Wisdom of the crowd

US University of Sussex   DEMOS   CASM Consulting   Ipsos MORI

#WhatAreYouDoingWithMyData: a framework for social media ethics

Steven Ginnis – Head of Digital Research, Social Research Institute, Ipsos MORI

SRA End of Year Conference 2015
14 December 2015
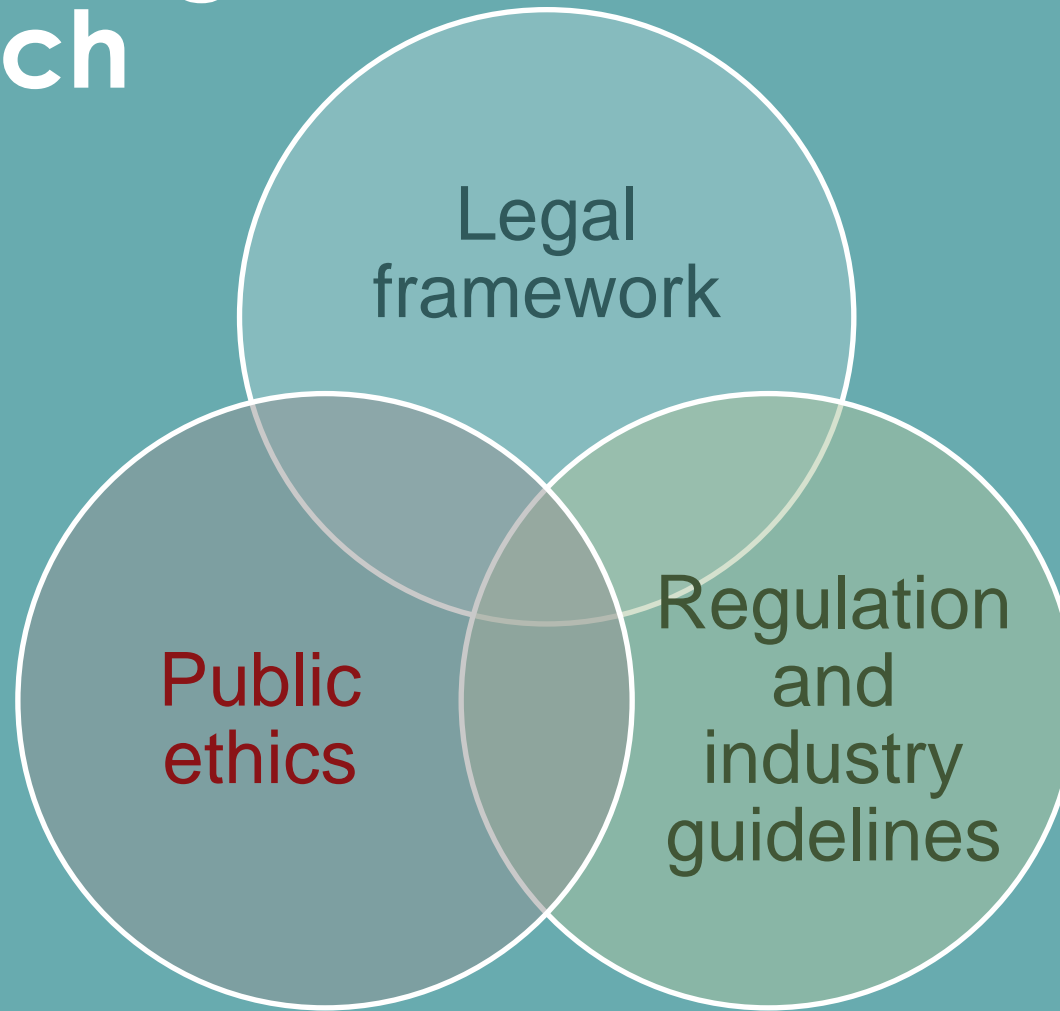
Innovate UK
Technology Strategy Board

Overview

| Current ethics landscape | Public Expectations | Public Acceptability | Guidelines for researchers, industry and platforms |

The building blocks of an ethical approach

Legal framework

Public ethics

Regulation and industry guidelines

How to solve a problem like social media ethics?

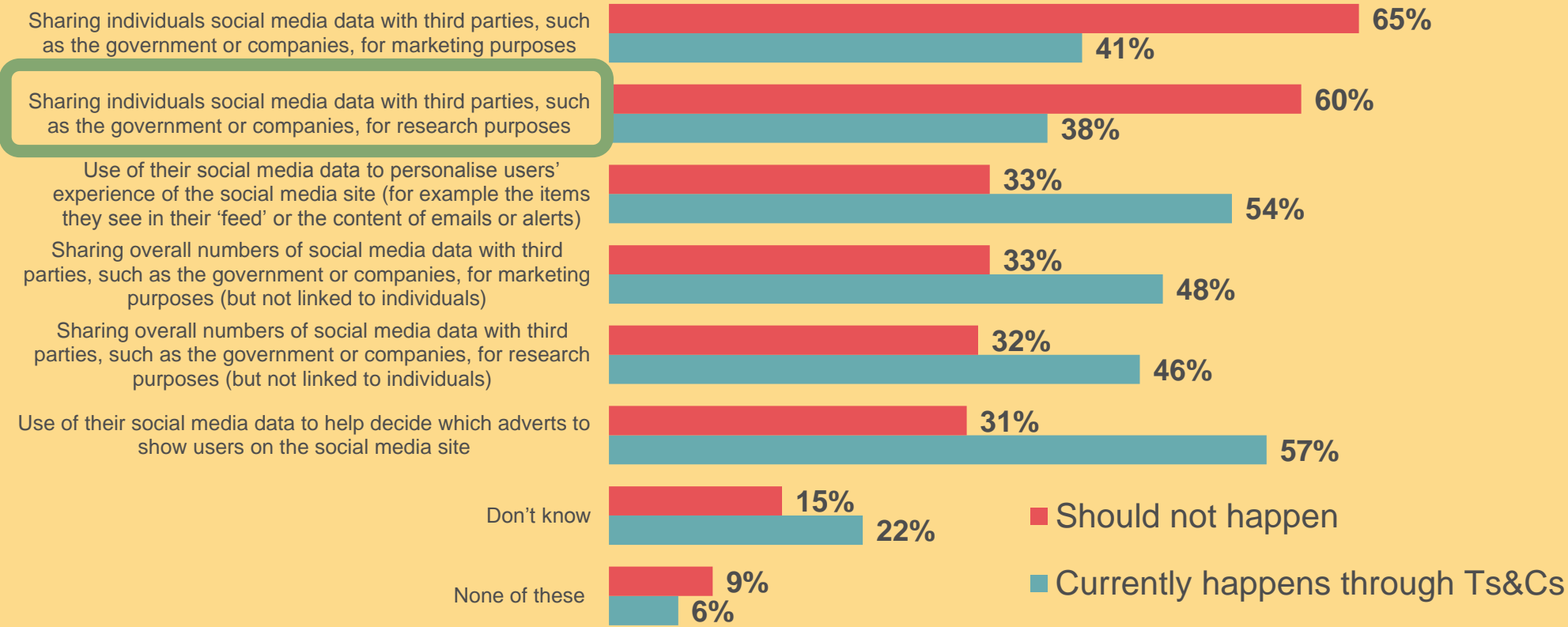|		
Children under the age of 16	**'Private' data**	**Data enrichment and personal sensitive information**
Anonymisation	**Informed consent**	**Key influencers**

Public have low awareness and little desire for current social media research framework through Ts&Cs

Sharing individuals social media data with third parties, such as the government or companies, for marketing purposes
- **65%** Should not happen
- **41%** Currently happens through Ts&Cs

Sharing individuals social media data with third parties, such as the government or companies, for research purposes
- **60%** Should not happen
- **38%** Currently happens through Ts&Cs

Use of their social media data to personalise users' experience of the social media site (for example the items they see in their 'feed' or the content of emails or alerts)
- **33%** Should not happen
- **54%** Currently happens through Ts&Cs

Sharing overall numbers of social media data with third parties, such as the government or companies, for marketing purposes (but not linked to individuals)
- **33%** Should not happen
- **48%** Currently happens through Ts&Cs

Sharing overall numbers of social media data with third parties, such as the government or companies, for research purposes (but not linked to individuals)
- **32%** Should not happen
- **46%** Currently happens through Ts&Cs

Use of their social media data to help decide which adverts to show users on the social media site
- **31%** Should not happen
- **57%** Currently happens through Ts&Cs

Don't know
- **15%** Should not happen
- **22%** Currently happens through Ts&Cs

None of these
- **9%** Should not happen
- **6%** Currently happens through Ts&Cs

- 🟥 **Should not happen**
- 🟦 **Currently happens through Ts&Cs**

Strong preferences for anonymised findings, for all

A. If one of my social media posts was used for research and was selected to be published, **I would want to remain anonymous** so that no one knew it was me

58% 16%

74%

4% 6%

10%
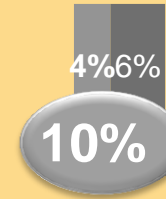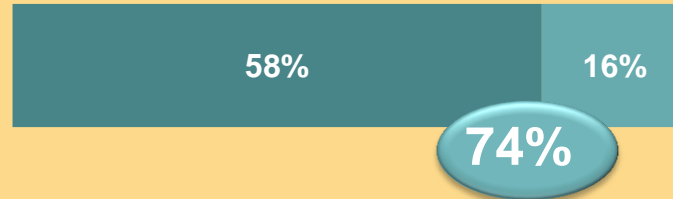
B. If one of my social media posts was used for research and was selected to be published, **I would like the post to be attributed** to me so that people could see what I said

54%

21%

A. All social media accounts should be **given the same rights to anonymity** when used in social media research, regardless of whether the account is held by a public institution, private company or high profile individual
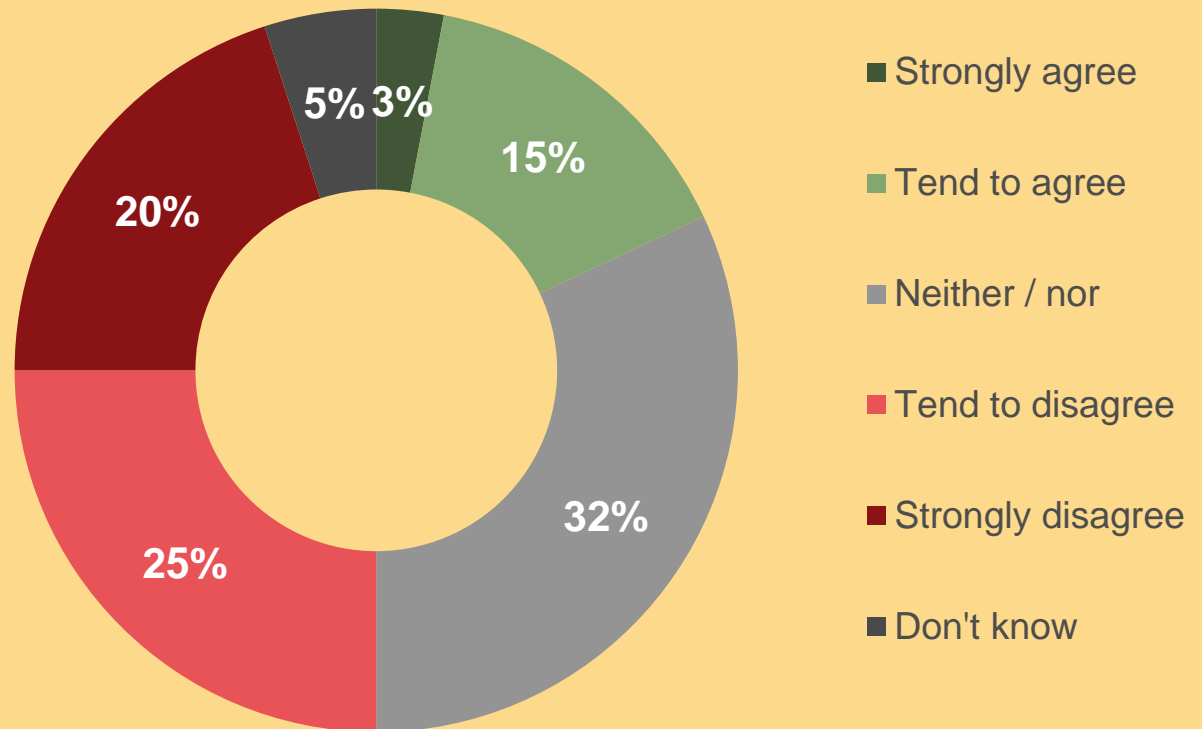
38% 16%

8% 10%
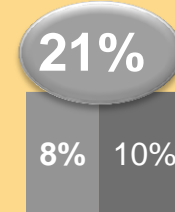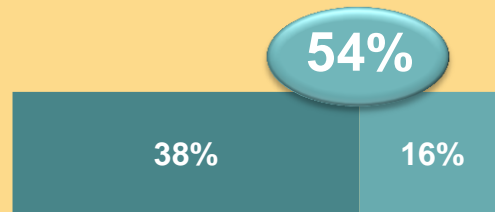
B. Social media accounts held by public institutions, private companies and high profile individuals **should be treated differently** to accounts held by members of the public, social media research involving these accounts should not be anonymous

■ Agree much more with A than B

■ Agree a little more with A than B

■ Agree a little more more with B than A

■ Agree much more with B than A

Uncertainty about new technology that 'derives' personal information

Q. To what extent do you agree or disagree with the following statements:

"It is acceptable for a researcher to use computer programmes to **estimate personal details** about an individual, such as gender or age, from other information such as their name, topics they have posted about and so on"

- ■ Strongly agree
- ■ Tend to agree
- ■ Neither / nor
- ■ Tend to disagree
- ■ Strongly disagree
- ■ Don't know

3%
15%
32%
25%
20%
5%

Q. How likely, if at all, would you be to approve the following research project on a scale of 1 to 10, where 1 is 'would definitely <u>not</u> approve' and 10 is 'definitely <u>would</u> approve'? SCENARIO A

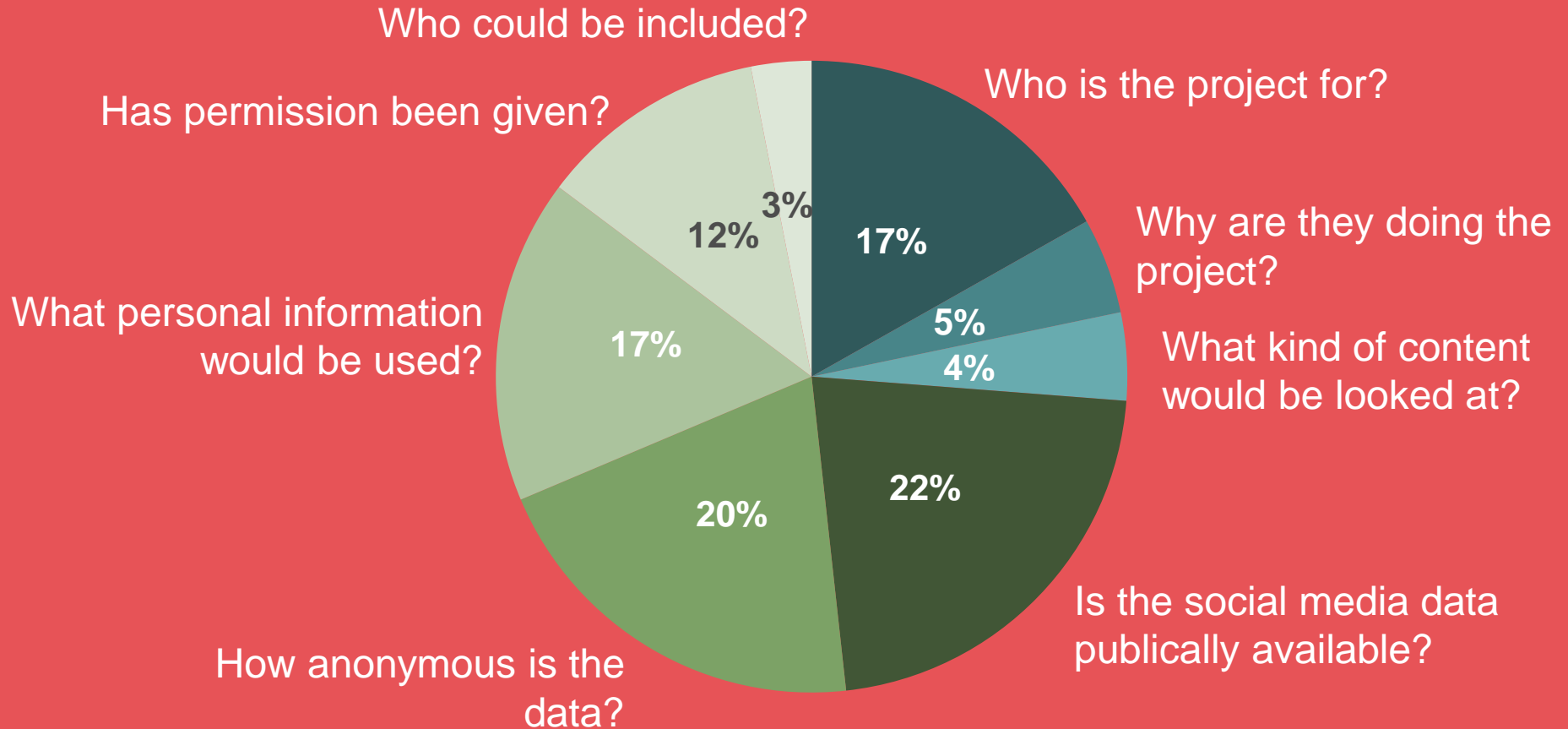|	
Who the project is for?	Researchers in universities and similar organisations
Why are they doing the project?	To review or act on comments about a product or service they deliver
Who could be included?	Anyone on social media who has used a word, hashtag or phrase relevant to the project
Has permission been given?	Only those who have opted in to their data being used for this specific project
Is the social media data publically available?	Collecting posts that have already been made public on a site where anyone can see contributions regardless of whether they have an account
What kind of content would be looked at?	Health behaviours
What personal information would be used?	Age, gender and broad location will be used to compare different groups of people
How anonymous is the data?	The researcher will not see names/locations and IDs during analysis; no posts will be published.

Q. How likely, if at all, would you be to approve the following research project on a scale of 1 to 10, where 1 is 'would definitely <u>not</u> approve' and 10 is 'definitely <u>would</u> approve'? SCENARIO B

Who the project is for?	A private company
Why are they doing the project?	To identify the most active or most well connected social media users in a network
Who could be included?	Anyone on social media who has been identified as visiting a broad location relevant to the project
Has permission been given?	All those who have agreed to the general terms and conditions of the social media site when they first signed up
Is the social media data publically available?	Collecting all types of public and private posts, including private messages between individuals.
What kind of content would be looked at?	Purchasing habits, or information on the products or brands people like
What personal information would be used?	Sensitive personal information (for example, sexuality and political affiliation) which is relevant to the project shall be used
How anonymous is the data?	Individual level posts will be seen by researchers, individual social media posts will be published unedited (including author name and any other details posted by the author)

How do your results compare?

	SCENARIO B	SCENARIO A
AVERAGE SCORE ON 1-10 SCALE	3.55	6.28
% giving 1-4 score of likelihood to approve	64%	23%
% giving 5-6 score of likelihood to approve	21%	28%
% giving 7-10 score of likelihood to approve	15%	50%

Context can be important, but we also need to get the basics right

Who could be included?

Has permission been given?

What personal information would be used?

How anonymous is the data?

3%

12%

17%

17%

5%

4%

22%

20%

Who is the project for?

Why are they doing the project?

What kind of content would be looked at?

Is the social media data publically available?

The devil is in the detail...

Index of relative importance in driving approvability

0.5		
	Posts that have already been made public on a site where anyone can see contributions (0.44)	The researcher will not see names/locations; no posts will be published. (0.41)
0.4	Publically available posts, and posts that can be seen by anyone with an account (0.38)	Only overall numbers are provided by the social media site. No raw data will be seen by the project. (0.37)
0.3	A charity / Researchers in universities (0.33)	
0.2	Public service/local councils (0.25) / All types of public/private posts (0.25)	Individual level posts will be seen by researchers, but posts will not be published. (0.22)
	Government department (0.22)	Individual level posts will be seen by researchers; individual social media posts will be published anonymously (0.17)
0.1		
0.0	A private company (0.0) / All types of public/private posts and private messages (0.0)	Individual level posts will be seen individual social media posts will be published unedited (0.0)

Who the project is for? **Is the social media data publically available?** **How anonymous is the data?**

Recommendations for embedding ethics in social media research

1	2	3	4	5
Boost awareness and build trust		**The option to opt-out**		
6	7	8	9	10
Minimising unnecessary personal data collection		**Removing under 16s from social media research**		
11	12	13	14	15
Permission for publication				
16	17	18	19	
	Defining 'private'		**Establishing an ethics review**	
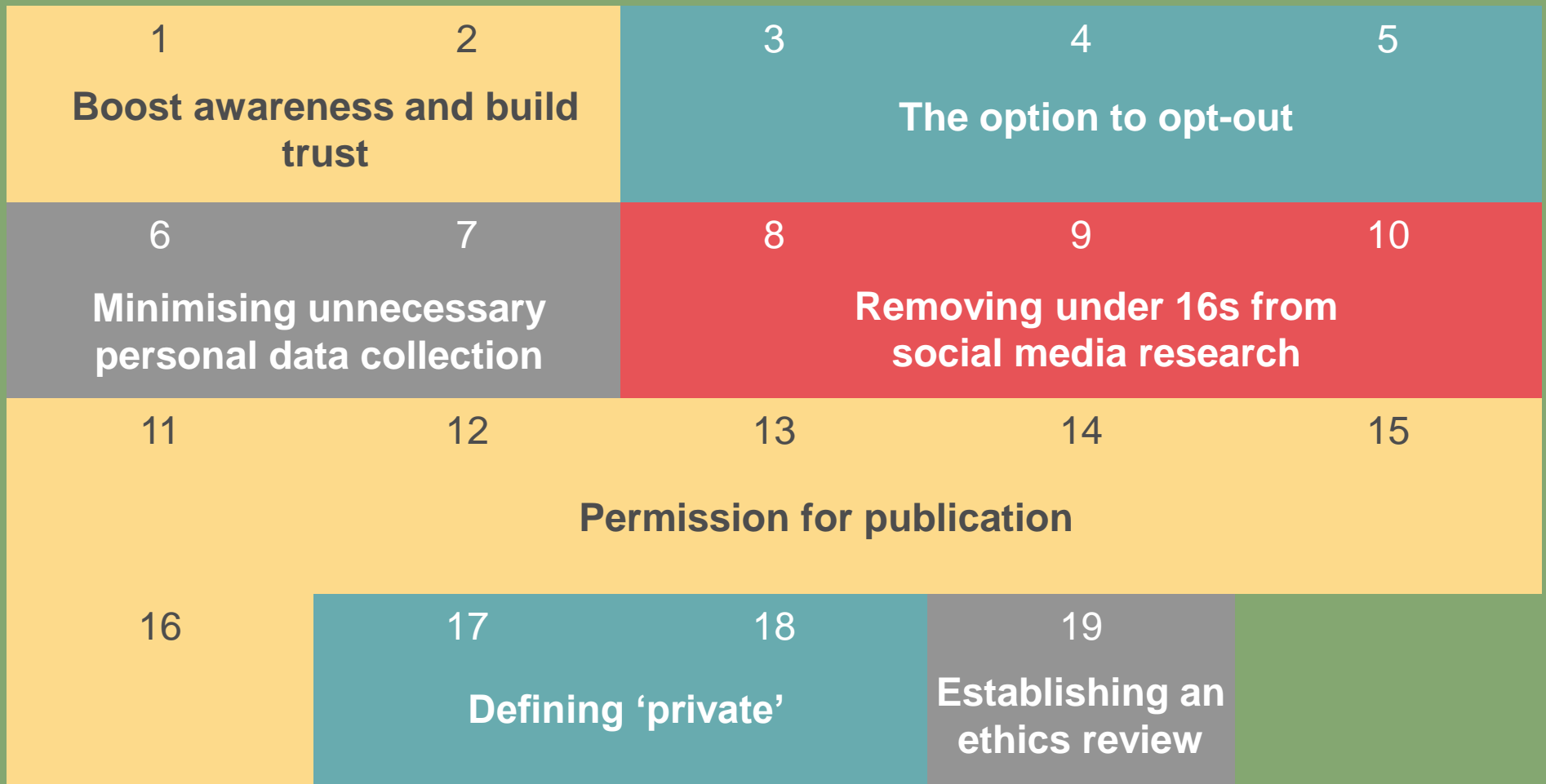
Conclusion: how do we move forward?

We need industry wide action, both to promote best practice but also to implement efficient improvements in ethics – e.g. industry opt out? Or making our case to platforms.

Everything starts with an internal ethics review

We need to keep the momentum, technology, platforms and methods constantly evolving.

For example: web scraping and images

But that doesn't mean individual researchers and companies can't act in the meantime

Challenge our starting assumptions and let ethics design the tech solutions

How do we work internationally to improve standards

UK, market research perspective but we're not the only ones doing social media research

Thank you!

Wisdom of the crowd

steven.ginnis@ipsos.com | 020 7347 3000
www.ipsos-mori.com/wisdomofthecrowd

US University of Sussex   DEM⊙S   CASM Consulting   Ipsos MORI

Innovate UK
Technology Strategy Board