DESIRED
ANALYSIS**

AIM Identify the common commentary of Australian's tweets about climate change

Thematic analysis to identify themes

**DESIRED
ANALYSIS**

PROBLEM Data set is exceeds the capabilities of traditional qualitative approaches

Framework

1. Compile a corpus

2. Compress corpus along a dimension of relevance

3. Extract a subset of data (from relevant spaces)

4. Perform qualitative analysis subset of data



Malcolm Roberts 🇦🇺

@MRobertsQLD

The most profound change we are going to witness is the unceremonious dumping of climate change scams under a Trump Presidency #auspol

11:25 PM - 8 Nov 2016

1. Australian location
2. Posted in 2016
3. References climate change

201,506 tweets

Framework

1. Compile a corpus

2. Compress corpus along
a dimension of relevance

3. Extract a subset of data
(from relevant spaces)

4. Perform qualitative
analysis subset of data

Framework

1. Compile a corpus

2. Compress corpus along a dimension of relevance

3. Extract a subset of data (from relevant spaces)

4. Perform qualitative analysis subset of data

- **Topics** are an abstract representations of semantically related words and concepts
- **Non-negative Matrix inter-joint Factorization (NMijF)**¹
 - Word co-occurrence
 - Social and temporal co-occurrence
- 41 batches, 5 topics per batch
- **Topic alignment algorithm**² to detect & group similar topics
- 22 groups of similar topics
- 102 topics were rare (similar to ≤ 1 topic)

Framework

1. Compile a corpus

2. Compress corpus along a dimension of relevance

3. Extract a subset of data (from relevant spaces)

4. Perform qualitative analysis subset of data

- **Topics** are an abstract representations of semantically related words and concepts
- **Non-negative Matrix inter-joint Factorization (NMijF)**¹
 - Word co-occurrence
 - Social and temporal co-occurrence
- 41 batches, 5 topics per batch
- **Topic alignment algorithm**² to detect & group similar topics
- 22 groups of similar topics
- 102 topics were rare (similar to ≤ 1 topic)

pari ratify agreement govern time
#parisagr serio storm amend action

Framework

1. Compile a corpus

2. Compress corpus along a dimension of relevance

3. Extract a subset of data (from relevant spaces)

4. Perform qualitative analysis subset of data

- **Topics** are an abstract representations of semantically related words and concepts
- **Non-negative Matrix inter-joint Factorization (NMijF)**¹
 - Word co-occurrence
 - Social and temporal co-occurrence
- 41 batches, 5 topics per batch
- **Topic alignment algorithm**² to detect & group (102 topics here, 5 topics per batch = 1 topic)
- 22 g
- 102

Prime Minister @TurnbullMalcolm, @JulieBishopMP and @JoshFrydenberg announce Australia has ratified Paris climate change agreement #auspol
♡ 3 9:05 AM - Nov 10, 2016

pari ratify agreement govern time
#parisagr serio storm amend action

Framework

1. Compile a corpus

2. Compress corpus along a dimension of relevance

3. Extract a subset of data (from relevant spaces)

4. Perform qualitative analysis subset of data

- **Topics** are an abstract representations of semantically related words and concepts
- **Non-negative Matrix inter-joint Factorization (NMiJF)**¹
 - Word co-occurrence
 - Social and temporal co-occurrence
- 41 batches, 5 topics per batch
- **Topic alignment algorithm**² to detect & group (1 topic)
- 22 g
- 102

Prime Minister @TurnbullMalcolm, @JulieBishopMP and @JoshFrydenberg announce Australia has ratified Paris climate change agreement #auspol
♡ 3 9:05 AM - Nov 10, 2016

pari ratify agreement govern time
#parisagr serio storm amend action

[1] (Nugroho, Zhao, Yang, Paris, & Nepal, 2017)

[2] (Chuang et al., 2015)

Framework

1. Compile a corpus

2. Compress corpus along a dimension of relevance

3. Extract a subset of data (from relevant spaces)

4. Perform qualitative analysis subset of data

- **Topics** are an abstract representations of semantically related words and concepts
- **Non-negative Matrix inter-joint Factorization (NMiJF)**¹
 - Word co-occurrence
 - Social and temporal co-occurrence
- 41 batches, 5 topics per batch
- **Topic alignment algorithm**² to detect & group (1 topic)
- 22 g
- 102

Prime Minister @TurnbullMalcolm, @JulieBishopMP and @JoshFrydenberg announce Australia has ratified Paris climate change agreement #auspol
♡ 3 9:05 AM - Nov 10, 2016

pari ratify agreement govern time
#parisagr serio storm amend action

[1] (Nugroho, Zhao, Yang, Paris, & Nepal, 2017)

[2] (Chuang et al., 2015)

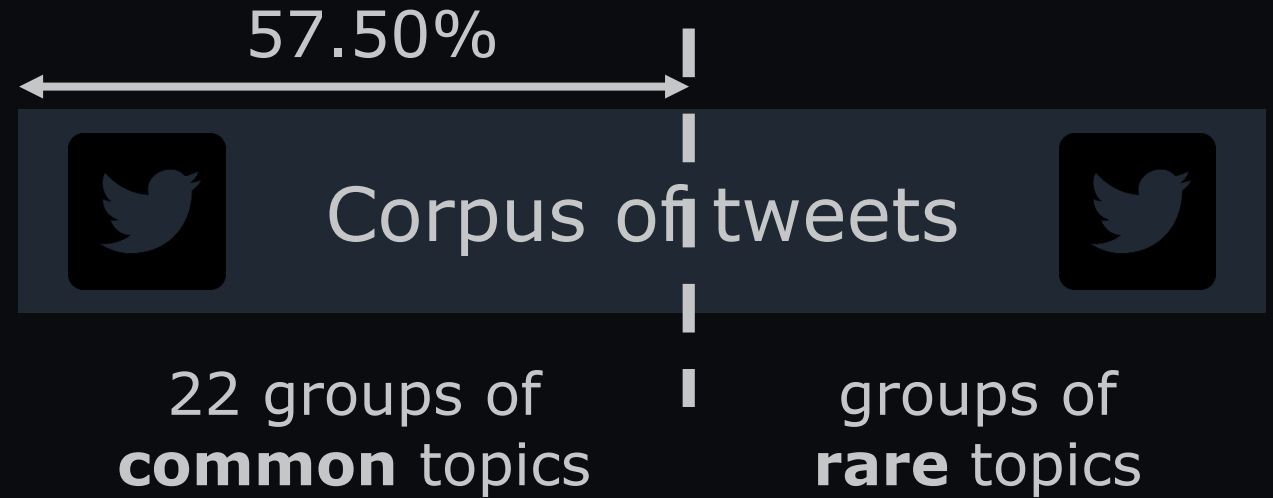
Framework

1. Compile a corpus

2. Compress corpus along
a dimension of relevance

3. Extract a subset of data
(from relevant spaces)

4. Perform qualitative
analysis subset of data



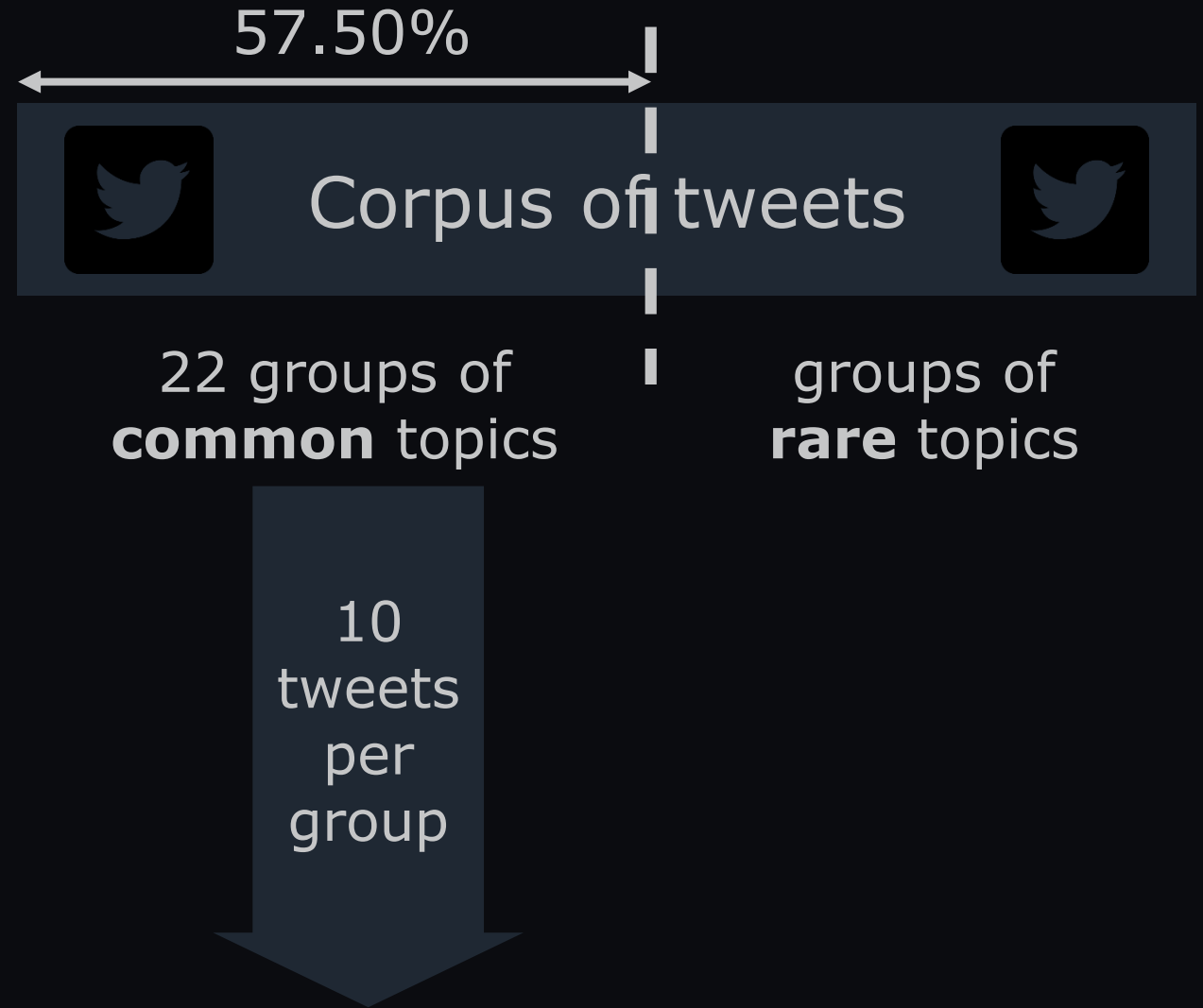
Framework

1. Compile a corpus

2. Compress corpus along
a dimension of relevance

3. Extract a subset of data
(from relevant spaces)

4. Perform qualitative
analysis subset of data



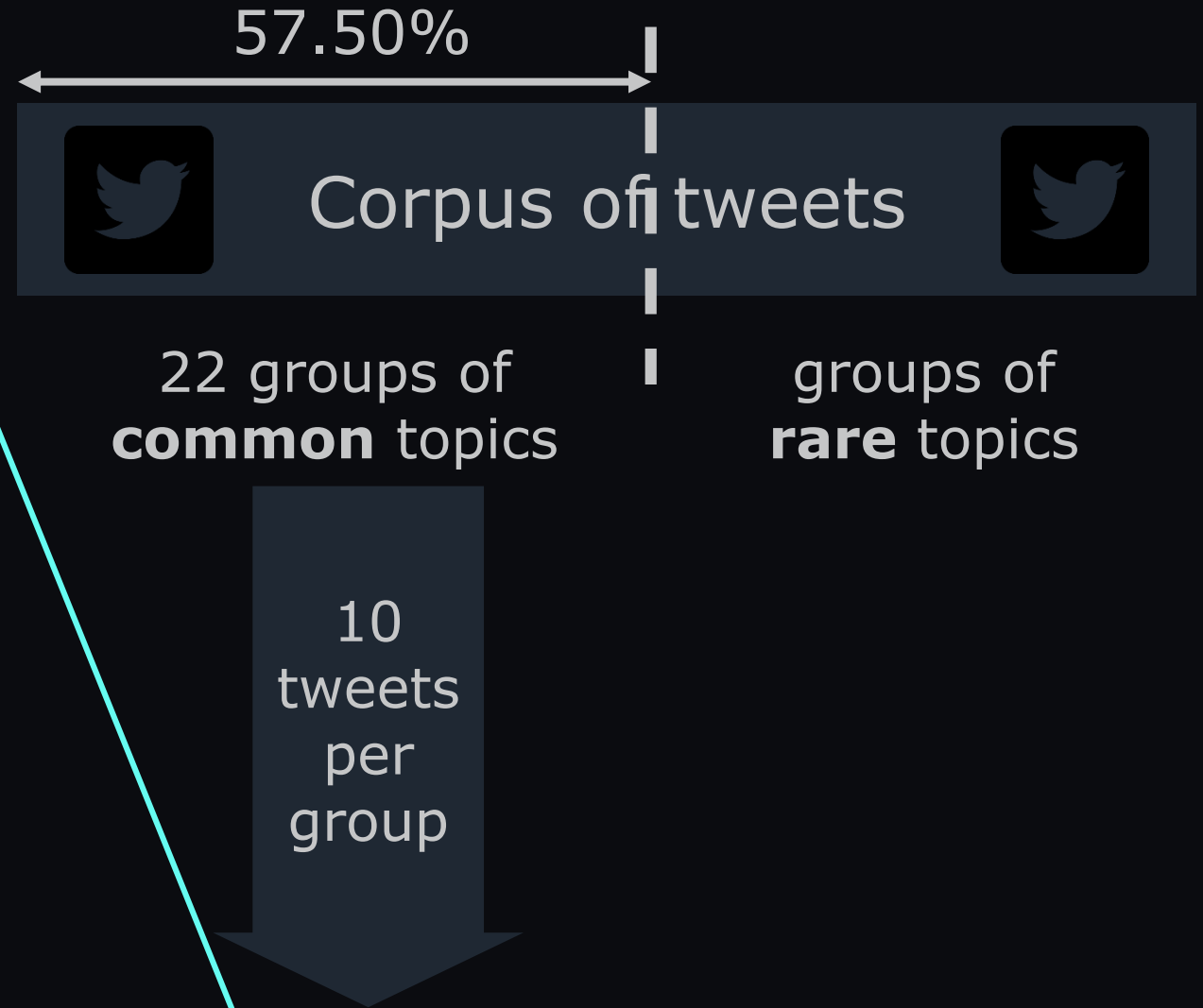
Framework

1. Compile a corpus

2. Compress corpus along
a dimension of relevance

3. Extract a subset of data
(from relevant spaces)

4. Perform qualitative
analysis subset of data



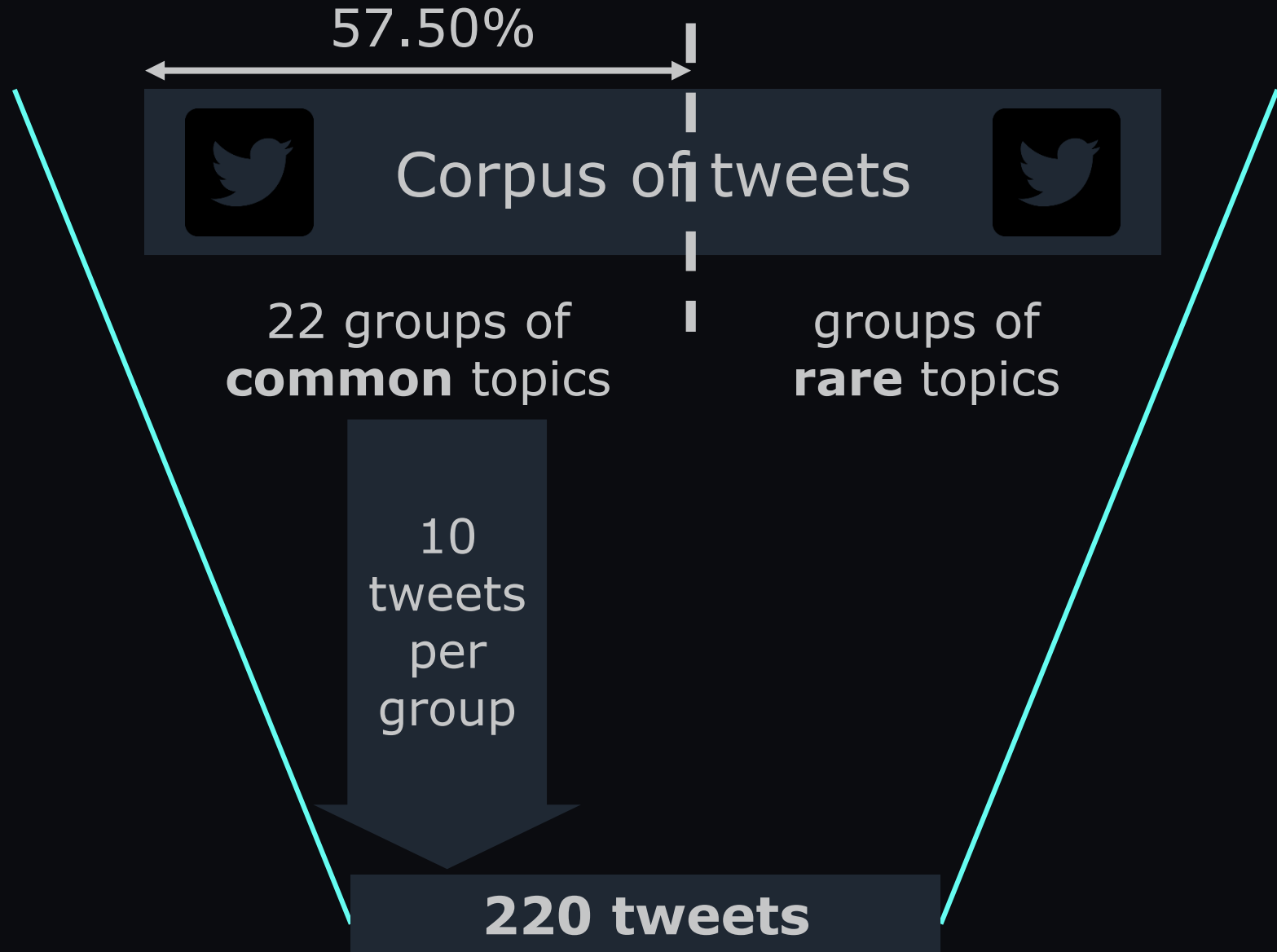
Framework

1. Compile a corpus

2. Compress corpus along
a dimension of relevance

3. Extract a subset of data
(from relevant spaces)

4. Perform qualitative
analysis subset of data



Framework

1. Compile a corpus

2. Compress corpus along
a dimension of relevance

3. Extract a subset of data
(from relevant spaces)

4. Perform qualitative
analysis subset of data



Results

Climate change action

Yes! Let's start working together for real solutions on climate change #QandA

Consequences of climate change

Reefs of the future could look like this if we continue to ignore #climatechange

Conversations on climate change

not so gripping from No Principles Malcolm. Not one mention of climate change in his pitch.

Climate change deniers

According to Senator elect Malcolm Roberts, NASA fiddles the figures on Climate Change.

The legitimacy of climate change and climate science

Do we have an international convention on 'Cloud Seeding'? Or it comes under United Nation's climate change agreement?

Discussion

- Qualitative researcher
 - Handle the diversity and scale of big data
- Quantitative researcher
 - Qualitative interpretations of model output are necessary
 - Framework can supplement models with a qualitative analysis

This research was supported by an Australian Government Research Training Program Scholarship from the University of Western Australia and a scholarship from the CSIRO Research Office and Data61



<https://psyarxiv.com/bynz4/>



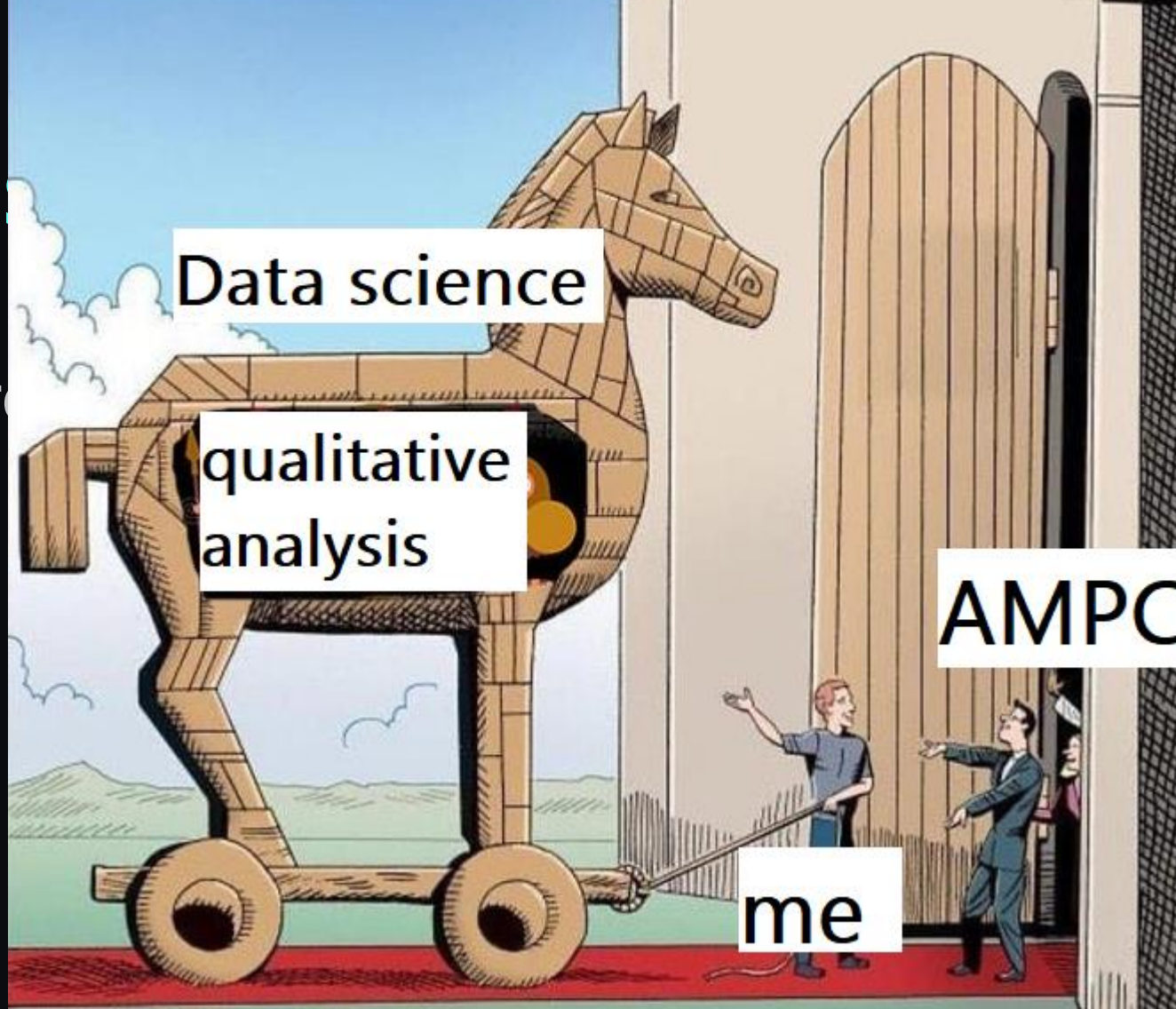
matthew.andreotta@research.uwa.edu.au



[@MattAndreotta](https://twitter.com/MattAndreotta)

Discuss

- Qualitative research
 - Handle the
- Quantitative research
 - Qualitative
 - Framework



sary
e analysis

This research was supported by an Australian Government Research Training Program Scholarship from the University of Western Australia and a scholarship from the CSIRO Research Office and Data61



<https://psyarxiv.com/bynz4/>



matthew.andreotta@research.uwa.edu.au



@MattAndreotta